DAUBERT RESPONSE APPENDIX TO DEFENDANT-INTERVENOR AND THE STATE OF WEST VIRGINIA'S JOINT MEMORANDUMS IN RESPONSE TO PLAINTIFF'S MOTIONS TO EXCLUDE EXPERTS' TESTIMONY

TABLE OF CONTENTS:

DAUBERT RESPONSE APPENDIX TO DEFENDANT-INTERVENOR AND THE STATE OF WEST VIRGINIA'S JOINT MEMORANDUMS IN RESPONSE TO PLAINTIFF'S MOTIONS TO EXCLUDE EXPERTS' TESTIMONY

Gregory Allen Brown, Ph.D. FACSM

Wellness Center 221, Cushing Building
Department of Kinesiology & Sport Sciences
University of Nebraska Kearney
1410 W 26th St
Kearney, NE 68849
(308) 865 - 8333
brownga@unk.edu

Academic Preparation

Doctor of Philosophy, Iowa State University. August 2002 -- Major in Health and Human Performance, Emphasis in the Biological Bases of Physical Activity, dissertation title: "Androgenic supplementation in men: Effects of age, herbal extracts, and mode of delivery."

Master of Science, Iowa State University, May 1999 -- Major in Exercise and Sport Science, Emphasis in Exercise Physiology, thesis title: "Oral anabolic-androgenic supplements during resistance training: Effects on glucose tolerance, insulin action, and blood lipids."

Bachelor of Science, Utah State University, June 1997 -- Major in Physical Education, Emphasis in Pre-physical Therapy.

Awards

College of Education Outstanding Faculty Teaching Award. University of Nebraska at Kearney 2019

Mortar Board Faculty Excellence Honors. Xi Phi Chapter, University of Nebraska at Kearney, Honored in 2006, 2007, 2008, 2012, 2013, 2015, and 2019

Profiled in New Frontiers, the University of Nebraska Kearney annual publication highlighting excellence in research, scholarship, and creative activity. 2009, 2017

College of Education Outstanding Scholarship / Research Award. University of Nebraska at Kearney 2009, 2014

College of Education Award for Faculty Mentoring of Undergraduate Student Research University of Nebraska at Kearney, 2007, 2010, & 2013

"Pink Tie" award from the Susan G. Komen Nebraska Affiliate, for outstanding service to the Central Nebraska Race for the Cure, 2013

Star Reviewer for the American Physiological Society and Advances in Physiology Education. 2010.

Fellow of the American College of Sports Medicine. Awarded April 23, 2008

UNK Senior Appreciation Program honoree, the University of Nebraska at Kearney

Iowa State University Research Excellence Award, Iowa State University, 2002

The Zaffarano Prize for Graduate Student Research, Iowa State University, 2002

Helen Hilton Lebaron Excellence in Research Award, Dept. of Health and Human Performance, Iowa State University, 2002

Best Paper Award, 2nd Annual Education Research Exchange. Iowa State University Education Research Exchange, 2001

Helen Hilton Lebaron Excellence in Research Award, Dept. of Health and Human Performance, Iowa State University, 2000

Professional Experience

Professor: University of Nebraska Kearney, Dept. of Kinesiology and Sport Sciences (2012-)

Associate Professor: University of Nebraska Kearney, HPERLS Dept. (2007-2012)

Assistant Professor: University of Nebraska Kearney, HPERLS Dept. (2004- 2007) Full Graduate Faculty status awarded on hire, 2004

Assistant Professor: Georgia Southern University, Jiann-Ping Hsu School of Public Health. (2002-2004) Full Graduate Faculty status awarded Nov. 26, 2002

Laboratory Director: Human Performance Laboratory, Georgia Southern University, Jiann-Ping Hsu School of Public Health. (2002-2004)

Research Assistant: Exercise Biochemistry and Physiology Laboratory, Iowa State University, Department of Health and Human Performance. (1997-2002)

Graduate Teaching Assistant: Iowa State University, Department of Health and Human Performance. (1997-2002)

Temporary Instructor: Iowa State University, Department of Health and Human Performance. (1999-2002)

Temporary Adjunct Faculty: Des Moines Area Community College. (2000)

Undergraduate Teaching Intern: Department of Biology, Utah State University. (1995-1996)

Refereed Publications

1. Schneider KM and Brown GA (as Faculty Mentor). What's at Stake: Is it a Vampire or a Virus? International Journal of Undergraduate Research and Creative Activities. 11, Article 4. 2019.
2. Christner C and Brown GA (as Faculty Mentor). Explaining the Vampire Legend through Disease. UNK Undergraduate Research Journal. 23(1), 2019. *this is an on campus publication
3. Schneekloth B and Brown GA. Comparison of Physical Activity during Zumba with a Human or Video Game Instructor. 11(4):1019-1030. International Journal of Exercise Science, 2018.
4. Bice MR, Hollman A, Bickford S, Bickford N, Ball JW, Wiedenman EM, Brown GA, Dinkel D, and Adkins M. Kinesiology in 360 Degrees. International Journal of Kinesiology in Higher Education, 1: 9-17, 2017

5. Shaw I, Shaw BS, Brown GA, and Shariat A. Review of the Role of Resistance Training and Musculoskeletal Injury Prevention and Rehabilitation. Gavin Journal of Orthopedic Research and Therapy. 1: 5-9, 2016
6. Kahle A, Brown GA, Shaw I, & Shaw BS. Mechanical and Physiological Analysis of Minimalist versus Traditionally Shod Running. J Sports Med Phys Fitness. 56(9):974-9, 2016
7. Bice MR, Carey J, Brown GA, Adkins M, and Ball JW. The Use of Mobile Applications to Enhance Learning of the Skeletal System in Introductory Anatomy & Physiology Students. Int J Kines Higher Educ 27(1) 16-22, 2016
8. Shaw BS, Shaw I, & Brown GA. Resistance Exercise is Medicine. Int J Ther Rehab. 22: 233-237, 2015.
9. Brown GA, Bice MR, Shaw BS, & Shaw I. Online Quizzes Promote Inconsistent Improvements on In-Class Test Performance in Introductory Anatomy & Physiology. Adv. Physiol. Educ. 39: 63-6, 2015
10. Brown GA, Heiserman K, Shaw BS, & Shaw I. Rectus abdominis and rectus femoris muscle activity while performing conventional unweighted and weighted seated abdominal trunk curls. Medicina dello Sport. 68: 9-18. 2015
11. Botha DM, Shaw BS, Shaw I & Brown GA. Role of hyperbaric oxygen therapy in the promotion of cardiopulmonary health and rehabilitation. African Journal for Physical, Health Education, Recreation and Dance (AJPHERD). Supplement 2 (September), 20: 62-73, 2014
12. Abbey BA, Heelan KA, Brown, GA, & Bartee RT. Validity of HydraTrend™ Reagent Strips for the Assessment of Hydration Status. J Strength Cond Res. 28: 2634-9. 2014
13. Scheer KC, Siebrandt SM, Brown GA, Shaw BS, & Shaw I. Wii, Kinect, & Move. Heart Rate, Oxygen Consumption, Energy Expenditure, and Ventilation due to Different Physically Active Video Game Systems in College Students. International Journal of Exercise Science: 7: 22-32, 2014
14. Shaw BS, Shaw I, & Brown GA. Effect of concurrent aerobic and resistive breathing training on respiratory muscle length and spirometry in asthmatics. African Journal for Physical, Health Education, Recreation and Dance (AJPHERD). Supplement 1 (November), 170-183, 2013
15. Adkins M, Brown GA, Heelan K, Ansorge C, Shaw BS & Shaw I. Can dance exergaming contribute to improving physical activity levels in elementary school children? African Journal for Physical, Health Education, Recreation and Dance (AJPHERD). 19: 576-585, 2013
16. Jarvi MB, Brown GA, Shaw BS & Shaw I. Measurements of Heart Rate and Accelerometry to Determine the Physical Activity Level in Boys Playing Paintball. International Journal of Exercise Science: 6: 199-207, 2013
17. Brown GA, Krueger RD, Cook CM, Heelan KA, Shaw BS & Shaw I. A prediction equation for the estimation of cardiorespiratory fitness using an elliptical motion trainer. West Indian Medical Journal. 61: 114-117, 2013.

18. Shaw BS, Shaw I, & Brown GA. Body composition variation following diaphragmatic breathing. African Journal for Physical, Health Education, Recreation and Dance (AJPHERD). 18: 787-794, 2012.
19. Shaw I, Shaw BS, & Brown GA. Concurrent Training and Pulmonary Function in Smokers. Int J Sports Med. 32:776-80, 2011
20. Nienhueser J, Brown, GA, Shaw BS & I Shaw. Effects of Energy Drinks on Metabolism at Rest and During Submaximal Treadmill Exercise in College Age Males. Int J Exerc Sci 4: 321-332, 2011
21. Shaw I, Shaw BS, & Brown GA. Relationship between Resistance Training and Self-Reported Habitual Nutrient Intake. South African Journal for Research in Sport, Physical Education and Recreation. 32: 109-116, 2010
22. Brown GA, Swendener AM, Shaw I, & Shaw BS. Comparison of anthropometric and metabolic responses to a short term carbohydrate restricted diet and exercise versus a traditional diet and exercise. African Journal for Physical, Health Education, Recreation and Dance (AJPHERD). 16: 535-544, 2010
23. Brown GA, Ray M, Abbey BA, Shaw BS, & Shaw I. Oxygen Consumption, Heart Rate and Blood Lactate Responses to an Acute Bout of Plyometric Depth Jumps in College Aged Men and Women. J Strength Cond Res. 24:275-82. 2010
24. Shaw I, Shaw BS, Brown GA, & Cilliers JF. Concurrent Resistance and Aerobic Training as Protection against Heart Disease. Cardiovasc J Afr 21: 196-199, 2010
25. Brown GA, Cook CM, Krueger RD, & Heelan KA Comparison of energy expenditure on a treadmill vs. an elliptical device at a self-selected exercise intensity. J Str Cond Res 24:1643-9, 2010
26. Shaw I, Shaw BS, & Brown GA. Role of Diaphragmatic Breathing and Aerobic Exercise in Improving Maximal Oxygen Consumption in Asthmatics. Science & Sports 25:139-145, 2010
27. Shaw I, Shaw BS, & Brown GA. Comparison of Resistance and Concurrent Resistance and Endurance Training Regimes in the Development of Strength. J Str Cond Res. 23: 2507-2514, 2009
28. Castell LM, Burke LM, Stear SJ, Wolfe RR, Newsholme EA, Trudeau F, Curi R, Brown GA, Vukovich MD, and DS King. BJSM reviews: A–Z of supplements: dietary supplements, sports nutrition foods and ergogenic aids for health and performance Part 2. Br. J. Sports Med. 43:807-810. 2009
29. Shaw BS, Shaw I, & Brown GA. Resistance Training and its Effect on Total, Central and Abdominal Adiposity. South African Journal for Research in Sport, Physical Education and Recreation. 31: 97-108. 2009
30. Shaw I, Shaw BS, & Brown GA. Influence of Strength Training on Cardiac Risk Prevention in Individuals without Cardiovascular Disease. African Journal for Physical, Health Education, Recreation and Dance (AJPHERD). 15: 424-432. 2009

31. Shaw BS, Shaw I, & Brown GA. Resistance Training and Predicted Risk of Coronary Heart Disease in Sedentary Males. African Journal for Physical, Health Education, Recreation and Dance (AJPHERD). Supplement: 247-257. 2009
32. Stahlnecker IV AC, Brown GA, Shaw BS, & Shaw I. Acute Effects of a Weight Loss Supplement on Resting Metabolic Rate and Anaerobic Exercise Performance. African Journal for Physical, Health Education, Recreation and Dance (AJPHERD). Supplement: 237-247. 2009
33. McWha JA, Horst S, Brown GA, Shaw I, & Shaw BS. Metabolic Changes Associated with Playing an Active Video Game Against a Human and Computer Opponent. African Journal for Physical, Health Education, Recreation and Dance (AJPHERD). Supplement: 219-228. 2009
34. Semin K, Stahlnecker IV AC, Heelan KA, Brown GA, Shaw BS, & Shaw I. Discrepancy between Training, Competition and Laboratory Measures of Maximum Heart Rate in NCAA Division 2 Distance Runners. J Sports Sci & Med. 7: 455 – 460, 2008
35. Brown GA, Rebok MP, Scott ML, Harris III J, Colaluca MK, Shaw I, & Shaw BS. Physiological and Biomechanical Responses of Running with and Without a Stroller. African Journal for Physical, Health Education, Recreation and Dance (AJPHERD). 14: 240-249, 2008
36. Brown GA, McFarland SP, Ray MW, Abbey BM, Shaw I, & Shaw BS. A Single Session of Brisk Walking Does Not Alter Blood Glucose Homeostasis in Overweight Young Men. African Journal for Physical, Health Education, Recreation and Dance (AJPHERD). 14: 250-264, 2008
37. Brown GA, Lynott F, & Heelan KA. A Service Learning Model for Teaching Fitness Assessment and Research Techniques to Undergraduate Exercise Science Students. Adv Physiol Educ. 32: 212-218, 2008
38. Carstensen C, Brown GA, Shaw I, & Shaw BS. Freely-Paced Walking in Healthy Adults Does Not Meet Minimum Intensity Guidelines for Health Improvement. African Journal for Physical, Health Education, Recreation and Dance (AJPHERD), 14: 178-187, 2008
39. Shaw BS, Shaw I, and Brown GA. Self-Reported Dietary Intake Following Endurance, Resistance And Concurrent Endurance And Resistance Training. J Sports Sci & Med 7: 255-259, 2008
40. Brown, GA. Teaching skeletal muscle adaptations to aerobic exercise using an APS classic paper by Dr. Philip Gollnick and colleagues. Adv Physiol Educ. 30: 113-118, 2006
41. Brown GA, Vukovich MD, & King DS. Testosterone Prohomone Supplements. Med. Sci. Sports Exerc. Med Sci Sports Exerc. 38: 1451-1461, 2006
42. Brown GA, & MacKenzie D. Resistance Exercise Does Not Change The Hormonal Response To Sublingual Androstenediol. Eur J Appl Physiol. 97:404-412, 2006
43. Brown GA, Vukovich MD, and King DS. Urinary excretion of steroid metabolites following chronic androstenedione ingestion. J. Clin. Endocrinol. Metab. 12:6235 – 6338, 2004

44. Brown GA, Dewey JC, Brunkhorst J, Vukovich MD, & King DS. Changes in serum testosterone and estradiol concentrations following acute androstenedione ingestion in young women. Horm Metab Res. 1:62-66, 2004

45. Kohut ML, Thompson JR, Campbell J, Brown GA, Vukovich MD, Jackson DA, & King DS. Ingestion of a Dietary Supplement Containing Dehydroepiandrosterone (DHEA) and Androstenedione Has Minimal Effect on Immune Function in Middle-Aged Men. J Am Coll Nutr. 22: 363-71, 2003

46. Brown GA, Martini ER, Roberts BS, Vukovich MD, & King DS. Acute hormonal responses to sublingual androstenediol intake in young men. J Appl Physiol. 92: 142-146, 2002.

47. Brown GA, Vukovich MD, Martini ER, Kohut ML, Franke WL, Jackson DA, & King DS. Effects of androstenedione-herbal supplements on serum sex hormone concentrations in 30-59 year old men. Int J Vitam Nutr Res. 71: 293-301, 2001

48. Brown GA, Vukovich MD, Martini ER, Kohut ML, Franke WL, Jackson DA, & King DS. Endocrine and lipid responses to chronic androstenediol-herbal supplementation in 30 to 58 year old men. J Am Coll Nutr. 20: 520-528, 2001.

49. Brown GA, Vukovich MD, Martini ER, Kohut ML, Franke ML, Jackson DA, & King DS. Endocrine response to chronic androstenedione intake in 30-56 year old men. J Clin Endocrinol Metab. 85: 4074-4080, 2000.

50. Brown GA, Vukovich MD, Reifenrath TA, Uhl NL, Parsons KA, Sharp RL, & King DS. Effects of anabolic precursors on serum testosterone concentrations and adaptations to resistance training in young men. Int J Sport Nutr Exerc Metab. 10: 342-362, 2000.

51. Brown GA, Vukovich MD, Sharp RL, Reifenrath TA, Parsons KA, & King DS. Effect of oral DHEA on serum testosterone and adaptations to resistance training in young men. J Appl Physiol. 87: 2274-2283, 1999.

52. King DS, Sharp RL, Vukovich MD, Brown GA, Reifenrath TA, Uhl NL, & Parsons KA. Effect of oral androstenedione on serum testosterone and adaptations to resistance training in young men: a randomized controlled trial. JAMA. 281: 2020-2028, 1999.

Refereed Presentations

1. Brown GA, Jackson B, Szekely B, Schramm T, Shaw BS, Shaw I. A Pre-Workout Supplement Does Not Improve 400 M Sprint Running or Bicycle Wingate Test Performance in Recreationally Trained Individuals. Med Sci Sport Exerc. 50(5), 2932. 65th Annual Meeting of the American College of Sports Medicine. Minneapolis, MN. June 2018.

2. Paulsen SM, Brown GA. Neither Coffee Nor A Stimulant Containing "Pre-workout" Drink Alter Cardiovascular Drift During Walking In Young Men. Med Sci Sport Exerc. 50(5), 2409. 65th Annual Meeting of the American College of Sports Medicine. Minneapolis, MN. June 2018.

3. Adkins M, Bice M, Bickford N, Brown GA. Farm to Fresh! A Multidisciplinary Approach to Teaching Health and Physical Activity. 2018 spring SHAPE America central district conference. Sioux Falls, SD. January 2018.

4. Shaw I, Kinsey JE, Richards R, Shaw BS, and Brown GA. Effect Of Resistance Training During Nebulization In Adults With Cystic Fibrosis. International Journal of Arts & Sciences' (IJAS). International Conference for Physical, Life and Health Sciences which will be held at FHWien University of Applied Sciences of WKW, at Währinger Gürtel 97, Vienna, Austria, from 25-29 June 2017.
5. Bongers M, Abbey BM, Heelan K, Steele JE, Brown GA. Nutrition Education Improves Nutrition Knowledge, Not Dietary Habits In Female Collegiate Distance Runners. Med Sci Sport Exerc. 49(5), 389. 64th Annual Meeting of the American College of Sports Medicine. Denver, CO. May 2017.
6. Brown GA, Steele JE, Shaw I, Shaw BS. Using Elisa to Enhance the Biochemistry Laboratory Experience for Exercise Science Students. Med Sci Sport Exerc. 49(5), 1108. 64th Annual Meeting of the American College of Sports Medicine. Denver, CO. May 2017.
7. Brown GA, Shaw BS, and Shaw I. Effects of a 6 Week Conditioning Program on Jumping, Sprinting, and Agility Performance In Youth. Med Sci Sport Exerc. 48(5), 3730. 63rd Annual Meeting of the American College of Sports Medicine. Boston, MA. June 2016.
8. Shaw I, Shaw BS, Boshoff VE, Coetzee S, and Brown GA. Kinanthropometric Responses To Callisthenic Strength Training In Children. Med Sci Sport Exerc. 48(5), 3221. 63rd Annual Meeting of the American College of Sports Medicine. Boston, MA. June 2016.
9. Shaw BS, Shaw I, Gouveia M, McIntyre S, and Brown GA. Kinanthropometric Responses To Moderate-intensity Resistance Training In Postmenopausal Women. Med Sci Sport Exerc. 48(5), 2127. 63rd Annual Meeting of the American College of Sports Medicine. Boston, MA. June 2016.
10. Bice MR, Cary JD, Brown GA, Adkins M, and Ball JW. The use of mobile applications to enhance introductory anatomy & physiology student performance on topic specific in-class tests. National Association for Kinesiology in Higher Education National Conference. January 8, 2016.
11. Shaw I, Shaw BS, Lawrence KE, Brown GA, and Shariat A. Concurrent Resistance and Aerobic Exercise Training Improves Hemodynamics in Normotensive Overweight and Obese Individuals. Med Sci Sport Exerc. 47(5), 559. 62nd Annual Meeting of the American College of Sports Medicine. San Diego, CA. May 2015.
12. Shaw BS, Shaw I, McCrorie C, Turner S., Schnetler A, and Brown GA. Concurrent Resistance and Aerobic Training in the Prevention of Overweight and Obesity in Young Adults. Med Sci Sport Exerc. 47(5), 223. 62nd Annual Meeting of the American College of Sports Medicine. San Diego, CA. May 2015.
13. Schneekloth B, Shaw I, Shaw BS, and Brown GA. Physical Activity Levels Using Kinect™ Zumba Fitness versus Zumba Fitness with a Human Instructor. Med Sci Sport Exerc. 46(5), 326. 61st Annual Meeting of the American College of Sports Medicine. Orlando, FL. June 2014.
14. Shaw I, Lawrence KE, Shaw BS, and Brown GA. Callisthenic Exercise-related Changes in Body Composition in Overweight and Obese Adults. Med Sci Sport Exerc. 46(5), 394. 61st Annual Meeting of the American College of Sports Medicine. Orlando, FL June 2014.

15. Shaw BS, Shaw I, Fourie M, Gildenhuys M, and Brown GA. Variances In The Body Composition Of Elderly Woman Following Progressive Mat Pilates. Med Sci Sport Exerc. 46(5), 558. 61st Annual Meeting of the American College of Sports Medicine. Orlando, FL June 2014.
16. Brown GA, Shaw I, Shaw BS, and Bice M. Online Quizzes Enhance Introductory Anatomy & Physiology Performance on Subsequent Tests, But Not Examinations. Med Sci Sport Exerc. 46(5), 1655. 61st Annual Meeting of the American College of Sports Medicine. Orlando, FL June 2014.
17. Kahle, A. and Brown, G.A. Electromyography in the Gastrocnemius and Tibialis Anterior, and Oxygen Consumption, Ventilation, and Heart Rate During Minimalist versus Traditionally Shod Running. 27th National Conference on Undergraduate Research (NCUR). La Crosse, Wisconsin USA. April 11-13, 2013
18. Shaw, I., Shaw, B.S., and Brown, G.A. Resistive Breathing Effects on Pulmonary Function, Aerobic Capacity and Medication Usage in Adult Asthmatics Med Sci Sports Exerc 45 (5). S1602 2013. 60th Annual Meeting of the American College of Sports Medicine, Indianapolis, IN USA, May 26-30 3013
19. Shaw, B.S. Gildenhuys, G.A., Fourie, M. Shaw I, and Brown, G.A. Function Changes In The Aged Following Pilates Exercise Training. Med Sci Sports Exerc 45 (5). S1566 60th Annual Meeting of the American College of Sports Medicine, Indianapolis, IN USA, May 26-30 2013
20. Brown, G.A., Abbey, B.M., Ray, M.W., Shaw B.S., & Shaw, I. Changes in Plasma Free Testosterone and Cortisol Concentrations During Plyometric Depth Jumps. Med Sci Sports Exerc 44 (5). S598, 2012. 59th Annual Meeting of the American College of Sports Medicine. May 29 - June 2, 2012; San Francisco, California
21. Shaw, I., Fourie, M., Gildenhuys, G.M., Shaw B.S., & Brown, G.A. Group Pilates Program and Muscular Strength and Endurance Among Elderly Woman. Med Sci Sports Exerc 44 (5). S1426. 59th Annual Meeting of the American College of Sports Medicine. May 29 - June 2, 2012; San Francisco, California
22. Shaw B.S., Shaw, I., & Brown, G.A. Concurrent Inspiratory-Expiratory and Aerobic Training Effects On Respiratory Muscle Strength In Asthmatics. Med Sci Sports Exerc 44 (5). S2163. 59th Annual Meeting of the American College of Sports Medicine. May 29 - June 2, 2012; San Francisco, California
23. Scheer, K., Siebrandt, S., Brown, G.A, Shaw B.S., & Shaw, I. Heart Rate, Oxygen Consumption, and Ventilation due to Different Physically Active Video Game Systems. Med Sci Sports Exerc 44 (5). S1763. 59th Annual Meeting of the American College of Sports Medicine. May 29 - June 2, 2012; San Francisco, California
24. Jarvi M.B., Shaw B.S., Shaw, I., & Brown, G.A. (2012) Paintball Is A Blast, But Is It Exercise? Heart Rate and Accelerometry In Boys Playing Paintball. Med Sci Sports Exerc 44 (5). S3503. 59th Annual Meeting of the American College of Sports Medicine. May 29 - June 2, 2012; San Francisco, California
25. Shaw, I., Shaw, B.S., and Brown G.A. Effort-dependent Pulmonary Variable Improvements Following A Novel Breathing Retraining Technique In Asthmatics. Med Sci Sports Exerc

43 (5). S617, 2011. 58th Annual Meeting of the American College of Sports Medicine. May 31-June 4, 2011 Denver, Colorado

26. Brown G.A. Shaw, B.S., and Shaw, I. Exercise and a Low Carbohydrate Diet Reduce Body Fat but Not PYY and Leptin Concentrations. Med Sci Sports Exerc 43 (5). S4627, 2011. 58th Annual Meeting of the American College of Sports Medicine. May 31-June 4, 2011 Denver, Colorado
27. Shaw, B.S., Shaw, I, and Brown G.A. Pulmonary Function Changes In Response To Combined Aerobic And Resistance Training In Sedentary Male Smokers. Med Sci Sports Exerc 43 (5). S492, 2011. 58th Annual Meeting of the American College of Sports Medicine. May 31-June 4, 2011 Denver, Colorado
28. Heiserman, K., Brown G.A., Shaw, I., and Shaw, B.S. Seated Weighted Abdominal Exercise Activates the Hip Flexors, But Not Abdominals, More Than Unweighted Crunches. A Med Sci Sports Exerc 43 (5). S277, 2011 58th Annual Meeting of the American College of Sports Medicine. May 31-June 4, 2011 Denver, Colorado
29. Brown, G.A., Nienhueser, J., Shaw, I., and Shaw, B.S. Energy Drinks Alter Metabolism at Rest but not During Submaximal Exercise in College Age Males. Med Sci Sports Exerc. 42 (5): S1930. 57th Annual Meeting American College of Sports Medicine, June 1-5, 2010. Baltimore, MD
30. Shaw, I, Shaw, B.S., and Brown G.A. Abdominal and Chest Wall Compliance in Asthmatics: Effects of Different Training Modes. Med Sci Sports Exerc. 42 (5): S1588. 57th Annual Meeting American College of Sports Medicine, June 1-5, 2010. Baltimore, MD.
31. Shaw, B.S., Shaw, I, and Brown G.A. Exercise Effects on Lipoprotein Lipids in the Prevention of Cardiovascular Disease in Sedentary Males Smokers. Med Sci Sports Exerc. 42 (5): S1586. 57th Annual Meeting American College of Sports Medicine, June 1-5, 2010. Baltimore, MD.
32. Brown, G.A. Collaborative Research at a Primarily Undergraduate University. Med Sci Sports Exerc. 42 (5): S424. 57th Annual Meeting American College of Sports Medicine, June 1-5, 2010. Baltimore, MD.
33. Nienhueser, J., Brown, G.A., Effects of Energy Drinks on Resting and Submaximal Metabolism in College Age Males. NCUR 24 (24th National Conference on Undergraduate Research). Missoula, MT. April 15-17, 2010
34. Brown, G.A., N. Dickmeyer, A. Glidden, C. Smith, M. Beckman, B. Malicky, B.S. Shaw and I. Shaw. Relationship of Regional Adipose Tissue Distribution to Fasting Plasma PYY Concentrations in College Aged Females. 56th Annual Meeting American College of Sports Medicine, May 27-30, 2009. Seattle, WA. Med Sci Sports Exerc. 41 (5): S1333
35. Shaw, B.S., I. Shaw, and G.A. Brown. Contrasting Effects Of Exercise On Total And Intra-abdominal Visceral Fat. 56th Annual Meeting American College of Sports Medicine, May 27-30, 2009. Seattle, WA. Med Sci Sports Exerc. 41 (5): S1718
36. Shaw, I., B.S. Shaw, and G.A. Brown. Role of Endurance and Inspiratory Resistive Diaphragmatic Breathing Training In Improving Asthmatic Symptomology. 56th Annual

Meeting American College of Sports Medicine, May 27-30, 2009. Seattle, WA. Med Sci Sports Exerc. 41 (5): S2713

37. McWha, J., S. Horst, G.A. Brown, B.S. Shaw, and I. Shaw. Energy Cost of Physically Active Video Gaming Against a Human or Computer Opponent. 56th Annual Meeting American College of Sports Medicine, May 27-30, 2009. Seattle, WA. Med Sci Sports Exerc. 41 (5): S3069
38. Horst, S., J. McWha, G.A. Brown, B.S. Shaw, and I. Shaw. Salivary Cortisol and Blood Lactate Responses to Physically Active Video Gaming in Young Adults. 56th Annual Meeting American College of Sports Medicine, May 27-30, 2009. Seattle, WA. Med Sci Sports Exerc. 41 (5): S3070
39. Glidden A., M. Beckman, B. Malciky, C. Smith, and G.A. Brown. Peptide YY Levels in Young Women: Correlations with Dietary Macronutrient Intake and Blood Glucose Levels. 55th Annual Meeting American College of Sports Medicine, May 28-31, 2008. Indianapolis, IN. Med Sci Sports Exerc. 40 (5): S741
40. Smith C., Glidden A. M. Beckman, B. Malciky, and G.A. Brown. Peptide YY Levels in Young Women: Correlations with Aerobic Fitness & Resting Metabolic Rate. 55th Annual Meeting American College of Sports Medicine, May 28-31, 2008. Indianapolis, IN. Med Sci Sports Exerc. 40 (5): S742
41. Brown, G.A. M. Holoubeck, B. Nylander, N. Watanabe, P. Janulewicz, M. Costello, K.A. Heelan, and B. Abbey. Energy Costs of Physically Active Video Gaming in Children: Wii Boxing, Wii tennis, and Dance Dance Revolution. 55th Annual Meeting American College of Sports Medicine, May 28-31, 2008. Indianapolis, IN. Med Sci Sports Exerc. 40 (5): S2243
42. McFarland, S.P. and G.A. Brown. One Session of Brisk Walking Does Not Alter Blood Glucose Homeostasis In Overweight Young Men. 53rd annual meeting of the American College of Sports Medicine, Denver, CO. Med Sci Sports Exerc 38: S205, 2006
43. Stahlnecker IV, A.C. and G.A. Brown Acute Effects of a Weight Loss Supplement on Resting Metabolic Rate and Anaerobic Exercise Performance. 53rd annual meeting of the American College of Sports Medicine, Denver, CO. Med Sci Sports Exerc 38: S403, 2006
44. Brown, G.A. and A. Swendener. Effects of Exercise and a Low Carbohydrate Diet on Serum PYY Concentrations 53rd annual meeting of the American College of Sports Medicine, Denver, CO.. Med Sci Sports Exerc 38: s461, 2006
45. Swendener, A.M. and G.A. Brown. Effects of Exercise Combined with a Low Carbohydrate Diet on Health. 53rd annual meeting of the American College of Sports Medicine, Denver, CO. Med Sci Sports Exerc 38: s460, 2006
46. Swendener, A.M. and G.A. Brown. Effects Of Exercise Combined With A Low Carbohydrate Diet On Health. NCUR® 20, 2006
47. Stahlnecker IV, A.C. and G.A. Brown. Acute Effects Of A Weight Loss Supplement On Resting Metabolic Rate And Anaerobic Exercise. NCUR® 20, 2006

48. Eck, L. M. and G.A. Brown. Preliminary Analysis of Physical Fitness Levels in Kinesiology Students. Southern Regional Undergraduate Honors Conference. March 31, 2005.

49. Brown, G.A., J.N. Drouin, and D. MacKenzie. Resistance Exercise Does Not Change The Hormonal Response To Sublingual Androstenediol. 52nd Annual Meeting of the American College of Sports Medicine, June 1-4, 2005, Nashville, TN. Med Sci Sports Exerc 37(5): S40, 2005

50. Brown, G.A., M.P Rebok, M.L. Scott, M.K. Colaluca, and J Harris III. Economy of Jogging Stroller Use During Running. 51st Annual Meeting of the American College of Sports Medicine, June 2-5, 2004, Indianapolis, IN. Med Sci Sports Exerc 36(5): S1714, 2004

51. M.P. Rebok, M.L. Scott, J. Harris III, M.K. Colaluca, and G.A. Brown. Economy of Jogging Stroller use During Running. Georgia Southern University Legislative Wild Game Supper, 2004.

52. M.P. Rebok, M.L. Scott, J. Harris III, M.K. Colaluca, and G.A. Brown. Energy cost of jogging stroller use during running. Annual Meeting of the Southeastern Chapter of the American College of Sports Medicine, 2004.

53. Brown, G.A., Effect of 8 weeks androstenedione supplementation and weight training on glucose tolerance and isokinetic strength. Annual Meeting of the Southeastern Chapter of the American College of Sports Medicine, 2004.

54. Brown, G.A., Vukovich, M.D., Kohut, M.L., Franke, W.D., Jackson, D.A., King, D.S., and Bowers, L.D. Urinary excretion of steroid metabolites following chronic androstenedione ingestion. 50th Annual Meeting of the American College of Sports Medicine, May 27-31 2003, San Francisco, CA. Med Sci Sports Exerc 35(5): S1835

55. Brown, G.A., E.R. Martini, B.S. Roberts, M.D. Vukovich, and D.S. King. Effects of Sublingual androstenediol-cyclodextrin on serum sex hormones in young men. 48th Annual Meeting American College of Sports Medicine, May 30 – June 2, 2001. Baltimore, MD. Med Sci Sports Exerc. 33(5): S1650

56. Kohut, M.L., J.R. Thompson, J. Campbell, G.A. Brown, and D.S. King. Ingestion of a dietary supplement containing androstenedione and dehydroepiandrosterone (DHEA) has a minimal effect on immune response. International Society of Exercise and Immunology, 3rd Annual Convention May 29-30, 2001. Baltimore, MD. Med. Sci. Sports Exerc. 33(5): SISEI12

57. Brown, G.A., E.R. Martini, B.S. Roberts, and D.S. King. Effects of Sublingual androstenediol-cyclodextrin on serum sex hormones in young men. Iowa State University Educational Research Exchange, March 24, 2001. Ames, IA.

58. Martini, E.R., G.A. Brown, M.D. Vukovich, M.L. Kohut, W.D. Franke, D.A. Jackson, and D.S. King. Effects of androstenedione-herbal supplementation on serum sex hormone concentrations in 30-59 year old men. Iowa State University Educational Research Exchange, March 24, 2001. Ames, IA.

59. King, D.S., G.A. Brown, M.D. Vukovich, M.L. Kohut, W.D. Franke, and D.A. Jackson. Effects of Chronic Oral Androstenedione Intake in 30-58 year Old Men. 11th International Conference on the Biochemistry of Exercise. June 4-7, 2000. Little Rock, Arkansas
60. Brown, G.A., M.L. Kohut, W.D. Franke, D. Jackson, M.D. Vukovich, and D.S. King. Serum Hormonal and Lipid Responses to Androgenic supplementation in 30 –59 year old men. 47TH Annual Meeting American College of Sports Medicine, May 31-June 3, 2000. Indianapolis, IN. Med Sci Sports Exerc. 32(5): S486
61. Brown, G.A., T.A. Reifenrath, N.L. Uhl, R.L. Sharp, and D.S. King. Oral anabolic-androgenic supplements during resistance training: Effects on glucose tolerance, insulin action, and blood lipids. 1999 Annual Meeting American College of Sports Medicine, Seattle, WA. Med Sci Sports Exerc. 31(5): S1293
62. Reifenrath, T.A., R.L. Sharp, G.A. Brown, N.L. Uhl, and D.S. King. Oral anabolic-androgenic supplements during resistance training: Effects on body composition and muscle strength. 1999 Annual Meeting American College of Sports Medicine, Seattle, WA. Med Sci Sports Exerc. 31(5): S1292
63. King, D.S., R.L. Sharp, G.A. Brown, T.A. Reifenrath, and N.L. Uhl. Oral anabolic-androgenic supplements during resistance training: Effects on serum testosterone and estrogen concentrations. 1999 Annual Meeting American College of Sports Medicine, Seattle, WA. Med Sci Sports Exerc. 31(5): S1291
64. Parsons, K.A., R.L. Sharp, G.A. Brown, T.A. Reifenrath, N.L. Uhl, and D.S. King. Acute effects of oral anabolic-androgenic supplements on blood androgen and estrogen levels in man. 1999 Annual Meeting American College of Sports Medicine, Seattle, WA. Med Sci Sports Exerc. 31(5): S1290

Book Chapters

Brown, G.A. Chapters on Androstenedione and DHEA. In: Nutritional Supplements in Sport, Exercise and Health an A-Z Guide. edited by Linda M. Castell, Samantha J. Stear, Louise M. Burke. Routledge 2015.

Brown, G.A. Evaluating a Nutritional Supplement with SOAP Notes to Develop Critical Thinking Skills. In: Teaching Critical Thinking and Clinical Reasoning in the Health Sciences, edited by Facione NC and Facione PA. Millbrae, CA: California Academic Press 2008

Non Refereed Publications

Brown, G.A. and King, D.S. Sport Dietary Supplement Update on DHEA supplementation. Human Kinetics Publishers, Inc. October, 2000.

Brown, G.A. Getting in Shape for Paintball in the Winter. Paintball Sports International, January, 1999

Invited Presentations

Brown G.A. Collaborative experiences with researchers in South Africa. Africa Summit 2019 (March 28, 2019). Presented by the University of Nebraska and the University of Nebraska Medical Center.

Peer Reviewer for the Following Journals

Advances in Physiology Education. http://www.the-aps.org/publications/advan/

African Journal For Physical, Health Education, Recreation and Dance (AJPHERD). ISSN: 1117-4315 http://www.ajol.info/journal_index.php?jid=153

Anatomical Sciences Education. http://www.asejournal.com

Asian Journal of Sports Medicine. http://asjsm.tums.ac.ir/index.php/asjsm

CardioVascular Journal of Africa. http://www.cvjsa.co.za/

Complementary Therapies in Medicine. http://ees.elsevier.com/ctim/

European Journal of Sport Science. http://www.tandf.co.uk/journals/titles/17461391.asp

Games for Health Journal. http://www.liebertpub.com/overview/games-for-health-journal/588/

Global Journal of Health and Physical Education Pedagogy. http://js.sagamorepub.com/gjhpep

Interactive Learning Environments. https://www.tandfonline.com/toc/nile20/current

International Journal of Exercise Science. http://digitalcommons.wku.edu/ijes/

Journal of Sports Sciences. http://www.tandf.co.uk/journals/titles/02640414.html

Journal of Strength and Conditioning Research. http://journals.lww.com/nsca-jscr/pages/default.aspx

Lung. http://www.springer.com/medicine/internal/journal/408

Pediatrics. http://pediatrics.aappublications.org/

Scandinavian Journal of Medicine and Science in Sports. http://www.blackwellpublishing.com/journal.asp?ref=0905-7188

South African Journal of Diabetes and Vascular Disease http://www.diabetesjournal.co.za/

The American Journal of Physiology - Endocrinology and Metabolism. http://ajpendo.physiology.org/

The American Journal of Physiology - Heart and Circulatory Physiology. http://ajpheart.physiology.org/

The American Journal of Physiology - Regulatory, Integrative and Comparative Physiology. http://ajpregu.physiology.org/

The International Journal of Sport Nutrition & Exercise Metabolism. http://www.humankinetics.com/IJSNEM/journalAbout.cfm

The Journal of Sports Science and Medicine (JSSM) http://www.jssm.org/

The International Journal of Nutrition and Metabolism ww.academicjournals.org/IJNAM

The Open Sports Sciences Journal. http://benthamscience.com/open/tossj/index.htm

The Journal of Applied Physiology. http://jap.physiology.org/

African Health Sciences. http://www.ajol.info/index.php/ahs

Menopause. http://journals.lww.com/menopausejournal/pages/default.aspx

Membership in Professional Organizations

American College of Sports Medicine

American Physiological Society

National Strength and Conditioning Association

Graduate Student Advisement/Mentoring

Kourtney Woracek. MAEd Thesis Committee. in progress

Marissa Bongers. MAEd Thesis Committee Director. Dietary Habits and Nutrition Knowledge in Female Collegiate Distance Runners. Degree Awarded Spring 2016.

Justin Thiel. MAEd Advisor. Degree Awarded Spring 2016.

Mitchell Sasek. MAEd Advisor. Degree Awarded Summer 2015

Chad Keller. MAEd Advisor. Degree Awarded Summer 2014

Faron Klingehoffer. MAEd Advisor. Degree Awarded Summer 2014

Joe Scharfenkamp. MAEd Internship Advisor. Degree Awarded Summer 2014

Andrew Hudson. MAEd Thesis Committee. Thesis Title. valuation of Weight Loss in Parents Participating in a Pediatric Obesity Treatment Intervention Degree Awarded Fall 2012

Megan Adkins. Doctoral Dissertation Committee. An Examination of Changes in Sedentary Time with the Integration of Technology for Children Participating in a Morning Fitness Program. Degree Awarded Summer 2011

Christopher Campbell. MAEd Advisor. Degree Awarded Spring 2011

Logan Brodine. MAEd Advisor. Degree Awarded Spring 2010

Megan Costello. MAEd Thesis Committee. Changes in the Prevalence of at risk of overweight or overweight in children. Degree Awarded Spring 2009

Pamela Janulewciz, MAEd Thesis Committee. Effects of Exercise Balls as Chair Replacements in a Fourth Grade Classroom. Degree Awarded Spring 2008

Melissa Shelden. MAEd Advisor.

Michael Bell. MAEd Advisor.

Karen DeDonder. MAEd Thesis Committee. Confidence Levels of Certified Athletic Trainers Regarding Female Athlete Triad Syndrome. Degree Awarded Spring 2008

Benjamin Nylander. MAEd Comprehensive Project Director. Degree Awarded Summer 2007

Eme Ferro. MAEd advisor. Degree Awarded Summer 2007

Julie McAlpin. MAEd Thesis Committee. Children Escorted to School; effect on Parental Physical Activity Degree awarded fall 2006

Michael Ray. MAEd Comprehensive Project Director. Degree Awarded Summer 2006

Seth McFarland. MAEd Thesis Committee Director. The Effects of Exercise Duration on Glucose Tolerance and Insulin Sensitivity in Mildly Overweight Men. Degree Awarded Summer 2005

Drew McKenzie. MS Academic Advisor. Degree Awarded Spring 2005

Matthew Luckie. MS Academic Advisor. Degree Awarded Spring 2005

Todd Lane. MS Academic Advisor

Leilani Lowery. MS Internship committee, Degree Awarded Spring 2003

Johnna Ware. MS Internship committee, Degree Awarded Spring 2003

David Bass. MS Internship committee, Degree Awarded Spring 2003

Crystal Smith. MS Internship committee, Degree Awarded Summer 2003

Undergraduate Student Research Mentoring

Cassidy Johnson. Project to be determined. Undergraduate Research Fellowship (Fall 2019 -)

Taylor Wilson. A comparison of High Intensity Interval Exercise on a bicycle ergometer to a treadmill on Resting Metabolic Rate the next day. Undergraduate Research Fellowship (Fall 2018 -)

Dakota Waddell. The effect of yoga versus mindful meditation on stress in physically active and non-physically active female college-aged students Undergraduate Research Fellowship (Fall 2018 -)

Dakota Waddell. A case study of the effects of the *osteostrong* program on bone mineral density and lean body mass in a paraplegic male. Undergraduate Research Fellowship (Fall 2017 – Spring 2018)

Andrew Fields. The effects of retraining running cadence on oxygen consumption in experienced runners. Undergraduate Research Fellowship. (Fall 2017 – Spring 2019)

Logan Engel. The effects of Tart Cherry Juice on Delayed Onset Muscle Soreness following Eccentric Exercise. Undergraduate Research Fellowship. Fall 2017 -

Stephanie Paulsen. Comparing the effects of coffee to a pre-workout drink on cardiovascular drift. Summer Student Research Program. University of Nebraska Kearney. Summer 2017.

Stephanie Paulsen. Comparing the effects of coffee to a pre-workout drink on resting and exercise metabolic rate. Undergraduate Research Fellowship. Spring 2017 - .

Rachael Ernest. Comparing the effects of coffee to a pre-workout drink on resting and exercise metabolic rate. Undergraduate Research Fellowship. Fall 2016 - Spring 2017.

Aleesha Olena. Evaluating the role of body composition on abdominal muscle definition. Undergraduate Research Fellowship. University of Nebraska Kearney. Fall 2016 - Spring 2017.

Marco Escalera. Evaluating the role of body composition on abdominal muscle definition. Undergraduate Research Fellowship. University of Nebraska Kearney. Fall 2015 - Spring 2017.

Trevor Schramm. Effects of "pre-workout' drinks on 400 m sprint performance and salivary cortisol concentrations. Undergraduate Research Fellowship. University of Nebraska Kearney. Spring 2016.

Taylor Turek. Evaluating the role of body composition on abdominal muscle definition. Undergraduate Research Fellowship. University of Nebraska Kearney. Fall 2015 - Spring 2016.

Brian Szekely. Effects of "pre-workout" drinks on Wingate test performance and blood lactate concentrations. Undergraduate Research Fellowship. University of Nebraska Kearney. Fall 2014 - Spring 2016.

Brianna Jackson. Effects of "pre-workout' drinks on 400 m sprint performance and salivary cortisol concentrations. Undergraduate Research Fellowship. University of Nebraska Kearney. Fall 2014 – Fall 2015.

Ashley Pearson. Changes in resting metabolic rate over a semester in undergraduate students. Undergraduate Research Fellowship. University of Nebraska Kearney. Fall 2013 - Spring 2015.

Tricia Young. Changes in resting metabolic rate over a semester in undergraduate students. Undergraduate Research Fellowship. University of Nebraska Kearney. Fall 2013 - Spring 2014.

Gavin Schneider. Effects of "pre-workout" drinks on resistance training performance. Undergraduate Research Fellowship. University of Nebraska Kearney. Fall 2013 - Spring 2014.

Bridgette Schneekloth. Physical Activity while engaging in a Zumba dance class or Microsoft Kinect Zumba. Summer Student Research Program. University of Nebraska Kearney. Summer 2013.

Bridgette Schneekloth. Physical Activity while engaging in Microsoft Kinect Track & Field running vs. free running on an indoor track. Undergraduate Research Fellowship. University of Nebraska Kearney. Fall 2012 - Spring 2014.

Adam Kahle. Evaluating changes in running mechanics with "barefoot" footwear. Summer Student Research Program. University of Nebraska Kearney. Summer 2012

Michelle Jarvi. Quantifying paintball as a form of physical activity in Boys. Undergraduate Research Fellowship. University of Nebraska Kearney. Fall 2011 - Spring 2012.

Benjamin Lentz, Krista Scheer, & Sarah Siebrandt. Wii, Kinect, and Move for Physical Activity: Analysis of Energy Expenditure, Heart Rate, and Ventilation. Undergraduate Research Fellowship. University of Nebraska Kearney. Fall 2010 - Spring 2012.

Katlyn Heiserman. Comparison of EMG activity in the rectus abdominis and rectus femoris during supine un-weighted abdominal crunch exercise and a seated abdominal crunch exercise weight machine. Summer Student Research Program. University of Nebraska Kearney. Summer 2010

Janae Nienhueser. Effects of Energy drink on resting and submaximal exercise metabolism in college age men. Summer Student Research Program. University of Nebraska Kearney. Summer 2009

Jessica McWha. Metabolic changes while playing active video gaming against a human and computer opponent. Summer Student Research Program and Undergraduate Research Fellowship. University of Nebraska Kearney. Summer 2008 – Spring 2009

Sarah Horst. Changes in blood lactate and salivary cortisol concentrations while "exergaming" against a human or computer opponent. Summer Student Research Program. University of Nebraska Kearney. Summer 2008

Craig Carstensen. Differences in the Physiological Response to Treadmill versus Freely Paced Walking. Summer Student Research Program. University of Nebraska Kearney. Summer 2006

Alvah Stahlnecker Acute effects of a weight loss supplement on resting metabolic rate and anaerobic exercise performance. Summer Student Research Program. University of Nebraska Kearney. Summer 2005

Allison Swendener. Effects of exercise combined with a low carbohydrate diet on health. Summer Student Research Program. University of Nebraska Kearney. Summer 2005

Kamilah Whipple. A measurement of the physical activity and fitness of undergraduate Georgia Southern University students. Ronald E. McNair Post-Baccalaureate Achievement Program. Georgia Southern University. Summer 2004.

Lindsey Eck. Preliminary Analysis of Physical Fitness Levels in Kinesiology Students. Independent undergraduate research project. Georgia Southern University. Summer 2004.

Description of Graduate Courses Taught

PE 870: Advanced Exercise Physiology Course presumes a student has had a basic course in exercise physiology. The content of cardiorespiratory fitness, body composition, muscular strength/flexibility, body fluids and metabolism is presented beyond the introductory level. (University of Nebraska at Kearney)

PE 866P: Nutrition for Health and Sport. (Dual listed/taught with PE 469) Metabolism and metabolic regulation, the influence of dietary practices on health and human performance, and mechanisms and consequences of weight loss and gain.. (University of Nebraska Kearney)

PE 861P: Physiology of Exercise. (Dual listed/taught with PE 461) Physiological processes of body as pertain to physical activity. How trained and untrained individuals differ, and importance of training. (University of Nebraska at Kearney)

TE 800: Education Research. This introductory web-based course in educational research focuses on evaluating and interpreting educational research and applying its findings to educational practice. (University of Nebraska at Kearney)

KINS 7230: Exercise Physiology. Focuses on the study of the effects of exercise on the physiological functions of the human organism with emphasis on theoretical orientations. (Georgia Southern University)

KINS 7231: Laboratory Techniques in Exercise Physiology. Acquaints the student with the use of typical laboratory equipment used in exercise physiology. (Georgia Southern University)

KINS 7238: Human Performance and Nutrition. Examines the interaction between nutrition and physical activity, including exercise and athletic performance. (Georgia Southern University)

KINS 7431: Applied Sport Physiology. Focuses on the study of exercise physiology principles applied to developing training and conditioning programs for enhancing health related fitness and performance (Georgia Southern University)

KINS 7899: Directed Independent Study. Provides the student with an opportunity to investigate an area of interest under the direction of faculty mentor (Georgia Southern University)

EXSP 551: Advanced Exercise Physiology 2. Analysis of factors affecting work capacity and performance. Human energy metabolism concepts and measurement. (Iowa State University)

Description of Undergraduate Courses Taught

PE 498: Special Topics. (University of Nebraska at Kearney)

PE 475: Research Methods in Exercise Science. This course is designed to introduce advanced undergraduate students to the processes of research in the field of Exercise Science including the processes of finding, reading and understanding Exercise Science research; data collection; data analysis; and data interpretation. (University of Nebraska at Kearney)

PE 469: Sports Nutrition. Metabolism and metabolic regulation, the influence of dietary practices on human performance. (University of Nebraska at Kearney)

PE 461: Physiology of Exercise. Physiological processes of body as pertain to physical activity. How trained and untrained individuals differ, and importance of training. (University of Nebraska at Kearney)

PE 388: General Studies Capstone - The Living Dead in Fact & Fiction. The Living Dead, such as Zombies and Vampires, are pervasive in fictional literature, television, and movies. During this course, novels, television episodes, and movies will be used to identify disease symptoms displayed by the living dead, and these symptoms will then be evaluated regarding what type of medical condition might cause the symptoms.

PE 310: Introduction to Exercise Physiology. Provides a foundation of scientific basis for understanding the body's anatomical structures and physiologic responses to acute exercise, as well as its adaptations to chronic exercise. (University of Nebraska at Kearney)

PE 107. This course is designed to introduce students to the field of Exercise Science as an area of academic study and as a professional career. Students majoring in Exercise Science should take this course in their first year. (University of Nebraska at Kearney)

KINS 4231: Fitness Evaluation and Exercise Prescription. Provides the student with an in-depth study of fitness appraisal and exercise prescription and the development, interpretation, implementation and management of fitness programs (with laboratory). (Georgia Southern University)

KINS 3133: Physiological Aspects of Exercise. Provides an in-depth perspective of physiological and biochemical responses of the human body when subjected to exercise (with laboratory). (Georgia Southern University)

GSU 1210: University Orientation 1. Designed to help first year students understand the purpose of a college education, learn about college requirements, explore values and interests, learn to make decisions and realistic choices, explore career objectives and programs of study, and establish supportive relationships with faculty and staff. Required of all new students during their first semester. (Georgia Southern University)

EX SP 462: Medical Aspect of Exercise. The role of exercise in preventive medicine. Impact of exercise on various diseases, and the effect of various medical conditions on the ability to participate in vigorous exercise and competitive sports. Principles of exercise testing and prescription for individuals with these conditions. Environmental and nutritional aspects of exercise. (Iowa State University)

EX SP 458: Principles of Exercise Testing and Prescription. Physiological principles of physical fitness; design and administration of fitness programs; testing, evaluation, and prescription; cardiac risk factor modification. (Iowa State University)

EX SP 455 (Renumbered as EX SP 358 for Fall 2001). Physiology of Exercise. Physiological basis of human performance; effects of physical activity on body functions (with laboratory). (Iowa State University)

EX SP 355: Biomechanics (Laboratory). Mechanical basis of human performance; application of mechanical principles to exercise, sport and other physical activities. (Iowa State University)

EX SP 258: Physical Fitness and Conditioning. Development of personal fitness using a variety of conditioning and exercise techniques such as aerobics, weight training, and aquatic fitness. Introduction to acute and chronic responses to exercise, and the role of exercise in health promotion and weight management. (Iowa State University)

EX SP 236: Fundamentals of Archery, Badminton, Bowling (Archery Segment). (Iowa State University)

EX SP 119: Archery 1. (Iowa State University)

EX SP 220: Physical Fitness and Conditioning. Development of personal fitness using a variety of conditioning and exercise techniques such as aerobics, weight training, and aquatic fitness. Introduction to acute and chronic responses to exercise, and the role of exercise in health promotion and weight management. (Des Moines Area Community College)

PE 157: Introduction to Athletic training. Introduction to methods of prevention and immediate care of athletic injuries. Basic information concerning health supervision of athletes, and some basic wrapping and strapping techniques for common injuries. (Des Moines Area Community College)

PE 144: Introduction to Physical Education. History and development of physical education as an academic discipline. Principles and current practices of teaching physical education. (Des Moines Area Community College)

PHYSL 130: Human Physiology. Principles of the regulation and maintenance of human physiology. (Utah State University; Volunteer Undergraduate TA)

PHYSL 103 Human Anatomy. Introduction to the structure and location of bones, muscles, and organs in the human body. (Utah State University; Volunteer Undergraduate TA)

Service

Service to the Profession

Associate Editor, Asian Journal of Sports Medicine (2019-).

Director, North American Chapter, International Physical Activity Projects (IPAP) (2009-)

Fellow, American College of Sports Medicine (2008-_)

National Research Foundation (South Africa) peer evaluator for grant applicants

National Research Foundation (South Africa) evaluator of applications for funding in Thuthuka Programme

External Evaluator for Master's Theses and Doctoral Dissertations, University of Johannesburg, Johannesburg South Africa.

Grant proposal reviewer for NASPE/ING Run for Something Better School Awards Program.

Session Chair. Special Event. Undergraduate Research Experiences in Exercise Science. ACSM Annual Meeting, 2010

Session Chair. 2nd Annual Education Research Exchange. Iowa State University Education Research Exchange, 2001

Current Service at the University of Nebraska at Kearney

University Wide

Faculty Senate Parliamentarian (April 2019 – April 2022)

Faculty Senate Oversight Committee Chair (April 2019 – April 2022)

Faculty Senate Executive Committee (April 2019 – April 2022)

Faculty Senate, At Large representative (Fall 2018-)

University Student Conduct Appeals Board (Fall 2019 - May 2020)

General Studies Council (fall 2013-)

University Safety Committee (Fall 2018 -)

University Student Travel Policy Committee (Fall 2019-)

University Retention Council (Fall 2019 -)

External Evaluator, Promotion Committee, Department of Social Work & Criminal Justice (Fall 2019-)

College of Education Dean Search Committee Member (Fall 2019 -)

College of Education

College of Education Promotion and Tenure Committee, Chair (Fall 2012 – present) Member (fall 2008 – spring 2012)

Department of Kinesiology and Sport Sciences

Kinesiology Lecturer Search Committee Member (Fall 2019 -)

Nebraska Kids Fitness and Nutrition Day, volunteer educator and student coordinator. (fall 2005-present)

Academic Advisor for Undergraduate exercise Science Students (Fall 2005 - present)

Previous Service at the University of Nebraska at Kearney

Recreation Faculty Search Committee Member (Spring 2019)

University Student Conduct Board (Fall 2016- May 2017, Fall 2018 – May 2019)

Faculty Senate Athletic Committee (Fall 2018-May 2019)

External Evaluator, Promotion & Tenure, Department of Social Work & Criminal Justice (Fall 2018)

External Evaluator, Faculty Annual Performance Reviews, Department of Social Work & Criminal Justice (Spring 2018)

University Graduate Council. (Fall 2014 – spring 2017)

University Graduate Council Standing Committee I: Policy & Planning Committee (fall 2014 – spring 2017)

Faculty Senate (April 2012- April 2016)

Faculty Senate Executive Council, (April 2014 – April 2016)

Faculty Senate representative to the Oversight Committee (September 2014 – April 2016)

Faculty Senate representative to the Grievance Committee (September 2014 – April 2016)

Faculty Senate representative to the Professional Conduct committee (September 2013 - April 2016)

Youth Agility Speed & Quickness program director (2011-2015)

Faculty Senate ad-hoc committee on best practices in peer evaluation (2013-2014)

Director of General Studies search committee, committee member (2013-2014)

Director of the Office of Sponsored Programs search committee member (2012-2013; 2013-2014)

College peer mentor for implementing Critical Thinking in the classroom (2013-2014)

Chair, Ad-hoc committee for the evaluation of a new Student Evaluation of Instruction survey (2012-2014 academic years)

Ad-hoc committee to enhance communication effectiveness within department faculty and staff (2013-2014)

Exercise Science faculty search (2012-2013)

Undergraduate Research and Creative Activity program review team (2011-2012)

Institutional Review Board for the protection of Human Research Subjects. (Service period 2006 - 2011)

Undergraduate Research Committee (Service fall 2008 – spring 2011)

University Graduate Council. (Service period 2006 - 2010)

Homecoming Hustle (HPERLS Fun Run) Race Director and Coordinator (Service period beginning Fall 2007 – fall 2009)

Ad-hoc Committee on Enhancing Enrollment and Course Offerings in PE 110 Dept. of HPERLS (Service period beginning fall 2006)

Graduate Council Standing Committee 1: Policy and Planning Committee. (Service period beginning fall 2006; Chair in 2007 – 2008 and 2009-2010)

General Studies Roundtable 2 (spring 2006-spring 2007)

Academic Affairs Committee on Teaching Continuity (Service period beginning fall 2006)

Health Science Program Assistant Director Search Committee, University of Nebraska at Kearney. (Service period summer 2006)

Graduate Program Chair, HPERLS Department, University of Nebraska at Kearney (Service period beginning summer 2006 - 2010)

Graduate Dean Search Committee. University of Nebraska at Kearney (Service period 2005 – 2006 academic year)

Assistant HPERLS Department Graduate Coordinator. (Service period 2005 – 2006 academic year)

University of Nebraska at Kearney Centennial Run committee. (Service period fall 2005)

Senior College of Central Nebraska, Fit after 50 course coordinator. (Service period 2005 – 2006 academic year)

Health Science Program Assistant Advisor Search Committee. (Service period summer 2005)

HPERLS Furniture Committee (Service period spring 2005)

Academic Advisor for Undergraduate exercise Science Students (Service period Beginning Fall 2005 academic year; ongoing)

Other Prior University Service

Institutional Review Board, Georgia Southern University (2003- 2004)

GSU Exercise Science undergraduate student advisor (2002 – 2004)

GSU Jiann-Ping Hsu School of Public Health extramural funding task force (2003-2004)

GSU Jiann-Ping Hsu School of Public Health Curriculum Committee (2003-2004)

GSU Jiann-Ping Hsu School of Public Health Assistant Graduate program director (2003-2004)

GSU Jiann-Ping Hsu School of Public Health Laboratory Director's Committee (2002-2004)

GSU Jiann-Ping Hsu School of Public Health Exercise Science Graduate program coordinator (2003-2004)

GSU Recreation and Athletic Center advisor to the personal training program (2003-2004)

Institutional Biosafety Committee, Georgia Southern University (2003-2004)

Kinesiology Cluster Area, Georgia Southern University, Jiann-Ping Hsu School of Public Health (2002-2004)

Biostatistics Faculty Search Committee. Georgia Southern University, Jiann-Ping Hsu School of Public Health (2002-2003, 2003-2004)

Computer Advisory Committee, Iowa State University, University-Wide, College of Education, and Dept. of Health and Human Performance (2000-2002)

Computer Fee Allocation Committee, Iowa State University (2000-2001)

Dept. of Health and Human Performance Graduate Student Association (Founding Officer and 1st President; 2001-2002)

Sport Management Faculty Search Committee, Iowa State University Dept. of Health and Human Performance (2001-2002)

Previous Community Involvement

Race Director, Central Nebraska Susan G. Komen Race for the Cure (2011, 2012, 2013 events)

Webelos Den Leader, Boy Scouts of America Pack 132, Kearney, NE. Chartered to the Church of Jesus Christ of Latter Day Saints

Scoutmaster, Boy Scouts of America Troop 132, Kearney, NE. Chartered to the Church of Jesus Christ of Latter Day Saints

Tiger Den Coach, Boy Scouts of America Pack 135, Kearney, NE. Chartered to Faith United Methodist Church.

Personal Fitness Merit Badge Counselor. Boy Scouts of America, Overland Trails Council Covered wagon District.

Certifications

American College of Sports Medicine: ACSM Certified Exercise Physiologist (05/21/1998 - 12/31/2021)

USA Track and Field: Level One Coach

American Red Cross: Community First Aid and CPR

Funding

Research Funding

Brown GA, Bice MR, Abbey BM. Shaw I, Shaw BS. Effects of aerobic exercise, resistance exercise, and combined aerobic & resistance exercise on food choices and endocrine signals of satiety in middle aged adults. Submitted 6/26/2017 to National Institutes of Health [PA16-200] - Academic Research Enhancement Award (Parent R15) (Application #1R15DK117436-01). Total Amount Requested: $367,708. (Resubmission of revised proposal; Pending Review.)

Brown GA, Bice MR, Abbey BM. Shaw I, Shaw BS. Effects of aerobic exercise, resistance exercise, and combined aerobic & resistance exercise on food choices and endocrine signals of satiety in middle aged adults. Submitted 6/26/2017 to National Institutes of Health [PA16-200] - Academic Research Enhancement Award (Parent R15) (Application #1R15DK117436-01). Total Amount Requested: $351,708. Pending Review.

Brown GA, Bice MR, Adkins MM, Hollman A, Bickford S, Bickford N, Ranglack D. HEAT it up (Health, Exercise, Aquaponics, Technology) summer camps to grow future health professionals in Rural Nebraska. Submitted 5/25/2017 to National Institutes of Health [PAR17-183] - NICHD Research Education Programs (R25) (Application # 1R25 HD094673-01) Total Amount Requested: $777,006. Pending Review.

Brown GA, Bice MR, Adkins MM, Hollman A, Bickford S, Bickford N, Ranglack D. Teaching Health, Exercise, Technology, & Aquaponics (THETA) Day Camps to Grow Future Health Professionals. University of Nebraska Rural Futures Institutes (RFI) $20,000 – Funded (July 1, 2017 – June 30, 2019)

Brown GA, Bice MR, Adkins MM, Hollman A, Bickford S, Bickford N, Ranglack D. Teaching Health, Exercise, Technology, & Aquaponics (THETA) Day Camps to Grow Future Health Professionals. University of Nebraska Rural Futures Institutes (RFI) and McCook Economic Development Council $11,400 – Funded (May 1, 2017 – August 30, 2017)

Brown GA, Abbey BM, Bice MR. **"**Is milk an effective rehydration beverage during repeated days of dehydrating exercise?" to the Dairy Research Institute® (DRI) $125,560 – Not funded.

Brown GA & Steele J. "Biochemistry Laboratory Experiences for Exercise Science Students" to the Kelly Fund, University of Nebraska. $23,947. Funded. August 2014- June 2016

Brown GA. "Horizon After School Quickness Program" to Blue Cross & Blue Shield of Nebraska for a Community Wellness grant. $14,106. Not funded

Brown GA. "Effects of chocolate milk taken immediately post exercise on the adaptations to strength training in men" to the Dairy Research Institute® (DRI) $123,192 – not funded.

Brown GA., Heelan KA, Bartee RT, & Maughan S. "Active Video Games as an Alternative to Traditional Group Exercise Classes" to the Robert Wood Johnson Health Games Research program. $297,201 – not funded

Brown GA., Nylander B, Heelan KA. Energy Expenditure for Active Video Game Systems: Dance Dance Revolution and Nintendo Wii. University of Nebraska at Kearney Research Services Council. $3,432. Funded

Brown G.A. Effects of green tea extract on fasting plasma insulin, glucose, leptin, and PYY concentrations in humans. University of Nebraska at Kearney Research Services Council. $3,822. Funded

Brown G.A. Dose response relationship between resistance exercise and changes in the hormonal regulation of blood glucose homeostasis. American Diabetes Association Junior faculty Award. $443,293. Not Funded.

Brown G.A., and K. Heelan. Health benefits of green tea extract in women. NIH NCCAM Exploratory/Developmental Grant for Clinical Studies (R21), PAR-03-153. $485,163. Not Funded.

Brown, G.A. Changes In Biomarkers Of Satiety, Aerobic Fitness, And Body Composition While On A Low Fat Or Low Carbohydrate Diet. University of Nebraska at Kearney Research Services Council. $3,750. Funded

Lynott, F., **Brown, G.A**., and K. Heelan. Health and Fitness of HPERLS Students. University of Nebraska at Kearney Research Services Council. $4,000. Funded

Brown G.A., K. Heelan and D.S. King. Pharmacokinetics & Efficacy of Sublingual Androstenediol for Treating Andropause. NIH NCCAM Exploratory/Developmental Grant for Clinical Studies (R21), PAR-03-153. $477,000. Not Funded.

Maughan S.L., D.P.Snider, and **G.A. Brown,** Physical Health and Social Factors Influencing Educational Success Among Hispanic Immigrant Children, University of Nebraska at Kearney Research Services Council. $4,214.60. Funded

McFarland S.P. and **G.A. Brown,** Effects of Exercise Duration on Glucose Tolerance In Mildly Overweight Men, University of Nebraska at Kearney Research Services Council. $750. Funded

Brown, G.A. Effects of Exercise Duration on Insulin Sensitivity In Mildly Overweight Men, University of Nebraska at Kearney Research Services Council. $2,000. Funded

McFarland S.P. and **G.A.Brown,** Effects of Exercise Duration on Glucose Tolerance In Mildly Overweight Men, Gatorade Sports Sciences Institute. $1,500. Not Funded

Brown, G.A. Effects of Exercise Duration on Glucose Tolerance and Insulin Sensitivity in Mildly Overweight Men. Life fitness Academy. $5,000. not funded

Brown, G.A. American College of Sports Medicine Foundation Grant. Endocrinology of weight lifting & androgen supplementation, $10,000. Not Funded.

Brown, G.A. and J.L. McMillan. Experimental and Applied Sciences. Effects of Green Tea Extract on Insulin Sensitivity and Adaptations to Exercise. $71,075. Not Funded.

Brown, G.A. American College of Sports Medicine Foundation Grant. Endocrinology of weight training & androgen supplementation, $10,000. Not Funded.

Brown, G.A. and J. Drouin. Georgia Southern University Faculty Research Grant. Effects of Resistance Training on the Hormonal response to Sublingual Androstenediol Intake. $5,000. Funded

King D.S. and **G.A. Brown.** *World Anti Doping Agency.* Effects of Testosterone Precursors on the Muscular and Hormonal Response to Resistance Training in Men. $464,634. Not Funded.

Brown, G.A. *American College of Sports Medicin*e Foundation Grant. Effect of Raisin Ingestion on Substrate Use During Exercise. $5,000. Not Funded.

King D.S. and **G.A. Brown.** *California Raisin Marketing Board.* The Glycemic Index Of Raisins Fed To Normal People And Non-Insulin Dependent Diabetics. $110,869. Not Funded.

King D.S. and **G.A. Brown.** *California Raisin Marketing Board.* The Effects Of Raisin Ingestion On Substrate Utilization and Endurance Exercise Performance In Trained Cyclists. $ 84,258. Not Funded.

Brown, G.A., E.R. Martini, and B.S. Roberts. Effect of Androstenediol on Serum Sex Hormone Concentrations. Iowa State University Professional Advancement Grant. Graduate Student Senate and Iowa State University Dept. of Health and Human Performance. $700. Funded

Instructional Development Funding

Brown G.A. and K.A. Heelan. University of Nebraska at Kearney. Proposal for the purchase of upgraded resistance exercise equipment in the Human Performance Laboratory. $21,100. Funded.

Brown G.A. and K.A. Heelan. University of Nebraska at Kearney. Proposal for the purchase of a new metabolic cart for the Human Performance Laboratory. $24,560. Funded

Brown, G.A. Georgia Southern University, Center for Excellence in Teaching Instructional Development Grant. Proposal for purchase of heart rate monitors, manual sphygmomanometers, and automated sphygmomanometers. $2,820. Funded.

Brown, G.A. Georgia Southern University, Center for Excellence in Teaching Innovative Teaching Strategies Retreat. Provides $2,000 in instructional technology funds to the participant. Funded.

Brown, G.A. Georgia Southern University, Center for Excellence in Teaching Travel Grant. $750. Funded.

Brown, G.A. Georgia Southern University student technology fee proposal. Proposal for purchase of Molecular Devices SpectraMax 250 plate reader. $17,000. Funded

Brown, G.A. Georgia Southern University student technology fee proposal. Proposal for purchase of Lode Excalibur Sport Bicycle Ergometer and Physiodyne Max 2 Metabolic Cart. $29,577. Funded

Brown, G.A. Georgia Southern University student technology fee proposal. Proposal for purchase of Packard Cobra 2 Automated Gamma Counter. $14,000. Not funded

References

Dr. Ina Shaw
+27 12 671 8810
ina.shaw@momentum.co.za
MMI Client Engagement Solutions
Visiting Professor - University of Johannesburg
Adjunct Professor - University of Venda
President: International Physical Activity Projects (IPAP)

Dr. Kenya Taylor
(308) 865-8843
taylorks@unk.edu
Dean, Graduate Studies & Research
University of Nebraska Kearney

Dr. Matthew R. Bice
(308) 865-8052
bicemr@unk.edu
Assistant Professor, Dept of Kinesiology & Sports Sciences
University of Nebraska Kearney

Search and select a Dictionary term

sex

n.

1. the traits that distinguish between males and females. Sex refers especially to physical and biological traits, whereas gender refers especially to social or cultural traits, although the distinction between the two terms is not regularly observed.

2. the physiological and psychological processes related to procreation and erotic pleasure.

Browse Dictionary

A B C D E F G H I J

K L M N O P Q R S T

U V W X Y Z Ω-#

May 23, 2022
Word of the Day

CAT

1. abbreviation for California Achievement Tests.

2. abbreviation for Children's Apperception Test.

3. abbreviation for the British version of the Cognitive Abilities Test.

4. abbreviation for computer adaptive testing.

5. acronym for computerized axial tomography (see computed tomography).

Also of Interest

Exploring Psychology

Psychology Careers Guide

Psychology in the News

The pursuit of happiness can end in pain

COVID took many in the prime of life, leaving families to pick up the pieces

Teens deal with mental health crisis that's overwhelming doctors

Dating coaches, who specialize in working with people with autism, are in demand

750 First St. NE, Washington, DC 20002-4242
Telephone: (800) 374-2721; (202) 336-5500
TDD/TTY: (202) 336-6123

San Francisco Examiner

FORUM

Opinion: When it comes to trans youth, we're in danger of losing our way

Fueled by isolation and social media, some youth rush toward gender identity as answer to distress

By Erica Anderson Special to The Examiner • January 3, 2022 8:30 am - Updated January 4, 2022 11:38 am

By Erica Anderson

Special to The Examiner

Through a grant from the San Francisco Department of Public Health, UCSF's Child and Adolescent Gender Center has for the past five years provided training and consultation on transgender kids to all youth-serving agencies and professionals in The City, including its public schools. The vision has been to make San Francisco a model in caring for its gender creative youth.

I was part of that effort, and for years worked at UCSF's Gender Center as one of its two psychologists. I provided consultation, taught in the professional schools and wrote about the work. It is well documented that LGBTQ youth are subject to minority stress and higher rates of almost every potential psychological and social problem.

As a trans woman and therapist to trans and gender creative people, I've worked hard to advance acceptance of trans identities, including those of trans youth. But increasingly I'm worried that in our zeal to identify and protect these special children and adolescents, we may have strayed from some core principles and we are in danger of losing our way.

The Latest

Will accidents increase as e-bikes become more popular?

How to Make NFT Art That Will Get Noticed

More than 75% of long COVID patients not hospitalized for initial illness

A fraudster is standing by to book you a flight at 50% off

Muni riders mask up despite inconsistent public transit safety requirements

San Francisco's next D.A.? Recall reform could limit Breed's options if Boudin is ousted

In this extraordinary time during a global pandemic, we have all been subject to extra stress to stay vigilant and avoid COVID and all its variants. Young people have pivoted to remote learning and stayed at home for in many cases more than an entire academic year, depriving them of ordinary social experiences. As a result, most adolescents have also depended upon social media and the internet to an extent never before seen.

We are learning some worrisome things about this massive, unplanned social experiment. Even the tech giants have conceded in their own research that there is a new kind of addiction/attraction to certain content and a kind of contagion among select groups, especially adolescent girls. Increased rates of depression and suicide, declines in dating and sexual activity, more reported loneliness and feelings of being left out, lower rates of involvement in extracurricular activities and surprisingly less sleep all characterize the current generation of adolescents. These trends seem to be accelerating in the era of the smartphone.

There is little question that reliance on screens and devices has isolated adolescents who may be most vulnerable and susceptible to peer and other influences, intensifying their usage of and reliance on whatever messages and images they see. I am concerned that our computer-mediated, always online environment is creating isolated echo chambers that can work on adolescents in an insidious way. And I believe that it's been worse during COVID.

For example, some content on YouTube and TikTok includes "influencers," who themselves are barely out of puberty. They dispense advice to other young people, specifically encouraging them to explore their gender identity freely.

On the one hand, I'm glad our society has evolved toward greater acceptance of all LGBTQ identities. On the other hand, some of the messaging has landed on vulnerable youth searching not just for keys to their own identity but solutions to other psychological and emotional problems, including serious psychiatric problems.

Here is where things may have gone wrong.

Some influencers are literally encouraging the idea that one's psychological distress may be because a young person is trans and is suffering from gender dysphoria. The remedy, they say, is to come out as trans or non-binary, which the influencers advise will alleviate their suffering. Welcomed into the company of other trans and gender creative persons, such young people may have found acceptance — though virtual acceptance, since much of this rapport is online.

They also may be coached on how to navigate and/or control these issues with their parents, who they are told may not "get it." Among the advice from these influencers is to make a quick social and gender transition, which may include a new chosen name and pronoun and access to gender-affirming hormones. Many of these influencers are literally dispensing medical advice.

Increasingly, I am contacted by parents whose child has come out to them as trans in recent weeks. Searching for help, they find me because they want to be affirming. But they

report that the newly asserted gender identity occurred during the pandemic, and they cannot recall any significant suggestion of gender creativity by their child prior to recent events, though many parents report previous psychological problems with their child.

For example, I received a recent inquiry from a San Francisco father whose 14 year old came out as trans in late October after a year of therapy for anxiety and depression. "We were pretty surprised by the news, as we'd had no indication that he had thoughts in this area," the father communicated.

Supportive, open-minded and conscientious parents like these have been contacting me at an accelerating rate in the past year. They also report that their initial contact with therapists leads to affirmation of their child's asserted trans identity and referral to gender clinics. The numbers of new cases at such clinics have exploded. (A recent Gallup poll found that 1 in 6 members of Gen Z identify as lesbian, gay, bisexual, transgender or queer.)

In some cases, well-meaning psychological and medical providers are allowing themselves to be "triangulated," pitting a child's wishes against parents who are reluctant to see their child quickly put on hormone blockers and/or cross-sex hormones. Minors need parental consent for gender-affirming medicines that can pause natal puberty and/or introduce physical changes concordant with the affirmed transgender identity. And minors, especially those between the ages of 12 and 17, often prefer not to heed the advice of their parents.

So instead of forging an alliance between child and parent to evaluate what is needed and drive consensus as the basis for gender affirmation, providers may challenge parents and fuel adolescent rebellion. Of course, virtually all young people need their parents and will for years. Some families experience a rupture from which it may be difficult to recover. In my experience, the vast majority of parents want to support their child whom they love. But they are overwhelmed with shock, grief and legitimate concern for their child's well-being.

In the hundreds of cases I have seen over the past half decade at UCSF and in my private practice, these types of cases are growing. Often by the time I get involved, there has been set up a pitched battle between a youth whose interest is to hurdle toward life-changing decisions with enduring consequences and parents who are bereft and torn between the acceptance and affirmation they want to give their child and their terror about consenting to medical interventions they fear are not right for their child at this time or at all. Ominously, such parents are worried that the child will later regret such decisions and blame parents for allowing it.

With some colleagues, I have been speaking and writing about these concerns. Unfortunately, we find the research on trans youth has not kept up with what is happening. The pandemic has turbo-charged these dynamic trends. Some deny the reality of peer influence upon identity formation. Others decry the methodological approach necessary and consistent with best practices, namely the World Professional Association for Transgender Health Standards of Care and the Endocrine Society and APA guidelines — which encourage an individualized, comprehensive biopsychosocial evaluation prior to initiation of gender-affirming medications and, of course, surgeries.

Research confirming the benefits of gender-affirming psychological and medical care has been done at university-based gender clinics like UCSF in the U.S. and Europe. The gender creative youth served by these clinics, which offer a careful methodical approach with the support of parents and professionals, go on to do well. In the case of a recently disclosed gender identity, the established clinics and best practices have encouraged gender exploratory therapy. This can be a matter of a few weeks to a year or more.

In my over 40 years as a psychologist, I've seen psychotherapeutic phenomena come and go. Eating disorders, multiple personality disorders and repressed memory syndrome have in retrospect spread through subgroups of adolescents and the professionals who have treated them. This spread is like wildfire through vulnerable underbrush, clearly borne in an environment of contagion.

Why is this phenomenon distinctly different from previous ones? How is it possible that gender identity formation constitutes the only area of development in adolescence that is immune from peer influence? Having gone to extraordinary lengths to make San Francisco

the best city in the world for trans acceptance and affirmation, let's not be derailed by zealous disregard for what we see happening in our own back, front and virtual yard.

The COVID pandemic doesn't appear to be going away anytime soon, nor are online dynamics detrimental to young people. So let's make sure that every young person questioning their gender gets what they need, not just what they want.

Erica E. Anderson, Ph.D. is the former president of the United States Professional Association for Transgender Health, former board member of WPATH and is writing a book on the evolution of the science, practice and culture dealing with transgender healthcare; she is based in Berkeley.

Fun, free, cheap: What to do in San Francisco this week
San Francisco Examiner

More than 75% of long COVID patients not hospitalized for initial illness
San Francisco Examiner

RELATED STORIES

Will accidents increase as e-bikes become more popular?

By Christopher Dolan and Nancy Villatoro

By Christopher Dolan and Nancy Villatoro

How to Make NFT Art That Will Get Noticed

In order to get noticed in the world of Non Fungible Films, you can take a look at their creator...

SPONSORED

More than 75% of long COVID patients not hospitalized for initial illness

By Pam Belluck

By Pam Belluck The New York Times

Journal of Athletic Training 2014;49(4):493–506
doi: 10.4085/1062-6050-49.2.19

www.natajournals.org

original research

Middle School Injuries: A 20-Year (1988–2008) Multisport Evaluation

Glenn Beachy, MS, ATC*; Mitchell Rauh, PhD, MPH, PT, FACSM†

*Punahou School, Honolulu, HI; †San Diego State University, CA

Context: Data on the incidence of injury in middle school sports are limited.

Objective: To describe overall, practice, and game injury rate patterns in 29 middle school sports.

Design: Descriptive epidemiology study.

Setting: Injury data collected over a 20-year period (1988–2008) at a single school.

Patients or Other Participants: Boy (n = 8078) and girl (n = 5960) athletes participating in 14 and 15 middle school sports, respectively.

Main Outcome Measure(s): Injury status and athlete-exposures (AEs) were collected by certified athletic trainers. Incidence rates per 1000 AEs (injuries/AEs) were calculated for overall incidence, practices and games, injury location, injury type, and injury severity (time lost from participation). Rate ratios (RRs) and 95% confidence intervals (CIs) were used to compare injury rates for sex-matched sports.

Results: Football had the highest injury rate for all injuries (16.03/1000 AEs) and for time-loss injuries (8.486/1000 AEs). In matched middle school sports, girls exhibited a higher injury rate for all injuries (7.686/1000 AEs, RR = 1.15, 95% CI = 1.1, 1.2) and time-loss injuries (2.944/1000 AEs, RR = 1.09, 95% CI = 1.0, 1.2) than boys (all injuries: 6.684/1000 AEs, time-loss injuries: 2.702/1000 AEs). Girls had a higher injury rate during practices (3.30/1000 AEs) than games (1.67/1000 AEs, RR = 1.97, 95% CI = 1.7, 2.4) for all sports. Only gymnastics (RR = 0.96, 95% CI = 0.3, 3.8) had a higher game injury rate for girls. Practice and game injury rates were nearly identical for boys in all sports (RR = 0.99, 95% CI = 0.9, 1.1). Only football (RR = 0.49, 95% CI = 0.4, 0.6) and boys' wrestling (RR = 0.50, 95% CI = 0.3, 0.8) reported higher game injury rates. Tendinitis injuries accounted for 19.1% of all middle school injuries.

Conclusions: The risk for sport-related injury at the middle school level was greater during practices than games and greater for girls than boys in sex-matched sports. Conditioning programs may be needed to address the high rate of tendinitis injuries.

Key Words: epidemiology, sports, athletes, adolescents

Key Points

- Injuries to middle school athletes were less frequent and less severe than those reported for secondary school and collegiate athletes.
- Injuries were more often sustained in practices than in games.
- In sex-matched sports, girls had a higher injury rate than boys. Football had the highest overall injury rate.

Understanding the incidence and risk factors for adolescent athletic injuries can lead to the implementation of injury-prevention strategies.[1] The National Federation of High Schools Association has documented participation data for high school athletics since 1971–1972, with more than 7.6 million athletes participating in secondary school athletics nationally during the 2009–2010 school year.[2] The risk of athletic injury at the high school level has become increasingly well documented.[3–11] The sport-related injury risk is also problematic at ages below the high school level. The National Center for Sports Safety has reported that more than 3.5 million children ages 14 and under receive medical treatment for sport injuries annually.[12] The observed increased popularity of boys' and girls' youth soccer participation has also resulted in a corresponding increase in injuries.[13–17] The actual participation and injury rates for middle school athletes are less clear. Although McEwin and Swain[18] stated that 96% of middle schools in the United States participated in interscholastic competition, participation data were not provided. The few studies of the incidence of injury among middle schoolers have focused on specific sports[1,19,20] and contact injuries among boy athletes only.[19,20]

Despite the increased reports of injury risk at the secondary school level,[3–7] a universally accepted injury definition is still lacking. Other limitations of prior studies include short-term injury tracking (eg, 1 to 2 years) and inconsistent use of an athlete-time denominator, such as athlete-exposures (AEs), that allows for comparison of injury rates by sex or across sports for severity of injury, practice and game settings, body part injured, and injury type.

McGuine[21] noted contradictory results in secondary school data reporting both sex-specific and practice versus game injury rates. Girls had a higher injury rate than boys in some matched sports[1,7,8] but similar risks in others.[7] Previous reports have indicated a higher game injury risk in contact sports,[20,22–24] whereas higher practice injury rates have been observed in noncontact sports.[8,25,26]

The purpose of our study was to describe the incidence and risk factors of injury for 29 middle school sports at a private school over a 20-year period. Studies of middle school populations are not comprehensive and do not

Table 1. Middle School Enrollment and Athletic Participation, 1988–2008

Sex	Grade	Overall Student Enrollment, Mean	Middle School Individual Athletes, Mean	Athlete Participation, %
Boys	7	185	93	50.30
Girls	7	187	87	48.70
Boys	8	179	116	65.00
Girls	8	188	107	57.30
Boys	9	221	45	20.30
Girls	9	216	14	6.20

address differences between boys and girls or practices and games. The exposure setting (ie, practice or game), injury location, injury type, and injury severity were evaluated for the overall sample and by sex.

METHODS

Participants and Setting

Punahou School, located in Honolulu, Hawaii, is a private, nonsectarian school with 3700 students in kindergarten through grade 12. Middle school athletes range in age from 12 to 15 years old and are in grades 7 to 9. The average enrollment for grades 7 and 8 is 740 students, with 57.4% of boys and 51.7% of girls participating in athletics. Although interscholastic league rules allow ninth graders (400 students) to participate at the middle school level, internal school policy limits ninth-grade participation in middle school sports (6.2% of ninth-grade girls, 21% of ninth-grade boys; Table 1). The ninth-grade boys may participate in middle school or high school football, depending on ability. Over 20 years, a total of 8081 boys participated in 14 sports and on 18 teams, with 5960 girls participating in 15 sports and on 19 teams (Table 2). Softball, baseball, basketball, soccer, and volleyball fielded 2 teams each for both boys and girls. During this study, 13 new sports were added to the athletic program. These consisted of boys' and girls' swimming (1994–1995), boys' and girls' diving and girls' water polo (1996), cheerleading and boys' and girls' tennis (1997), boys' and girls' golf (1998), girls' wrestling (1999), and boys' and girls' judo (2005). Gymnastics for girls was introduced in 1993 but dropped in 2007 (Table 2). This study was approved by the University of Hawaii Institutional Review Board.

Data Collection

Five certified athletic trainers were on staff during the study period, ensuring consistency of evaluation and treatment. Of the 3 full-time athletic trainers, 1 had been on staff for the full 20 years of this study and the others for 15 and 5 years, respectively. Two sports medicine physicians (physiatrists) were contracted by the school as team physicians for all levels and sports during the time of the study. They provided coverage for all football games and evaluated students referred by the athletic trainers in a weekly clinic. Athletes were then further referred to their primary care physicians for follow-up evaluation and treatment.

Injuries were evaluated by the athletic training staff and recorded daily into a database program designed by the principal author (G.B.). The data were initially recorded using AppleWorks integrated software (versions 2 and 3; Apple Inc, Cupertino, CA), but all injury data are now recorded and stored on FileMaker Pro (versions 5 to 11; FileMaker, Inc, Santa Clara, CA). The physician's diagnosis and time lost from activity were updated as the athlete returned to play. In 2001, the Vienna concussion guidelines[27] were implemented and then revised to align with both the Prague[28] and Zurich[29] concussion guidelines.

The Athletic Training Room Evaluation Form was used to collect all injury data, including name, grade, sex, sport, coach, date of injury, date of return to activity, team session, body part, category (eg, strain, sprain), evaluation results, severity of injury, activity status, and the evaluator's impressions and proposed treatment and rehabilitation. The player's position, activity at the time of injury, and field surface and conditions were also recorded.

Injuries

Throughout the study, *injury* was defined as any athlete complaint that required the attention of the athletic trainer, regardless of the time lost from activity.[3] Evaluation forms were completed for all athlete complaints. Lacerations requiring sutures or advanced medical care were recorded, but a nosebleed or simple abrasion was not recorded. Five injury-severity classifications were used: (1) minor, no time lost from activity; (2) mild, 1 to 7 days lost; (3) moderate, 8 to 21 days lost; (4) severe, 22 or more days lost; and (5) catastrophic, permanent disability, dismemberment, or death.[3]

Athlete-Exposures

An *AE* was defined as 1 athlete participating in 1 practice or game. Participation was not recorded daily. Exposures were estimated using team rosters and number of practice and game dates. The average middle school athletic season was 11 to 12 weeks of training and competition.

Data Analysis

The total injury rate was calculated as the total number of injuries divided by the total number of AEs and was expressed as injuries per 1000 AEs (injuries/1000 AEs). Total injury rates were calculated separately for all injuries and for time-loss (at least 1 day) injuries.

Sex-Matched Sports

The rate ratio (RR) was used to compare the injury rates in girl and boy athletes competing in the same sport (Table 2):

$$\text{Girls' to boys' RR} = \frac{(\text{Total girls' injuries}/\text{total girls' exposures}) \times 1000}{(\text{Total boys' injuries}/\text{total boys' exposures}) \times 1000}$$

If the girls' to boys' sport RR was above 1.0, the girl was more likely to sustain an injury than the boy. If the girls' to boys' sport RR was below 1.0, the boy was more likely to sustain a higher injury rate.

Exposure Setting (Practices Versus Games)

Practice and game time-loss injury-rate comparisons were based on the following RR (Table 3):

Table 2. Injury-Risk Sex Comparisons by Sport Among Middle School Athletes, 1988–2008

			Overall Injuries					Time-Loss Injuries			
Sport	Participation, y	Total Athletes	n	AEs	Injury Rate (/1000 AEs)	Rate Ratio	95% Confidence Interval	n	Injury Rate (/1000 AEs)	Rate Ratio	95% Confidence Interval
Softball (girls)	20	555	302	42 594	7.090	0.96	0.8, 1.1	109	2.559	0.78	**0.6, 0.9**[a]
Baseball (boys)	20	712	340	46 116	7.373	1	Referent	152	3.296	1.00	Referent
Basketball											
Girls	20	553	362	39 558	9.151	1.03	0.9,1.2	137	3.463	1.02	0.8, 1.3
Boys	20	592	378	42 630	8.867	1	Referent	145	3.401	1.00	Referent
Cheerleading (girls)	11	135	35	8775	3.989	NA	NA	21	2.393	NA	NA
Cross-country											
Girls	20	756	411	37 830	10.864	1.36	**1.2, 1.6**	147	3.886	1.23	0.9, 1.3
Boys	20	710	285	35 630	7.999	1	Referent	113	3.171	1.00	Referent
Diving											
Girls	13	47	11	3055	3.601	3.04	0.9, 130.9	6	1.964	0.00	NA
Boys	13	13	1	845	1.183	1	Referent	0	0.000	0.00	Referent
Football (boys)	20	1435	1600	99 810	16.030	NA	NA	847	8.486	NA	NA
Golf											
Girls	11	39	0	2340	0.000	0.00	NA	0	0.000	0.00	NA
Boys	11	96	0	5760	0.000	0	Referent	0	0.000	0.00	Referent
Gymnastics (girls)	15	156	65	9390	6.922	NA	NA	29	3.088	NA	NA
Judo											
Girls	4	21	7	1030	6.796	1.50	0.5, 4.2	4	3.883	1.05	0.2, 3.8
Boys	4	51	11	2435	4.517	1	Referent	9	3.696	1.00	Referent
Soccer											
Girls	20	666	412	51 732	7.964	1.07	0.9, 1.2	165	3.190	1.10	0.9, 1.4
Boys	20	706	410	55 068	7.445	1	Referent	159	2.887	1.00	Referent
Swimming											
Girls	14	386	33	25 090	1.315	1.86	**1.0, 3.8**	17	0.678	3.36	**1.1, 13.7**
Boys	14	305	14	19 825	0.706	1	Referent	4	0.202	1.00	Referent
Tennis											
Girls	12	266	27	15 960	1.692	1.24	0.7, 2.3	15	0.940	1.01	0.5, 2.2
Boys	12	268	22	16 080	1.368	1	Referent	15	0.933	1.00	Referent
Track											
Girls	20	1537	878	72 165	12.167	1.46	**1.3, 1.6**	352	4.878	1.46	**1.2, 1.7**
Boys	20	1589	619	74 505	8.308	1	Referent	249	3.342	1.00	Referent
Volleyball											
Girls	20	572	191	41 184	4.638	0.90	0.7, 1.1	54	1.311	0.72	0.5, 1.0
Boys	20	529	196	37 932	5.167	1	Referent	69	1.819	1.00	Referent
Water polo											
Girls	13	232	23	13 920	1.652	1.34	0.8, 2.4	6	0.431	1.45	0.4, 4.8
Boys	20	447	33	26 820	1.230	1	Referent	8	0.298	1.00	Referent
Wrestling											
Girls	10	39	24	2340	10.256	1.03	0.7, 1.6	15	6.410	1.49	0.8, 2.5
Boys	20	625	370	37 170	9.954	1	Referent	160	4.305	1.00	Referent
Total											
Girls[b]		5669	2681	348 798	7.686	1.15	**1.1, 1.2**	1027	2.944	1.09	**1.0, 1.2**
Boys[c]		6643	2679	400 816	6.684	1	Referent	1083	2.702	1.00	Referent
All		12312	5360	749 614	7.150	NA		2110	2.815	NA	
Sex matched											
Girls		5960	2781	366 963	7.578	0.89	0.8, 0.9	1077	2.935	0.76	**0.7, 0.8**
Boys		8078	4279	500 626	8.547	1	Referent	1930	3.855	1.00	Referent
All		14038	7060	867 589	8.137	NA		3007	3.466	NA	

Abbreviations: AE, athlete-exposures; NA, not available.

[a] Boldface indicates significance.
[b] Minus cheerleading and gymnastics.
[c] Minus football.

$$\text{Practice to game RR} = \frac{(\text{Total practice injuries}/\text{total practice exposures}) \times 1000}{(\text{Total game injuries}/\text{total game exposures}) \times 1000}$$

If the practice/game RR was above 1.0, the athlete was more likely to sustain an injury in practice than during a game. If the practice/game RR was below 1.0, the individual was more likely to sustain an injury during a game than in a practice.

For time-loss injuries only, RR comparisons were also calculated for body location injured, type of injury, severity of injury, and specific injury types. The 95% confidence interval (95% CI) was used to determine the statistical

Table 3. Injury-Risk Estimates During Practices and Games by Sport and Sex, 1988–2008

	Practices			Games				
Sport	Exposures	Time-Loss Injuries	Injury Rate (/1000 AEs)	Exposures	Time-Loss Injuries	Injury Rate (/1000 AEs)	Practice/Game Rate Ratio	95% Confidence Interval
Softball								
Girls	31 610	98	3.100	10 984	11	1.001	3.10	**1.7, 6.4**[a]
Baseball								
Boys	32 168	121	3.762	13 948	31	2.223	1.69	**1.1, 2.6**
Basketball								
Girls	28 690	104	3.625	10 868	33	3.036	1.19	0.8, 1.8
Boys	30 966	119	3.843	11 664	26	2.229	1.72	**1.1, 2.8**
Cheerleading (girls)	6075	21	3.457	2700	0	0.000	0.00	NA
Cross-country								
Girls	31 776	129	4.060	6054	18	2.973	1.37	0.8, 2.4
Boys	29 924	105	3.509	5706	8	1.402	2.50	**1.2, 6.0**
Diving								
Girls	2585	6	2.321	470	0	0.000	0.00	NA
Boys	715	0	0.000	130	0	0.000	0.00	NA
Football (boys)	85 834	637	7.421	13 976	210	15.026	0.49	**0.4, 0.6**
Golf								
Girls	1989	0	0.000	351	0	0.000	0.00	NA
Boys	4896	0	0.000	864	0	0.000	0.00	NA
Gymnastics (girls)	8136	25	3.073	1254	4	3.190	0.96	0.3, 3.8
Judo								
Girls	925	4	4.324	105	0	0.000	0.00	NA
Boys	2180	9	4.128	255	0	0.000	0.00	NA
Soccer								
Girls	38 548	124	3.217	13 184	41	3.110	1.03	0.7, 1.5
Boys	41 076	119	2.897	13 992	40	2.859	1.01	0.7, 1.5
Swimming								
Girls	21 230	17	0.801	3860	0	0.000	0.00	NA
Boys	16 775	4	0.238	3050	0	0.000	0.00	NA
Tennis								
Girls	10 640	15	1.410	5320	0	0.000	0.00	NA
Boys	10 720	12	1.119	5360	3	0.560	2.00	0.5, 11.0
Track								
Girls	59 269	325	5.483	12 896	27	2.094	2.62	**1.8, 4.0**
Boys	61 193	212	3.464	13 312	37	2.779	1.25	0.9, 1.8
Volleyball								
Girls	29 984	52	1.734	11 200	2	0.179	9.71	**2.6, 82.3**
Boys	27 404	64	2.335	10 528	5	0.475	4.92	**2.0, 15.7**
Water polo								
Girls	11 600	6	0.517	2320	0	0.000	0.00	NA
Boys	22 350	7	0.313	4470	1	0.224	1.40	0.2, 63.1
Wrestling								
Girls	2028	14	6.903	312	1	3.205	2.15	0.3, 91.1
Boys	32 302	123	3.808	4868	37	7.601	0.50	0.3, 0.8
Total								
Girls	285 085	940	3.297	81 878	137	1.673	1.97	**1.7, 2.4**
Boys	398 503	1532	3.844	102 123	398	3.897	0.99	0.9, 1.1
All	683 588	2472	3.616	184 001	535	2.908	1.24	**1.1, 1.4**

Abbreviations: AEs, athlete-exposures; NA, not available.

[a] Boldface indicates significance.

significance for all rate ratios. Confidence intervals that did not include 1.0 indicated statistical significance ($P \leq .05$).

All analyses were conducted with the STATA (version 5.0; STATA Corporation, College Station, TX) and SPSS (version 18.0; SPSS, Inc, Chicago, IL) statistical packages.

RESULTS

Sport Participation

Over the 20-year period, 14 038 athletes were included in this analysis, with a greater number of boys (n = 8078) than girls (n = 5960; Table 2). By sport, track had the greatest number of athletes participating (n = 3126; 1537 girls, 1589 boys), followed by cross-country (n = 1466; 756 girls, 710 boys), and football (n = 1435). Boys had a higher participation level (n = 500 626 AEs) than girls (n = 366 963 AEs; Table 2). By sport, track had the highest participation level (n = 146 670 AEs) followed by soccer (n = 106 800 AEs) and football (n = 99 810 AEs).

Injury Rate

Boy athletes reported 4279 injuries, with 1930 injuries (45.1%) resulting in at least 1 day lost from activity

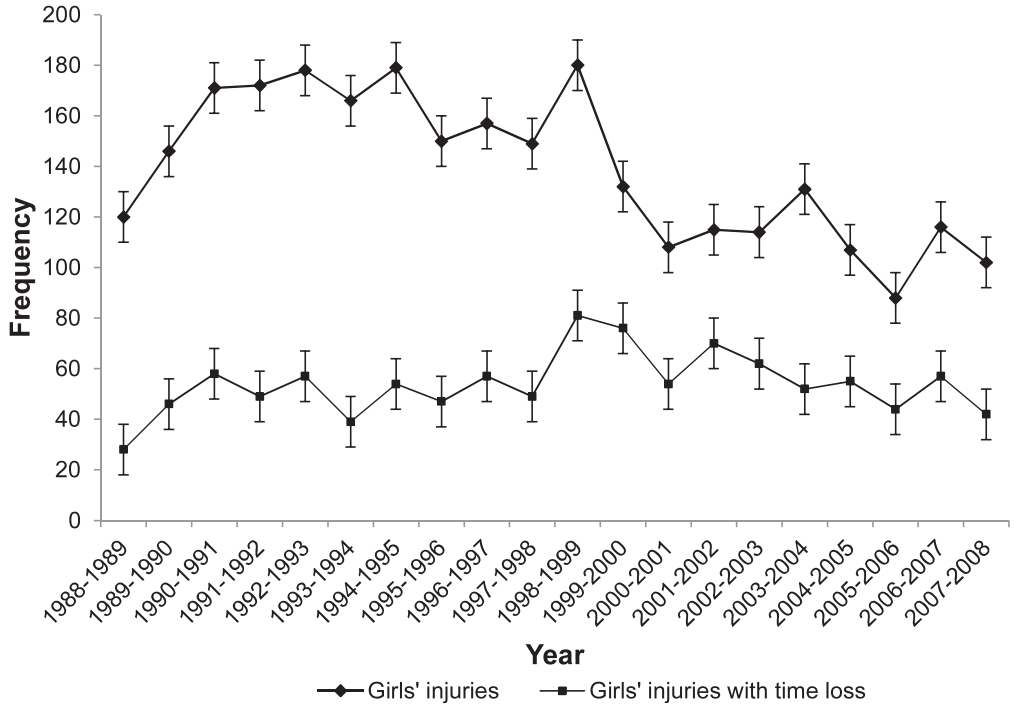


Figure 1. Annual non–time-loss and time-loss injuries for middle school girls (1988–2008).

(Table 2). Girl athletes reported 2781 injuries, with 1077 injuries (38.7%) resulting in at least 1 day lost from activity. Most injuries (4053) were classified as minor (no time lost from participation) and accounted for 54.9% and 61.3% of all injuries for boys and girls, respectively.

The injury rate for girls was highest in 1990 and 1998 and gradually declined over the following 10-year period (Figure 1), even with the addition of 10 new sports for girls. The injury rate for boys remained relatively the same (Figure 2). The time-loss injury rate for girls rose gradually until 1998 and then gradually declined (Figure 1), whereas

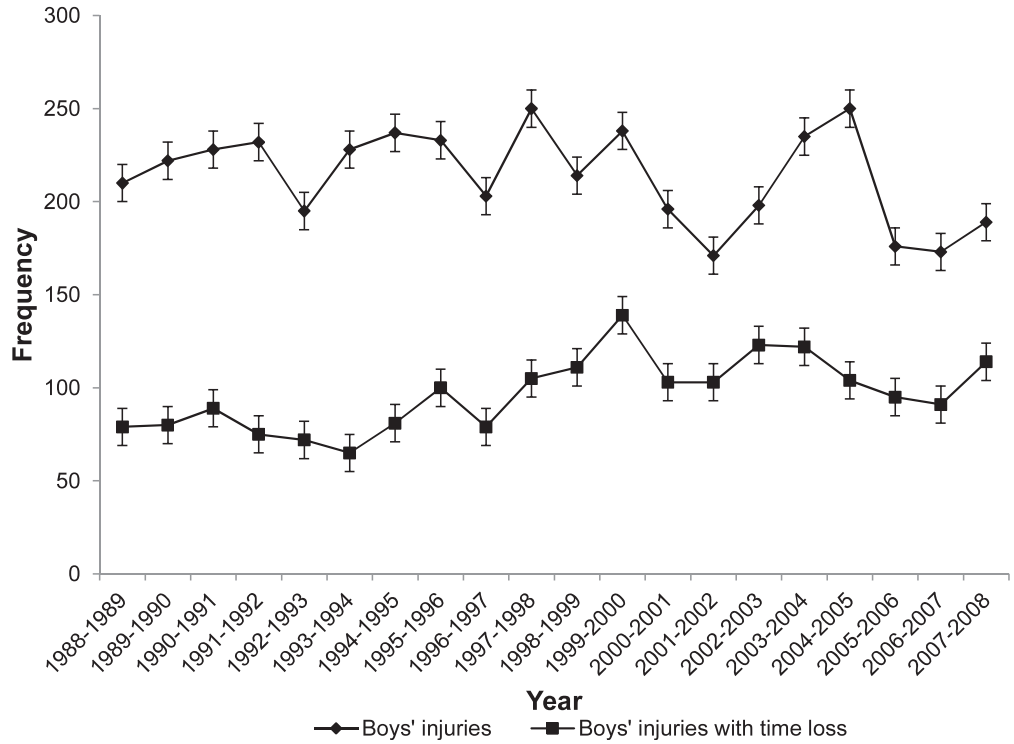


Figure 2. Annual non–time-loss and time-loss injuries for middle school boys (1988–2008).

Table 4. Injury by Severity in Practice and Game Settings by Sex, 1988–2008

			Injuries											
			Mild[a]				Moderate[b]				Severe[c]			
Sex	Setting	AEs	Number	Injury Rate	Practice/ Game Rate Ratio	95% Confidence Interval	Number	Injury Rate	Practice/ Game Rate Ratio	95% Confidence Interval	Number	Injury Rate	Practice/ Game Rate Ratio	95% Confidence Interval
Girls	Practices	285 085	843	2.957	2.05	**1.7, 2.5**[d]	62	0.217	1.19	0.7, 2.3	31	0.109	1.11	0.5, 2.8
	Games	81 878	118	1.441	1.00	Referent	15	0.183	1.00	Referent	8	0.098	1.00	Referent
	Total	366 963	961	2.619			77	0.210			39	0.106		
Boys	Practices	398 503	1305	3.275	1.12	0.9, 1.3	132	0.331	0.89	0.6, 1.3	99	0.248	0.45	**0.3, 0.6**
	Games	102 123	299	2.928	1.00	Referent	38	0.372	1.00	Referent	57	0.558	1.00	Referent
	Total	500 626	1604	3.204			170	0.340			156	0.312		
Overall	Practices	683 588	2148	3.160	1.35	**1.2, 1.5**	194	0.284	0.99	0.7, 1.4	130	0.190	0.54	**0.4, 0.7**
	Games	184 001	417	2.266	1.00	Referent	53	0.288	1.00	Referent	65	0.353	1.00	Referent
	Total	867 589	2565	2.956			247	0.285			195	0.225		

Abbreviation: AEs, athlete-exposures.

[a] Mild injury: 1 to 7 days lost from sport participation.
[b] Moderate injury: 8 to 21 days lost from sport participation.
[c] Severe injury: 22 or more days lost from sport participation.
[d] Boldface indicates significance.

the time-loss injury rate for boys has gradually increased over the 20-year period (Figure 2).

Sex Comparisons

In sex-matched sports, middle school girls were more likely to sustain any injury (RR = 1.15, 95% CI = 1.1, 1.2) or a time-loss injury (RR = 1.09, 95% CI = 1.0, 1.2) than middle school boys (Table 2). Although middle school girls were more likely to sustain an injury in practice (RR = 1.97, 95% CI = 1.7, 2.4) than during a game, the risk of injury was similar during practices and games for middle school boys (RR = 0.99, 95% CI = 0.9, 1.1; Table 3).

Middle school girls were twice as likely to sustain a mild injury during a practice than in a game (RR = 2.05, 95% CI =1.7, 2.5), but no differences were noted during practices or games for moderate or severe injuries. For boys, the risk of a severe injury was 55% less during practices than during games (RR = 0.45, 95% CI = 0.3, 0.6; Table 4). We found no differences between practices and games for mild or moderate injuries.

Specific sex-matched sport differences and specific injury differences are reported under these topics.

Injury Rate: All Injuries. For all injuries, the overall injury rate was 8.137/1000 AEs. Football recorded the highest injury rate for all sports at 16.030/1000 AEs, followed by girls' track (12.167/1000 AEs), girls' cross-country (10.864/1000 AEs), girls' wrestling (10.256/1000 AEs), and boys' wrestling (9.954/1000 AEs; Table 2). Diving, tennis, water polo, swimming, and golf reported the lowest injury rates. However, in sex-matched sports, girl athletes were more likely to incur an injury than boy athletes in cross-country (RR = 1.36, 95% CI = 1.2, 1.6), swimming (RR = 1.86, 95% CI = 1.0, 3.8), and track (RR = 1.46, 95% CI = 1.3, 1.6).

Injury Rate: Time-Loss Injuries. For time-loss injury rates, the overall injury rate was 3.466/1000 AEs. Football recorded the highest injury rate (8.486/1000 AEs), followed by girls' wrestling (6.410/1000 AEs), girls' track (4.878/1000 AEs), boys' wrestling (4.305/1000 AEs), and girls' (3.883/1000 AEs) and boys' judo (3.696/1000 AEs; Table 2). When all sex-matched sports were compared, girls were more likely to sustain a time-loss injury than their male counterparts (RR = 1.09, 95% CI = 1.0, 1.2). Girl track athletes (RR = 1.46, 95% CI = 1.2, 1.7) and girl swimmers (RR = 3.36, 95% CI = 1.1, 13.7) were at higher risk for time-loss injury than their male counterparts. Softball players were at lower injury risk than baseball players (RR = 0.78, 95% CI = 0.6, 0.9).

Exposure Setting (Practices Versus Games)

For the total sample, practice injury rates were higher than game injury rates (RR = 1.24, 95% CI = 1.1, 1.4; Table 3). For all girls' sports, a higher injury rate was recorded in practices (3.297/1000 AEs) than in games (1.673/1000 AEs) (RR = 1.97, 95% CI = 1.7, 2.4). For boys, the risk of injury was nearly identical during practices and games (RR = 0.99, 95% CI = 0.9, 1.1).

For softball (RR = 3.10, 95% CI = 1.7, 6.4), baseball (RR = 1.69, 95% CI = 1.1, 2.6), boys' cross-country (RR = 2.50, 95% CI = 1.2, 6.0), girls' track (RR = 2.62, 95% CI = 1.8, 4.0), boys' basketball (RR = 1.72, 95% CI = 1.1, 2.8), and girls' (RR = 9.71, 95% CI = 2.6, 82.3) and boys' volleyball (RR = 4.92, 95% CI = 2.0, 15.7), the injury risk was greater during practices than in games. Only in boys' wrestling (RR = 0.50, 95% CI = 0.3, 0.8) and football (RR = 0.49, 95% CI = 0.4, 0.6) was the likelihood of sustaining an injury greater during a game than during a practice.

Injury Severity

Minor injuries accounted for 57% of all reported middle school injuries. Middle school girls reported that 61.2% of all injuries involved no time loss, compared with 54.8% of all boys' injuries. Although mild injuries were more likely to occur during practices than during games (RR = 1.35, 95% CI = 1.2, 1.5), severe injuries were more likely to occur during games (RR = 0.54, 95% CI = 0.4, 0.7; Table 4). Girls were twice as likely to report a mild injury during practice than during a game (RR = 2.05, 95% CI = 1.7, 2.5), whereas boys were less likely to incur a severe injury

Table 5. Injury by Sex and Body Part in Practice and Game Settings, 1988–2008

Sex	Body Part	Total Time-Loss Injuries	Percentage of Total Injuries[a]	Time-Loss Practice Injuries	Practice Injury Rate (/1000 AEs)[b]	Time-Loss Game Injuries	Game Injury Rate (/1000 AEs)[c]	Practice/Game Rate Ratio	95% Confidence Interval
Girls	Ankle	265	24.61	224	0.786	41	0.501	1.59	**1.1, 2.3**[d]
	Shin/calf	182	16.90	175	0.614	7	0.085	7.18	**3.4, 18.1**
	Knee	156	14.48	135	0.474	21	0.256	1.85	**1.2, 3.1**
	Thigh	146	13.56	124	0.435	22	0.269	1.62	**1.1, 2.7**
	Wrist/hand/finger	56	5.20	46	0.161	10	0.122	1.32	0.7, 2.9
	Foot/toe	52	4.83	49	0.172	3	0.037	4.50	**1.5, 22.6**
	Hip	43	3.99	36	0.126	7	0.085	1.44	0.6, 3.8
	Back	42	3.90	33	0.116	9	0.110	1.02	0.5, 2.4
	Shoulder/upper arm	35	3.25	29	0.102	6	0.073	1.39	0.6, 4.1
	Head/face	33	3.06	27	0.095	6	0.073	1.29	0.5, 3.8
	Elbow/forearm	22	2.04	18	0.063	4	0.049	1.29	0.4, 5.3
	Other	20	1.86	19	0.067	1	0.012	5.17	0.8, 215.4
	Neck	13	1.21	11	0.039	2	0.024	1.58	0.4, 14.7
	Trunk/abdomen	12	1.11	12	0.042	0	0.000	0.00	NA
Boys	Ankle	316	16.37	251	0.630	65	0.636	0.99	0.8, 1.3
	Knee	228	11.81	179	0.449	49	0.480	0.94	0.7, 1.3
	Thigh	210	10.88	169	0.424	41	0.401	1.06	0.8, 1.5
	Wrist/hand/finger	198	10.26	158	0.396	40	0.392	1.01	0.7, 1.5
	Back	135	6.99	111	0.279	24	0.235	1.19	0.7, 1.8
	Hip	129	6.68	104	0.261	25	0.245	1.07	0.7, 1.7
	Neck	129	6.68	98	0.246	31	0.304	0.81	0.6, 1.4
	Shin/calf	118	6.11	108	0.271	10	0.098	2.77	**1.7, 8.2**
	Shoulder/upper arm	118	6.11	92	0.231	26	0.255	0.91	0.6, 1.5
	Foot/toe	115	5.96	103	0.258	12	0.118	2.20	**1.2, 4.4**
	Head/face	82	4.25	51	0.128	31	0.304	0.42	**0.3, 07**
	Elbow/forearm	63	3.26	43	0.108	20	0.196	0.55	**0.3, 0.9**
	Trunk/abdomen	55	2.85	38	0.095	17	0.166	0.57	**0.3, 0.8**
	Other	34	1.76	28	0.070	6	0.059	1.20	0.6, 4.1
Total	Ankle	581	19.32	475	0.695	106	0.576	1.21	0.9, 1.5
	Knee	384	12.77	314	0.459	70	0.380	1.21	0.9, 1.6
	Thigh	356	11.84	293	0.429	63	0.342	1.25	0.9, 1.7
	Shin/calf	300	9.98	283	0.414	17	0.092	4.48	**3.0, 9.2**
	Wrist/hand/finger	254	8.45	204	0.298	50	0.272	1.10	0.8, 1.5
	Back	177	5.89	144	0.211	33	0.179	1.17	0.8, 1.7
	Hip	172	5.72	140	0.205	32	0.174	1.18	0.8, 1.8
	Foot/toe	167	5.55	152	0.222	15	0.082	2.73	**1.6, 4.9**
	Shoulder/upper arm	153	5.09	121	0.177	32	0.174	1.02	0.7, 1.6
	Neck	142	4.72	109	0.159	33	0.179	0.89	0.6, 1.5
	Head/face	115	3.82	78	0.114	37	0.201	0.57	**0.4, 0.9**
	Elbow/forearm	85	2.83	61	0.089	24	0.130	0.68	0.4, 1.1
	Trunk/abdomen	67	2.23	50	0.073	17	0.092	0.79	0.4, 1.1
	Other	54	1.80	47	0.069	7	0.038	1.81	0.9, 5.1

Abbreviations: AEs, athlete-exposures; NA, not available.

[a] Injuries: girls, 1077; boys, 1930; total, 3007.
[b] Practice exposures: girls, 285 085; boys, 398 503; total, 683 588.
[c] Game exposures: girls, 81 878; boys, 102 123; total, 184 001.
[d] Boldface indicates significance.

in a practice than in a game (RR = 0.45, 95% CI, 0.3, 0.6; Table 4).

Body Part

Ankle, knee, thigh, shin/calf, and wrist/hand/finger injuries were the most frequently reported time-loss injuries by middle school athletes. Ankle and knee injuries were the time-loss injuries reported most overall (n = 581, rate = 0.669/1000 AEs, and n = 384, rate = 0.442/1000 AEs, respectively) and during practices (0.695/1000 AEs and 0.459/1000 AEs, respectively) and games (0.576/1000 AEs and 0.380/1000 AEs, respectively; Table 5).

Shin/calf (RR = 4.48, 95% CI = 3.0, 9.2) and foot/toe (RR = 2.73, 95% CI = 1.6, 4.9) injury rates were greater during practices than during games. Only the injury rate for head/face injuries (RR = 0.57, 95% CI = 0.4, 0.9) was higher during games than during practices.

For girl athletes, higher injury rates for the ankle (RR = 1.59, 95% CI = 1.1, 2.3), knee (RR = 1.85, 95% CI = 1.2, 3.1), thigh (RR = 1.62, 95% CI = 1.1, 2.7), shin/calf (RR = 7.18, 95% CI = 3.4, 18.1), and foot/toe (RR = 4.50, 95% CI = 1.5, 22.6) were reported during practice sessions than during games. For boys, although the risk of sustaining a head/face (RR = 0.42, 95% CI = 0.3, 0.7), elbow/forearm (RR = 0.51, 95% CI = 0.3, 0.9), or trunk/abdomen (RR = 0.46, 95% CI = 0.3, 0.8) injury was higher during games than during practices, shin/calf (RR = 3.46, 95% CI = 1.7, 8.2) and foot/toe injuries (RR = 2.20, 95% CI = 1.2, 4.4)

Table 6. Injury Rates by Sex and Injury Type in Practice and Game Settings, 1988–2008

Sex	Injury Type	Total Time-Loss Injuries	Percentage of Total Injuries[a]	Time-Loss Practice Injuries	Practice Injury Rate (/1000 AEs)[b]	Time-Loss Game Injuries	Game Injury Rate (/1000 AEs)[c]	Practice/Game Rate Ratio	95% Confidence Interval
Girls	Tendinitis	309	28.69	302	1.059	7	0.085	12.39	**5.8, 30.5**[d]
	Sprain	299	27.76	248	0.870	51	0.623	1.40	**1.2, 2.4**
	Strain	293	27.21	251	0.880	42	0.513	1.72	**1.1, 1.9**
	Contusion	69	6.41	47	0.165	22	0.269	0.61	0.4, 1.1
	Other	48	4.46	45	0.158	3	0.037	4.31	**1.4, 21.7**
	Fracture	44	4.09	36	0.126	8	0.098	1.29	0.6, 3.2
	Neurotrauma	15	1.39	10	0.035	5	0.061	0.57	0.2, 2.1
Boys	Strain	541	28.03	461	1.157	80	0.783	1.48	**1.2, 1.9**
	Sprain	532	27.56	415	1.041	117	1.146	0.91	0.7, 1.1
	Tendinitis	269	13.94	255	0.640	14	0.137	4.67	**2.7, 8.7**
	Contusion	245	12.69	167	0.419	78	0.764	0.55	**0.4, 0.7**
	Fracture	149	7.72	101	0.253	48	0.470	0.54	**0.4, 0.8**
	Other	114	5.91	87	0.218	27	0.264	0.83	0.5, 1.4
	Neurotrauma	80	4.15	49	0.123	31	0.304	0.41	**0.2, 0.6**
Total	Strain	834	27.74	709	1.037	131	0.712	1.46	**1.3, 1.9**
	Sprain	831	27.64	666	0.974	159	0.864	1.13	0.9, 1.3
	Tendinitis	578	19.22	557	0.815	21	0.114	7.14	**4.6, 11.5**
	Contusion	314	10.44	214	0.313	100	0.543	0.58	**0.4, 0.7**
	Fracture	193	6.42	146	0.214	51	0.277	0.77	**0.5, 0.9**
	Neurotrauma	95	3.16	123	0.180	35	0.190	0.95	0.8, 1.9
	Other	162	5.39	59	0.086	36	0.196	0.44	**0.3, 0.7**

Abbreviation: AEs, athlete-exposures.

[a] Injuries: girls, 1077; boys, 1930; total, 3007.
[b] Practice exposures: girls, 285 085; boys, 398 503; total, 683 588.
[c] Game exposures: girls, 81 878; boys, 102 123; total, 184 001.
[d] Boldface indicates significance.

were more likely to be reported during a practice than a game.

Injury Type

Overall, strains, sprains, and tendinitis were the injury types reported most frequently by middle school athletes (Table 6). The rates for strains (RR = 1.46, 95% CI = 1.3, 1.9), sprains (RR = 1.13, 95% CI = 0.9, 1.3), and tendinitis (RR = 7.14, 95% CI = 4.6, 11.5) were higher during practices than during games, whereas contusions (RR = 0.58, 95% CI = 0.4, 0.7) and fractures (RR = 0.77, 95% CI = 0.5, 0.9) were more likely to occur during games.

For girl athletes, the injury rate of reporting tendinitis was 12 times higher (RR = 12.39, 95% CI = 5.8, 30.5) during practices than during games. Strains (RR = 1.72, 95% CI = 1.1, 1.9) and sprains (RR = 1.40, 95% CI = 1.2, 2.4) were also more likely to occur during practices than games. Tendinitis (RR = 4.67, 95% CI = 2.7, 8.7) and strain (RR = 1.48, 95% CI = 1.2, 1.9) rates were higher for boys during practice than during games, but the rates for contusions (RR = 0.55, 95% CI = 0.4, 0.7), fractures (RR = 0.54, 95% CI = 0.4, 0.8), and neurologic injuries (RR = 0.41, 95% CI = 0.2, 0.6) were greater during games than during practices.

Specific Injury Types

Concussions. The overall rate for concussion was 0.067/1000 AEs, with the rate almost 3 times higher (RR = 2.83, 95% CI = 1.5, 5.9) for boy athletes (0.092/1000 AEs) than girl athletes (0.033/1000 AEs) (Table 7). Middle school athletes were at higher risk of concussion during games than during practices (RR = 0.33, 95% CI = 0.2, 0.6). Football accounted for 35 of the 58 concussions (2.2% of all football injuries, rate = 0.35/1000 AEs) with 28 resulting in time lost from activity. The remaining concussions were spread among 6 sports, with basketball (n = 7, rate = 0.085/1000 AEs), soccer (n = 7, rate = 0.066/1000 AEs), and wrestling (n = 6, rate = 0.15/1000 AEs) participants recording the most injuries. The boys' concussion rate was 0.063/1000 AEs for practices and 0.206/1000 AEs for games (RR = 0.31, 95% CI = 0.2, 0.6). Concussions were more likely to occur during competition for boys. Girls participating in soccer (n = 6) and basketball (n = 5) accounted for the majority of concussions. Girls reported a higher rate of concussion during games (0.061/1000 AEs) than during practices (0.025/1000 AEs; RR = 0.40, 95% CI = 0.1, 1.6), but the difference was not significant.

Rotator Cuff Injuries. The rate of rotator cuff injury was nearly twice as high for boys (0.033/1000 AEs) as for girls (0.019/1000 AEs). Although girls were more likely to report a rotator cuff injury during practices (RR = 1.72, 95% CI = 0.2, 79.3) and boys during games (RR = 0.83, 95% CI = 0.3, 3.5), the differences were not significant.

Anterior Cruciate Ligament Injuries. Only 8 anterior cruciate ligament (ACL) injuries (0.009/1000 AEs) were reported during the 20-year recording period (Table 7), with no differences between practices and games (RR = 0.81, 95% CI = 0.1, 8.2). Football players accounted for all 5 of the boys' injuries; 4 of these were practice injuries. The 3 girls' injuries occurred during basketball, softball, and wrestling: 2 during practice and 1 during a game. No significant differences in ACL risk were seen during games or practices for boys (RR = 1.03, 95% CI = 0.1, 50.5) or girls (RR = 0.57, 95% CI = 0.1, 33.9). All 8 ACL injuries required surgical repair.

Table 7. Specific Injuries by Sex in Practice and Game Settings, 1988–2008

Sex	Injury	Total Time-Loss Injuries	Total Injury Rate (/1000 AEs)[a]	Practice Time-Loss Injuries	Practice Injury Rate (/1000 AEs)[b]	Game Time-Loss Injuries	Game Injury Rate (/1000 AEs)[c]	Practice/Game Rate Ratio	95% Confidence Interval
Girls	Concussion	12	0.032	7	0.025	5	0.061	0.40	0.1, 1.6[d]
	Rotator cuff injury	7	0.019	6	0.021	1	0.012	1.72	0.2, 79.3
	Anterior cruciate ligament	3	0.008	2	0.007	1	0.012	0.57	0.1, 33.9
	Medial tibial stress syndrome	112	0.305	110	0.386	2	0.024	15.80	**4.3, 132.1**
	Ankle sprain	210	0.573	173	0.607	37	0.452	1.34	0.9, 2.0
Boys	Concussion	46	0.092	25	0.063	21	0.206	0.31	**0.2, 0.6**
	Rotator cuff injury	17	0.034	13	0.033	4	0.039	0.83	0.3, 3.5
	Anterior cruciate ligament	5	0.010	4	0.010	1	0.010	1.03	0.1, 50.5
	Medial tibial stress syndrome	52	0.104	51	0.128	1	0.010	13.07	**2.2, 526.2**
	Ankle sprain	223	0.445	176	0.442	47	0.460	0.96	0.7, 1.4
Total	Concussion	58	0.067	32	0.047	26	0.141	0.33	**0.2, 0.6**
	Rotator cuff injury	24	0.028	19	0.028	5	0.027	1.02	0.4, 3.5
	Anterior cruciate ligament	8	0.009	6	0.009	2	0.011	0.81	0.1, 8.2
	Medial tibial stress syndrome	164	0.189	161	0.236	3	0.016	14.45	**4.9, 70.8**
	Ankle sprain	433	0.499	349	0.511	84	0.457	1.12	0.9, 1.4

Abbreviation: AEs, athlete-exposures.

[a] Total AEs: girls, 366 963; boys, 500 626.
[b] Practice exposures: girls, 285 085; boys, 398 503; total, 683 588.
[c] Game exposures: girls, 81 878; boys, 102 123; total, 184 001.
[d] Boldface indicates significance.

Medial Tibial Stress Syndrome. The overall injury rate for medial tibial stress syndrome (MTSS) was 0.189/1000 AEs, and the risk was almost 3 times higher (RR = 2.93, 95% CI = 2.1, 4.2) for girl athletes (0.30/1000 AEs) than boy athletes (0.10/1000 AEs). Middle school athletes were 14 times more likely to sustain an MTSS injury during practice (RR = 14.45, 95% CI = 4.9, 70.8; Table 7). The injury rate for MTSS among middle school girls was almost 16 times higher in practices than in games (RR = 15.80, 95% CI = 4.3, 132.1). Similar findings were noted for boys (RR = 13.07, 95% CI = 2.2, 526.2).

Ankle Sprains. Ankle sprains had the highest specific injury rate for boys and girls (Table 7). Girls had a higher injury rate for ankle sprains during practices (0.607/1000 AEs) than boys (0.442/1000 AEs), but boys had a slightly higher game injury rate (0.460/1000 AEs) than girls (0.452/1000 AEs). Although boys had a higher risk of ankle sprains during games and girls a higher rate of ankle sprains during practices, the differences were not significant.

Tendinitis. Tendinitis accounted for 19% of all reported injuries, with higher rates during practices (0.808/1000 AEs) than during games (0.114/1000 AEs; RR = 13.07, 95% CI = 4.6, 11.5). Patterns were similar for middle school boy and girl athletes.

DISCUSSION

Main Findings

Our results indicate that (1) the overall risk of injury was greater for boys for non–time-loss or time-loss injuries, (2) football had the highest rate of injury for both non–time-loss and time-loss injuries, (3) the risk of time-loss injuries varied for sex-matched sports, (4) the overall risk of injury was greater during practices than in games, especially for girls, (5) most injuries were mild in nature, (6) the ankle and knee had the highest incidences of injury, (7) strains, sprains, and tendinitis were the most common injury types, (8) ankle sprains were the most frequent specific injury, and (9) boy athletes were more likely to suffer concussions than girl athletes.

To our knowledge, this is the largest prospective, longitudinal study of middle school sport injuries to date. The compilation of 20 years of injury data at the same site and with the same athletic trainers is unique for both middle and secondary school programs. The duration of this study is also substantially longer than other studies at the secondary or middle school levels.[3,5,7,30,31]

All Injury Versus Time-Loss–Injury Rates

The reporting of non–time-loss versus time-loss injury rates remains debatable. Presently, rates based on time-loss injuries[6,7,9,11] have been advocated because injuries that cause time loss may be more accurately recalled.[30,32] However, a time-loss injury definition that requires a coach, parent, athletic trainer, or physician report may underestimate the true injury burden.[3,30,32] Even though many injuries are not serious in nature, injuries considered minor in nature and that do not cause immediate impairment may still have long-term consequences.[30] Additionally, reporting non–time-loss injuries may provide a better perspective of the daily workload of the athletic trainer. In reporting both non–time-loss and time-loss injuries, we noted that the rates for non–time-loss injuries were higher than for time-loss injuries. These findings support those of other youth[3,30] and collegiate[32] studies in which rates have been reported using both injury definitions. Our results suggest that non–time-loss injury rates should be used to reflect the true extent of the problem, but the difficulty of collecting accurate data in most settings (eg, lack of enough athletic trainers at the middle school level to identify and record the non–time-loss injuries) may preclude this recommendation. Given that our study may be the first to describe injuries in most middle school girls' and boys' sports, additional research is needed to confirm these findings.

Although middle school sport participation levels increased over the 20-year period, the total number of injuries decreased. However, the number and percentage of day-loss injuries increased over the same time period. The greater participation by girls in athletics, both interscholastic and club sports, may explain the change in the occurrence of injuries. Over the 20-year period, the girl athletes may have become more competitive and less likely to report a minor injury to the athletic trainer or coach in order to continue participation in the sport. The reasons for these changes in severity are not clear. We speculate that the athletes may be reporting only those injuries they consider more problematic, rather than the nuisance injuries that may have little effect on their participation level.

Because data for middle school sport injuries are limited, we used reports from secondary school and some collegiate studies for comparison. Based on our data, middle school girls were more likely to sustain an injury during athletic participation than their male counterparts in the same sport. This finding is consistent with secondary school injury surveillance reports by Comstock et al,[11] Rechel et al,[33] and Powell and Barber-Foss.[7] However, similar to reports at the high school and collegiate levels, injury rates varied by sport.[11,14,32,34–36]

Sex Comparisons

Overall, we observed a difference between practice and game injury rates for girls but no such difference for boys. This is consistent with findings of secondary school studies.[8,15,37] Middle school cross-country girl athletes reported a higher injury rate than their male counterparts, which is also consistent with other studies.[7,8,37] Rauh et al[8] suggested that girl runners might be more apt to report minor pain complaints than boys. Girls' and boys' soccer players reported higher injury-incidence rates in games than in practices, which is consistent with previous studies.[15,38]

Practices Versus Games

Our findings indicated that the injury risk was greater during practices than during competition for all sports except football, gymnastics, and male wrestling. This finding is in contrast to previous observations at the secondary school level,[7,9,10,16,33,39] where higher rates occurred during competitions than during practices in all sports. Our findings are also noteworthy because practices are conducted daily in the competitive interscholastic middle school setting, compared with community-based programs in which practices may take place only twice per week. Dompier et al[30] noted that youth football players sustain more game injuries, which is consistent with our findings. In a study of school children, Backx et al[40] reported that basketball and soccer players were at 3 times greater risk of injury during games than during practices. This result is in direct contrast to our finding that middle school boy basketball players were at more risk of injury during practices than during games. Thus, except for a few sports, such as football, gymnastics, and wrestling, in which coverage is known to be needed at competitions, we suggest that middle school administration and athletic departments also ensure that appropriate medical coverage is available at practices for those middle school sports with greater injury risks.[41]

Team Contact Sports

Basketball. In this study, middle school girl and boy basketball players reported similar injury rates for all injuries and time-loss injuries. Both girl and boy basketball players had higher injury rates during practices than games, but the rates were only different for boys. Our findings are in contrast to those at the high school level demonstrating greater injury rates in boys than girls.[5,7,23,33,39] Furthermore, authors[7,33,39] of several high school studies have noted that girl basketball players were more likely to sustain an injury in games than in practices.

Football. Similar to all levels of play, middle school football players had the highest injury rate of all sports.[3,4,7,11,42–44] Our findings for all injuries, overall (16.03/1000 AEs) and time loss (8.49/1000 AEs), are similar to those reported by Dompier et al[30] (17.8/1000 AEs and 10.7/1000 AEs, respectively) among youth football players. Regarding time-loss injuries only, our rate was similar to reports of youth[45] (10.4/1000 AEs) and high school[7] (8.1/1000 AEs) players but higher than in other youth and high school (3.54/1000 AEs[43] and 4.36/1000 AEs,[44] respectively) football studies. Our findings that game injury rates were greater than practice injury rates are similar to those of other youth football studies[20,30,45] and suggest that the incidence of injury is likely to be higher in games than in practices at this level.

Soccer. Although middle school girl soccer players sustained a 10% higher risk of injury than middle school boy soccer players, comparison with secondary school soccer studies is somewhat difficult due to considerable variations in reported injuries for girls' and boys' secondary school soccer players.[7,9,33,34] In our study, girl and boy middle school soccer players were equally likely to sustain a practice injury as a competition injury. This result is in contrast to secondary school studies[7,15,33,34] in which competition injury rates have been consistently higher than practice rates for both boys and girls.

Individual Sport: Wrestling

Contemporary published data on middle school–aged boy wrestlers are limited to tournament competition data,[46] with no practice injury data available. To our knowledge, injury data for female wrestlers at any level have not been reported. In our study, only 1 match injury was reported, which likely reflects the low number of interscholastic participants and matches. The overall injury rate was nearly identical for boys' and girls' wrestling. Girl wrestlers reported higher practice injury rates than boy wrestlers, and boy wrestlers' injuries were more likely to occur during matches, but the limited participation numbers for girls did not allow for appropriate statistical comparisons. The higher rate during competition for middle school boy wrestlers is consistent with that of high school[7,33] and collegiate[47] male wrestlers. Knowles et al[34] and Rechel et al[33] noted that game injuries were more frequent but reported lower injury rates in secondary school wrestlers. Agel et al[47] reported a much higher match injury rate for collegiate male wrestlers.

Noncontact Sports: Team

Softball/Baseball. Epidemiologic comparisons of softball and baseball injury rates are a standard practice

at all levels. In our study, the middle school baseball players had a higher injury rate than the middle school softball players. Athletes in both sports were more likely to sustain a practice injury than a game injury. This finding is in direct contrast to data presented by Radelet et al,[33] who noted a higher game injury rate in a community youth baseball program. Studies of secondary school baseball[5,7,10,33] and collegiate baseball[25,26,48] have shown similar higher injury rates for games than for practices. Secondary school softball injury data, although less well reported,[7] also revealed a greater game injury rate.

Volleyball. Middle school volleyball is one of 3 sex-matched sports in which boys had a higher injury rate than girls. Both sexes were more likely to be injured during practices, with girl and boy volleyball players at nearly 10- and 5-fold greater risk, respectively. In the only age group–related comparison compiled from a national registry, Kujala et al[14] reported on volleyball injuries in the under-15-years age group: Girls reported more injuries than boys, but no comparison between practice and games was provided. Authors[33] of a secondary school girl's volleyball injury study indicated a higher injury rate during games than during practices, and the only sex-matched secondary school report[3] noted a nearly identical injury risk per 100 athletes. At the collegiate level, Lanese et al[25] observed a higher injury rate for men than women; however, the rates were based on athletes rather than actual exposure data. Using exposure data, over a 15-year period, female collegiate volleyball players had a higher rate of injury during games than practices.[49]

Noncontact Sports: Individual

Cross-Country. Middle school girl cross-country runners reported a higher injury rate for all injuries than middle school boy cross-country runners. However, even though girl runners also had a higher injury rate for time-loss injuries than boy runners, the findings were not significant. Comparative middle school cross-country injury data are unavailable, but our findings agree with studies of high school[8,31] and collegiate[42] cross-country runners in which girl runners had a higher risk of injury than boy runners. Rauh et al[8,31] reported that girl runners were at greater risk than boy runners based on time-loss injuries; our results showed a higher but nonsignificant rate. Other studies of high school cross-country runners have shown that boys had a higher[6] or similar[3] injury risk to girl runners, so additional studies are needed. Boy and girl cross-country runners had higher injury rates during practices than during games, yet the higher risk during practices was only significant for the boys. These results support similar findings reported by Rauh et al[8,31] that cross-country runners may be at greater risk of injuries during practices than during games.

Track and Field. For middle school track-and-field athletes, girls were at higher risk for injuries as compared with boys. Although girl and boy track-and-field athletes were more likely to sustain an injury during practice than during meets, the risk was only different for the girls. Event specialist information was lacking for these middle school athletes. Published injury-rate data for secondary school track and field are varied and limited. In 1 secondary school study,[6] girls had a higher injury risk than boys, whereas another study[3] reported identical injury rates (per 100 athletes). Comparisons with our study are limited because none of the authors provided rates per AEs. Collegiate track and field data are equally sparse but in contrast to our findings. In 2 investigations,[25,26] male track runners had a greater injury rate than female track runners. None of the authors reported practice or game injury rates; thus, comparisons cannot be made.

Severity of Injury

The reporting of minor, no–time-loss injuries may provide a more accurate picture of the athletic trainer's workload. Minor injuries accounted for more than half (57%) of all reported middle school injuries for both girl and boy athletes. To date, only 3 published studies[3,30,32] have reported no–time-loss injuries, and those percentages are similar to our findings. In the only secondary school study,[3] 60% of injuries did not involve time loss. Youth football[30] (58.6%) and collegiate[32] (78%) studies have also reported similar estimates for no–time-loss injuries. Although mild injuries (1–7 days lost) for girls and for the overall sample were more likely to occur during practices, severe injuries (22 or more days lost) were more likely to occur during games, especially for boys.

Injury Location

Lower extremity injuries accounted for 70% of all days-lost injuries for middle school athletes. This finding is comparable with data reported at the high school[7,11] and collegiate levels.[35] Ankle and knee injuries have been the most commonly reported injuries in published literature,[3,5,7,34] and our results indicate that these body parts are also the most likely to be affected at the middle school level. Ankle injuries had the highest injury rate for both sexes in practices and games. Knee injuries were the second highest rate for boys and third highest rate for girls. Shin/calf injuries were the second most frequently reported injury for girls and third most frequently reported for boys. Wrist/hand/finger injuries were more common for boys. In studies assessing multiple high school sports, Beachy et al[3] observed that shin injuries were the third most reported injury for secondary school athletes, with girls reporting a higher injury frequency than boys. At the collegiate level, Lanese et al[25] noted that 7% to 8% of injuries involved the shin/calf. In summary, the high incidence of shin/calf injuries reflects the need to monitor conditioning and running programs for the middle school–aged athlete.

Specific Injury Types

Small injury numbers may warrant cautious data interpretation, regardless of injury site and sex. This is true for the rotator cuff and ACL injuries reported in this study because the small numbers limit the comparative value for middle school athletes. We will try to retain that objectivity when discussing these injuries.

Concussions

Reports on the incidence and effects of concussion have primarily focused on football players at all competitive levels. We found that 2.2% of middle school football players incurred a concussion, a value consistent with that noted by

Dompier et al[30] (2.7% of youth football players) but unlike that of Turbeville et al[20] (none among middle school–aged football players over 2 years). Our concussion injury rates of 0.067/1000 AEs and 4.0 per 100 football players are lower than those reported at the high school and collegiate levels. Gerberich et al[50] reported 19 injuries per 100 secondary school football players (24% of all injuries reported). Guskiewicz et al[51,52] stated that 5.6% of secondary school and 6.3% of collegiate football players surveyed sustained at least 1 concussion. Our lower rate may be attributed to the lack of physical maturity and the associated reduced collision intensity, but football continues to have the highest concussion rate at the middle school level.

In this study, concussions accounted for 0.4% of all injuries to girls and 1.0% of all injuries for boys. Overall, boy athletes were 3 times more likely to incur a concussion than girl athletes. In a study of youth soccer players reporting to the emergency room, Adams and Schiff[53] also found that boys were more likely (7.8%) to sustain a concussion injury than girls (0.5%). However, when we excluded concussion injuries that occurred in football players, the risk was similar. At the high school level, the risk of concussion injury by sex appears equivocal, with some studies indicating a greater risk for girls[9,39] and others[54] showing a greater risk for boys. At the collegiate level, however, the risk of concussion for sex-matched sports appears to be higher for female than male athletes.[35]

We noted that the incidence of concussion injuries was 3 times higher during games than during practices. This finding is in agreement with studies of middle school,[20] high school,[7,11,32] and collegiate athletes[35,55] and suggests that the intensity of impact and risk taking may be greater in games.

Concussion guidelines changed during the 20-year period of this study. During the 1980s and 1990s, the Colorado Medical Society[56] and the American Academy of Neurology guidelines[57] allowed return to activity if symptoms resolved within 15 minutes of the injury. The 7 athletes with no–time-loss concussions were in this category.

Anterior Cruciate Ligament Injuries

Injuries to the ACL are concerning due to the severity and the time lost from participation and because females are at greater risk of injury than males in matched sports. Only 8 ACL injuries (0.6% of all injuries) occurred during the 20-year period. The overall injury rate of 0.009/1000 AEs was much lower than the rates reported in high school studies[7,9,11] and the 0.15/1000 AEs observed by Hootman et al[35] for all collegiate sports. In collegiate football, ACL injuries accounted for 3% of all collegiate sport injuries, with an injury rate of 0.018/1000 AEs.[35] We found similar risks of ACL injuries for girl and boy middle school athletes but different risks than those reported at the high school,[36] military,[58] and collegiate[59] levels. It may be that the female middle school athletes had not reached the pubertal or maturational levels that have been suggested to affect their hormonal, structural, and neuromuscular traits and possibly put them at greater risk of injury.[36,58]

Medial Tibial Stress Syndrome

Our finding that the risk of MTSS was higher for girl athletes than boy athletes is consistent with values reported at the high school level.[60,61] For both girl and boy athletes, MTSS was more likely to be reported in practices than in games. We are unaware of any studies of MTSS by participation setting. The higher occurrence of MTSS during practices may be related to training errors, including overly repetitive activities or training regimens with short recovery times.[31,60] Further research is needed to help identify these factors.

Limitations

The primary limitation of our study was that all injury and participation data were collected from only 1 private school that has sufficient resources to employ more than 1 full-time athletic trainer. Thus, the extent to which our findings are generalizable is unknown. Therefore, these results await comparison with the results of future investigations among middle school athletes. Because only a few athletic trainers were involved in this study, evaluating and reporting of the data were performed consistently and with little variation. An additional potential limitation of our study was the possibility of nonreporting bias. Again, the injury-reporting system at Punahou School is longstanding, and the procedures for middle school athletes to report any injury, regardless of severity, have been fostered throughout the 20-year period. Some sports had too few injuries or participation numbers by sex for ample statistical comparisons. These sports may be new to the middle school population, resulting in limited participation, or the sport may have a lower injury risk. Some athletes may have had self-treated injuries (eg, tendinitis, minor contusions) that were not reported to the athletic training staff. Finally, most middle schools in the United States do not allow ninth-grade participation in middle school sports.

We hope that our initial findings for these sports will provide the impetus for all middle schools to report their data in these sports for comparative purposes. We used an injury-surveillance design, so detailed data on each injured athlete were not available. Thus, efforts should be focused toward implementing epidemiologic studies designed to identify risk factors in these athletes at the middle school level.

CONCLUSIONS

Middle school athletes who participated in an extensive interscholastic program sustained a wide variety of sport injuries. The frequency and severity of those injuries were less severe than injuries reported for the secondary school and collegiate athlete. Injuries were more likely during practices than during games. In matched middle school sports, girls exhibited a higher injury rate than their male counterparts. Football continues to have the highest injury rate for all sports. Concussions and ACL injuries were less common than at the secondary level.

With injury assessment and the cooperation of coaches, athletic trainers, and parents, modifications to training and conditioning programs can enable the middle school athlete to compete successfully with limited time lost from activity. Additional middle school injury-surveillance data are needed from a variety of settings, both public and private, and for all sports to better understand the injury patterns of the middle school athlete.

ACKNOWLEDGMENTS

For their assistance in injury management and care over the past 20 years, special thanks to Beth Ann Young, ATC; Darryl Funai, ATC; and Matt Martinson, ATC. Thanks also to Cedric Akau, MD; Timothy Olderr, MD; and Ron Hanagami, RPT, for their guidance and input while working with our many athletes.

REFERENCES

1. Emery CA. Risk factors for injury in child and adolescent sport: a systematic review of the literature. *Clin J Sport Med*. 2003;13(4):256–268.
2. The National Federation of State High School Associations. http://www.nfhs.org. Accessed June 15, 2009.
3. Beachy G, Akau CK, Martinson M, Olderr TF. High school sports injuries: a longitudinal study at Punahou School, 1988 to 1996. *Am J Sports Med*. 1997;25(5):675–681.
4. Chambers RB. Orthopaedic injuries in athletes (ages 6 to 17): comparison of injuries occurring in six sports. *Am J Sports Med*. 1979;7(3):195–197.
5. Garrick JG, Requa RK. Injuries in high school sports. *Pediatrics*. 1978;61(3):465–469.
6. McLain LG, Reynolds S. Sports injuries in a high school. *Pediatrics*. 1989;84(3):446–450.
7. Powell JW, Barber-Foss KD. Injury patterns in selected high school sports: a review of the 1995–1997 seasons. *J Athl Train*. 1999;34(3):277–284.
8. Rauh MJ, Margherita AJ, Rice SG, Koepsell TD, Rivara FP. High school cross country injuries: a longitudinal study. *Clin J Sport Med*. 2000;10(2):110–116.
9. Yard EE, Schroeder MJ, Fields SK, Collins CL, Comstock RD. The epidemiology of United States high school soccer injuries, 2005–2007. *Am J Sports Med*. 2008;36(10):1930–1937.
10. Collins CL, Comstock RD. Epidemiological features of high school baseball injuries in the United States, 2005–2007. *Pediatrics*. 2008;121(6):1181–1187.
11. Comstock RD, Knox C, Yard E, Gilchrist J. Sports-related injuries among high school athletes—United States, 2005–06 school year. *MMWR Morb Mortal Wkly Rep*. 2006;55(38):1037–1040.
12. National Center for Sports Safety. http://www.sportssafety.org/content/Home.aspx. Accessed September 21, 2009.
13. Committee on Sports Medicine and Fitness. Injuries in youth soccer: a subject review. *Pediatrics*. 2000;105(3):659–661.
14. Kujala UM, Taimela S, Antti-Poika I, Orava S, Tuominen R, Myllynen P. Acute injuries in soccer, ice hockey, volleyball, basketball, judo, and karate: analysis of national registry data. *BMJ*. 1995;311(7018):1465–1468.
15. Le Gall F, Carling C, Reilly T. Injuries in young elite female soccer players: an 8-season prospective study. *Am J Sports Med*. 2008;36(2):276–284.
16. Leininger RE, Knox CL, Comstock RD. Epidemiology of 1.6 million pediatric soccer-related injuries presenting to US emergency departments from 1990 to 2003. *Am J Sports Med*. 2007;35(2):288–293.
17. Soderman K, Adolphson J, Lorentzon R, Alfredson H. Injuries in adolescent female players in European football: a prospective study over one outdoor soccer season. *Scand J Med Sci Sports*. 2001;11(5):299–304.
18. McEwin CK, Swain J. *Clearing the Hurdles: Issues and Answers in Middle School Sports*. Westerville, OH: National Middle School Association; 2007:13–32.
19. Collins HR. Contact sports in junior high school. *Texas Med*. 1967;63(10):67–69.
20. Turbeville SD, Cowan LD, Asal NR, Owen WL, Anderson MA. Risk factors for injury in middle school football players. *Am J Sports Med*. 2003;31(2):276–281.
21. McGuine T. Sports injuries in high school athletes: a review of injury-risk and injury-prevention research. *Clin J Sport Med*. 2006;16(6):488–499.
22. Pasque CB, Hewett TE. A prospective study of high school wrestling injuries. *Am J Sports Med*. 2000;28(4):509–515.
23. Messina DF, Farney WC, DeLee JC. The incidence of injury in Texas high school basketball: a prospective study among male and female athletes. *Am J Sports Med*. 1999;27(3):294–299.
24. Yang J, Marshall SW, Bowling JM, Runyan CW, Mueller FO, Lewis MA. Use of discretionary protective equipment and rate of lower extremity injury in high school athletes. *Am J Epidemiol*. 2005;161(6):511–519.
25. Lanese RR, Strauss RH, Leizman DJ, Rotondi AM. Injury and disability in matched men's and women's intercollegiate sports. *Am J Public Health*. 1990;80(12):1459–1462.
26. Sallis RE, Jones K, Sunshine S, Smith G, Simon L. Comparing sports injuries in men and women. *Int J Sports Med*. 2001;22(6):420–423.
27. Aubry M, Cantu R, Dvorak J, et al. Summary and agreement statement of the First International Conference on Concussion in Sport, Vienna 2001: recommendations for the improvement of safety and health of athletes who may suffer concussive injuries. *Br J Sports Med*. 2002;36(1):6–10.
28. McCrory P, Johnston K, Meeuwisse W, et al. Summary and agreement statement of the 2nd International Conference on Concussion in Sport, Prague 2004. *Br J Sports Med*. 2005;39(4):196–204.
29. McCrory P, Meeuwisse W, Johnston K, et al. Consensus statement on concussion in sport: the 3rd International Conference on Concussion in Sport, held in Zurich, November 2008. *J Clin Neurosci*. 2009;16(6):755–763.
30. Dompier TP, Powell JW, Barron MJ, Moore MT. Time-loss and non–time-loss injuries in youth football players. *J Athl Train*. 2007;42(3):395–402.
31. Rauh MJ, Koepsell TD, Rivara FP, Margherita AJ, Rice SG. Epidemiology of musculoskeletal injuries among high school cross-country runners. *Am J Epidemiol*. 2006;163(2):151–159.
32. Caine D, Caine C, Maffulli N. Incidence and distribution of pediatric sport-related injuries. *Clin J Sport Med*. 2006;16(6):500–513.
33. Radelet MA, Lephart SM, Rubinstein EN, Myers JB. Survey of the injury rate for children in community sports. *Pediatrics*. 2002;110(3):e28.
34. Powell JW, Dompier TP. Analysis of injury rates and treatment patterns for time-loss and non–time-loss injuries among collegiate student-athletes. *J Athl Train*. 2004;39(1):56–70.
35. Rechel JA, Yard EE, Comstock RD. An epidemiologic comparison of high school sports injuries sustained in practice and competition. *J Athl Train*. 2008;43(2):197–204.
36. Knowles SB, Marshall SW, Bowling MJ, et al. Risk factors for injury among high school football players. *Epidemiology*. 2009;20(2):302–310.
37. Hootman JM, Dick R, Agel J. Epidemiology of collegiate injuries for 15 sports: summary and recommendations for injury prevention initiatives. *J Athl Train*. 2007;42(2):311–319.
38. Ingram JG, Fields SK, Yard EE, Comstock RD. Epidemiology of knee injuries among boys and girls in US high school athletics. *Am J Sports Med*. 2008;36(6):1116–1122.
39. Borowski LA, Yard EE, Fields SK, Comstock RD. The epidemiology of US high school basketball injuries, 2005–2007. *Am J Sports Med*. 2008;36(12):2328–2335.
40. Backx FJ, Beijer HJ, Bol E, Erich WB. Injuries in high-risk persons and high-risk sports: a longitudinal study of 1818 school children. *Am J Sports Med*. 1991;19(2):124–130.
41. Committee on Sports Medicine and Fitness. Medical conditions affecting sports participation. *Pediatrics*. 2001;107(5):1205–1209.
42. Buckley WE, Powell J. NAIRS: an epidemiological overview of the severity of injury in college football 1975–80 seasons. *J Athl Train*. 1982;17(4):279–282.

43. Knowles SB, Marshall SW, Bowling JM, et al. A prospective study of injury incidence among North Carolina high school athletes. *Am J Epidemiol*. 2006;164(12):1209–1221.
44. Shankar PR, Fields SK, Collins CL, Dick RW, Comstock RD. Epidemiology of high school and collegiate football injuries in the United States, 2005–2006. *Am J Sports Med*. 2007;35(8):1295–1303.
45. Malina RM, Morano PJ, Barron M, Miller SJ, Cumming SP, Kontos AP. Incidence and player risk factors for injury in youth football. *Clin J Sport Med*. 2006;16(3):214–222.
46. Lorish TR, Rizzo TD Jr, Ilstrup DM, Scott SG. Injuries in adolescent and preadolescent boys at two large wrestling tournaments. *Am J Sports Med*. 1992;20(2):199–202.
47. Agel J, Ransone J, Dick R, Oppliger R, Marshall SW. Descriptive epidemiology of collegiate men's wrestling injuries: National Collegiate Athletic Association Injury Surveillance System, 1988–1989 through 2003–2004. *J Athl Train*. 2007;42(2):303–310.
48. Dick R, Sauers EL, Agel J, et al. Descriptive epidemiology of collegiate men's baseball injuries: National Collegiate Athletic Association Injury Surveillance System, 1988–1989 through 2003–2004. *J Athl Train*. 2007;42(2):183–193.
49. Agel J, Palmieri-Smith RM, Dick R, Wojtys EM, Marshall SW. Descriptive epidemiology of collegiate women's volleyball injuries: National Collegiate Athletic Association Injury Surveillance System, 1988–1989 through 2003–2004. *J Athl Train*. 2007;42(2):295–302.
50. Gerberich SG, Priest JD, Boen JR, Straub CP, Maxwell RE. Concussion incidences and severity in secondary school varsity football players. *Am J Public Health*. 1983;73(12):1370–1375.
51. Guskiewicz KM, Weaver NL, Padua DA, Garrett WE Jr. Epidemiology of concussion in collegiate and high school football players. *Am J Sports Med*. 2000;28(5):643–650.
52. Guskiewicz KM, McCrea M, Marshall SW, et al. Cumulative effects associated with recurrent concussion in collegiate football players: the NCAA Concussion Study. *JAMA*. 2003;290(19):2549–2555.
53. Adams AL, Schiff MA. Childhood soccer injuries treated in U.S. emergency departments. *Acad Emerg Med*. 2006;13(5):571–574.
54. Schulz MR, Marshall SW, Mueller FO, et al. Incidence and risk factors for concussion in high school athletes, North Carolina, 1996–1999. *Am J Epidemiol*. 2004;160(10):937–944.
55. Dick RW. Is there a gender difference in concussion incidence and outcomes? *Br J Sports Med*. 2009;43(suppl 1):i46–i50.
56. Colorado Medical Society Sports Medicine Committee. Guidelines for the management of concussion in sports. *Proceedings of the Mild Brain Injury in Sports Summit*. Dallas, TX: National Athletic Trainers' Association; 1994:106–109.
57. Kelly JP, Rosenberg JH. The development of guidelines for the management of concussion in sports. *J Head Trauma Rehabil*. 1998; 13(2):53–65.
58. Mountcastle SB, Posner M, Kragh JF Jr, Taylor DC. Gender differences in anterior cruciate ligament injury vary with activity: epidemiology of anterior cruciate ligament injuries in a young, athletic population. *Am J Sports Med*. 2007;35(10):1635–1642.
59. Agel J, Arendt EA, Bershadsky B. Anterior cruciate ligament injury in national collegiate athletic association basketball and soccer: a 13-year review. *Am J Sports Med*. 2005;33(4):524–530.
60. Plisky MS, Rauh MJ, Tank RT, Heiderscheit BC, Underwood FB. Medial tibial stress syndrome among high school cross country runners: incidence and risk factors. *J Orthop Sports Phys Ther*. 2007; 37(2):40–47.
61. Bennett JE, Reinking MF, Pluemer B, Pentel A, Seaton M, Killan C. Factors contributing to the development of medial tibial stress syndrome in high school runners. *J Orthop Sports Phys Ther*. 2001; 31(9):504–510.

Address correspondence to Glenn Beachy, MS, ATC, Punahou School, 1601 Punahou Street, Honolulu, HI 96822. Address e-mail to gbeachy@punahou.edu.

38th International Society of Biomechanics in Sport Conference, Physical conference cancelled, Online Activities: July 20-24, 2020

SEX-BASED ANALYSIS OF THE BIOMECHANICS OF PITCHING

Madison J. Blankenship, Hunter L. Frisk, Evan M. Martin, and William P. Ebben

Biomechanics Research Unit, Exercise Science Research Laboratory, Lakeland University, Plymouth, WI

This study assessed sex-based differences in the lower extremity kinetics and ball velocity during pitching. Fifteen men baseball players and fifteen women softball players threw fastballs on two force platforms, to assess propulsive and landing biomechanics. Doppler radar was used to assess ball velocity. Kinetic and kinematic data comparing men and women were analyzed with independent samples *t*-test. Paired samples *t*-test were used to assess difference between the propulsive and landing phases. Pearson's bivariate correlations were used to assess the relationship between study variables and ball velocity. Few sex-based difference in the magnitude and rate of propulsive force development exist. Sex based differences ($p < 0.05$) were found for all but one landing phase variable. None of the biomechanical variables assessed were related to ball velocity.

KEYWORDS: baseball, softball, gender, ground reaction force, Doppler radar

INTRODUCTION: Athlete biomechanics is believed to have an important influence on pitching performance. Previous research examined a variety of biomechanical variables and their effect on softball pitching. Similarly, biomechanics studies assessed baseball pitching. Only one study focused on the sex-based differences in the kinematics and kinetics variable associated with pitching a baseball.

Research focused on the biomechanics of softball pitching included assessment of the vertical plane (Nimphius et al., 2016; Oliver & Plummer, 2011) or vertical, medio-lateral and anterior-posterior planes (Guido et al., 2009) of motion. These studies evaluated only the push-off phase (Nimphius et al., 2016) or landing phase (Guido et al., 2009; Oliver & Plummer, 2011) of pitching. In the process, the role of kinetic variables on softball pitching mechanics such as lower limb angles and stride length have been studied (Guido et al., 2009; Oliver and Plummer, 2011).

In addition to studies assessing softball pitching, biomechanics studies of baseball pitching have used one force platform to assess the push-off leg (Elliot et al., 1988; Oyama & Myers, 2017), the landing leg (Guido et al., 2009) or two platforms to assess each limb (Kageyama et al., 2015; MacWilliams et al., 1998). Vertical and anterior-posterior kinetics (Elliott at al., 1988; Kageyama et al., 2015; Oyama & Myers 2017) or vertical, medio-lateral, and anterior posterior kinetics (Guido & Warner, 2012; Kageyama et al., 2015; MacWilliams et al., 1998; McNally et al., 2015) were assessed. These studies sought to evaluate the role of ground reaction forces on upper body movements and pitching mechanics (Elliott et al., 1988; Guido & Werner, 2012; Kageyama et al., 2015), or to compare the kinetic and kinematic differences based on subject age and level (Kageyama et al., 2015).

Baseball pitching research assessed wrist or pitch velocity as well. Propulsive phase kinetics were correlated to wrist velocity (MacWilliams et al., 1998). In contrast, ground reaction forces were not correlated with the ball velocity (Oyama & Myers, 2017).

Only one study included the sex-based analysis of pitching a baseball, demonstrating a number of similarities and few differences between the sexes when using motion analysis (Chu et al., 2009). Another study compared the analysis of softball players (Guido et al., 2009) to baseball players (MacWilliams et al., 1998). No previous study has examined the biomechanics of men and women baseball and softball pitchers, respectively within the same study. Therefore, the purpose of this study was to assess the relationship between propulsive and landing phase kinetics, athlete whole body velocity, and pitched ball speed as well as assess sex-based differences in these kinetic variables and the relationship between these variables and ball speed.

METHODS: Subjects included 15 men (mean ± SD, age = 19.47 ± 1.18 yr; body mass = 84.96 ± 10.75 kg; height = 179.83 ± 8.70 cm) and 15 women (mean ± SD, age = 20.07 ± 2.17 yr; body mass = 80.32 ± 22.94 kg; height = 169.33 ± 5.97 cm). Subjects included current or former high school and college baseball and fast-pitch softball pitchers. The subjects were informed of the risks associated with the study and provided informed written consent. The study was approved by the institution's internal review board.

Subjects performed a general, dynamic, specific, and sport-specific warm-up. The general warm-up included low intensity jogging for approximately four minutes. The dynamic warm-up included exercises performed sport specific planes of motions, with increasing intensity. The specific warm-up included activating all of the muscles used in the throwing motion. The sport specific warm-up included a range of warm up pitches in increasing intensity from approximately 50-100% of the subject's maximum velocity.

During testing, all subjects threw six fastballs from the full wind-up motion with at least fifteen seconds rest between pitches. The subjects pitched off of a pitching rubber that was bolted to a force platform and threw into a net with a strike zone ten meters away. The test pitches were performed on two force platforms (Accupower, Advanced Mechanical Technologies Incorporated, Watertown, MA, USA) in series, which were countersunk and mounted flush to the floor. The first force platform captured the subject's propulsive phase and the second captured the landing phase of the pitch. The force platforms were calibrated prior to the testing session. Data were acquired at 1000 Hz and analyzed in real time with proprietary software (Accupower, Advanced Mechanical Technologies Incorporated, Watertown, MA, USA).

Velocity of each pitch was determined by Doppler radar (Speedster III, Bushnell Outdoor Products, Overland Park, KS). The three highest velocity pitches were included for analysis consistent with previous research (Nimphius et al, 2016).

Data were analyzed with a statistical software program (SPSS 26.0, International Business Machines Corporation, Armonk, New York) using independent samples *t*-tests to assess the differences in subjects background, pitch velocity, propulsive phase biomechanics, and landing phase biomechanics. A paired samples *t*-test was used to determine the differences between propulsive phase and landing phase ground reaction forces. Pearson's bivariate correlations were used to assess the relationship between the biomechanical variables and pitched ball velocity. Intraclass correlation coefficients (ICC) and coefficients of variation (CV) were determined for all dependent variables. The *a priori* alpha level was set at $p \leq 0.05$. All data are expressed as means ± SD.

RESULTS: Subject age, weight, height, years of high school and college pitching experiences were not statistically different ($p \geq 0.05$). Men were taller than women ($p = 0.001$). Fastball velocity was significantly greater ($p = 0.001$) for men (33.17 ± 2.21 $m \cdot s^{-1}$) than women (22.52 ± 1.47 $m \cdot s^{-1}$). Table 1 shows the time, distance, and velocity of the subject from propulsive to landing phase. Results of the analysis of the propulsive phase and landing phases are shown in Tables 2 and 3, respectively. Table 4 shows the comparison of the propulsive and landing phase kinetics. There was no correlation between any of the biomechanical variables assessed and ball velocity for either men or women ($p \geq 0.05$). The trial-to-trial reliability of the dependent variables were assessed using average measures Intraclass correlation coefficients (ICC) and coefficients of variation (CV). The ICC's for the test exercises and all dependent variables ranged from 0.77 to 0.96 for the horizontal GRF data, and 0.87 to 0.98 for the vertical GRF data. Coefficients of variation for all data ranged from 13.9% to 28.5%.

Table 1. Mean ± SD data for the baseball and windmill softball for time, distance, and whole body velocity from peak V GRF during propulsion to peak V GRF during landing (N = 30).

	Men (N=15)	Women (N=15)	Significance
Distance (m)	1.57 ± 0.10	1.57 ± 0.14	$p = 0.27$
Time (ms)	0.41 ± 0.09	0.38 ± 0.04	$p = 0.86$
Velocity ($m \cdot s^{-1}$)	4.08 ± 0.86	4.17 ± 0.47	$p = 0.71$

Table 2. Propulsive phase kinetic data for the pitched baseball and softball (N = 30).

	Men (N=15)	Women (N=15)	Significance
V-GRF (N)	1124.48 ± 150.86	1101.83 ± 263.45	$p = 0.78$
V-GRF/BW	1.36 ± 0.16	1.42 ± 0.10	$p = 0.23$
H-GRF (N)	419.32 ± 83.65	352.83 ± 85.03	$p = 0.039$
H-GRF/BW	0.51 ± 0.10	0.46 ± 0.10	$p = 0.24$
H:V	0.37:1 ± .05:1	0.32:1 ± .07:1	$p = 0.04$
V-RFD ($N \cdot s^{-1}$)	13240.45 ± 1767.76	13060.70 ± 3107.07	$p = 0.85$
V-RFD/BW ($N \cdot s^{-1}$)	16.00 ± 1.80	16.83 ± 1.18	$p = 0.15$
H-RFD ($N \cdot s^{-1}$)	4617.50 ± 931.18	3883.98 ± 945.51	$p = 0.039$
H-RFD/BW ($N \cdot s^{-1}$)	5.59 ± 1.12	5.09 ± 1.15	$p = 0.24$

V = vertical; H = horizontal anterior; GRF = ground reaction force; GRF/BW = ground reaction force normalized to body weight; H:V = ratio of the vertical to horizontal anterior ground reaction force; RFD = rate of force development.

Table 3. Landing phase kinetic data for the pitched baseball and softball (N = 30).

	Men (N=15)	Women (N=15)	Significance
V-GRF (N)	1190.20 ± 184.08	1477.56 ± 325.07	$p = 0.006$
V-GRF/BW	1.43 ± 0.13	1.91 ± 0.13	$p \leq 0.001$
H-GRF (N)	366.06 ± 108.15	288.28 ± 65.76	$p = 0.024$
H-GRF/BW	0.44 ± 0.10	0.38 ± 0.10	$p = 0.14$
H:V	0.31:1 ± 0.07:1	0.20:1 ± 0.05:1	$p = 0.001$
V-RFD ($N \cdot s^{-1}$)	14522 ± 2231.73	18038.27 ± 3962.85	$p = 0.006$
V-RFD/BW ($N \cdot s^{-1}$)	17.45 ± 1.53	23.36 ± 2.59	$p \leq 0.001$
H-RFD ($N \cdot s^{-1}$)	6925.94 ± .2043.39	5452.83 ± 1245.65	$p = 0.024$
H-RFD/BW ($N \cdot s^{-1}$)	8.25 ± 1.90	7.22 ± 1.88	$p = 0.14$

V = vertical; H = horizontal anterior; GRF = ground reaction force; GRF/BW = ground reaction force normalized to body weight; H:V = ratio of the vertical to horizontal anterior ground reaction force; RFD = rate of force development.

Table 4. Comparison of the kinetics of the propulsive and landing phases (N = 30).

	Propulsive Phase	Landing Phase	Significance
V-GRF (N)	1113.05 ± 211.26	1333.88 ± 297.88	$p \leq 0.001$
V-GRF/BW	1.39 ± 0.13	1.67 ± 0.30	$p \leq 0.001$
H-GRF (N)	386.07 ± 89.51	327.17 ± 96.43	$p = 0.013$
H-GRF/BW	0.48 ± 0.10	0.41 ± 0.10	$p = 0.014$
V-RFD ($N \cdot s^{-1}$)	13150.58 ± 2485.44	16280.61 ± 3630.65	$p \leq 0.001$
V-RFD/BW ($N \cdot s^{-1}$)	16.41 ± 1.55	20.41 ± 3.66	$p \leq 0.001$
H-RFD ($N \cdot s^{-1}$)	4250.74 ± 994.65	6189.38 ± 1823.74	$p \leq 0.001$
H-RFD/BW ($N \cdot s^{-1}$)	5.34 ± 1.15	7.74 ± 1.93	$p \leq 0.001$

V = vertical; H = horizontal anterior; GRF = ground reaction force; GRF/BW = ground reaction force normalized to body weight; H:V = ratio of the vertical to horizontal anterior ground reaction force; RFD = rate of force development.

DISCUSSION: This is the first study to compare the biomechanics of men and women baseball and softball players, respectively. Results show a difference in propulsive H:V when data were normalized to body weight. Additionally, almost all of the landing phase kinetic demands are different between men and women. Others showed more knee flexion upon landing for women compared to men, when assessing baseball players (Chu, et al., 2009).

In the current study, men demonstrated vertical ground reaction forces during landing that were 1.43 times body weight, compared to women who produced 1.9 times body weight. Others showed that men produced 1.5 times body weight (MacWilliams et al., 1998), and women produced 1.39 (Guido et al., 2009) to 1.79 (Oliver & Plummer, 2011) times body weight.

The posteriorly directed horizontal ground reaction forces during landing in the current study were 0.44 times body weight for men and 0.38 times body weight for women. This was lower

for men, but similar to women when compared to other studies, which found these values to be 0.72 and 0.36 times body weight for men (MacWilliams et al., 1998) and women (Oliver & Plummer, 2011), respectively. In the current study, men generated higher propulsive H:V and horizontal landing ground reaction force, whereas women develop more vertical force and a higher rate of force development during landing. Thus, the propulsive H:V in this study was significantly higher for men compared to women, with values similar to those previously shown for men (Elliott et al., 1998; Oyama & Myers, 2017).

The current study demonstrated no differences in whole body velocity measured from propulsive to landing phase peak vertical ground reaction forces. The only other sex-based analysis of pitching showed that compared to men, women had a greater time from stride foot contact to ball release when throwing a baseball (Chu et al., 2009).

Men threw with more velocity than women in the current study, consistent with previous research comparing men and women baseball players (Chu et al., 2009). The men in this study threw approximately 33.2 $m \cdot s^{-1}$, compared to others who threw 34.87 $m \cdot s^{-1}$ (Guido & Werner, 2012) and 35.2 $m \cdot s^{-1}$ (Kageyama et al., 2015). The women in the current study threw at approximately 22.5 $m \cdot s^{-1}$ compared to 24 $m \cdot s^{-1}$ (Oliver & Plummer, 2011) and 25 $m \cdot s^{-1}$ (Guido et al., 2009).

None of the biomechanical variables in the current study were correlated to ball velocity consistent with some reports (Guido et al., 2009). Others found a small number of variables were related to pitched softball velocity such as peak propulsive phase vertical ground reaction force and time between peak forces (Nimphius, et al., 2016) as well as vertical ground reaction force during the landing phase (Oliver and Plummer, 2011).

CONCLUSION: When normalized to body weight, there are few sex-based differences in the propulsive phase of pitching. Men rely more on a greater H:V during this phase, and their training strategies should emphasize horizontal more than vertical force production. Compared to men, women demonstrate higher vertical ground reaction forces and rates of force development during the landing phase. Training strategies for women should increase their capability to manage larger magnitudes and rates of vertically directed force.

REFERENCES:

Chu, Y., Fleisig, G.S., Simpson, K.J. & Andrews, J.R. (2009). Biomechancial comparison between elite female and male baseball pitchers. *Journal of Applied Biomechanics*, 25, 22-31.

Elliott, B., Grove, J.R. & Gibson, B. (1988). Timing of lower limb drive and throwing limb movement in baseball pitching. *International Journal of Sport Biomechanics*, 4, 59-67.

Guido, J.A. & Werner, S.L. (2012). Lower-extremity ground reaction forces in collegiate baseball pitchers. *Journal of Strength and Conditioning Research*, 26 (7), 1782-1785.

Guido, J.A., Werner, S.L. & Meister, K. (2009). Lower-extremity ground reaction forces in youth windmill softball pitchers. *Journal of Strength and Conditioning Research*, 23 (6),1873-1876.

Kageyama, M., Sugiyama, T., Kanehisa, H. & Meada, A. (2015). Difference between adolescent and collegiate baseball pitchers in the kinematics and kinetics of lower limb and trunk during pitching motion. *Journal of Sports Science and Medicine*, 14, 246-255.

MacWilliams, B.A., Choi, T., Perezous, M.K., Chao, E.Y. & McFarland, E.G. (1998). Characteristic ground reaction forces in baseball pitching. *The American Journal of Sports Medicine,* 26 (1), 66-71.

Nimphius, S., McGuigan, T.J. Suchomel, T.J. & Newton, R.U. (2016). Variability of a force signature during windmill softball pitching and relationship between discrete force variables and pitch velocity. *Human Movement Science,* 47, 151-158.

Oliver, G.D. & Plummer, H. (2011). Ground reaction forces, kinematics, and muscle activation during the windmill softball pitch. *Journal of Sports Sciences*, 29 (10), 1071-1077.

Oyama, S. & Myers, J.B. (2017). The relationship between the push off ground reaction force and ball speed in high school baseball pitchers. *Journal of Strength and Conditioning Research* 32 (5), 1324-1328.

ACKNOWLEDGEMENTS: This study was funded by a Clifford D. Feldmann Foundation research grant.

Research Article

Handgrip Strength: A Comparison of Values Obtained From the NHANES and NIH Toolbox Studies

Richard W. Bohannon, Ying-Chih Wang, Sheng-Che Yen, Kimberly A. Grogan

Importance: Handgrip dynamometry is probably the most commonly used method to characterize overall human muscle strength.

Objective: To compare and summarize grip strength measurements obtained from two population-based studies.

Design: Secondary data analysis.

Setting and Participants: Data from (1) the 2011–2014 National Health and Nutrition Examination Survey (NHANES) with 13,918 participants and (2) the 2011 normative phase of the National Institutes of Health (NIH) Toolbox project with 3,594 participants.

Outcomes and Measures: The NHANES values used were the mean and best of three trials; the NIH Toolbox value used was the one maximum trial after a practice trial.

Results: General linear model analysis revealed that values obtained from the NIH Toolbox differed from NHANES best values but not from NHANES mean values. The analysis also indicated, regardless of the values used, that grip strength differed significantly between dominant and nondominant sides, males and females, and age groups. We provide updated reference values for handgrip strength.

Conclusions and Relevance: On the basis of these analyses, we summarize grip strength measures obtained from the NHANES and NIH Toolbox for side, gender, and age group strata. Reference values are essential to assist in the interpretation of testing results and clinical decision making.

Muscle strength, defined as the force or torque brought by muscle to bear on the environment, is essential for maintaining and changing the position of body segments and the body as a whole. Muscle strength, therefore, plays an important role in everyday function (Bohannon, 2015; Vaapio et al., 2011). It also serves as an indicator of present health status and a predictor of future health status (Bohannon, 2015).

There are numerous options for measuring muscle strength, but handgrip dynamometry is probably used most often to characterize overall human muscle strength (Bohannon, 2008). Measurements of grip strength obtained by dynamometry are easily procured and have been shown to be reliable (Bohannon, 2017), valid (Turner & Ebrahim, 1992), and responsive (Kim et al., 2014). In addition, normative reference values have been published that can be used to interpret grip strength measurements obtained from individuals and groups of interest.

Optimally, reference values used for interpreting status should have been obtained from a population representative sample (Ritchie & Palomaki, 2004) within the past 15 or 20 yr (Strauss et al., 2006). Numerous peer-reviewed studies provide such values. Specifically, reference values for grip strength that fulfill these criteria have been reported for residents of several countries, among them Great Britain (Dodds et al., 2014), Australia (Massy-Westropp

et al., 2011), Canada (Wong, 2016), Korea (Shim et al., 2013), Germany (Günther et al., 2008), and Japan (Seino et al., 2014). The importance of these reference values notwithstanding, caution must be exercised in applying them to residents of other countries (Dodds et al., 2016; Massy-Westropp et al., 2011).

Our impetus in undertaking the current study was the limited availability of current reference values over the age span for grip strength of U.S. residents. We were aware of Perna et al.'s (2016) recent publication addressing such values, but it focused on the combined strength of both hands. We examined the literature and identified two large-scale, population-based studies that provide sufficiently current data to generate such reference values for U.S. residents: the National Health and Nutrition Examination Survey (NHANES) and the National Institutes of Health (NIH) Toolbox.

The NHANES (Centers for Disease Control and Prevention [CDC], 2011; Perna et al., 2016) is a survey research program conducted by the National Center for Health Statistics (NCHS), housed within the CDC, to assess the health and nutritional status of U.S. adults and children and to track changes over time. Although the NHANES data collection has been ongoing for decades, handgrip dynamometer assessment was not included in the data collection until 2011. The handgrip component was added to provide nationally representative data on muscle strength; prevalence estimates of children with poor muscle strength; and data to study the association between muscle strength and other health conditions and risk factors, such as obesity, physical activity, and dietary patterns.

The NIH Toolbox (Beaumont et al., 2013; Gershon et al., 2013; Reuben et al., 2013) was a research study that aimed to develop a multidimensional set of brief royalty-free measures that researchers can use to assess cognitive, sensory, motor, and emotional function in people ages 3–85 yr. This entire set of measures can be administered to study participants in 2 hr or less, across diverse study designs and settings. Motor domain experts identified handgrip dynamometry as a proxy for upper extremity muscle strength. Toolbox measures have been normed and validated in a broad sample of the U.S. population in a cross-sectional data collection in 2011 (Beaumont et al., 2013).

The primary purpose of this project, therefore, was to compare and summarize grip strength measurements obtained from two population-based studies: the NHANES and the NIH Toolbox. We hypothesized that grip strength measurements from the two studies would differ significantly, as would measurements obtained from males and females, dominant and nondominant sides, and different age groups.

Method

The NHANES providing data for our study was approved by the NCHS research ethics review board; the NIH Toolbox study was approved by the Northwestern University institutional review board. Further approval was not sought for our study because the data used were free of personal identifiers.

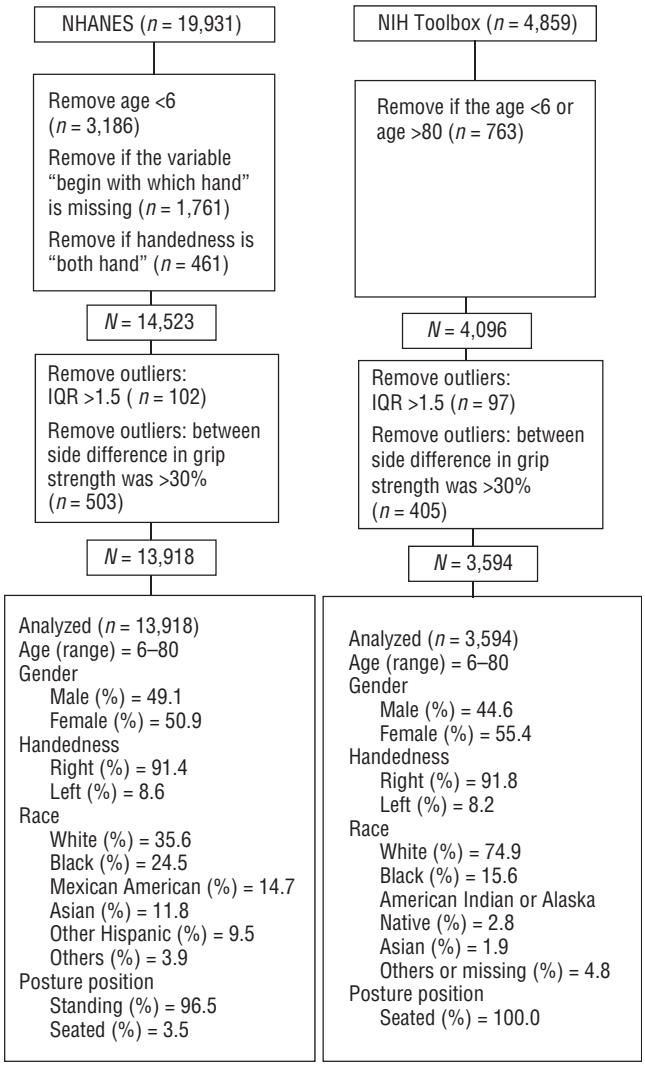
NHANES Data

We used NHANES data from 2011 to 2014. The data were acquired from a stratified multistage probability sample of civilian noninstitutionalized U.S. residents.

Participants.

Of the initial 19,931 data records, 3,186 participants were removed because they were younger than age 6 yr (no handgrip data), and 2,222 participants were removed because of missing values (e.g., begin with which hand, handedness). In addition, 605 participants were excluded as outliers because their grip strength values were greater than 1.5 interquartile range (IQR) of the same sex and age group or because their between-sides difference in grip strength was ≥ 30%. Thus, data from 13,918 remaining participants (ages between 6 and 80) were included in the final analysis with a comparable representation of males (49.1%) and females (50.9%). By self-report, 91.4% of the sample was right-hand dominant. Although the largest percentage of the sample was White (35.6%), other races were represented. Figure 1 presents the flowchart of the data management process.

Figure 1. Data cleaning and merging flowchart.



Note. IQR = interquartile range; NHANES = National Health and Nutrition Examination Survey; NIH = National Institutes of Health.

Procedures.

A detailed description of testing procedures can be found in the NHANES *Muscle Strength Procedures Manual* (CDC, 2011). In brief, the muscle strength/grip test involved the measurement of isometric grip strength with a calibrated Takei Digital handgrip dynamometer (Takei Scientific Instruments, Niigata City, Japan). Participants were randomly assigned to start the test with their dominant or nondominant hand. The grip test was performed in the standing position unless the participant was physically limited. A practice trial was performed with the hand opposite the hand tested first, unless the participant had only one hand eligible for the test.

Before the test, a test administrator adjusted the grip size of the dynamometer until the second joint of the participant's index finger was at a 90° angle on the handle. Participants were asked to squeeze the dynamometer as hard as possible with each hand with the elbow fully extended at the side. The head was straight, the wrist was neutral, and the feet were hip width apart and even. There were three testing trials for each hand. Best values, expressed in kilograms, were determined for each hand. For this study, the mean values, averaged from three trials, were also calculated for comparison purposes. Under these two conditions, data were labeled *best* and *mean* values, respectively.

NIH Toolbox

We used NIH Toolbox data from the muscle strength/grip test component of the Motor Domain, obtained from August to November 2011.

Participants.

The norming sample included people with the following characteristics: community-dwelling and noninstitutionalized; ages 3–85 yr; capable of following test instructions (in English or Spanish); and able to give informed consent or, in the case of children, give assent with accompanying informed consent by proxy (i.e., parent or guardian). Data were collected at 10 sites (Atlanta, Chicago [Oak Brook], Cincinnati, Columbus, Dallas, Los Angeles, Minneapolis, Philadelphia, Phoenix, St. Louis) from primarily urban- and suburban-dwelling participants, and sampling was stratified by age, gender, and primary language (English or Spanish). Age was stratified into 21 age bands, within which target quotas were set relative to the U.S. population distribution of race, ethnicity, and level of education (parents' education for children). Detailed norming plans for the NIH Toolbox have been described (Beaumont et al., 2013).

Of the initial 4,859 data records, 763 were removed because the age of the participant was younger than 6 yr or older than 80 yr. Moreover, 502 were excluded as outliers because their grip strength values were greater than 1.5 IQR from

the mean for other participants of the same sex and age group or because their between side difference in grip strength was ≥ 30% (Figure 1). Thus, data from 3,594 remaining participants (ages between 6 and 80) were included in the final analysis with somewhat fewer males (44.6%) than females (55.4%). By self-report, 91.8% of the sample was right-hand dominant. Although the majority of the sample was White (74.9%), other races were represented. All participants either consented to the study or assented and participants' parents or guardians provided written consent after being informed about the study's purpose and procedures.

Procedures.

Detailed descriptions of the protocol are provided in the *NIH Toolbox Administration Manual* (NIH, 2012). Briefly, a calibrated digital Jamar® dynamometer (Patterson Medical Ltd., Warrenville, IL) with its handle in its second position was used. Participants squeezed the dynamometer while seated with their arms by their sides, elbows flexed 90°, and forearms in a neutral position. A single submaximal practice trial was completed with each hand, followed after at least 30 s by a single maximal trial of 3 to 4 s from each hand. Participants were encouraged by the examiner who chanted "harder, harder, harder." Data were recorded in pounds but later converted to kilograms.

Statistical Analysis

All analyses were performed using IBM SPSS Statistics (Version 23; IBM Corp., Armonk, NY). To compare handgrip measures, age was stratified into 18 age bands: each year of age from 6 through 17, 18–29, 30–39, 40–49, 50–59, 60–69, and 70–80 yr old, consistent with the NIH sampling plan. Two (study: NHANES vs. NIH Toolbox) × 2 (gender: male vs. female) × 2 (side: dominant vs. nondominant) × 18 (age group) general linear model (GLM) analyses were conducted: (1) NHANES best values versus NIH handgrip forces and (2) NHANES mean values versus NIH handgrip forces. For independent variables found to have a significant main or interactive effect on grip strength, pairwise post hoc comparisons between NHANES and NIH forces were conducted using GLM. Separate comparisons were completed for the dominant and nondominant sides. Descriptive statistics were tabulated. On the basis of the numerous hypothesis tests conducted and a desire to reduce the risk of Type 1 error, a significant level of $p < .005$ was adopted as an indicator of statistical significance.

Results

The GLM comparing NHANES best values to NIH handgrip values demonstrated that, overall, grip strength values were higher for the NHANES group than the NIH Toolbox group ($F = 98.6$, $p < .001$), for males than for females ($F = 3{,}967.6$, $p < .001$), for the dominant side than the nondominant side ($F = 9{,}497.9$, $p < .001$), and for some age groups than others ($F = 1{,}406.9$, $p < .001$). Pairwise comparisons showed significant differences in grip strength ($p < .005$) between the NHANES and NIH Toolbox studies in five age groups of the dominant side of males, six age groups of the dominant sides of females, four age groups of the nondominant side of males, and four age groups of the nondominant side of females.

The GLM comparing NHANES mean values to NIH handgrip values demonstrated that, overall, grip strength values were not higher for the NHANES group than the NIH Toolbox group ($F = 2.6$, $p = .105$). In contrast, grip strength values were higher for males than for females ($F = 3{,}984.0$, $p < .001$), for the dominant sider than the nondominant side ($F = 1{,}869.5$, $p < .001$), and for some age groups than others ($F = 1{,}522.5$, $p < .001$). Pairwise comparisons showed significant differences in grip strength ($p < .005$) between the NHANES and NIH Toolbox studies in one age group of the dominant side of males, two age groups of the dominant side of females, one age group of the nondominant side of males, and three age groups of the nondominant side of females.

Tables 1 and 2 provide summary grip strength statistics for each study stratified by side, gender, and age. These statistics have potential as a reference for interpreting individual handgrip performance.

Table 1. Summary of Handgrip Strength Measurements (kg) Obtained From the Dominant Hand of Male and Female Participants

	M (*SD*), *n*			*F* (*p*)	
Gender (age, yr)	NHANES Mean	NHANES Best	NIH Tool Box	NHANES Mean vs. Toolbox	NHANES Best vs. Toolbox
Male (6)	10.0 (2.2), 211	10.7 (2.3), 213	9.6 (2.8), 95	1.6 (.206)	13.5 (.000)
Male (7)	11.8 (2.5), 213	12.6 (2.6), 213	11.2 (3.4), 113	2.6 (.106)	17.1 (.000)
Male (8)	13.2 (2.9), 202	14.1 (2.9), 202	13.4 (3.5), 98	0.2 (.636)	3.7 (.055)
Male (9)	15.3 (3.0), 196	16.4 (3.0), 196	16.5 (3.8), 101	8.0 (.005)	0.0 (.878)
Male (10)	17.5 (3.5), 193	18.6 (3.6), 193	17.9 (3.6), 102	1.2 (.277)	2.6 (.106)
Male (11)	20.2 (4.4), 176	21.4 (4.7), 176	20.5 (4.1), 97	0.2 (.697)	2.6 (.105)
Male (12)	22.9 (4.9), 152	24.1 (5.1), 152	23.6 (5.3), 91	1.0 (.307)	0.5 (.501)
Male (13)	28.2 (6.2), 157	29.7 (6.4), 157	28.2 (6.8), 105	0.0 (.973)	3.4 (.066)
Male (14)	32.1 (6.2), 151	34.0 (6.5), 151	33.5 (7.5), 102	2.6 (.110)	0.3 (.611)
Male (15)	36.8 (6.8), 143	38.8 (7.2), 143	36.7 (7.8), 96	0.0 (.870)	4.9 (.028)
Male (16)	39.3 (7.1), 166	41.4 (7.4), 166	39.1 (7.8), 89	0.0 (.849)	5.6 (.019)
Male (17)	40.9 (7.1), 133	43.2 (7.3), 133	44.3 (9.8), 104	9.6 (.002)	1.1 (.291)
Male (18–29)	44.5 (8.1), 1,129	47.2 (8.6), 1,129	46.3 (8.2), 68	2.9 (.088)	0.7 (.397)
Male (30–39)	47.2 (8.6), 801	49.9 (9.0), 801	43.6 (10.7), 56	9.2 (.003)	25.3 (.000)
Male (40–49)	45.5 (7.6), 756	47.8 (7.9), 757	43.2 (9.3), 75	6.1 (.014)	22.6 (.000)
Male (50–59)	42.1 (7.6), 721	44.0 (7.8), 724	42.7 (10.2), 77	0.4 (.538)	1.9 (.164)
Male (60–69)	38.9 (8.1), 695	40.8 (8.3), 696	37.3 (9.9), 57	2.0 (.160)	8.7 (.003)
Male (70–80)	33.2 (7.5), 637	34.6 (7.8), 637	33.0 (9.6), 77	0.0 (.833)	2.8 (.092)
Female (6)	9.6 (2.1), 193	10.4 (2.2), 194	8.8 (2.8), 99	6.4 (.012)	26.1 (.000)
Female (7)	10.8 (2.2), 187	11.6 (2.4), 187	11.8 (3.4), 107	9.2 (.003)	0.3 (.607)
Female (8)	12.8 (2.9), 177	13.7 (2.9), 177	12.8 (3.3), 91	0.0 (.950)	4.6 (.032)
Female (9)	14.6 (3.1), 190	15.5 (3.2), 190	14.9 (3.3), 99	0.8 (.385)	2.0 (.154)
Female (10)	17.0 (3.3), 171	18.0 (3.4), 171	17.4 (4.0), 106	0.9 (.331)	1.9 (.171)
Female (11)	19.7 (4.1), 219	20.8 (4.4), 219	19.9 (3.9), 95	0.3 (.612)	3.0 (.084)
Female (12)	22.8 (4.8), 154	24.0 (4.9), 154	23.4 (4.5), 108	1.1 (.303)	1.1 (.305)
Female (13)	24.5 (5.3), 150	25.8 (5.3), 150	25.2 (5.4), 94	1.1 (.299)	0.6 (.454)
Female (14)	26.2 (4.7), 156	27.6 (4.9), 156	26.6 (6.1), 113	0.2 (.618)	2.3 (.132)
Female (15)	27.4 (4.5), 146	28.8 (4.8), 146	28.2 (5.0), 100	1.8 (.177)	1.0 (.323)
Female (16)	27.4 (4.7), 177	28.8 (4.8), 177	28.5 (5.3), 99	3.6 (.058)	0.1 (.739)
Female (17)	27.7 (5.6), 129	29.2 (5.7), 129	28.6 (5.6), 104	1.7 (.198)	0.5 (.478)
Female (18–29)	29.1 (5.2), 1132	30.7 (5.4), 1,134	29.0 (7.3), 166	0.1 (.780)	13.5 (.000)
Female (30–39)	30.0 (5.4), 830	31.6 (5.7), 831	29.1 (6.1), 204	4.6 (.033)	31.8 (.000)
Female (40–49)	29.4 (5.5), 836	31.0 (5.7), 838	29.4 (6.5), 143	0.0 (.957)	9.1 (.003)
Female (50–59)	27.5 (5.4), 780	28.9 (5.6), 782	27.1 (6.4), 98	0.5 (.473)	8.8 (.003)
Female (60–69)	25.3 (5.1), 738	26.6 (5.2), 740	23.0 (6.5), 88	14.6 (.000)	35.3 (.000)
Female (70–80)	21.0 (5.2), 700	22.0 (5.3), 704	20.7 (5.2), 77	0.2 (.682)	4.4 (.036)

Note. *M* = mean; NHANES = National Health and Nutrition Examination Survey; NIH = National Institutes of Health; *SD* = standard deviation.

Discussion

Much of the research proposing to characterize the grip strength of people in the United States across the age span is outdated and used small convenience samples. Both NHANES and NIH Toolbox projects have remedied these issues by recently examining large population-based samples of U.S. residents. Both of the projects have rigorous sampling plans, data collection procedures, and ongoing quality control practices. For the NHANES, NCHS staff and field supervisors regularly monitored examiners. Retraining sessions are conducted periodically with the examiners to reinforce the proper protocols and techniques. All data are reviewed systematically for logical or operational inconsistencies and examiner errors. For the NIH Toolbox, examiners were trained on standardized procedures before data collection, and an administration manual (NIH, 2012) was provided in both English and Spanish for all sites.

Table 2. Summary of Hand Grip Strength Measurements (kg) Obtained From the Nondominant Hand of Male and Female Participants

	M (*SD*), *n*			*F* (*p*)	
Gender (age, yr)	NHANES Mean	NHANES Best	NIH Tool Box	NHANES Mean vs. Toolbox	NHANES Best vs. Toolbox
Male (6)	9.7 (2.1), 210	10.4 (2.2), 213	9.5 (3.0), 95	0.2 (.623)	8.8 (.003)
Male (7)	11.3 (2.5), 213	12.1 (2.6), 213	11.0 (3.4), 113	0.8 (.358)	11.5 (.001)
Male (8)	12.8 (2.9), 202	13.7 (3.0), 202	13.2 (3.5), 98	0.8 (.361)	1.7 (.200)
Male (9)	14.6 (2.8), 196	15.6 (3.0), 196	15.7 (3.7), 101	6.9 (.009)	0.0 (.862)
Male (10)	16.5 (3.3), 193	17.6 (3.5), 193	17.0 (3.5), 102	1.4 (.234)	1.5 (.221)
Male (11)	19.5 (4.0), 176	20.7 (4.1), 176	19.3 (4.0), 97	0.1 (.700)	7.2 (.008)
Male (12)	21.7 (4.9), 152	22.8 (5.1), 152	22.5 (4.2), 91	1.8 (.183)	0.3 (.589)
Male (13)	26.7 (5.8), 157	28.1 (6.0), 157	26.6 (6.5), 105	0.0 (.867)	3.8 (.053)
Male (14)	30.5 (6.2), 151	32.0 (6.4), 151	31.0 (6.7), 102	0.5 (.501)	1.4 (.237)
Male (15)	34.7 (6.5), 143	36.7 (6.7), 143	35.0 (7.8), 96	0.1 (.761)	3.1 (.082)
Male (16)	36.9 (7.1), 166	39.2 (7.3), 166	36.6 (7.2), 89	0.1 (.741)	7.1 (.008)
Male (17)	38.0 (6.6), 133	40.2 (6.7), 133	41.5 (8.6), 104	12.6 (.000)	1.7 (.196)
Male (18–29)	42.3 (8.2), 1,129	44.7 (8.5), 1,129	43.9 (7.9), 68	2.3 (.127)	0.6 (.433)
Male (30–39)	44.8 (8.4), 800	47.4 (8.7), 801	42.5 (9.8), 56	4.0 (.045)	16.4 (.000)
Male (40–49)	43.5 (7.6), 754	45.7 (7.9), 757	41.7 (9.4), 75	3.4 (.064)	16.6 (.000)
Male (50–59)	40.2 (7.3), 723	42.2 (7.5), 724	41.0 (10.4), 77	0.8 (.365)	1.7 (.196)
Male (60–69)	37.0 (7.8), 693	38.8 (8.0), 696	36.2 (9.1), 57	0.5 (.465)	5.4 (.020)
Male (70–80)	31.5 (7.2), 635	32.9 (7.3), 637	32.0 (10.0), 77	0.3 (.592)	1.0 (.317)
Female (6)	9.2 (2.0), 193	10.0 (2.2), 194	8.5 (2.7), 99	5.2 (.024)	23.1 (.000)
Female (7)	10.5 (2.2), 187	11.2 (2.3), 187	11.4 (3.2), 107	8.3 (.004)	0.3 (.573)
Female (8)	12.3 (2.7), 177	13.1 (2.8), 177	12.0 (2.8), 91	0.4 (.529)	8.0 (.005)
Female (9)	13.8 (3.0), 190	14.6 (3.1), 190	14.1 (3.3), 99	1.0 (.327)	1.6 (.213)
Female (10)	15.9 (3.0), 171	16.9 (3.1), 171	16.4 (3.9), 106	1.2 (.274)	1.4 (.241)
Female (11)	18.4 (3.9), 219	19.5 (4.1), 219	19.2 (3.9), 95	2.5 (.116)	0.4 (.553)
Female (12)	21.2 (4.5), 154	22.5 (4.7), 154	22.1 (4.1), 108	2.3 (.131)	0.6 (.449)
Female (13)	22.7 (4.9), 150	24.0 (5.1), 150	24.2 (4.7), 94	5.8 (.017)	0.1 (.761)
Female (14)	24.1 (4.7), 156	25.4 (4.9), 156	24.7 (5.5), 113	0.9 (.338)	1.2 (.277)
Female (15)	24.9 (4.6), 146	26.3 (4.8), 146	26.2 (5.0), 100	3.9 (.049)	0.1 (.817)
Female (16)	25.2 (4.4), 177	26.6 (4.5), 177	27.0 (5.0), 99	8.9 (.003)	0.4 (.532)
Female (17)	26.0 (5.2), 128	27.4 (5.4), 129	27.1 (6.1), 104	2.1 (.152)	0.2 (.667)
Female (18–29)	27.1 (5.0), 1,132	28.6 (5.2), 1134	27.3 (6.8), 166	0.2 (.645)	8.2 (.004)
Female (30–39)	28.1 (5.1), 830	29.6 (5.4), 831	28.0 (6.0), 204	0.1 (.739)	14.9 (.000)
Female (40–49)	27.6 (5.2), 838	29.0 (5.4), 838	28.3 (6.7), 143	2.1 (.152)	2.3 (.127)
Female (50–59)	25.7 (5.1), 781	27.0 (5.2), 782	25.5 (6.5), 98	0.1 (.739)	7.1 (.008)
Female (60–69)	23.9 (4.8), 739	25.1 (4.9), 740	22.2 (6.5), 98	8.8 (.003)	25.6 (.000)
Female (70–80)	19.7 (5.0), 700	20.7 (5.1), 704	19.6 (5.1), 77	0.0 (.891)	3.6 (.060)

Note. *M* = mean; NHANES = National Health and Nutrition Examination Survey; NIH = National Institutes of Health; *SD* = standard deviation.

These controls notwithstanding, there are important differences in how grip strength was measured in the two studies. These include, but are not limited to, differences in the hand dynamometer used (e.g., Takei vs. Jamar; Amaral et al., 2012), dynamometer handle position (Trampisch et al., 2012), upper limb position (e.g., shoulder, elbow; Desrosiers et al., 1995; Oxford, 2000; Su et al., 1994), and posture (e.g., sitting vs. standing; Balogun et al., 1991). All of these factors have the potential to influence grip strength measures. The NIH Toolbox study, unlike the NHANES study, used a protocol very similar to that recommended by the American Society of Hand Therapists (1992) and more recently proposed by Roberts et al. (2011).

Our finding that gender, side, and age group affect grip strength confirms the results of a legion of previous studies. The confirmation was necessary, however, to justify our stratification of grip strength values. The values provided in

Tables 1 and 2 can be used as rough normative references for grip strength measured using the same procedures. That noted, we do not recommend use of the NHANES values because they involve a protocol considerably different from those recommended by the American Society of Hand Therapists (1992) and by Roberts et al. (2011). Moreover, the NHANES protocol calls for adjustment of the handle position and three trials, both of which require additional time and are not necessary (Trampisch et al., 2012). The NIH Toolbox values presented can be interpreted in light of those presented in a 2006 meta-analysis (Bohannon et al., 2006). Such an interpretation, however, must take into account that NIH Toolbox data were summarized for dominant and nondominant sides and the meta-analysis summarized data for the left and right sides. In any case, the NIH Toolbox values tend to be lower than those in the meta-analysis, except perhaps for the oldest adults.

There were several limitations of the study. Because this study included secondary data sources, the researchers were not in control of the data collection procedures. In both studies, handgrip values were collected as part of a larger set of survey questions or measures, so fatigue may have had an effect. Some variables, such as handedness in both studies, relied on self-report. Missing data and extreme responses were encountered in both large-scale studies across multiple sites. When reading the results of this study, clinicians should be aware that different data collection protocols and devices were used in these two studies.

The value of handgrip strength as an indicator of overall strength and as a predictor of important outcomes notwithstanding (Bohannon, 2008, 2015; Vaapio et al., 2011), current nationally relevant reference values are needed if the grip strength of individuals and groups i to be interpreted. Herein, we provide information toward that end. Nevertheless, more specific reference values are required—specific to subpopulations and to performance within the spectrum of measured scores.

Implications for Occupational Therapy Practice

The results of this research indicate the need for further study to develop handgrip strength reference norms. Occupational therapy practitioners need such norms for interpreting the performance of hand function, evaluating treatment effects, and formulating treatment goals.

- Handgrip strength is an indicator of overall strength and is a predictor of important outcomes such as functional independence in daily activities.
- Handgrip strength reference values are essential to assist in the interpretation of testing results and clinical decision making.
- To make patient-centered treatment plans and interpret grip strength of individuals and groups correctly, occupational therapy practitioners need reference values stratified by age group and gender.

Conclusion

Some of the stratified grip strength values from the NHANES and NIH Toolbox studies differ. On the basis of this finding and the lack of conformity of the NHANES protocol with current recommendations, we cannot recommend the values for broad application as reference norms.

References

Amaral, J. F., Mancini, M., & Novo Júnior, J. M. (2012). Comparison of three hand dynamometers in relation to the accuracy and precision of the measurements. *Brazilian Journal of Physical Therapy*, *16*, 216–224. https://doi.org/10.1590/S1413-35552012000300007

American Society of Hand Therapists. (1992). *Clinical assessment recommendations* (2nd ed.). Chicago: Author.

Balogun, J. A., Akomolafe, C. T., & Amusa, L. O. (1991). Grip strength: Effects of testing posture and elbow position. *Archives of Physical Medicine and Rehabilitation*, *72*, 280–283.

Beaumont, J. L., Havlik, R., Cook, K. F., Hays, R. D., Wallner-Allen, K., Korper, S. P., . . . Gershon, R. (2013). Norming plans for the NIH Toolbox. *Neurology*, *80*(Suppl. 3), S87–S92. https://doi.org/10.1212/WNL.0b013e3182872e70

Bohannon, R. W. (2008). Is it legitimate to characterize muscle strength using a limited number of measure? *Journal of Strength and Conditioning Research*, *22*, 166–173. https://doi.org/10.1519/JSC.0b013e31815f993d

Bohannon, R. W. (2015). Muscle strength: Clinical and prognostic value of hand-grip dynamometry. *Current Opinion in Clinical Nutrition and Metabolic Care*, *18*, 465–470. https://doi.org/10.1097/MCO.0000000000000202

Bohannon, R. W. (2017). Test–retest reliability of measurements of hand-grip strength obtained by dynamometry from older adults: A systematic review of research in the PubMed database. *Journal of Frailty and Aging*, *6*, 83–87. https://doi.org/10.14283/jfa.2017.8

Bohannon, R. W., Peolsson, A., Massy-Westropp, N., Desrosiers, J., & Bear-Lehman, J. (2006). Reference values for adult grip strength measured with a Jamar dynamometer: A descriptive meta-analysis. *Physiotherapy*, *92*, 11–15. https://doi.org/10.1016/j.physio.2005.05.003

Centers for Disease Control and Prevention. (2011). *National Health and Nutrition Examination Survey (NHANES): Muscle strength procedures manual*. Atlanta: Author. Retrieved from https://wwwn.cdc.gov/nchs/data/nhanes/2011-2012/manuals/Muscle_Strength_Proc_Manual.pdf

Desrosiers, J., Bravo, G., Hébert, R., & Mercier, L. (1995). Impact of elbow position on grip strength of elderly men. *Journal of Hand Therapy*, *8*, 27–30. https://doi.org/10.1016/S0894-1130(12)80153-0

Dodds, R. M., Syddall, H. E., Cooper, R., Benzeval, M., Deary, I. J., Dennison, E. M., . . . Sayer, A. A. (2014). Grip strength across the life course: Normative data from twelve British studies. *PLoS One*, *9*, e113637. https://doi.org/10.1371/journal.pone.0113637

Dodds, R. M., Syddall, H. E., Cooper, R., Kuh, D., Cooper, C., & Sayer, A. A. (2016). Global variation in grip strength: A systematic review and meta-analysis of normative data. *Age and Ageing*, *45*, 209–216. https://doi.org/10.1093/ageing/afv192

Gershon, R. C., Wagster, M. V., Hendrie, H. C., Fox, N. A., Cook, K. F., & Nowinski, C. J. (2013). NIH Toolbox for assessment of neurological and behavioral function. *Neurology*, *80*(Suppl. 3), S2–S6. https://doi.org/10.1212/WNL.0b013e3182872e5f

Günther, C. M., Bürger, A., Rickert, M., Crispin, A., & Schulz, C. U. (2008). Grip strength in healthy Caucasian adults: Reference values. *Journal of Hand Surgery*, *33*, 558–565. https://doi.org/10.1016/j.jhsa.2008.01.008

Kim, J. K., Park, M. G., & Shin, S. J. (2014). What is the minimum clinically important difference in grip strength? *Clinical Orthopaedics and Related Research*, *472*, 2536–2541. https://doi.org/10.1007/s11999-014-3666-y

Massy-Westropp, N. M., Gill, T. K., Taylor, A. W., Bohannon, R. W., & Hill, C. L. (2011). Hand grip strength: Age and gender stratified normative data in a population-based study. *BMC Research Notes*, *4*, 127. https://doi.org/10.1186/1756-0500-4-127

National Institutes of Health. (2012). *NIH Toolbox administration manual*. Retrieved from http://www.healthmeasures.net/images/nihtoolbox/Training-Admin-Scoring_Manuals/NIH_Toolbox_Administration_Manual-English_9-25-12.pdf

Oxford, K. L. (2000). Elbow positioning for maximum grip performance. *Journal of Hand Therapy*, *13*, 33–36. https://doi.org/10.1016/S0894-1130(00)80050-2

Perna, F. M., Coa, K., Troiano, R. P., Lawman, H. G., Wang, C. Y., Li, Y., . . . Kraemer, W. J. (2016). Muscular grip strength estimates of the U.S. population from the National Health and Nutrition Examination Survey 2011–2012. *Journal of Strength and Conditioning Research*, *30*, 867–874. https://doi.org/10.1519/JSC.0000000000001104

Reuben, D. B., Magasi, S., McCreath, H. E., Bohannon, R. W., Wang, Y. C., Bubela, D. J., . . . Gershon, R. C. (2013). Motor assessment using the NIH Toolbox. *Neurology*, *80*(Suppl. 3), S65–S75. https://doi.org/10.1212/WNL.0b013e3182872e01

Ritchie, R. F., & Palomaki, G. (2004). Selecting clinically relevant populations for reference intervals. *Clinical Chemistry and Laboratory Medicine*, *42*, 702–709. https://doi.org/10.1515/CCLM.2004.120

Roberts, H. C., Denison, H. J., Martin, H. J., Patel, H. P., Syddall, H., Cooper, C., & Sayer, A. A. (2011). A review of the measurement of grip strength in clinical and epidemiological studies: Towards a standardised approach. *Age and Ageing*, *40*, 423–429. https://doi.org/10.1093/ageing/afr051

Seino, S., Shinkai, S., Fujiwara, Y., Obuchi, S., Yoshida, H., Hirano, H., . . . Takahashi, R.; TMIG-LISA Research Group. (2014). Reference values and age and sex differences in physical performance measures for community-dwelling older Japanese: A pooled analysis of six cohort studies. *PLoS One*, *9*, e99487. https://doi.org/10.1371/journal.pone.0099487

Shim, J. H., Roh, S. Y., Kim, J. S., Lee, D. C., Ki, S. H., Yang, J. W., . . . Lee, S. M. (2013). Normative measurements of grip and pinch strengths of 21st century Korean population. *Archives of Plastic Surgery*, *40*, 52–56. https://doi.org/10.5999/aps.2013.40.1.52

Strauss, E., Sherman, E. M., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary*. New York: Oxford University Press.

Su, C. Y., Lin, J. H., Chien, T. H., Cheng, K. F., & Sung, Y. T. (1994). Grip strength in different positions of elbow and shoulder. *Archives of Physical Medicine and Rehabilitation*, *75*, 812–815.

Trampisch, U. S., Franke, J., Jedamzik, N., Hinrichs, T., & Platen, P. (2012). Optimal Jamar dynamometer handle position to assess maximal isometric hand grip strength in epidemiological studies. *Journal of Hand Surgery*, *37*, 2368–2373. https://doi.org/10.1016/j.jhsa.2012.08.014

Turner, D. P., & Ebrahim, S. (1992). Relation between handgrip strength, upper limb disability and handicap among elderly women. *Clinical Rehabilitation*, *6*, 117–123. https://doi.org/10.1177/026921559200600205

Vaapio, S., Salminen, M., Vahlberg, T., & Kivelä, S. L. (2011). Increased muscle strength improves managing in activities of daily living in fall-prone community-dwelling older women. *Aging Clinical and Experimental Research*, *23*, 42–48. https://doi.org/10.1007/BF03337743

Wong, S. L. (2016). Grip strength reference values for Canadians aged 6 to 79: Canadian Health Measures Survey, 2007 to 2013. *Health Reports*, *27*, 3–10.

Richard W. Bohannon, DPT, EdD, PT, is Professor, Department of Physical Therapy, College of Pharmacy and Health Sciences, Campbell University, Buies Creek, NC.

Ying-Chih Wang, PhD, OTR/L, is Associate Professor, Department of Occupational Science and Technology, University of Wisconsin–Milwaukee; wang52@uwm.edu

Sheng-Che Yen, PhD, PT, is Assistant Professor, Department of Physical Therapy, Northeastern University, Boston, MA.

Kimberly A. Grogan, MS, OTR/L, is Clinician, Northern Suburban Special Education District, Highland Park, IL.

Acknowledgments

This study was funded in whole or in part with federal funds from the Blueprint for Neuroscience Research and the Office of Behavioral and Social Science Research, National Institutes of Health, under Contract No. HHS-N-260-2006-00007-C, with additional support from the National Children's Study, under Contract No. HHS-N-267-2007-00027-C. The authors declare no conflict of interest.

Citation: Bohannon, R. W., Wang, Y.-C., Yen, S.-C., & Grogan, K. A. (2019). Handgrip strength: A comparison of values obtained from the NHANES and NIH toolbox studies. *American Journal of Occupational Therapy, 73,* 7302205080. https://doi.org/10.5014/ajot.2019.029538

PECOP Blog

May 16th, 2022

The Olympics, sex, and gender in the physiology classroom

Are there sex based difference in athletic performance before puberty?

In the pas few years mos sate legislatures have considered laws sating that only members of the female sex can participate in girl's and women's sports (37 sates in 2021 alone), and as of April 20, 2022 ffteen sates have adopted such legislation (1). There have also been several well publicized insances of transwomen competing for championships in women's sports (for example see 2, 3, 4). The International Olympic Committee, the NCAA, and other sports governing bodies have also recently revised their policies regarding the inclusion of transwomen in women's sports (5, 6). All of this has resulted in sudents in my exercise physiology classes commonly asking quesions about sex-based diferences in sports performance and the inclusion of transwomen in women' s sports.

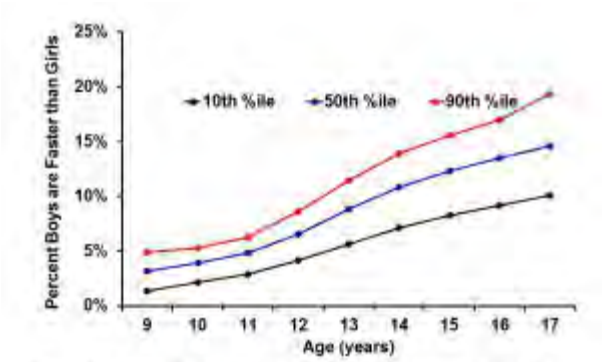


Figure 1. Percent faster running speed for boys compared to girls aged 9-17 years during the last stage of a 20 m shuttle run for the 10th, 50th, and 90th percentile (adapted from Tomkinson et al. 17)

In a previous PECOP Blog (7) I briefy summarized the sex-based advantages men have in athletic performance in adults, and the research evaluating the efects of tesosterone suppression and cross sex hormone use on factors that infuence athletic performance. In this PECOP Blog, I will briefy summarize the sex based prepubertal diferences in athletic performance and touch on puberty blockers.

A 2012 report from the CDC indicated there were no diferences between 6–1 1-year-old boys and girls in performance on physical ftness tess (8). Many sports leagues for pre-pubertal children are not separated by sex since the focus is developing basic sports skills rather than competition (9). Furthermore, some scholars have sated that there are no diferences in athletic performance between boys and girls prior to the onset of puberty, and that it is only the increased tesosterone secretion during puberty that causes males to outperform females in athletic competition (10, 11).

On the other hand, evaluations of ftness tesing in children as young as 3 years old shows that boys perform better than girls of the same age on tess of muscular srength, muscular endurance, and aerobic ftness (12-17). For example, Tomkinson et al. (17) observed that at age 9 boys are running an average of 3.2% faser than girls of the same age during the las sage of a 20 m shuttle run (Figure 1). In a separate evaluation Tomkinson et al. (16) reported that at age 9 boys have a bent arm hang time that is an average of

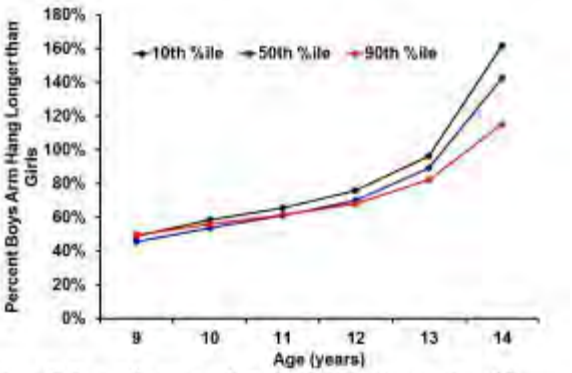


Figure 2. Percent longer arm hang time for boys compared to girls aged 9-14 years in the 10th, 50th, and 90th percentile (adapted from Tomkinson et al. 16)

48.1% longer than girls of the same age (Figure 2).

Furthermore, youth records from USA Track & Field (18) in the 8-and-under age group and in the 9-10-years-old age group (who can reasonably be assumed to be pre-pubertal) show that boys outperform girls in all events (Table 1).

Table 1. American Youth Outdoor Track & Field Records for boys and girls in the 8-and-under and 9-10 year-old age groups (11).

	8 and Under			9-10		
	Boys	Girls	% Difference	Boys	Girls	% Difference
100 m (sec)	13.65	13.78	0.94%	11.69	12.07	7.73%
200 m (sec)	27.52	28.21	5.15%	25.64	26.19	1.76%
400 m (sec)	62.48	65.19	5.48%	58.31	59.97	1.12%
800 m (sec)	148.58	158.11	6.62%	138.51	141.96	1.77%
1500 m (sec)	308.52	314.72	1.97%	278.62	289.87	4.24%
Average			3.51%			3.33%
Long Jump (m)	4.48	3.99	11.78%	5.08	4.78	6.26%
Shot Put (m)	10.41	8.96	15.86%	13.48	10.53	27.28%
Javelin Throw (m)	33.29	24.07	38.42%	44.18	27.13	16.82%

sec = seconds, m = meters. The percent difference between the boys and girls was calculated using the same equation as Miland-Stafford (20)

The smalles diference in track and feld records between boys and girls is 0.94% in the 8-and-under 100 m run, with the larges diference being 38.42% in the 8-and-under javelin throw. We recently analyzed top 10 data for national performance from Athletic.net in 100 m, 200 m, 400 m, 800 m, 1500 m, and 1600 m running events for children in the 7-8 and 9–10 year-old age groups for the years 2019-2021 and found that across all events 7-8-year-old boys were 4.4 ± 1.9% faser than girls, and 9-10-year -old boys were 5.4 ± 1.8% faser than girls (fgure 3; not yet published data). Youth records from USA swimming also show that in 19 out of 23 events the national records for 10 and under boys are faser than girls by an average of 1.72% (19). It is important to note that in competition the diference between frs and second place often comes down to as little as 0.02% diference in speed (Data to be presented at the 2022 ACSM Annual Meeting).

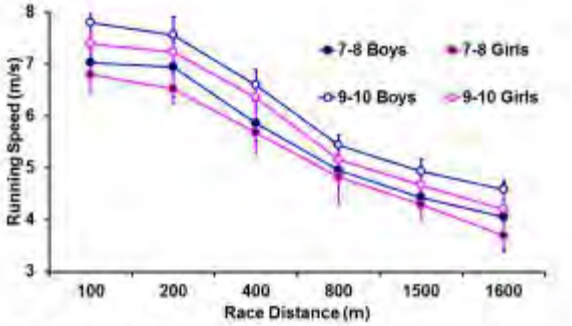


Figure 3. Running velocity in 7-8 and 9-10 year-old age groups in 100 m, 200 m, 400 m, 800 m, 1500 m, and 1600 m track events based on national data for top 10 performances during the years 2019-2021 (data are means ± standard deviation).

There is no quesion that the diferences in running performance between prepubertal boys and girls is less than the 10-13% diference in running performance observed between pos-pubertal boys and girls, and between adult men and women (10, 11, 20). And there is no quesion that the large increases in circulating tesoserone experienced by boys during puberty is responsible for mos of the diferences in athletic performance between pos-pubertal boys and girls, and between adult men and women (21). But the exisence of diferences in athletic performance between prepubertal boys and girls is well demonsrated (12-19). Juxtaposing the satements of no pre-pubertal athletic diferences between boys and girls (8, 10, 1 1) and the evidence demonsrating that there are pre-pubertal athletic diferences between boys and girls (12-19) can facilitate an interesing discussion about data collection, sample size, data analysis, and other factors that may contribute to these contradictory fndings.

When explaining the biological causes of the prepubertal athletic advantages in boys, a good sarting point is to discuss the diferences in growth and development between boys and girls and to explain the processes of sex determination and sex diferentiation (22). Sex determination occurs at conception with the conferral of sex chromosomes. Six weeks later, sex diferentiation begins to become apparent and during the remainder of development the gonads and genitalia acquire male or female characterisics. During sexual diferentiation, the presence of the SR Y gene on the Y chromosome along with androgen exposure and anti-Müllerian hormone cause the internal and external genitalia to follow the male developmental pathway. In the absence of the SRY gene on the Y chromosome, lack of androgen exposure, and lack of anti-Müllerian hormone the female developmental pathway occurs. Of course these few brief sentences fail to cover the myriad of complex interactions of genes, primordial sem cells, and hormones that regulate sex development, and the possible diferences and disorders that can occur . But it is remarkable that with all of the possible misseps that can happen during sexual diferentiation and development, sex can be accurately and easily identifed at birth 99.83% of the time (23).

Further subsantiating the important role of sex in growth and development are the World Health Organization fetal growth charts (24), which indicate small but meaningful sex-based diferences with male fetuses being consisently larger than female fetuses. Similarly, subsantiating the important role of sex in growth and development, the Centers for Disease Control and Prevention have diferent growth charts for boys and girls from birth through adolescence with boys having consisently higher values for body mass and body height (25).

With an eye towards physical ftness and athletic performance, sarting at birth and continuing throughout youth girls have more body fat and less fat-free mass than boys. For example, Davis et al. (26) in an evaluation of 602 infants reported that at birth and age 5 months, infant boys have larger total body mass, body length, and fat-free mass while having lower percent body fat than infant girls. In an evaluation of 20 boys and 20 girls ages 3-8 years old, matched for age, height, and body weight Taylor et al. (27) reported that the boys had less body fat, lower percent body fat, and a higher bone free lean body mass than the girls, such that the girls' fat mass was 52% higher than the boys, while the bone-free lean tissue mass was 9% lower. In an evaluation of 376 prepubertal [Tanner Stage 1] boys and girls, Taylor et al. (28) observed that the boys had ~22% more lean mass, and ~13% less body fat (when expressed as percent of total body mass) than did the girls. In a review of 22 peer reviewed publications on the topic, Staiano and Katzmarzyk (29) concluded that girls have more total body fat than boys throughout childhood and adolescence. It is a tenet of exercise science that having more lean body mass provides athletic advantages, so it is reasonable to conclude that having more lean body mass contributes to the prepubertal sex-based male athletic advantages.

It is worth noting that serum tesoserone concentrations in boys are higher for the frs 5 months after birth than in girls (30). Tesoserone concentrations are then similar between boys and girls until the onset of puberty , when tesoserone concentrations increase 10-20-fold in boys. Given the well know anabolic and androgenic efects of tesoserone, the higher tesoserone levels in newborn boys likely contributes to the sex related diferences in body size and composition in newborns. It is unknown how much the lingering sex-linked diferences in body size, body composition, physical ftness, and athletic performance are due lasing efects of the higher tesoserone levels in newborns, and how much the diferences are due to Y chromosome or other sex-linked efects.

Strongly suggesing that sex linked diferences in physical ftness and athletic performance in children before puberty

are due to biological factors, Eiberg et al. (13) measured body composition, VO_2max, and physical activity in 366 Danish boys and 332 Danish girls between the ages of 6 and 7 years old. Their observations indicated that absolute VO_2max was 11% higher in boys than girls, while relative to body mass the boys' VO_2max was sill 8% higher than the girls. Accelerometry based measurements of physical activity indicated that when boys and girls regularly participated in the same amount and intensity of physical activity, the boys had higher measured physical ftness than the girls. When the fndings of Eiber g (13) are taken collectively with the fndings of lar ge scale school based physical ftness tesing in children that also shows pre-pubertal boys outperforming girls in measurements of aerobic ftness, muscular srength, and muscular endurance (12, 14-17), the youth records from USA Track & Field (18) showing that pre-pubertal boys outperform girls in all events, and the 10 and under records from USA Swimming showing that boys outperform girls in 19 out of 23 events (19), there exiss srong evidence that there are diferences in physical ftness and athletic performance between boys and girls before puberty.

And fnally , this discussion arising from laws sating that only members of the female sex can participate in girl' s and women's sports can lead to quesions about the efects of puberty blockers on physical ftness and athletic performance in prepubertal children. Puberty blockers are correctly known as gonadotropin-releasing hormone agoniss (GnRHa), which disrupt the normal pattern of secretion of as gonadotropin-releasing hormone causing the pituitary gland to sop producing follicle-simulating hormone and luteinizing hormone. Unfortunately , there is minimal research on the efects of puberty blockers on factors that infuence physical ftness and athletic performance.

To the bes of my knowledge, there is no research on the efects of puberty blockers on muscle srength, running speed, or other measures of athletic performance. Indeed, Klaver et al. (31) is the only published research that I am aware of that has evaluated the use of puberty blockers on any athletic performance related factor, and this is only on body composition. Klaver et al. (31) demonsrated that the use of puberty blockers in Tanner sage 2-3 teenagers increased body fat and decreased lean body mass in transgirls, but the use of puberty blockers did not eliminate the diferences in body composition between transgirls and comparable female teenagers. Roberts and Carswell (32), concluded that there is no published research that sufciently characterizes the impact of puberty blockers on growth or fnal adult height. Thus, the efect of prescribing puberty blockers to a male child before the onset of puberty on the physical components of athletic performance is almos entirely unknown. This becomes a great point in a discussion to remind sudents of the ever -evolving nature of science. Any further discussion on this topic becomes speculation or can be removed from the realm of physiology and into metaphysical discussions of what is or is not fair. Such metaphysical discussions can be fascinating, and also heated, so caution is advisable when proceeding outside of the realm of physiology in a physiology classroom.

In summary, there is srong evidence that even before puberty there are sex-based diferences in physical ftness and athletic performance with boys running faser , jumping farther and higher, and demonsrating greater muscle srength than girls of the same age. These pre-pubertal sex based diferences are smaller than the diferences in pos pubertal teens and adults, but the diferences are likely meaningful in terms of competition. There is currently insufcient evidence to determine what efects puberty blockers have on physical ftness and athletic performance in children.

References

1. **Lavietes M**. (April 13, 2022) Kentucky Legislature overrides governor's veto of transgender sports ban

[online]. NBCNews.com https://www.nbcnews.com/nbc-out/out-politics-and-policy/kentucky-legislature-overrides-governors-veto-transgender-sports-ban-rcna24303 [Accessed April 20, 2022]

2. **Barnes K**. (March 17, 2022) Amid protess, Penn swimmer Lia Thomas becomes frs known transgender athlete to win Division I national championship. [online]. espnW.com. https://www.espn.com/college-sports/sory/_/id/33529775/amid-protess-pennsylvania-swimmer_-lia-thomas-becomes-frs-known-transgender_-athlete-win-division-national-championship [Accessed April 20, 2022]
3. **Ellingworth J, Ho S.** (Augus 2, 2021) Transgender weightlifter Hubbard makes hisory at Olympics. [online]. APNews.com https://apnews.com/article/2020-tokyo-olympics-sports-weightlifting-laurel-hubbard-e721827cdaf7299f47a9115a09c2a162 [Accessed April 20, 2022]
4. **Morton V.** (June 3, 2019) CeCe Telfer, Franklin Pierce transgender hurdler, wins NCAA women's national championship [online]. Washingtontimes.com https://www.washingtontimes.com/news/2019/jun/3/cece-telfer-franklin-pierce-transgender-hurdler-wi/ [Accessed April 20, 2022]
5. **Yurcaba C.** (January 22, 2022) NCAA's new trans athlete guidelines sow confusion amid Lia Thomas debate [online]. NBCnews.com https://www.nbcnews.com/nbc-out/out-news/ncaas-new-trans-athlete-guidelines-sow-confusion-lia-thomas-debate-rcna13073 [Accessed April 20, 2022]
6. **Nair A, Nair R, Davis T.** (April 8, 2022) Transgender women unable to compete in British Cycling events as policy suspended [online]. Reuters.com https://www.reuters.com/lifesyle/sports/british-cycling-suspend-transgender-participation-policy-2022-04-08/[Accessed April 20, 2022]
7. **Brown G.** (Augus 18, 2021). The Olympics, sex, and gender in the physiology classroom [online]. PECOP Blog. https://blog.lifescitrc.org/pecop/2021/08/18/the-olympics-sex-and-gender-in-the-physiology-classroom/ [Accessed April 20, 2022]
8. **Ervin RB, Wang CY, Fryar CD, Miller IM, and Ogden CL.** [online] Measures of Muscular Strength in U.S. Children and Adolescents, 2012. NCHS Data Brief No. 139, December 2013. (https://www.cdc.gov/nchs/products/databriefs/db139.htm; accessed April 6, 2022)
9. **Wells MS, Arthur-Banning SG.** The Logic of Youth Development: Consructing a Logic Model of Youth Development through Sport. J Pakr & Rec Admin. 26: 189-202, 2008
10. **Handelsman DJ.** Sex diferences in athletic performance emer ge coinciding with the onset of male puberty. Clin Endocrinol (Oxf). 87:68-72, 2017
11. **Handelsman DJ, Hirschberg AL, Bermon S.** Circulating Tesoserone as the Hormonal Basis of Sex Diferences in Athletic Performance. Endocr Rev. 39:803-829, 2018
12. **Catley MJ, and Tomkinson GR.** Normative health-related ftness values for children: analysis of 85347 tes results on 9-17-year -old Ausralians since 1985. Br J Sports Med 47: 98-108, 2013.
13. **Eiberg S, Hasselstrom H, Gronfeldt V, Froberg K, Svensson J, and Andersen LB.** Maximum oxygen uptake and objectively measured physical activity in Danish children 6-7 years of age: the Copenhagen school child intervention sudy . Br J Sports Med 39: 725-730, 2005.
14. **Latorre Roman PA, Moreno Del Castillo R, Lucena Zurita M, Salas Sanchez J, Garcia-Pinillos F, and Mora Lopez D.** Physical ftness in preschool children: association with sex, age and weight satus. Child Care Health Dev 43: 267-273, 2017.

15. **Tambalis KD, Panagiotakos DB, Psarra G, Daskalakis S, Kavouras SA, Geladas N, Tokmakidis S, and Sidossis LS.** Physical ftness normative values for 6-18-year -old Greek boys and girls, using the empirical disribution and the lambda, mu, and sigma satisical method. Eur J Sport Sci 16: 736-746, 2016.
16. **Tomkinson GR, Carver KD, Atkinson F, Daniell ND, Lewis LK, Fitzgerald JS, Lang JJ, and Ortega FB.** European normative values for physical ftness in children and adolescents aged 9-17 years: results from 2 779 165 Euroft performances representing 30 countries. Br J Sports Med 52: 1445-14563, 2018.
17. **Tomkinson GR, Lang JJ, Tremblay MS, Dale M, LeBlanc AG, Belanger K, Ortega FB, and Leger L.** International normative 20 m shuttle run values from 1 142 026 children and youth representing 50 countries. Br J Sports Med 51: 1545-1554, 2017.
18. (December 19, 2018) American Youth Outdoor Track & Field Records. [online] USATF http://legacy.usatf.org/satisics/records/view .asp? division=american&location=outdoor%20track%20%26%20feld&age=youth&sport=TF (accessed April 20, 2022)
19. (2022) National Age Group Records [online]. USA Swimming. https://www.usaswimming.org/times/popular-resources/national-age-group-records (accessed April 20, 2022)
20. **Millard-Stafford M, Swanson AE, Wittbrodt MT**. Nature Versus Nurture: Have Performance Gaps Between Men and Women Reached an Asymptote? Int J Sports Physiol Perform. 13:530-535, 2018
21. **Levine BD, Joyner MJ, Keith NR, Bagish AL, Pedersen BK, Schmidt W, Stachenfeld N, Girard O, Nagatomi R, Foster C, Okazaki K, Stellingwerf T, Jiexiu Z, Robson SJ, Bailey DM, Bosch A, Murphy RM, Qiu J, Lollgen H, Mitchell J, Kearney J, Scott JM, Lundby C, Steinacker J, Trappe S, La Gerche A, Masuki S, Roach R, Schneider S, Millet G, Kohrt WM, Roberts WO, Kraus WE, Benjamin HJ, Koning JJ, Gatterer H, Wehrlin JP, Charkoudian N, Lawley JS, Hopman MTE, Hawley J.** The role of tesoserone in athletic performance. [online] https://web.law.duke.edu/sites/default/fles/centers/sportslaw/Experts_T_Statement_2019.pdf (accessed April 6, 2022).
22. **Rey R, Josso N, Racine C. Sexual Differentiation.** 2020 May 27. In: Feingold KR, Anawalt B, Boyce A, Chrousos G, de Herder WW, Dhatariya K, Dungan K, Hershman JM, Hofand J, Kalra S, Kaltsas G, Koch C, Kopp P, Korbonits M, Kovacs CS, Kuohung W, Laferrère B, Levy M, McGee EA, McLachlan R, Morley JE, New M, Purnell J, Sahay R, Singer F, Sperling MA, Stratakis CA, Trence DL, Wilson DP, editors. Endotext [Online]. South Dartmouth (MA): MDText.com, Inc.; 2000–. PMID: 25905232. (Accessed April 6, 2022)
23. **Sax L.** How common is intersex? a response to Anne Fauso-Sterling. J Sex Res. 39:174-8, 2002
24. **Kiserud T, Piaggio G, Carroli G, Widmer M, Carvalho J, Neerup Jensen L, Giordano D, Cecatti JG, Abdel Aleem H, Talegawkar SA, Benachi A, Diemert A, Tshefu Kitoto A, Thinkhamrop J, Lumbiganon P, Tabor A, Kriplani A, Gonzalez Perez R, Hecher K, Hanson MA, Gülmezoglu AM, Platt LD.** The World Health Organization Fetal Growth Charts: A Multinational Longitudinal Study of Ultrasound Biometric Measurements and Esimated Fetal Weight. PLoS Med. 14:e1002220, 2017

25. **Centers for Disease Control and Prevention.** Clinical Growth Charts [online] https://www.cdc.gov/growthcharts/clinical_charts.htm; (Accessed April 6, 2022)
26. **Davis SM, Kaar JL, Ringham BM, Hockett CW, Glueck DH, and Dabelea D**. Sex diferences in infant body composition emerge in the frs 5 months of life. J Pediatr Endocrinol Metab 32: 1235-1239, 2019.
27. **Taylor RW, Gold E, Manning P, and Goulding A.** Gender diferences in body fat content are present well before puberty. Int J Obes Relat Metab Disord 21: 1082-1084, 1997.
28. **Taylor RW, Grant AM, Williams SM, and Goulding A** . Sex diferences in regional body fat disribution from pre- to pospuberty . Obesity (Silver Spring) 18: 1410-1416, 2010.
29. **Staiano AE, Katzmarzyk PT.** Ethnic and sex diferences in body fat and visceral and subcutaneous adiposity in children and adolescents. Int J Obes (Lond). 36:1261-9. (2012).
30. **Senefeld JW, Lambelet Coleman D, Johnson PW, Carter RE, Clayburn AJ, Joyner MJ.** Divergence in Timing and Magnitude of Tesoserone Levels Between Male and Female Youths. JAMA. 324:99-101, 2020
31. **Klaver M, de Mutsert R, Wiepjes CM, Twisk JWR, den Heijer M, Rotteveel J, Klink DT** . Early Hormonal Treatment Afects Body Composition and Body Shape in Young Transgender Adolescents. J Sex Med 15: 251-260, 2018.
32. **Roberts SA, Carswell JM**. Growth, growth potential, and infuences on adult height in the transgender and gender-diverse population. Andrology. 9:1679-1688, 2021.

Dr. Greg Brown is a Professor of Exercise Science in the Department of Kinesiology and Sport Sciences at the University of Nebraska at Kearney where he has been a faculty member since 2004. He is also the Director of the General Studies program at the University of Nebraska at Kearney. He earned a Bachelor of Science in Physical Education (pre-Physical Therapy emphasis) from Utah State University in 1997, a Maser of Science in Exercise and Sport Science (Exercise Physiology Emphasis) from Iowa State University in 1999, and a Doctorate of Philosophy in Health and Human Performance (Biological Basis of Health & Human Performance emphasis) from Iowa State University in 2002. He is a Fellow of the American College of Sports Medicine and an American College of Sports Medicine Certifed Exercise Physiologis.

Posed in Classroom Content, Diversity, engagement, Teaching Strategies on May 16, 2022 by **Margaret Stieben** (Poss |Profle) Comment
Leave a comment

May 3rd, 2022

Cultivating Belonging through Asynchronous Discussion Assignments and "State Your Perspective"

Advancing diversity, equity, and inclusion (DEI) within college classrooms, whether virtual or in-person, has perhaps never been as high a priority as now. One outcome of pandemic teaching has been critical evaluation of hisoric teaching practices, placing the onus on insructors to provide inclusive learning environments that are responsive and adaptive to a wide range of individualized circumsances. At the same time, some sudents have expressed feeling isolated and disconnected from peers, reducing motivation and academic persisence. Cultivating a sense of community and belonging in educational spaces, for all learners, is a current hot topic in higher education. In fact, two recent PECOP blogs have centered around the related idea of incorporating team-building practices to enrich learning in physiology education (From a Group to a Team: Medical Education Orientation Curriculum for Building Efective Teams and Developing a Community of Practice in an A&P Course)

Belonging, or the belief that one's individual abilities and attributes are valued, respected, and on par with others' abilities, is a srong driving force for persisence in STEM felds (1, 2, see also the Iowa State University Center for Excellence in Learning and Teaching webpage: Foser a Sense of Belonging and the Indiana School of Education Building a Sense of Community for All resources). I am not an expert on this, yet I care about supporting the community of learners within the courses that I teach. This led me to ask: What can I do to build sudents' undersanding of physiology while also deepening their belief that they belong here, in my classroom, which in turn may foser resilience, persisence, and improved satisfaction within college-level coursework?

Collaborative work is included in all courses I teach. These collaborations take diferent forms based on the learning goals for the course, learner characterisics (frs year versus fourth year sudents, for example), and topic complexity . Summarized below is one course activity I have used which aims to: (1) help sudents maser challenging physiology concepts through peer-to-peer interactions, (2) develop communication skills related to expressing ideas about human function (a highly-valued professional skill), and (3) build community and a sense of belonging.

Asynchronous Discussion Assignments and "State Your Perspective". One course I teach is an in-person, large lecture-syle Human Physiology service course for second, third, and fourth year under graduate sudents (as well as a handful of graduate sudents) from biomedical sciences, biomedical engineering, pharmacy interes, public health, and other STEM programs. Many sudents express trouble "learning how to learn" human physiology , which can be quite diferent compared with the academic work typical for their varied primary programs of sudy . They also report feeling isolated in a large classroom and that they have trouble fnding sudy groups, which they value while preparing for exams.

Traditionally, exams in this Human Physiology course were comprised predominantly of multiple choice quesions and a few short answer quesions (e.g., 3-4 sentences in length). I recently found myself asking: WHAT IF sudents moved

from providing short written explanations on exams that lacked detail due to time consraints to having sufcient time to carefully think through how to explain a physiological process? And, WHAT IF this activity could be designed in such a way to help sudents recognize what they undersand (and what they don’ t undersand) in advance of an exam, giving them the opportunity to review course materials and try again? And, WHAT IF groups of sudents were working through this together, leveraging peer-to-peer learning?

These quesions, along with experiences from the online and blended insruction I have been doing for many years, gave rise to incorporating asynchronous, online discussion assignments that sudents would complete in small groups (6-8 sudents per group). The goal was to give sudents an opportunity to practice using appropriate anatomical and physiological terminology to precisely describe how the human body functions in a relatively low-sakes setting that supported peer interactions. Students were given a discussion prompt (see below for examples) to which they posed an initial response in the LMS-based virtual discussion forum. Next, all group members were responsible for reviewing their peers’ initial poss and providing two follow-up responses, adding to and building upon the initial physiological descriptions. There were a total of four sets of discussion assignments, one per unit, across the semeser . While the discussion assignment sructure remained similar from unit to unit, the expectation to communicate increasingly complex ideas was inherent within the discussion prompts.

Specifcally to address DEI and belonging, sudents were to begin their initial responses with a “State Your Perspective” satement. “State Your Perspective” entailed providing a 1-2 sentence summary satement to describe the context by which the topic at hand was viewed. In Human Physiology, this might be knowledge based on prior coursework, the focus of the lab in which they worked, practical clinical experiences for those who work in health care settings, and such. While ice-breaker introductions are frequently incorporated into group work, the use of bolder “State Your Perspective” language is intentional. It helps to move from a generic introduction that generally alludes to difering background experiences to an explicit and purposeful satement intended to summarize the specifc context for the way a particular physiological function is undersood.

Here are excerpts of the discussion prompts and how “State Your Perspective” is modeled for sudents.

UNIT 1 Discussion Prompt: One theme for UNIT 1 has been to develop connections between new information and previously-known concepts in order to undersand how the human body works: What have you learned in prior courses that apply to human physiology? Specify (1) the prior knowledge/what you knew before this course, and (2) the new ideas presented UNIT 1 that expands upon your background knowledge and therefore your undersanding of human function.

- **“State Your Perspective”:** Include a 1-sentence introduction at the beginning of your initial pos that includes your major and anything else important for your group members to know that provides context for your perspective. For example “I am a third year biomedical sciences sudent and I work in a research

lab that sudies RNA, therefore I have learned ….".

- As you will see, some of your group members may have academic backgrounds that are diferent from yours, and they might present concepts in a diferent way . This is great! We hope the discussions become more interesing from sharing multiple ways to view the same physiological concept.

UNIT 2 Discussion Prompt: Prepare an answer to one of the Exam 2 Study Guide prompts to share with your group members. Include at leas one type of conceptual model within your response: how one "Core Concept of Physiology" can be used to remember this process [see Reference 3 for information about the Core Concepts of Physiology], an originally-created concept map, an analogy, an annotated fgure, or another self-generated sudy tool.

- Begin your response with a 1-sentence "State Your Perspective" that provides context for your response. For example "I am a pharmacy interes sudent, and it is important for me to learn about neurotransmitters and receptors because …."

UNIT 3 Discussion Prompt: Summarize one physiology concept presented in UNIT 3 for your group members, in your own words and including the appropriate anatomy and physiology terminology. Suggesed length: 4-6 sentences. NEXT: Create four diferent 1-sentence satements about your topic, including two satements that are TRUE and two satements that are F ALSE (but don't identify which is which, see below).

- Begin with a 1-sentence introduction, similar to previous discussion forums so that your new group members undersand something about your perspective. Example: "I am an interdisciplinary sudies sudent interesed in healthcare; therefore, I found the lecture on hypertension really interesing …."
- For your responses to classmates: Carefully review each satement. Select one that you think is false and provide a physiological rationale to support your reasoning. Next, make the appropriate corrections to turn it into a TRUE satement.

Teaching Hint #1: This is manageable in a large lecture course of 150-250 sudents because I have teaching assisants who undersand their primary responsibility is to regularly engage directly with sudents in the small-group discussions and provide feedback for correct and incorrect descriptions (this is a high priority for sudents. Practically speaking, this equates to each TA managing 6-10 groups of ~8 sudents each.

Teaching Hint #2: Once the grading is completed, I ask the TAs to summarize what they learned about how sudents learn physiology. This has been a good way to mentor TAs and prompt thoughts about their own teaching philosophies. I sometimes ask them what they would change (nothing like grading 50+ discussion assignments based on a poorly-worded prompt…). In fact, this is how the UNIT 3 true/false satements came to be; a graduate sudent proposed it as a way to incorporate greater critical thought and reasoning within discussion assignments.

So what did sudents think about this type of discussion assignment? Here are examples of comments provided on the

end-of-class evaluation forms, paraphrased and in aggregate form (i.e., these are not actual sudent comments but represent themes in responses):

- The discussion assignments were a good way for me to think critically about one idea then communicate my undersanding of human function to my peers.
- Discussions were a great way to see what my classmates were doing to learn human physiology that I could apply to my own learning—my group members proposed sudy srategies and ways of thinking about the human body that I hadn't thought of before.
- I enjoyed learning from my peers, who might know something more than me based on their experiences outside of class.
- Even though this was a large lecture course with quite a bit of content presented online, I enjoyed interacting with my peers, the professor, and TAs in the discussions. I felt like everyone was there to support my learning.

Despite initial skepticism, very few sudents conveyed negative comments about the discussion assignments or described them as "busy-work".

Beyond sudent feedback, here are a few subjective comments conveying my personal observations about classroom dynamics that arose from this course activity.

- By design, one aim of "State Your Perspective" satement was to help sudents recognize that they hold certain views on a topic based on their background experiences. For some 20-something year-olds, it might not be intuitive that they, in fact, have certain perspectives and attitudes that they bring into group work. "State Your Perspective" has the potential to be afrming—when articulating prior experiences it can become more explicit, to ourselves and others, that we all have something unique to contribute to group work.
- Sharing perspectives, along with the underlying narrative (but briefy , in 1-2 sentences), seemed to normalize the idea that we all have diferent backgrounds and experiences so OF COURSE we may hold diferent perspectives, or ways of viewing things.
- Because the context for why discussion prompts were answered with a particular focus was evident, it seemed to reduce the pressure that every sudent should know "everything". Insead, over time and through several rounds of discussions, sudents became more comfortable talking about what they undersood and what they didn' t undersand. Clarifcations could be made and misperceptions could be corrected by peers, who almos always demonsrated remarkable diplomacy and kindness toward their classmates.
- In some cases, the online and asynchronous nature of these discussions seemed to reduce barriers with regard to asking for help. It seemed to move sudents from a mindset of "I should know this but I don't/everyone knows this but me" to the non-threatening "This is a topic maybe I need to ask about." Students seemed less self-conscious when asking quesions.

In summary, collaboration during small group, asynchronous discussion assignments seemed to promote a sense of community and belonging among sudents in a Human Physiology for non-majors course. As the insructor , it was

rewarding to see improvement in sudents' abilities to explain physiological processes across the semeser . It was also extremely rewarding to see the great care exhibited by sudents to be inclusive and supportive of their peers.

References:

1. Herman J, Hilton M. Supporting Students' College Success The Role of Assessment of Intrapersonal and Interpersonal Competencies (Consensus Study Report of the National Academies of Sciences, Engineering, and Medicine). Washington, DC: The National Academies Press, 2017.
2. Wilton M, Gonzalez-Nino E, McPartlan P, Terner Z, Chrisofersen RE, Rothman JH. Improving academic performance, belonging, and retention through increasing sructure of an introductory biology course. CBE Life Sci Educ, 18:1-13, 2019.
3. Michael J, Clif W, McFarland J, Modell H, Wright A. The Core Concepts of Physiology A New Paradigm for Teaching Physiology. New York: Springer, 2017.

My Perspective: I am an Associate Professor of Insruction in the Department of Health and Human Physiology at the University of Iowa. I am the Program Director for the B.S. Human Physiology program, which serves approximately 625 majors. I am also an active participant in several undergraduate sudent success initiatives at the collegiate level. The mos rewarding part of my job is learning about how sudents learn physiology , in their own words. I solicit sudent feedback for their academic experiences regularly.

Jennifer Rogers, PhD

Associate Professor of Insruction

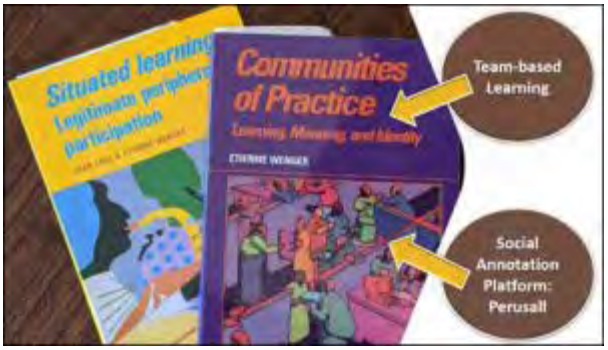
Department of Health and Human Physiology

University of Iowa

Posed in Collaboration, Diversity, Inclusion, Non-majors on May 3, 2022 by **Margaret Stieben** (Poss |Profle) Comment
Leave a comment

April 7th, 2022

Developing a Community of Practice in an A&P Course



This blog is about sriving to create a Community of Practice (CoP) to engage sudents in Situated/Social Learning by using Team-based activities and assessments along with the web-based social learning annotation platform, Perusall.

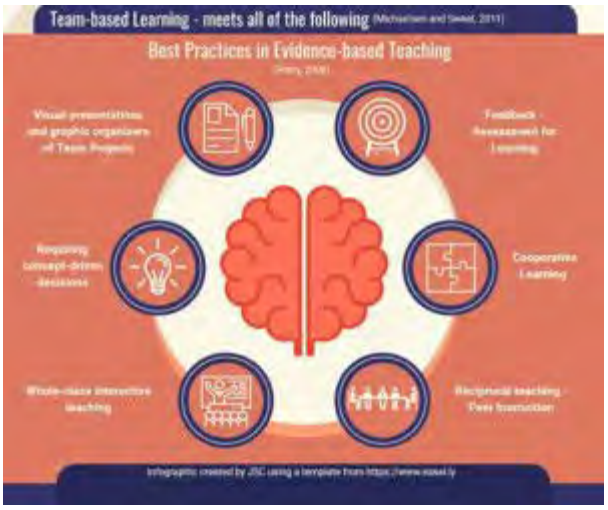
We have all experienced those "Aha" moments when something we were sruggling with suddenly becomes clear . Think back to a time when you experienced real/durable learning. When I did that, three things popped into my mind: a hallway discussion in graduate school with classmates in my neurophysiology class about the Goldman-Hodgkin-Katz equation; American Physiological Society – Insitute of Teaching and Learning (APS-ITL) conferences/interaction with Physiology Educators Community of Practice (PECOP); and the Community of Practice at HCC via the Insructional Development Center (IDC) which or ganizes and facilitates Bes Practices and Faculty Academy. And what this made me realize was that I learned bes in a social setting with peers rather than isolated in my room/ofce tackling a topic by myself. Although this was new to me, Lave and Wenger realized this long ago.

Lave and Wenger put forth the social learning theory of situated learning and communities of practice (CoP) in the early 90s. Core ideas of their theory are that learning is identity formation through social participation and that communities of practice are groups of people (communities), brought together by a need for shared learning (domain) for something they do together (practice) and learn how to do it better as they interact regularly (Lave and Wenger, 1991; Wenger, 1998). And I believe, in a classroom setting, this should be framed within a signifcant learning environment. See Fink (2003) for explicit seps that can be taken to create an environment conducive to learning.

While a CoP is often discussed relative to professional societies, I believe that a CoP can develop within an A&P course and bring about durable learning through social interaction. In this case, then, the domain includes the sudents who are in the course to learn A&P – shared learning needs; the community includes the class as well as the community within sudent groups/teams; and the practice includes interactions and participation in evidence-based teaching bes practices from the resources those produce.

The following infographic is a summary of bes practices in evidence-based teaching (Petty , 2006) which Michaelson and Sweet (2011) sugges can be met by and are a part of Team-based Learning (TBL). These include Visual presentation and graphic organizers which are met in my classes by team projects; feedback and assessment for learning; cooperative learning; reciprocal teaching e.g., peer insruction; whole-class interactive teaching; requiring concept-driven decisions e.g., concept quesions and higher -order thinking levels for summative assessments. This provides a very srong rationale for using TBL. And TBL, by its very nature, promotes social learning.

Michaelsen and Richards (2005) identifed the four key components of TBL: group formation; meaningful team



assignments; routine feedback; and accountability. The following infographic includes the components of TBL and summarizes some of the ways they are addressed in my courses. I will go into more detail on some of these throughout the blog.

Formation of diverse teams is very important for the successful use of team-based learning. In the physical classroom, I used a 'show of hands' to quesions asked on the frs day of class and had the sudents line up, then I counted them of into the appropriate number of groups. Quesions used were: "How many have ……had me as an insructor before?; had medical terminology?; a college degree or certifcate?; been born outside of IL?" etc. This provides transparency in how the teams are formed and lets sudents know what things the insructor thinks are important to include in each team.

For the virtual, online-synchronous classroom, I use the web-based platform, CATME Smarter Teamwork, Team-maker tool. Team-maker tool page can be found at this link. The Team-maker tool simplifes the team-assignment process, for the virtual classroom, and creates diverse teams. Insructors decide which criteria will be used to form teams/sudent groups. For example, it is helpful for team members to have similar work schedules to facilitate group work. It is also helpful for team members to have dissimilar GPAs. Insructors can also write cusom quesions and criteria to add to the Team Maker Tool survey. CATME Smarter Teamwork platform is a product of and adminisered by Purdue University. General information about the CATME Smarter Teamwork platform can be found at this link.

In addition to properly forming teams, teams mus be properly managed. Team members should receive feedback regarding their efectiveness in the team early and often. I use Peer Evaluation (PE) Surveys adminisered by the CATME Smarter Teamwork platform to help teams and team members become more efective. The TBL community uses the phrase "forming, sorming, and norming," to describe phases teams go through during the semeser . PE surveys helps teams to progress to the norming phase more quickly. Team members are evaluated in 5 areas: contributing to the team's work, interacting with teammates, keeping the team on track, expecting quality, and having relevant knowledge, skills, and abilities.

Three PE surveys are adminisered over the course of the semeser . The frs two PE surveys (week 5 and week 10) were formative and the third one (week 15) was summative. Students' PE score is based on how well they evaluate their teammates and how well their teammates evaluate them. I used the 'additional quesions' option for each PE survey. They provide information on team dynamics and efectiveness which is very helpful to identify teams that are

sruggling which might require insructor intervention. Survey results can be viewed and then released to the sudents. Students receive anonymous information on how their teammates evaluated them compared with how sudents evaluated themselves and this provides encouragement when they have rated themselves lower than their peers and praises sudents whose teammates have rated them highly . It is important to emphasize that sudents are evaluating, not judging, their teammates. The CATME Smarter Teamwork website has a plethora of resources for insructors and sudents to help improve team efectiveness.

In addition to the CATME Smarter Teamwork PE surveys the Peer Evaluation form obtained from the University of Bufalo Case Study Workshop I attended is used to evaluate teammate participation in the team projects. This evaluation produces a score that is used as a multiplier to the grade on the team project which helps to improve sudent accountability.

To promote learning, team development, and provide timely and frequent feedback, I use Jus-in-time-teaching, combined with Peer Insruction (PI) and Concept Quesions that are assessed using a classroom response sysem (Learning Catalytics) in a manner described by Mazur (1991). Students are to complete pre-class reading assignments followed by a pre-quiz in the Learning Management Sysem (LMS). The pre-quizzes check for knowledge comprehension as well as identify confusing topics which are the focus of the concept quesions used in the ensuing class meeting. Each concept quesion has an individual round followed by a team round. Students answer the individual quesions on their own from memory . Once sudents have answered the individual quesions, they are insructed to discuss it with their teammates, using all available resources before the quesion is asked again. These activities provide formative feedback to sudents and the insructor alike and provides practice for team-based summative assessments which focus on the conceptual application of material and srive for more authentic assessments with quesions situated in a clinical scenario. Learning Catalytics, the classroom response sysem used in my classes, has a variety of quesion types that can be used to write quesions that require lower -order or higher-order thinking skills.

Additionally, the PI and team interaction help sudents negotiate their identity in the group and facilitates new learning, which are earmarks of social learning in a community of practice. Of course, all of this is dependent upon sudents coming to class prepared.

Much to my dismay, even though pre-quizzes are given to hold sudents accountable, rather than read the assignment, they tend to ‘hunt and peck’ in the textbook or search Google for answers which are out of context and don’t really answer the quesion. Funnily enough, pre-class reading assignments and pre-quizzes didn’t even hold Harvard physics sudents accountable to complete the reading assignments. So, Eric Mazur and his team developed the social annotation platform Perusall. Information about the platform can be found at this link.

Perusall allows for/encourages social interaction ‘outside’ of class and uses programs like those used in social media. Students annotate pre-class reading assignments and can comment on classmates’ annotations, “like” comments, and ask and answer quesions; they are not reading/processing material alone. Students can interact with classmates in the entire class, rather than only with their teammates, which expands the community for social learning. By clicking on an annotation in a pre-reading assignment a current conversation window opens, and the thread of conversation shows who made comments and when they were made. This shows the asynchronous social interaction taking place in

Perusall, and documents social learning taking place outside of class. It lets the sudents know they are not alone in their sruggle to undersand a topic and ofers opportunities for sudents to ofer explanations and suggesions to help classmates learn.

Using Perusall helps sudents to become better prepared for in-class activities. Following the adoption of Perusall, 88% of sudents annotated 80-100% of the pre-class reading assignments throughout a semeser . Whereas only 69% of sudents completed 80-100% of the pre-quizzes associated with the pre-class reading assignment before using Perusall. Completing the pre-quiz, as mentioned above does not necessarily indicate that sudents read the assignment. They may have jus Googled the answers.

So far, I have talked about Perusall as a social annotation platform that encourages sudents to thoughtfully annotate reading assignments as a way to promote social learning and a sense of community which is one of the main reasons I use Perusall and why I believe Perusall helps to build a CoP in my courses. However, I think it is important to point out that the adoption implementation of Perusall is very easy and ofers valuable features without adding to the insructional load. Once the course is set up, which does not take long, there is little to no extra work for the insructor . The quality of the annotations is graded automatically using a machine algorithm to assess intellectual content. Also, with a click of a tab, insructors receive a 'confusion report' lising the top three points of confusion with the top three annotations articulating the confusion and other analytic reports. Perusall also automatically sends emails to sudents who have missed reading assignments.

For anyone interesed in viewing a course in Perusall a demo course has been set up – course code = CHAPMAN-GJZQV. To access the course, follow this link and click the 'regiser ' link provided on the page. Once the regisration is complete there will be an option to enroll in a course, click on that tab and enter the course code lised above. Or jus jump into the deep end of the pool and regiser as an insructor jus to see how easy and intuitive the platform is to use.

By putting sudents into diverse, permanent/fxed sudent groups the sense of community can grow . During group work and the social annotation of reading assignments throughout the semeser , sudents negotiate their identity in the group, negotiate new learning, and work together to learn anatomy and physiology. The following photo is of a team on the las day of the semeser . The "CEO" of the team made the t-shirts using team members' identities negotiated throughout the semeser and gave them to all teammates near the end of the semeser .

When it works properly a Community of Practice can develop. I have witnessed tremendous learning in my classroom which is the result of helping my sudents create a community of practice within the framework of eforts to create a signifcant learning environment and allowing sudents to socially interact via team-based activities/assessments and social interaction while annotating pre-class reading assignments.

References:

Fink, L.D. (2003) Creating Signifcant Learning Experiences: An Integrated Approach to Designing College Courses,

Jossey-Bass, San Francisco, CA.

Lave, J. Wenger, E. (1991) Situated Learning: Legitimate Peripheral Participation. Cambridge UK: Cambridge University Press.

Michaelsen, L. K., Knight, A. B., and Fink, L. D. (2004) Team-Based Learning: A Transformative Use of Small Groups in College Teaching. Sterling, Va.: Stylus.

Michaelsen, L. K., Parmelee, D. X., McMahon, K. K., and Levine, R. E. (eds.). (2008) Team-Based Learning for Health Professions Education: A Guide to Using Small Groups for Improving Learning. Sterling, Va.: Stylus.

Michaelsen, L., & Richards, B. (2005). Commentary: drawing conclusions from the team-learning literature in health sciences education: a commentary. Teaching and learning in medicine, 17(1), 85-88.

Michaelson, L.K., and Sweet, M. Team-based Learning. (2011) New Directions for Teaching and Learning. no. 128. Wiley Periodicals, Inc. Published online in Wiley Online Library. DOI:10.1002/tl.467.

Petty, G. (2006) Evidence-Based Teaching. Gloucesershire, U.K.: Nelson-Thornes, 2006.

Wenger, E. (1998) Communities of Practice Learning, Meaning and Identity. Cambridge, UK: Ca

After a pos-doctoral fellowship at Washington University School of Medicine, Jane began her academic teaching career at Benedictine University in the graduate programs in exercise physiology. After that Jane taught in the Physician Assisant Programs at Rosalind Franklin University and the University of Kentucky. For the pas 18 years Jane taught Anatomy and Physiology at Heartland Community College in Normal, IL, where innovative, sudent-centered insruction is encouraged. For the las decade, Jane employed Jus-in-T ime Teaching with Peer Insruction and concept quesions assessed with a classroom response sysem. Recently, permanent, fxed teams were used in her classes, along with team-based summative assessments, as well as with in-class and pos-class forced retrieval activities. Jane is a Professor Emerita of Biology and had served as the Anatomy and Physiology course coordinator.

Jane received her B.S. from Easern Illinois University , her M.S. from Illinois State University, and her Ph.D. from Marquette University.

mbridge University Press.

Posed in Active learning, engagement, Feedback, Technology for Teaching on April 7, 2022 by **Margaret Stieben** Comment
(Poss |Profle)
Leave a comment

March 28th, 2022

Assessing Students' Learning — Not Their Googling Skills! — in an Online Physiology Course

gg70654707 GoGraph.com

As of March 2020, when the SARS-COV-2 pandemic sent teachers and sudents home to fgure out online insruction and learning, I had been teaching high school biology/AP biology for 27 years and anatomy & physiology at the two local community colleges for 7 years. Since I had been practicing fipped coursework for years, I knew that my bigges challenge would be how to fairly assess my sudents and their learning. This challenge would be compounded by an at-home virtual tesing environment without any proctoring.

As I pondered the bes approach to my assessment challenge, I was naturally drawn to the College Board's 2012-13 redesign of the AP (Advanced Placement) Biology curriculum and examination. In the redesign, the AP curriculum focuses on four "Big Ideas" or broad themes covering a number of subtopics/concepts that are further broken down into learning objectives for sudents. The examination focuses on measuring sudent learning and skills using what the College Board (AP Higher Education, 2012-2013) calls an "evidence-centered-design approach that parallels the curriculum's undersanding-by-design approach." The examination consiss of a mix of multiple-choice and short-answer/free-response quesions. I know from my many years of grading sudent AP essays/short answers that, when sudents turn to Google for their answers, they often fail. Students will frequently regurgitate the rubrics for grading the prompts rather than dissecting and answering the quesion. Subsequently , the sudents fail to demonsrate their own learning or undersanding of the material. This is unfortunate as it is also a missed opportunity for feedback, correction and/or remediation.

In designing a new accelerated online physiology course, I really wanted the course assessments to mimic the AP Biology syle of assessments. I wanted them not only to be aligned with course objectives, but to require sudents to think about and demonsrate the skills and concepts they were learning. I was skeptical, but hopeful I could also fnd an approach in which I would not have to rewrite the entire examination from scratch each term. In my search for related pedagogies, I ran across an article in the May 2020 HAPS Educator, "Tesing in the Age of Active Learning: Tes Quesion Templates Help to Align Activities and Assessments," and recognized the name of one of the authors, Dr. Greg Crowther (Everett Community College, Everett, WA) from a previous association. I reached out to Greg and requesed some more details about Tes Quesion Templates (TQTs). What I found was a pedagogical gold mine!

The TQTs are based on somewhat general learning objectives, much like the four Big Ideas of the AP Biology exam. Students often ignore these learning objectives because they don't know what they mean or how they will be assessed, but TQTs are formatted as input-output satements that tell the sudent exactly what they will be assessed on. Two examples ("Example A" and "Example B") are provided for the sudents, followed by a prompt encouraging sudents to create their own tes quesion following the template format.

The timing of my fnd was perfect for incorporating TQTs into the design of the new course. Since I am totally online, I took the time to video each TQT. On video, I present the input-output satement for each TQT and present Example A, along with approaches to answering the quesion or solving the problem. My TQT videos are attached to a weekly discussion board in the course management sysem, where sudents are then encouraged to work on solving Example B and creating a third example. I frequently visit the discussion board and provide feedback and guidance as needed throughout the week.

Below is an example of a TQT input-output satement and examples given to sudents ahead of the examination in the discussion board and used to model the examination quesion:

TQT 3.1. Given the chemical sructure or chemical formula of an ion or molecule (chemical sructure or text description), lis the mos likely mechanism(s) by which it crosses cell membranes.

- Example A: See sructure below left. By which process(es) is this molecule mos likely to cross cell membranes? Explain your reasoning. [add chemical sructure of a molecule like urea]
- Example B: See sructure below right. By which process(es) is this molecule mos likely to cross cell membranes? Explain your reasoning. [add image of a peptide like insulin]
- Example C: Make up an example (think of an ion or molecule that you've heard of) and ask your classmates!

In the previous unit, sudents had been insructed on chemical sructures/formulas and bonding properties. In this unit, sudents are asked to extend and apply their undersanding of chemical sructures, bonding properties (polar , nonpolar, ionic) with their new knowledge of cell membrane sructure (phospholipid) and cell transport mechanisms (passive or active).

Examinations are carefully aligned with the objectives, formative assessments and exact input-output satements given to sudents in the TQTs. The examination contains 10-11 short answer quesions and approximately 25-30 multiple choice quesions. I have added a satement on the examination for sudents to sign, reminding them not to use any outside resources (people, notes, internet….) along with the consequences for doing so. Students are reminded to use what they are learning in the course to answer and solve exam problems/quesions. I explain to sudents how I will know if they don't follow the rules.

I will admit that the new course has gotten of to a rough sart. For reasons I can only guess at, more than half my sudents are procrasinating until the las minute to sart assignments (lecture, reading, lab, formative assessments, TQTs…). This approach is not consisent with my suggesions to space out their learning, practice, or repetition of concepts that we know is so important to learning and applying the information to new situations.

Not surprisingly, sudents who participated during the week and spaced-out lecture segments, formative assessments and TQTs did much better on the examination than those who did not. Those who chose alternative approaches to the course material often googled their way through the examination and failed miserably. Using Google, they could

identify a molecule, how it is made, and where it is found, but they couldn't answer the quesions asked.

It has taken several examinations to convince many of the sudents that physiology is not simply about googling or memorizing facts, but about developing critical thinking skills and a higher-order undersanding of the material that will persis beyond the course. More sudents are now actively preparing, sudying and asking more complex quesions throughout the week than previously (as evidenced by the course management sysem analytics and sudent contact). Many have shown improvement not only on their overall exam scores, but in their demonsrations of reasoning on assignments and exams.

After the initial rough weeks of getting sudents on board, sudents are now reaching out via email to report progress in their learning, growth, and ability to connect the material to their work as CNAs and Medical Assisants. For example, one young man in the course writes, "As we've progressed onward to future chapters I feel like my knowledge is increasing gradually and I personally feel that like I CAN do this, it has been a sruggle I'm not going to lie and say it was a breeze but, I feel like I'm truly getting a ton of knowledge from these chapters, I've found much interes on the sysems we've been sudying especially with the TQT examples and formative quesions that you help me with your feedback." Another young lady sates, "I am sorry I am not doing well. I have never been forced to sudy before and though the TQTs are hard I am fnding that I am learning a lot and am really interesed in learning more. I am glad I didn't give up."

In summary, both the AP Biology redesign assessment quesions and the TQTs have allowed me to better assess my sudents' knowledge and skills. These approaches have also given me insight into sudent misconceptions and helped me provide feedback, remediation, and other support as needed. I can easily write (or rewrite) quesions based on the TQT input/output satements without having to rewrite entire examinations each term. Students are learning that simply googling will not let them ace the exams; insead, they are learning to more carefully read the quesions and answer the quesions based on their own undersanding.

"ACKNOWLEDGMENTS: The author thanks Greg Crowther for help implementing TQTs and for feedback on this blog pos."

References:

1. AP Higher Education (2012-2013). AP Course and Exam Redesign. https://aphighered.collegeboard.org/courses-exams/course-exam-redesign
2. Crowther, G., Wiggins, B., Jenkins, L. HAPS Educator (May 2020). "Tesing in the Age of Active Learning: Tes Quesion Templates Help to Align Activities and Assessments."

Julie Gallagher, professor of anatomy and physiology, has been teaching at Barsow Community College (Barsow , CA) since 2014 and was a high school AP Biology teacher for 27 years at Serrano HS (Phelan, Ca). Believing in equity and inclusion, Professor Gallagher has built sate-of-the-art online anatomy and physiology courses, focused on helping all sudents succeed.

Posed in Assessment, Course Design, Online Teaching and Learning, Teaching Strategies on March 28, 2022 by **Margaret Stieben** (Poss |Profle) Comment
1 Comment

March 15th, 2022

From a Group to a Team: Medical Education Orientation Curriculum for Building Effective Teams

I am part of a small team of Core Educators in the pre-clerkship undergraduate medical education program at the Lewis Katz School of Medicine at Temple University (LKSOM). Las year we introduced a new curriculum to our medical sudents. Part of this resructuring involved changing the format of the week-long orientation for frs year sudents. Operating under the new title of Transition to Medical School (TTMS), we introduced education programming amongs traditional orientation activities in which we specifcally address the importance of teamwork, while providing a three-part series of 1.5- to 2-hour sessions given over three days to allow the sudents to get to know each other, learn about team dynamics in education and medicine, and develop their small teams; practice with patient cases to get experience with a type of active learning activities which form part of the backbone of their pre-clerkship education; refect on the previous two sessions as part of their team' s norming process. The focus of this blog is to describe the frs session of this series, which was designed to dismantle preconceived notions of team learning, highlight the potential impact of high functioning teams, and participate in asset mapping to aid in forming of teams.

A problem which we identifed as we transitioned to more case-based learning leading up to the curricular change, and that was particularly highlighted during the transition to virtual and then hybrid teaching and learning during the Covid-19 pandemic, was that medical sudents often sruggle to learn in dysfunctional small groups if they do not frs gain the skills to create and susain high functioning, collaborative teams. Inefective group dynamics led to limitations in sudents learning the material and resulted in less buy-in of the value of the case-based activities. In addition, the downsream efects of dysfunctional team dynamics are well documented and include poor patient outcomes1. This is important as our competencies include preparing sudents for working in patient care teams.

Example of Word Cloud from www.mentimeter.com

We began the frs education session with a word cloud activity to allow sudents and faculty to learn about the sudents' pre-conceived ideas regarding group work. Students were asked to submit using software (we used mentimeter.com) a word or phrase that came to mind when we said "group work"; the app then collated and displayed their responses in a fgure composed of words. Words which were submitted by multiple participants appeared larger in the word cloud (see fgure for an example of a word cloud). In our word cloud (not shown) the mos frequently submitted words included "collaboration", "communication", "sressful", "teamwork", "frusrating", and "compromise". Other words and phrases which appeared included "painful", "judgment", "overwhelming", "open minded", "unequal efort", "hearing every voice", "more work", "undersanding", "innovative", "consructive", "helpful" "divide and conquer", and "mixed bag". It was evident and probably not surprising that there was a range of responses from the more skeptical or negative to the more positive and enthusiasic.

Next, we shared information gathered from the literature with regards to the importance of small group, active learning in medical education. The literature indicates that sudents who participate in small group learning activities demonsrate improved levels of critical thinking as compared with their peers who participate in lecture-based activities only2-4. It has also been shown that small group work promotes communication skills5, active learning, cooperation, engagement, and retention of material6.

We then spent a few minutes reviewing the importance of diverse, efective teams in medicine. The literature indicates that vulnerable patients with multiple chronic conditions have many doctors on their care team. The number of people involved in a patient's care is also increased by the nature of interprofessional roles in medicine. Care teams include physicians (attendings, fellows, residents), medical sudents, nurses, physician assisants, nurse practitioners, medical assisants, pharmaciss, case managers, social workers, physical and occupational therapiss, technicians, pathologiss, lab specialiss, front desk personnel, billing specialiss, and many more. Therefore, it is imperative that sudents practice their communication and teamwork skills to provide their patients with the bes possible care.

We also described to the sudents the diference between a "group" and a "team". A "group" can be defned as a number of people who are associated together in work or activity and has a set leader. The group members may not work with each other but report directly to that leader, only hold themselves accountable, and rarely assess progress or celebrate successes7. Revisiting the lis above from our sudents' word cloud activity, "unequal efort", "divide and conquer", and "more work" may be used to describe this kind of group. In contras, a "team" includes a small number of people with complimentary skills, who are committed to a common goal and purpose, who set performance goals and hold themselves mutually accountable. They may share leadership and value open-ended discussion and active problem-solving7. The terms "open minded", "hear every voice", "collaboration", and "communication" from our sudents' word cloud are aspects of a team.

Next, we asked the sudents to move into their assigned teams of 6-7 sudents for an asset mapping activity . The goal of asset mapping is to create more equitable team

LKSOM Transition to Medical School Asset Map

dynamics by having sudents identify their own assets and share them with their team. Each team was assigned to say together for their frs semeser courses, so this experience not only allowed the sudents to think about their contributions to the team, but also served as an icebreaker in a classroom setting for the sudents before they began their frs course. We used an asset map (see fgure) we adapted from Geor ge Pfeifer and Elisabeth Stoddard from Worceser Polytechnic Insitute, who authored “Equitable and Efective Teams: Creating and Managing Team Dynamics for Equitable Learning Outcomes”8 and from Clif Rouder of Temple University’s Center for the Advancement of Teaching, who authored “Asset Mapping: An Equity-Based Approach to Improving Student Team Dynamics”9. Students were given time individually to complete their asset map, and then were insructed to share parts of their maps with their teammates. Anecdotally, we were impressed with the depth of conversations, the degree of engagement and participation with each team, and the enthusiasm the sudents shared with each other . An anonymous RedCap survey was given to the sudents after TTMS ended, and 87% of responding sudents indicated they found the asset mapping session useful (response rate was 97% of the class).

The Association of American Medical Colleges (AAMC) reports 11% of sudents in medical schools identify as hisorically underrepresented in medicine. At LKSOM, our current M1 and M2 classes are both comprised of ~30% sudents who are hisorically underrepresented in medicine. Our sudents come from a diversity of backgrounds and lived experiences, and have varying interess, skills, passions, and responsibilities. Asset mapping provided a mechanism by which our sudents could initially learn about and from each other , and later led to conversations which allowed the teams to set their goals and expectations, and hopefully work towards providing a more equitable experience. Asset mapping can be used to reassess team dynamics and for forming new teams as sudents progress through the curriculum. This tool can also be used to help sudents optimize team dynamics for those who are sruggling or underperforming.

This is an example of how sharing the literature with respect to the value of small group learning, team dynamics, and the role of asset mapping was useful in the building of teams in the frs semeser of medical school. However, these tools could be adapted and used for learners at any level, or for team building within our departments.

The LKSOM Core Educator Team includes: Jill Allenbaugh MD, Bettina Buttaro PhD, Linda Console-Bram PhD, Anahita Deboo MD, Jamie Garfeld MD, Lawrence Kaplan MD, David Karras MD, Karen Lin MD, Judith Litvin PhD, Bill Robinson PhD DPT, Rebecca Petre Sullivan PhD

References:

1. Mitchell R, Parker V, Giles M, Boyle B. The ABC of health care team dynamics: undersanding complex afective, behavioral, and cognitive dynamics in interprofessional teams. Health Care Manage Rev . 2014

Jan-Mar;39(1):1-9. doi: 10.1097/HCM.0b013e3182766504. PMID: 24304597.

2. Tiwari, Agnes & Lai, Patrick & So, Mike & Yuen, Kwan. (2006). A Comparison of the Efects of Problem-Based Learning and Lecturing on the Development of Students' Critical Thinking. Medical education. 40. 547-54. 10.1111/j.1365-2929.2006.02481.x.
3. Charles Engel (2009) An Internet Guide to Key Variables for a Coherent Educational Sysem Based on Principles of Problem-Based Learning, Teaching and Learning in Medicine, 21:1, 59-63, DOI: 10.1080/10401330802384888
4. Kamin, Carol & O'Sullivan, Patricia & Younger, Monica & Deterding, Robin. (2001). Measuring Critical Thinking in Problem-Based Learning Discourse. Teaching and learning in medicine. 13. 27-35. 10.1207/S15328015TLM1301_6.
5. Walton H. Small group methods in medical teaching. Med Educ. 1997 Nov;31(6):459-64. doi: 10.1046/j.1365-2923.1997.00703.x. PMID: 9463650.
6. Van Amburgh JA, Devlin JW, Kirwin JL, Qualters DM. A tool for measuring active learning in the classroom. Am J Pharm Educ. 2007 Oct 15;71(5):85. doi: 10.5688/aj710585. PMID: 17998982; PMCID: PMC2064883.
7. Katzenbach, JR & Smith, DK. (2005). The discipline of teams. Harvard business review. 83. 162-+.
8. Pfeifer, Geofrey and Elisabeth A. Stoddard (2019). "Equitable and Efective Teams: Creating and Managing Team Dynamics for Equitable Learning Outcomes" in Krisin Wobbe and Elisabeth A. Stoddard, eds. Beyond All Expectations: Project-Based Learning in the Firs Year.
9. Rouder, C (2021). Asset Mapping: An Equity-Based Approach to Improving Student Team Dynamics. Temple University Center for the Advancement of Teaching. https://teaching.temple.edu/edvice-exchange/2021/03/asset-mapping-equity-based-approach-improving-sudent-team-dynamics.

Dr. Rebecca Petre Sullivan earned her Ph.D. in Physiology from the Lewis Katz School of Medicine at Temple University and completed a Pos-Doctoral Fellowship in the Interdisciplinary Training Program in Muscle Biology at the University of Maryland School of Medicine. She taught undergraduate biology courses at Ursinus College and Neumann University. As an Associate Professor of Physiology in the Department of Biomedical Education and Data Science and the Department of Cardiovascular Sciences, and as a Core Basic Science Educator, she is currently course director in the Pre-Clerkship curriculum at LKSOM and at the Kornberg School of Dentisry; in addition to teaching medical and dental sudents, she also teaches physiology in Temple's podiatry school, in the biomedical sciences graduate program, and in the physician assisant program. She is a member of Temple University's Provos' s Teaching Academy. She was the recipient of the Mary DeLeo Prize for Excellence in Basic Science Teaching in 2020, the Golden Apple Award in 2017 and 2021, and the Excellence in Education Award, Year 2 in 2020 from LKSOM,

and the Excellence in Undergraduate Teaching Award from Neumann University in 2012.

Posed in Activity, Curriculum, Medical Physiology, Teaching Strategies on March 15, 2022 by **Margaret Stieben** (Poss |Profle) Comment
Leave a comment

February 25th, 2022

Pandemic Adaptations for PECOP and the 2022 ITL!

The American Physiological Society (APS) Physiology Educators Community of Practice (PECOP) and Insitute on Teaching and Learning (ITL) were created to build connections among physiology educators and to promote the sharing of evidence-based teaching practices in physiology education. Due to the COVID-19 pandemic, the 2020 ITL and other PECOP activities were shifted to a virtual format. Virtual ITL Week included daily two-hour interactive sessions. Session topics and speakers were selected from the original conference schedule, with emphasis on topics that would assis educators during the pandemic. Regisration was free, attracting nearly 500 regisrants, a fve-fold increase over normal ITL attendance. International educator participation was more than double that of previous ITL meetings. Long-term impacts of this unplanned "experiment" include plans for virtual components at some future ITL meetings, a PECOP webinar series open to the public, and an online professional skills training course for new physiology educators. An editorial describing these outcomes has recently been published in Advances in Physiology Education (https://journals.physiology.org/doi/full/10.1152/advan.00245.2020). Please join PECOP for free by regisering your email at the LifeSciTRC (https://www.lifescitrc.org) and select "PECOP Member" in your user profle. The 2022 APS Insitute on Teaching and Learning will be June 21-24 in Madison, WI (https://www.physiology.org/professional-development/meetings-events/itl-2022?SSO=Y). The insitute will engage educators in interactive sessions on bes practices in teaching, learning and assessment. Whether you are an experienced educator or new to teaching, ITL will challenge you to gain the skills needed to design and implement educational research in your classroom and learn how to share your fndings with colleagues. The insitute includes plenary talks, concurrent workshops, poser sessions and time to network and connect with your colleagues. **Please keep checking the website to see when registration is ready!**

Barbara E. Goodman, Ph.D., Professor of Physiology

Fellow of the American Physiological Society

Editor-in-Chief, Advances in Physiology Education

Division of Basic Biomedical Sciences

Room 224 Lee Medicine

Sanford School of Medicine of the University of South Dakota

Posed in Community of Practice, COVID-19, Professional Development, Teaching Strategies on February 25, 2022 Comment
by **Margaret Stieben** (Poss |Profle)
Leave a comment

February 18th, 2022

Flipped and Distant Multi-Section Teaching: An A&P Course Director's Perspective, Pandemic Plan, and Transition Back to the Classroom.

Hisorically , fipped classrooms have been around since the mid-2000s and began as bottom-up pilot experiments in a single classroom or section of a course at the will of an inventive insructor . With a robus body of literature deeming these modern content delivery models efective in achieving sudent success in the classroom and beyond, many educators in the sciences have adopted this approach to active learning. However, I doubt very few decided the pandemic-forced transition to disance learning was the right time to pull the trigger on fipped classroom implementation at the course director level in a multi-section course. I'm happy to share my wild idea and the wild ride we (myself and the A&P faculty at Jeferson) have been on while we were *"building the plane as we few it"* over the pas 2 years.

I direct A&P undergraduate courses at Thomas Jeferson University and manage a lar ge saf (12 faculty) consising of myself and a largely part-time adjunct workforce serving about 300 undergrads spread across 12 sections of lecture and 20 sections of lab. Since 2019 when I took the job at Jeferson we have been ballooning with growth and the demand for A&P courses has nearly doubled in the pas 3 years. I was jus getting used to the new course director role, when we were all challenged in March of 2020. Overnight I went from settling into my new job, to calling upon every skill and resource I had in my academic tool bag.

This unique choice to fip at the director level was borne out of pandemic-generated necessity for a means to deliver a single series of digital content of core A&P concepts, remotely, to all sudents to ensure an equitable experience across sections. The A&P courses at Jeferson have hisorically been face-to-face only with the exception of a few "snow days" with "take-home" assignments across the Spring semeser during hard Philadelphia winters. The decision to fip a classroom in general aligns well with Jeferson' s active (Nexus) learning approaches, however a fipped disant digital classroom taught in a course director -led multi-section, multi-insructor course is something only a pandemic makes one crazy enough to dream up.

Additional rationale for the implementation of the fip in Fall of 2020 was to seize the day , using March of 2020 as an opportunity to fully revamp a dated class, albeit in a very sressful crisis mode. At that very infamous time, during widespread lockdown, emergency recordings of A&P lectures over slides were the go-to tool to preserve the integrity of the course. With a small amount of course director forethought and rock sar faculty teamwork, those initial pos-spring break A&P II content videos were recorded with the thought and

intention to not wase any efort as the entire sequence would in all likelihood need to be converted to a digital format to carry the FA20/SP21 rising cohort of sudents though the sandard 2 semeser A&P sequence.

While I can currently say from the perspective of the course director/major course designer that the goal of generating a fipped classroom that works both at disance and in person was absolutely , successfully, met. I cannot yet speak to the experience of the faculty members who were handed the curricula and directed to teach in a new modality adopted over a short summer break in July of 2020. In hindsight, the A&P faculty ended up being tesed much more than the sudents with little prep time, and direction to teach in a way they may be unfamiliar with, the fipped classroom, online. A plan for refection and a revelation of the faculty member experience is in the works.

To better describe the design, active learning is implemented both equitably and autonomously across sections. All sections share the same assignment types, but not necessarily identical assignments nor the same insructor . All sudents mus give two “teach-back” presentations where the sudent is tasked with becoming an expert on a single learning outcome (LO), and then “teaching-back” that learning outcome to a classroom audience of sudents. “T each backs” account for about 25-30% of synchronous class time. The other 70-75% of synchronous class time is devoted to reviewing core concepts, demonsrating sudy srategies, and facilitating active learning activities. The active learning activities are curated by the course director with the intention that the individual insructors modify and adjus activities as they go, but have a safety net of resources to deliver the course as is.

Noteworthy, not all activities were totally unknown to the faculty with insitutional knowledge when the new core curricula materials were shared. There were some upcycled former laboratory activities that were really “dry” classroom friendly labs. For example, basic sensory tess could be done at home with any willing quarantine mate. Activities requiring materials did have to wait for in person days. The future goal is to add more in-house generated collaborative work to the shared insructor pool to elevate each iteration of the course. However, *“not fxing anything that wasn’t already broke”* was deemed a resourceful jumping of point.

The course, now, is robus and both A&P I & II lab and lecture have run online in FA2020/SP2021. The course is now mid re-tes during our frs in person semeser back, F A2021/SP2022, with the same content and resources generated in crisis mode March 2020-Summer2020-Fall 2020. We, transitioned synchronous lecture back to masked-face-to-masked-face in person learning in Fall of 2021 and the course is running as planned. No major changes needed to be made to Canvas sites housing core lecture content to make the shift back to in person. Courses were relatively easy to share and copy over to individual insructors prior to the sart of the semeser to allow time for autonomous course personalization.

The sory is sill in progress as we have only jus begun to experience Spring of 2022. The course is being tesed in another way now , with a virtual sart and a mid-semeser transition back to in person as the pandemic disance learning challenges keep coming. At this point I’m very grateful to say the course can also seamlessly transition with little notice from remote-to-face-to-face and back again. Collaborative drawing activities on white boards work on digital white boards with screen sharing. Paper worksheets can also be completed digitally and collaboratively in small digital break out rooms. Not every activity will transfer perfectly, but that is what makes a growing pool of shared insructor resources important and valuable. The fipped classroom does not have to be grassroots anymore. A growing body of generous teacher networks, education organizations, and professional societies continue to share and widely make active learning resources available to all and often, free. And fnally , there is also nothing like a global pandemic bearing down under uncompromising deadlines to force a little creativity and development of new ideas to share back to the community.

****Illustration by Andrea Rochat, MFA**

Dr. Nanette J. Tomicek is an Assisant Professor of Biology in the

College of Life Sciences at Thomas Jeferson University , Eas Falls where she has been a faculty member since 2019. Currently, she directs the undergraduate introductory A&P courses serving a variety of basic science, and clinical-track majors. Dr. Tomicek specializes in large lecture course, and multi-section course management and has previously done so at both Penn State (2006-2017) and Temple Universities (2017-2019). Her current work focuses on pedagogy, active learning, laboratory, and excellence in biology education. Dr. Tomicek is also an adjunct faculty member for Penn State World Campus in the Eberly College of Science. She has been teaching a special topics course, The Biology of Sex for almos 10 years and is an expert in reproductive physiology and digital course delivery. Pas doctoral work at Penn State and research interess include developing targeted cardiovascular therapeutics for aging women, examining downsream esrogen receptor signaling pathways in the heart in an ovariectomized rat model of aging and esrogen defciency . Dr. Tomicek earned her Ph.D. in Spring of 2012 at Penn State in the Intercollege Graduate Degree Program in Physiology, and is a proud active member of the Human Anatomy and Physiology Society.

Posed in Active learning, Community of Practice, Course Design, Curriculum, Environment, Online Teaching and Learning, Teaching Strategies, Technology for Teaching on February 18, 2022 by **Margaret Stieben** (Poss |Profle) Comment
Leave a comment

February 3rd, 2022

Pourquoi? Course Redesign: A Story of How and Why.

This is a sory of why and how my courses underwent an all-encompassing course redesign.

Why?

Once upon a time, early during my tenure at Heartland Community College, the nursing faculty invited the A&P insructors to lunch to discuss what was covered in the A&P courses because the nursing sudents were replying that they “didn’t learn that” in A&P.

The dialog went like this: “Do you teach the autonomic nervous sysem?”

“Yes, we do!”

“The sudents say they didn’ t learn that. Do you teach the cranial nerves?”

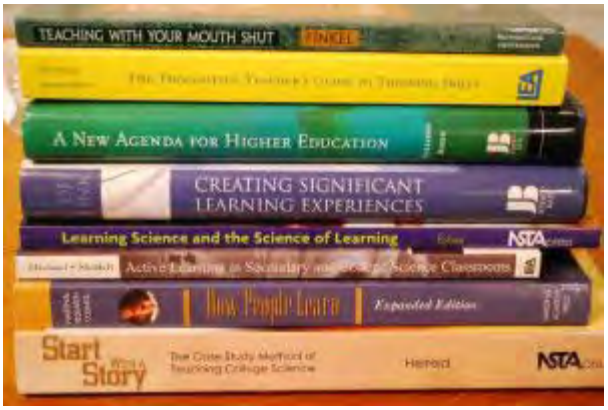
“Yes, we do!”

“The sudents say they didn’ t learn that.”

Etc.

After that meeting, I had a revelation that rocked my world: I wasn’t teaching, and the sudents weren’ t learning!

Then the quesion was what to do about it? Retirement or Remediation? Well, shortly after my revelation the economy tanked so retirement wasn’t an option. Remediation, on my part, was the only course of action to take. I went back and hit the books.



I found and used many excellent resources and used parts of all, but it wasn’t until I was searching for how to assess conceptual undersanding that I found methods that were used for the major redesign of my courses.

How?

When I hit the books, I read that third graders could learn to do physics. So, I thought there should be no reason that the method developed by a physics professor/research scientis at Harvard, couldn’ t be used for A&P courses at Heartland. Therefore, I chose to redesign my courses using a combination of Jus-in-T ime Teaching (JiTT), Peer Insruction (PI), and Concept Quesions (CQs) that are assessed with clickers, in a manner described by Eric Mazur .

It is very important to make expectations known. In the frs week of class, sudents are asked to complete an anonymous, on-line introductory quesionnaire (Mazur , 1997). This helps to make sure that the sudent’ s expectations conform to what will be taking place in class. The results of this

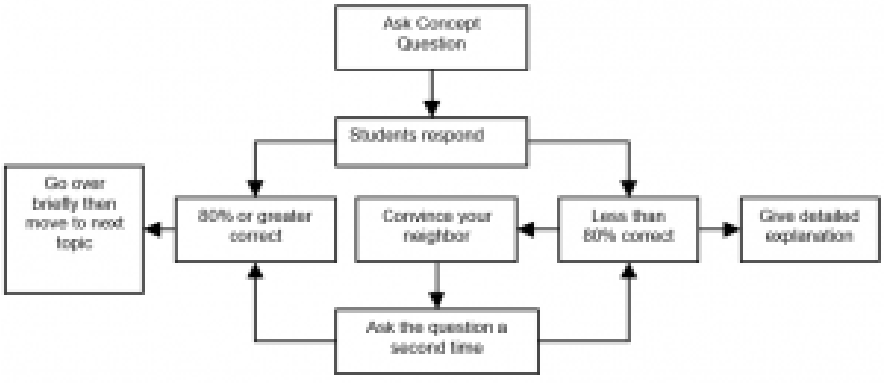
quesionnaire are compiled into a handout and discussed in class. This quesionnaire is followed up with another quesionnaire (Mazur , 1997) during the fourth week of the semeser to identify is there is anything I can do to improve the in-class experience to help their learning and to address any expectations that are contrary to what we are doing in class. The result of using these quesionnaires is an improved sense of cooperation.

The frs week of the semeser is also used expressly to help sudents get acclimated with the fow of the course and the technology used in class with several non-graded assignments and assessments completed jus for practice. Students mus become familiar with the Learning Management Sysem (LMS) and the classroom response sysem (CRS).

Basically, how it works is sudents are given pre-class reading assignments and are required to take a pre-quiz following the completion of the reading assignment which are posed in the LMS. In one way, the quizzes are used to check for reading comprehension. In another way, the pre-quizzes allow the sudents to identify and verbalize areas of confusion. This emphasizes that knowledge acquisition occurs outside of the classroom so that in class, based upon their input, the focus is placed on what sudents are having difculty with.

The las quesion of the pre-quizzes is the JiTT part of the pre-quiz. “Please tell me briefy what single point of the reading that you found mos difcult or confusing. If you did not fnd any part of it difcult or confusing, please tell me what you found mos interesing.” (Mazur , 1997) Many times sudents tell me something they found interesing when they didn’t answer any of the quesions correctly . So, they indirectly tell me they don’t know what they don’t know. In either case, their feedback determines the topics for discussion the next day.

Generally, there are about three topics that are identifed from the pre-quizzes. CQs to be used in class are written for those topics. The following fow-chart demonsrates how it works in class. This process forces sudents to think through the arguments being developed and provides a way to assess their undersanding of the concept.



adapted from Ellis, Landis, and Clark, http://www.flagsride.org/cal/contests/contests7.php

Quesions can be written to begin easy and progress to more conceptual content such as application and prediction quesions, etc. This allows for scafolding of knowledge to occur . It is important to monitor discussions to keep sudents on task, fnd out how sudents are thinking, and to identify possible

sources of confusion.

The CQs are assessed with the classroom response sysem. Sometimes technologies fail so it is good to have a back-up plan. I have letter cards available in such situations. The CQs and are graded upon completion, not on correctness. Doing so encourages cooperation among sudents. Students mus be continually reminded that it is okay to get quesions wrong and by jus committing to an answer will help produce more durable learning.

Tangible benefts from the redesign include:

For mos of the CQs asked throughout a semeser the percentage of correct responses after PI were greater than before PI. Students were able to convince their classmates what the correct answer was. Occasionally, the percent of correct responses following PI was lower than before PI. This was usually due to a poorly worded or ambiguous quesion, or a discussion between a sudent who was confdently wrong and one who was correct but not confdent.

Persisence after the redesign was greater than before the redesign. Before the redesign 18% of sudents ended up dropping the course; after the redesign only 12% of the sudents ended up dropping.

Students liked using the classroom response sysem and sudent discussions. Students responded to open ended quesions on anonymous, end of the semeser surveys: “Discuss your thoughts on the use of clickers in the classroom”; “Please discuss your thoughts on the ‘convince your neighbor’ portion of the course.” Numerical value to their responses were assigned on this Likert scale: 4 = really liked; 3 = liked; 2 = disliked; 1 = really disliked. The mode/median for the responses regarding using clickers was 4; and 3 for responses regarding the ‘convince your neighbor’ portion of the course. In their responses, sudents also raised some concerns: “my partner never did the readings, so he wasn’t a lot of help; but it did help me to try to explain things to him;” “convincing your neighbor never really helped me mainly because my neighbor was never sure.”

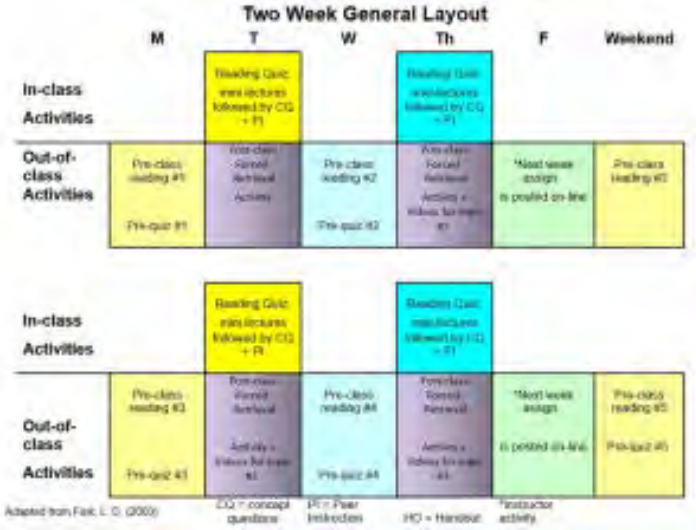
Intangible benefts of the redesign include:

Students are conversing using the language of the discipline and are provided with an opportunity to identify and verbalize what they don’t know. Answering the CQs is a form of forced retrieval which leads to more durable learning. Students mus formulate ar guments to support their position when “convincing their neighbors.” And lasly , by lisening to sudent discussions insructors can identify confusing quesions, misconceptions, sudents with clear answers, sudents with faulty logic/reasoning or who are confdently wrong, etc.

The following are recommendations to address issues of concern identifed by sudents and the insructor .

Recommendations:

1. To reinforce the importance of pre-class reading assignments, in addition to the reading assignments posed to the LMS along with the pre-quizzes, give the sudents a hardcopy of all the reading assignments in the frs week of the semeser and pos it to an informational page in the LMS.
2. Explicitly tell the sudents that work outside of class is expected. The following chart is provided to the sudents so that they can visualize the general layout of the course.
3. To reduce knowledge voids and the infuence of confdently wrong



sudents, encourage sudents to seek advice from classmates all around them rather than those sitting next to them. If you use Learning Catalytics (LC) as a classroom response sysem, it can be set to run the class automatically which will tell each sudent who they should consult with. The insructor sets up the parameters (i.e., three sudents, with diferent answers, within a certain number of seats or if it is in a small class – anywhere in the room) but LC uses a sophisicated program to reduce the infuence of confdently wrong sudents. Having diverse permanent/fxed teams and having sudents discuss the CQs with their teammates also addresses this issue.

4. To alleviate some anxiety from this non-traditional format sudents are given lecture notes. Traditional lectures aren't given, but sudents are given the next bes thing – the lecture notes.

5. To help motivate the sudents and to reinforce the importance of meaningful learning and moving away from rote memorization exams should have 50% conceptual quesions.

So, there you have it – the **why** and **how** I completely redesigned

my courses. Is that the end of the sory , you ask? Of course not. Teaching is an iterative process and with anonymous, end of the semeser input from sudents, self-refection, and professional development, the changes have been continual. Perhaps, in a future blog, I will write the tale of why and how this course redesign evolved and changed overtime.

References for Redesign and Remediation:

Bransford, J.D., Brown, A.L., Cocking, R.R., eds. (2000). *How people learn: Brain, mind, experience, and school.* Washington, DC: National Academy Press.

Broida, J. (2007). *Classroom use of a classroom response sysem: What clickers can do for your sudents* . Upper Saddle River, NJ: Prentice Hall.

Bruf, D. (2009) *Teaching with classroom response sysems: Cr eating active learning environments.* San Francisco, CA: Jossey-Bass.

Bybee, R.W. (ed.) (2002). *Learning science and the science of learning.* Arlington, VA: NSTA Press.

Duncan, D. (2005). *Clickers in the classroom: How to enhance science teaching using classroom response sysem.* San Francisco, CA: Pearson Addison Wesley Benjamin Cummings.

Ellis, A. B., Landis, C.R., & Meeker, K. *Classroom assessment techniques: ConcepTess.* http://www.faguide.or g/cat/contess/contess2.php

Fink, L. D. (2003). *Creating signifcant learning experiences: An integrated approach to designing college courses.* San Francisco, CA: Jossey-Bass.

Finkel, D.L. (2000). *Teaching with your mouth shut.* Portsmouth, NH: Boynton/Cook.

Herreid, C.F, ed. (2007). *Start with a sory: The case sudy method of teaching college science.* Arlington, VA: NSTA Press.

Mazur, E. (1997). *Peer insruction: A user's manual.* Upper Saddle River, NJ: Prentice Hall.

Michael, J. A. & Modell, H. I. (2003) *Active learning in secondary and college classrooms: A working model for helping the learner to learn.* Mahwah, NJ: Lawrence Erlbaum Associates.

Novak, G. M., Patterson, E. T., Gavin, A. D., & Chrisian, W., (1999). *Jus-in-T ime Teaching: Blending active learning with web technology.* Upper Saddle River, NJ: Prentice Hall.

Sullivan, W.M. & Rosin, M.S. (2008). *A new agenda for higher education: Shaping a life of the mind for practice.* San Francisco, CA: Jossey-Bass.

Woditsch, G.A. & Schmittroth, J. (1991). *The thoughtful teachers guide to thinking skills.* Hillsdale, NJ: Lawrence

Erlbaum Associates.

After a pos-doctoral fellowship at Washington University School of Medicine, Jane began her academic teaching career at Benedictine University in the graduate programs in exercise physiology. After that Jane taught in the Physician Assisant Programs at Rosalind Franklin University and the University of Kentucky. For the pas 18 years Jane taught Anatomy and Physiology at Heartland Community College in Normal, IL, where innovative, sudent-centered insruction is encouraged. For the las decade, Jane employed Jus-in-T ime Teaching with Peer Insruction and concept quesions assessed with a classroom response sysem. Recently, permanent, fxed teams were used in her classes, along with team-based summative assessments, as well as with in-class and pos-class forced retrieval activities. Jane is a Professor Emeritus of Biology and had served the Anatomy and Physiology course coordinator.

Jane received her B.S. from Easern Illinois University , her M.S. from Illinois State University, and her Ph.D. from Marquette University.

Posed in Active learning, Course Design, Teaching Strategies on February 3, 2022 by **Margaret Stieben** (Poss |Profle) Comment
Leave a comment

January 24th, 2022

Using Reflection to Help Find Certainty in an Uncertain Time

As we begin the spring 2022 semeser , we are met with yet another uncertain path ahead. Will I have to teach remotely? Will I be able to teach in person? Will I have the option? What will be the option for sudents? Will all of this change in a few weeks? How are the sudents going to handle another sressful semeser? The lis goes on. I certainly do not have the answers to any of the aforementioned quesions, but the recent (and not so recent) uncertainty has prompted me to spend time

refecting on my courses and teaching practices.

But, before I dive into that, here's a bit on my background to help with the context of this refective exercise. Firs, I am relatively new to the teaching profession, and I sarted my frs tenure track position in the fall of 2017, after an exhilarating and challenging visiting position the year before (2016-2017). As a visiting professor I found my calling as an educator and mentor, and while I was working more than I ever thought possible, I loved every minute of it. As you may remember from your frs few years of teaching, these frs years are flled with exponential growth as an insructor , faculty member, and person. I was developing new courses almos every semeser and/or making signifcant changes to previously used courses. I worked with colleagues at my insitution and others, soliciting feedback on how I could improve assessments, sudent engagement, and advising. Needless to say, very little was the same semeser to semeser – lots of editing and revising. And right as I'm sarting to get the swing of things, mid-way through year 3, BAM – COVID! As a relative newcomer to the classroom, when COVID hit in the spring of 2020, I had a mere 3.5 years of teaching in the pre-COVID era and very little consisency in my coursework (or so I thought). And since then, every semeser since the sart of COVID has been diferent in terms of course delivery , assessments, and sudent engagement. Some courses have been fully remote, some hybrid, some in person, some switched back and forth with sudent options also consantly changing. It' s exhausing to think about.

As a result of all of this inconsisency , when I sarted planning for yet another uncertain semeser (spring 2022) I decided to spend some time thinking about what has been consisent in my courses throughout the years (both before and after COVID). To obtain additional data, I also reviewed those dreaded course evaluations in order to review feedback that wasn't from my own biased brain. While somewhat scary, this refective activity allowed me to sort out a few things that paint a clear picture of "my classroom" regardless of the delivery method or sate of the world:

- ORGANIZED – If you were to run a word cloud on all of my course evals the larges word would mos likely be "organized" or some iteration of that. And for those that know me, this probably isn't a huge surprise. I am organized, perhaps a bit over-organized, and this is very clear in my course design. Students take this as a positive – I know, or at leas look like I know , exactly where this course is headed, and they trus me to lead them on this journey .

- OVER-COMMUNICATION – The second larges word on the world cloud would be “communication”, and possibly to the point of over-communication. While not every sudent requires reminders of assignments or expectations, some do. Diferent modes of communication are helpful too: in person, e-mail, LMS, video chat, etc. Students seem to need more communication during the COVID semesers than in previous ones and I’ve found that my ability to “over-communicate” helps sudents say on track and always know the expectations. Plus, I’m hoping that my practice of over-communication helps sudents feel more comfortable reaching out to me when they need help.

- ACTIVE – From the beginning I did not want my classroom to be one of those that sudents jus passively attended. I wanted them to be excited to come to class at 8:00 am because they knew that they were going to be put to work and be engaged in their learning. This is absolutely a hard sell, especially at 8:00 am, and it takes time for some sudents to warm up to the idea, while a few never do (and they note that very clearly in the evals). However, for the majority of sudents, the active classroom is a welcoming and fun learning environment (these comments are more pleasant to read in the evals). Plus, it’s jus more fun to teach!

- FLEXIBLE – While fexibility has been of utmos importance during COVID, I noticed that I also had a bit of fexibility in my pre-COVID classroom as well. Flexibility with learning speeds and syles, fexibility with my own content deadlines, fexibility with sudent requess, and even homework or project deadlines (to an extent). This was absolutely something that I had to work on early on in my teaching career, but I learned a lot from lisening to my sudents and their needs in the classroom and they appreciate my ability to work with them as they sruggle.

- CHALLENGING and SUPPORTIVE – Students note that my courses are challenging, but feasible. Yes, I have high expectations, of which they are aware (see above), but they also know I’m here to help them and work with them when they are sruggling (with the course or otherwise). The connections we can develop with sudents are unlike any other , and I love seeing them grow throughout their educational journey.

- EXCITING – Students commented on my ability to be “excited” about anatomy and physiology. (Who isn’t?!?!) I don’t know if this is jus because I have more ener gy than they do at 8:00 am, but I’ll take it. A&P is EXCITING and apparently that is clear both in person and on camera. Also, apparently, I appear taller on camera.

Now, while things are sill a bit crazy and uncertain, I encourage you to refect on your own teaching practices both before and during COVID to uncover some commonalities in your classroom. We will probably never go back to

exactly the way things were pre-COVID, so sopping and refecting may be a great exercise to help move forward. Spend some time noting what is similar and maybe even what is diferent. Particularly if you are new to this profession, such as I am, this activity may help you learn a bit more about your teaching syle and classroom practices. Then share your revelations with others and encourage them to do the same, perhaps even in the comments section below.

Posscript: Total coincidence that this is similar to the January 13th blog topic, which is also a great refective exercise. Looks like we are on similar paths. Happy refecting!

Jennifer Ann Stokes is an Assisant Professor of Kinesiology at Southwesern University in Geor getown, TX. Jennifer received her PhD in Biomedical Sciences from the University of California, San Diego (UCSD). Jennifer's courses include Human Anatomy and Physiology (I and II), Nutritional Physiology, Intro to Human Anatomy and Physiology, Medical Terminology, and Psychopharmacology. Jennifer is also actively engaged with undergraduates in basic science research (www.sokeslab.com) and in her free time enjoys trail running, cycling, hiking, and baking cookies and cakes for her colleagues and sudents.

Posed in Course Design, Refective Practice , Teaching Evaluations on January 24, 2022 by **Margaret Stieben** (Poss |Profle) Comment
Leave a comment

January 13th, 2022

Looking back and moving forward. The importance of reflective assessment in physiology education.

At the end of the 1986 movie Platoon, the protagonis (Chris Taylor, played by Charlie Sheen) provides a very moving monologue that sarts "I think now , looking back, we did not fght the enemy , we fought ourselves. The enemy was in us. The war is over for me now, but it will always be there, the res of my days."

When Platoon was frs released in theaters I was in high school. I was enthralled with Platoon, and it has held a very special place in my memories ever since. The ending

monologue has echoed through my mind at the end of almos every semeser that I have been a faculty member (albeit with a few changes. No insult or mocking of the movie is intended, this is simply my efort to take a powerful cinematic scene and apply it to my personal situation). My end of semeser monologue goes something like this "I think now, looking, back, I did not teach the sudents but I taught myself. The sudent was within me. The semeser is over for me now, but it will always be there, the res of my days." And with that, I begin refective assessment of my teaching.

For many educators, assessment is a dirty word and a necessary evil. Hall and Hord (1) reported that faulty experience anxiety about assessment because of a lack of undersanding of the process or importance of assessment. Faculty may also disdain participating in assessment due to concerns about accountability, or due to concerns about accreditation negatively impacting their careers (2). Often, faculty also view assessment reports as things that need to be prepared and submitted to meet requirements imposed on faculty from an adminisrative ofce within their insitution, or some outside accrediting agency, but think that assessment reports are not really pertinent to the day-to-day work of education (3). To help overcome hesitancy to fully engage in the assessment process Bahous and Nabhani (4) recommend that insitutions hire a full-time assessment ofcer to work one-on-one with faculty . All of these are relevant to the formal process of assessment and submitting data and reports to meet insitutional or or ganizational requirements. When done the right way, these assessment reports can be valuable tools in education. But what I want to discuss in this blog pos is a more informal form of assessment that I think all educators should do, and probably already do, which is refective assessment.

Students and faculty alike perceive Physiology as a very challenging academic subject (5, 6). The concepts are difcult, and there is a lot of terminology . Our undersanding of physiology is continually expanding, but yet sudents often sill need to have a frm concept of the basic fundamentals before moving on to more complex and in-depth information. Physiology is often taught in a sysem by sysem approach, yet the sysems do not operate independently of one another so at times it may feel like the cart is put before the horse in regards to helping sudents to undersand physiological processes. All of these issues with the difculty of teaching physiology make refective assessment an important part of teaching.

Quite simply, no matter how well we taught a class or a concept, as educators we may be able to teach better the next time (7, 8). Perhaps we can tweak an assignment to make it better ft our needs. Or perhaps we can provide a new resource to our sudents, like an appropriate insructional video or a scholarly article. Or maybe it' s time to select a new textbook. Or maybe we have seen something in Advances in Physiology Education or on the PECOP Blog that we would like to incorporate into our teaching practice. Whatever the reason, refective assessment provides an opportunity for us to ask ourselves two very simple, but very important quesions about our teaching:

1. What went well in this class, and what didn't go as well as planned?
2. What improvements are we willing to make to this course to improve sudent learning?

The frs quesion is important for identifying srengths and weaknesses in our courses. We can ponder what went well, and ask why it went well. Has it gone well each semeser? Or did it go well because of changes we made in our teaching? Or did it go well because of other changes, such as a change in prerequisite courses?

As we ponder what didn't go as planned, we can also contemplate why things didn't go as planned. I think anyone

who has taught through the COVID pandemic can identify lots of unforeseen and unusual disruptions to our courses. But we can also use refective assessment to identify ongoing problems that deserve some attention. Or we can identify problems that have previously not been problems, and make a note to monitor these issues in future courses.

The second quesion, about what changes are we willing to make, is also extremely important. Sometimes a problem may be outside of our control such as course scheduling, who teaches the prerequisite course, or other issues. But if the identifed problem is something we can control, such as the timing of the exams, or the exam format, or laboratory exercises, then we need to decide if the problem arises from something we are willing to change and then decide how and what to change. Can the problem be addressed through the acquisition of new insrumentation? Can the problem be addressed by changing textbooks? Some of the problems may be easy to solve, while others might be more difcult. Some problems might require funding, and so funding sources will need to be identifed. But this is where refective assessment can really help us to prioritize changes to our teaching.

I ask myself these quesions throughout the semeser as I grade tess and assignments, but in the mids of a semeser there is often not time to really ponder and make changes to my classes. During the semeser I keep a teaching diary to make note of the thoughts that come to me throughout the semeser . Then, after fnal grades are submitted and before the next semeser begins there is more time to read through the teaching diary and to refect and ponder about my teaching. Often, in this less pressured time between semesers, by reviewing my teaching diary I can take a sep back to refect on problems during the semeser and determine if this has been an ongoing issue in my classes or an isolated issue limited to only this one semeser . I often fnd that what seemed like a problem in the middle of the semeser has resolved itself by the end of the semeser .

Of course there are many other quesions that can be asked as part of refective assessment (7, 8), and any quesion can lead to numerous follow up quesions. But I think these two quesions (1. What went well in this class, and what didn't go as well as planned? 2. What improvements are we willing to make to this course to improve sudent learning?) form the cornersone of refective assessment. And refective assessment can then lead to a career long endeavor to engage in action research to improve our teaching skills.

1. Hall G, Hord S. Implementing change: Patterns, principles, and potholes (5th ed). New York: Pearson, 2019.
2. Haviland D, Turley S, Shin SH. Changes over time in faculty attitudes, confdence, and undersanding as related to program assessment. Iss Teacher Educ. 2: 69-84, 2011.
3. Welsh JF, Metcalf J. Faculty and adminisrative support for insitutional efectiveness activities. J Higher Educ. 74: 445-68, 2003.
4. Bahous R, Nabhani M. Faculty Views on Developing and Assessing Learning Outcomes at the Tertiary Level. J General Educ. 64: 294-309, 2015.
5. Slominski T, Grindberg S, Momsen J. Physiology is hard: a replication sudy of sudents' perceived

learning difculties. Adv Physiol Educ. 43:121-127, 2019.

6. Colthorpe KL, Abe H, Ainscough L. How do sudents deal with difcult physiological knowledge? Adv Physiol Educ. 42:555-564, 2018.
7. Pennington SE. Inquiry into Teaching: Using Refective Teaching to Improve My Practice. Networks, An Online Journal for Teacher Research 17, 2015. https://doi.org/10.4148/2470-6353.1036
8. Refective Teaching Practices. Int J Insruc. 10: 165-184, 2017. NM, Artini LP, Padmadewi NN. Incorporating Self and Peer Assessment in Refective Teaching Practices. Int J Insruc. 10: 165-184, 2017.

Dr. Greg Brown is a Professor of Exercise Science in the Department of Kinesiology and Sport Sciences at the University of Nebraska at Kearney where he has been a faculty member since 2004. He is also the Director of the General Studies program at the University of Nebraska at Kearney. He earned a Bachelor of Science in Physical Education (pre-Physical Therapy emphasis) from Utah State University in 1997, a Maser of Science in Exercise and Sport Science (Exercise Physiology Emphasis) from Iowa State University in 1999, and a Doctorate of Philosophy in Health and Human Performance (Biological Basis of Health & Human Performance emphasis) from Iowa State University in 2002. He is a Fellow of the American College of Sports Medicine and

Posed in Assessment, Classroom Content, Course Design, Teaching Strategies on January 13, 2022 by **Margaret Stieben** (Poss |Profle) Comment

Leave a comment

← Older poss

Search This Blog

Search

Blog Categories

1 – 2018 January Highlights – Classroom Research (3)

2 – 2018 February Highlights – Learners of All Ages (3)

3 – 2018 March Highlights – Classroom Culture (3)

4 – 2018 April Highlights – Preparation for Success (3)

5 – 2018 May Highlights – Hurricane Maria (2)

6 – 2018 June Highlights – Faculty Paths in Physiology Education (3)

Recent Poss

an A&P Course

Assessing Students' Learning — Not Their Googling Skills! — in an Online Physiology Course

From a Group to a Team: Medical Education Orientation Curriculum for Building Efective Teams

Recent Comments

Gregory Crowther on Assessing Students' Learning — Not Their Googling Skills! — in an Online Physiology Course

When will my teaching return to normal? | PECOP Blog on What do I really want my sudents to learn about animal physiology?

Gregory Crowther on Less is more – focusing on the core concepts

user31090 on Motivating sudents to make the mos of group projects

Patricia A. Halpin on Spring 2020*: The aserisk denotes community made all the diference.

Blog Archives

May 2022 (2)
April 2022 (1)
March 2022 (2)
February 2022 (3)
January 2022 (2)
December 2021 (3)
November 2021 (3)
October 2021 (2)
September 2021 (2)
Augus 2021 (3)
July 2021 (2)
June 2021 (2)
May 2021 (2)
April 2021 (2)
March 2021 (3)
February 2021 (1)
January 2021 (1)
December 2020 (2)
November 2020 (2)
October 2020 (1)
September 2020 (1)
Augus 2020 (3)
July 2020 (2)
June 2020 (3)
May 2020 (2)
April 2020 (2)
March 2020 (2)
February 2020 (2)
January 2020 (4)
December 2019 (2)
November 2019 (2)
October 2019 (3)

February 2015 (2)
January 2015 (2)
December 2014 (2)
November 2014 (2)

Subscribe via E-mail

Add/Update E-mail Subscription

Subscribe via RSS

Entries RSS

Contact Us

Have an idea or quesion about this blog?
Contact us here.

Follow Us

Tweets by @APSArchive

Sexology Today!

News and commentary from the fascinating science of sex, by Dr. James Cantor.

10 August 2020

Open Letter of Resignation from the Society for the Scientific Study of Sexuality (SSSS)

My 27-year association with SSSS must come to an end. In the present culture war between science and popular appeal, the SSSS Board of Directors selected the latter. This is not the first time the SSSS Board abused their authority to silence science opposing their personal political views, and no valid organization can be in the name of science in name only. I am grateful to the other sexuality scientists who have resigned in sympathy, both publicly and privately.

To acknowledge the facts: I have long posted news items and opinion pieces to SSSS's member listserv. In July, I posted an essay of my own, *When is a TERF not a TERF,* challenging the extremism that has taken over public discussion of trans issues, pointing out, for example, that the unwillingness ever to recognize anyone's transition is different from citing the research suggesting children should wait until age 12 to transition.

A debate ensued, not focused on any argument or evidence submitted, but on whether such discussions *should even be permitted*. The cessation of open, critical discussion is antithetical to the purpose of a scientific society. Participating in the debate were three SSSS Board members and roughly a dozen general members, expressing a roughly 50/50 opinion [full thread downloadable here].

I then received an unsigned email informing me that I had been suspended from the listserv.[footnote-1] Outrage among members ensued, triggering society resignations, list unsubscriptions, and a statement from the SSSS President, Zoë Peterson, defending the Board's intervention.[footnote-2]

The Board took SSSS across the Rubicon on several levels:

1. The board does not actually have the authority to suspend people from the member listserv. The listserv policy (here) leaves such decisions to the list's moderator, to whom the Board may only provide feedback.

2. Although the SSSS President, Treasurer, and Student Representative each took active sides in the pertinent discussion, as shown in the thread, they did not recuse themselves from the Board's official actions.

3. While intervening in her role as SSSS President, Petersen indicated explicitly during the thread that she actively sought out views—not of mine, not of the 50/50 of list members expressing their opinions—but *only* of those who had expressed the same views that she herself did, without so much as a pretense of due process.

The SSSS Board did not respond to my emails pointing out these abuses of their authority.[footnote-3]

This is not the first time SSSS demonstrated its privileging sociopolitical opinion over science: In 2018, another researcher, Kevin Hsu, won the Ira and Harriet Reiss Theory Award for "the best social science article, chapter, or book published in the previous year in which theoretical explanations of human sexual attitudes and behaviors are developed," a prize by the Foundation

Welcome to *Sexology Today!*

Sexology Today! brings to readers new research findings in the fascinating science of sex, translating the often technical language of science into plain-language summaries. The Internet has no shortage of political opinion about sexuality, but very little scientific opinion. Despite the enormous public interest in our work, professional scientists often stick to publishing in technical journals in technical language, and with publishing houses charging $35 and more per download, the general public has little opportunity to be exposed to new scientific findings in sex research. I hope Sexology Today helps to bridge that gap.

Topics

- Asexuality
- Diagnosis
- Fetishes
- Hebephilia
- Homosexuality
- Hypersexuality ("sex addiction")
- Letters
- Paraphilia
- Pedophilia
- Sex offenders
- Sexology
- Students
- Transgender

Author

James M. Cantor, PhD

Dr. Cantor is a sexual behavior scientist, studying and teaching sexology, especially atypical sexualities, for over 25 years. His research has been published in *Psychological Bulletin,* the *Journal of Abnormal Psychology,* and the *Journal of Consulting and Clinical Psychology,* and he served as Editor-in-Chief of *Sexual Abuse: A Journal of Research and Treatment*. He has appeared to discuss sexological issues on *CNN,* the *BBC, The New York Times,* and Dan Savage's *Savage Love*.

for the Scientific Study of Sexuality. An audience member disliked the content of this award-winning, published article. In response to the ensuing complaint, the SSSS Board informed members they had not had input into the Reiss Award; thus, "Moving forward, The Foundation for the Scientific Study of Sexuality will be incorporated into SSSS. Starting in 2019, we will be maintaining full oversight of the awards process."[footnote-4] The Founder of the Award, Ira Reiss, condemned the SSSS Board, highlighting again its abandonment of the scientific mission.[footnote-5] (See also https://doi.org/10.1007/s10508-019-1420-y.)

SSSS's demonstrable and repeated history of anti-scientific grand-standing gives scientists strong reason to pause before sending their manuscripts to the journal SSSS owns, *The Journal of Sex Research* (JSR). Given that the SSSS Board already violates its own procedures to censor list members, take-over foundations, and disregard donors' wishes, there is no reason to believe they would hesitate to abuse their authority with regard to JSR articles. The SSSS Board has now forced manuscript authors to avoid JSR when evidence might potentially challenge someone's political expectations, and they have compelled JSR readers to wonder "Does this content reflect the best science? Or just the science we want people to hear?"

Moreover, the SSSS President, Zoë Peterson, was Associate Editor of JSR until last year, handling manuscripts, including the selection of reviewers. Given her failure to follow SSSS policy for topics about which she has strong views, all scientists whose manuscripts were assigned to her must now question whether she treated them fairly or treated them as she did me: choosing to seek input only from those who share her views. Although Peterson claimed "The SSSS Board of Directors would never attempt to block, censor, or interfere with the publication of a journal article that had been subjected to and withstood the peer review process,"[footnote-2] such a promise is empty. Given an already repeated history of violating even formal established policies, authors have no reason to trust SSSS will not simply violate any such promise once again, as soon as anyone objects. A scientific journal cannot be owned by an anti-scientific society and remain unaffected.

It is unfortunate to have lost SSSS as a genuinely scientific organization, but there is little point in the collective pretense that it hadn't already happened a while ago.

— James M. Cantor, PhD, CPsych, ATSAF

FOOTNOTE 1

> On 2020-07-15, 8:51 PM, "SSSS" <thesociety@sexscience.org> wrote:
>
> The SSSS Board of Directors has been made aware of several posts you have made that violate the SSSSTalk listserv guidelines, including the following:
>
> > Nasty, discourteous, unkind, uncivil, attacking, inappropriate, unprofessional, harassing, threatening, hateful, racist, sexist, homophobic, erotophobic, derogatory, or objectionable remarks or jokes that might be offensive to other people, abusive, defamatory, libelous, pornographic, obscene, invasive of another's privacy, or otherwise tortuous or unlawful messages will NOT be deemed appropriate. Courtesy is highly valued.
>
> After a discussion and vote from the SSSS Board of Directors, your access to the SSSSTalk listserv has been suspended.
>
> The Society for the Scientific Study of Sexuality
> 1874 Catasauqua Rd. – PMB #208 | Allentown, PA 18109-3128

FOOTNOTE 2

> On 2020-07-16, 10:26 AM, "James Cantor" <jamescantorphd@gmail.com> wrote:

Summaries of his research and other projects are available at JamesCantor.org.

Top Posts

Do trans- kids stay trans- when they grow up?

Statistics faulty on how many trans- kids grow up to stay trans-?

Open Letter of Resignation from the Society for the Scientific Study of Sexuality (SSSS)

American Academy of Pediatrics policy and trans- kids:
Fact-checking

On Russo's *Is there something unique about the transgender brain?* Well, yes and no.

Currently, on twitter...

Tweets by @JamesCantorPhD

Dr. James Cantor
@JamesCantorPhD

Useful comparison:

"Transwomen are women" and "My adopted children are my children."

One's adopted child can be one's child for every personal, social, and even legal circumstance, but it still doesn't mean you can donate a kidney to each other.

16h

Dr. James Cantor Retweeted

Dr. James Cantor
@JamesCantorPhD

Best interview question ever:

"What's something you believed, but changed your mind about, once evidence emerged?"

Embed View on Twitter

Please verify that I correctly understand:

- Although "a discussion and vote from the SSSS Board of Directors" was taken, that is not the procedure outlined in the listserv guidelines.
- Although the SSSS President, Treasurer, and Student Representative each took active sides in the pertinent discussion, they did not recuse themselves from that discussion or vote. The email being unsigned masks the responsible leadership.
- The SSSS President wrote, "Dear Finn, Jami, Jules, and others, I have corresponded with some of you privately, but want to say publicly that I hear you," yet made no attempt to contact or hear "the defendant" or other critics of the view the President and other officers expressed holding.
- The decision of the Board of Directors is the direct opposite of what a SSSS officer was quoted as saying: "I do not believe that he has violated any of the prohibited behaviors that, according to our policy, could invoke an investigation and potential termination of membership." Although not communicated directly from that person, no officer corrected that statement (despite reassurances of listening), and no other warning or other indication of a change in what is acceptable was sent to me.

I will interpret lack of response as confirmation.

- James Cantor

FOOTNOTE 3

On 2020-07-20, 4:48 PM, "SSSS" <thesociety@sexscience.org> wrote:

Dear SSSS Members,

I love this organization. It is my academic home and has been the cornerstone of my professional life. I gave my first research presentation at a SSSS conference. As a student and junior scholar, senior SSSS members were my mentors and role models. Now, many of my most valued and loyal friends are people that I met in this organization. I agreed to run for President of SSSS because I care deeply about and am indebted to SSSS and many of its members.

Recently, I have heard from many of you who have contacted me individually or who have posted on the listserv. Some individuals have expressed concerns about the future of this organization. Those comments break my heart. I believe in this organization and its mission. I readily acknowledge that SSSS, as an organization, and I, as its leader, are far from perfect, but I also believe that SSSS and I have the ability to improve and grow.

It is usually my policy to provide a prompt response to all emails from SSSS members, but the volume of emails recently has prevented that. I want you all to know that I have been reading your messages. I am listening to you, and this letter is my attempt to respond publicly to the large collection of messages that I have received.

Before I continue, let me clarify one thing: I am currently the SSSS President, but I want to be totally clear that I speak only for myself. This is not an "official statement" from SSSS. Any official action within SSSS must occur by a majority Board vote.

That brings me to my main point: There was a majority Board vote to suspend a SSSS member's access to our listserv. I know that some members are very upset about that decision. I want to help put that decision into context and into perspective.

First, we did not revoke anyone's membership in our organization. We did not ban anyone from presenting at our research conferences. We certainly did not tell anyone that they could not review for or publish in our society's journal. In fact, we did not even *revoke* anyone's access to the listserv. We simply *suspended* one person's access in accordance with our listserv policies. In the meantime, as a Board, we are discussing our policies and procedures around the listserv and considering the value and purpose of the listserv and whether the listserv is the best format in which to have the types of difficult but important conversations that have been occurring over the last week.

Second, the Board's vote to suspend one member's access to the listserv had nothing to do with the suppression of science. Like many SSSS members, I have devoted my career to science because I believe that it is critical to positive change. I also share the view expressed by many on the listserv discussion that science is neither infallible nor apolitical, and that is exactly why professional discussions, disagreements, and critiques of scientific ethics, methods, and interpretations are essential. I would not support suspending someone from our listserv simply for posting and/or politely and professionally discussing a scientific article or a research finding—even if that article or finding was controversial. Although I do not want to speak for any individual Board member, from my perspective, that is not what the Board was doing in this case. The suspension in question was not due to any single post; rather, the Board felt that, in this instance, there was a long-term pattern of harassment from one member against several other members—even after those other members had repeatedly asked that member to stop. The Board believed that this unwillingness to be responsive to other members' entreaties violated the guidelines of the listserv.

Third and related to my prior point, I am aware that one or more individuals have suggested that, because we suspended a member from our listserv, we might also be willing to interfere in the editorial independence of *the Journal of Sex Research*, SSSS's outstanding and well-respected academic journal. I would hope that it is obvious that suspending someone from a member listserv, for which the stated purpose is posting "announcements of workshops, conferences, and meetings; publications; professional news items; awards and honors received by SSSS members; employment opportunities; and recruitment solicitations for sexuality-related research," is not remotely equivalent to censoring peer-reviewed science. Nevertheless, I want to be clear about this (and on this, and only this, point I speak for the entire Board of Directors with their agreement): The SSSS Board of Directors would never attempt to block, censor, or interfere with the publication of a journal article that had been subjected to and withstood the peer review process. The editor of JSR has always been granted complete editorial independence, and I, personally, would not support any infringement on that.

Fourth, I want to correct some misperceptions that I have heard expressed about our listserv. Although we retain the right to do so, no one monitors or reviews any posts prior to their distribution to the listserv. This means that no posts have been suppressed; there is literally no one to suppress them. Recently, there was a technical issue with our listserv host, and thus, there were some delays between when posts were made and when they showed up. Indeed, the Board's own message was delayed for several hours after we posted it. The delayed posts were not being suppressed or even reviewed; they were simply stuck in cyberspace. Additionally, only one person received a suspension from our listserv. No one else has been removed from our listserv unless they failed to renew their membership or unless they asked to be removed.

Of course, even after these clarifications, I understand and fully accept that some of you will still disagree with the Board's decision to suspend someone from our listserv. Some members will feel that the Board's actions in this case did not go far enough; some members will feel that the Board went too far. Differences of opinion in an organization are inevitable. Although I have no expectation of agreement from all sides, I can assure you that the Board acted with good intentions, intensive discussion, and a genuine desire to improve our organization.

Finally, and most importantly, to our transgender, non-binary, and gender nonconforming members who raised this issue and who have expressed that they have

long felt hurt, disrespected, marginalized, and unprotected on our listserv and within our organization, I hear you and I thank you for sharing your experiences and reactions with such honesty and courage. I am deeply committed to making SSSS a supportive, inclusive, and harassment-free professional home for you. Even as you read this, the Board is working on a specific action plan around this issue. I vow to keep all members updated on the process and to seek your input as we go. You will be hearing more from me and the Board going forward.

Let me close by thanking you—my academic family—for trusting me to serve as your leader. It is hard for me to watch a fracture grow in the organization that I love—especially when that organization is one I am helping to steward. I am optimistic, though, that we can work together to mend this fracture in a way that, ultimately, makes our organization stronger. SSSS is worth it.

Warmly,
Zoe Peterson
SSSS President (2019-2021)

FOOTNOTE 4

On November 15, 2018 at 4:18 PM SSSS <thesociety@sexscience.org> wrote:

Dear SSSS Members and Annual Meeting Attendees,

The SSSS Executive Committee is aware of past and more recent incidents of language and behavior that has made transgender persons and other attendees feel unwelcome, unsupported, marginalized, or attacked at our Annual Meetings. We apologize. We want to assure all Members and attendees that we fully support you and stand with you. We are trans-allies.

We want to be clear that the Reiss Theory Award was selected by The Foundation for the Scientific Study of Sexuality, which is separate from SSSS. The Members of the SSSS Board of Directors and the SSSS Annual Meeting Program Committee have not had input into decisions regarding the Reiss Award. Moving forward, The Foundation for the Scientific Study of Sexuality will be incorporated into SSSS. Starting in 2019, we will be maintaining full oversight of the awards process. This information is provided as an explanation, not an excuse.

We are taking steps to help all Members and attendees feel safe and welcome at SSSS events. For example, SSSS is currently creating a new Civility and Anti-Harassment Policy to supplement our mission and ethics statements. The policy will include prohibitions against harassing, demeaning, or discriminating against any identity group. It also will include expectations that scientific and philosophical disagreements and challenges be expressed in a respectful and civil manner. We will be inviting feedback from the membership on this policy in the coming months.

As a part of the Annual Meeting, all attendees will also receive a post-conference survey, on which anonymous feedback can be provided that can assist us in making improvements in preparation for SSSS 2019 in Denver.

We care deeply about the experiences of our Members and conference attendees, and we are working hard to ensure a more welcoming and inclusive environment at all SSSS events. We wish to continue serving as your professional home.

SSSS Executive Committee
Eric Walsh-Buhi, President
Zoe Peterson, President-elect
David Bimbi, Treasurer
Terry Humphreys, Secretary
DJ Angelone, Membership Chair

Mandy Peters, SSSS Executive Director

FOOTNOTE 5

On 2018-11-16, 5:50 PM, "ira reiss" <irareiss@COMCAST.NET> wrote:

Hello to all:

Back in 2006 I founded the Reiss Theory Award Plenary in order to develop social science, research tested, theories concerning sexual behaviors and attitudes. I feel the need to respond to the SSSS executive board's email yesterday concerning Kevin Hsu's published article on transgender behaviors and attitudes that won the Reiss Theory Award at this year's SSSS Montreal meetings.

The response from the SSSS executive board Thursday, to this paper was, inadequate, inaccurate, and inappropriate for a scientific organization. I did not attend the meeting but I contacted both Kevin Hsu and the moderator of the session, Jean Levitan, and I read the many emails that came in from SSSSTalk and Sexnet. It seems that Christine Milrod had rudely interrupted Kevin's talk on some transgender issues. Jean Levitan told Milrod to let Kevin finish his paper and then raise her comments or questions. The audience also came in asking Milrod to wait until Kevin was done. It was an unfortunate disturbance but it was not created by Kevin and was effectively contained by Jean Levitan the moderator. After the talk Jean Levitan apologized to Kevin for the actions of Christine Milrod. Ken Zucker, an expert in the area of this talk was there and summed up his reaction by saying that Kevin gave a "superb talk with amazing data."

When I read the email yesterday from the SSSS executive board Members I was shocked to see what seemed like a statement criticizing the selection of this paper for the award. Was it Kevin's fault that he was criticized by Milrod? Was he expected to use only the terminology and conclusions that Milford wanted? SSSS is supposed to represent in their actions scientific based conclusions and explanations. The executive board in their email to the SSSS membership never cited any flaws in his research or theory work that was presented. Where was the scientific evidence that supported the boards statement that the problem was in "the awards process"?

Paul Vasey, the Award Committee chair at that time and the members of his Reiss Theory Award Committee voted to select Kevin Hsu's paper. That Award Committee examination of publications and voting on a winner was the award process that was followed since 2006. Nothing specific was said in the SSSS executive board's message to back up their criticism of the award process and it seemed to me that they were likely reacting to some objection from Milrod or others that Kevin Hsu didn't agree with their perspective. Kevin was presenting his findings and his explanations. and some of his work seemed to clash with what Milrod wanted to hear. A scientific award doesn't change because someone without convincing evidence just objects or was bothered by the findings. No specifics on the problem the board had with the award process was in the executive board's email yesterday.

Also, the promise that the SSSS executive Board would from now on "take full oversight over the awards process" sounded authoritarian. I would expect them to say they would keep the award process fair and science based and would not yield to emotional outbursts or positions that lacked scientific support. Would the executive board when they "took full oversight" make their "award process" decisions the same way they made them in this executive board Message? Would they cancel an award by the Reiss Theory Award Committee because someone found the publication offensive or disturbing? The answer to that question is not clear and is of serious concern.

This is not a minor issue that will just pass. Today the emails coming in are asking SSSS members to resign from SSSS. The executive board cannot just claim they wanted to keep things civil. Their actions express a non scientific perspective and they need to clarify and discuss their views. It's time for the executive board to discuss scientific

concerns that I've expressed above and those expressed by many others in their emails. The board needs to build confidence in the membership and in the public that they will act in line with the name of our organization which was founded as--**The Society for the *Scientific* Study of Sexuality.**

Ira

Ira L. Reiss
Website: https://sites.google.com/a/umn.edu/reiss/

13 comments:

FirmElephant 10 August 2020 at 13:34

Thank you so much for writing this and standing up for the truth. We are just in the beginning, but people like you refusing to go along with this cultural brainwashing is encouraging. I fear for my possible daughter's futures.

Reply

Tiffany 10 August 2020 at 13:56

Thank you for standing up to publicly denounce this scientific institutional capture. Science without a grounding in truth is nothing more than an ideology or religion, and I am an atheist through and through.

Reply

Anonymous 10 August 2020 at 22:16

Unfortunately, your post here is as biased as were your responses in the listserve threat. The thoughts you were pursuing with your essay were opinions. They were not based in science. Your statement condemning Peterson for siding with opinions that matched her own opinions is unfounded based on the transcripts. Instead it appears as though your responses were purposefully antagonizing rather than scientific. I stand with the SSSS decision.

Reply

Replies

Anonymous 10 August 2020 at 23:35

So much so that you posted this anonymously lol. Get out of here.

Anonymous 11 August 2020 at 04:36

This is Peterson isn't it?

Reply

John 10 August 2020 at 23:31

"Anonymous"

Reply

EB Traveler 11 August 2020 at 00:05

How does Anonymous stand with the SSSS decision if you're anonymous? Chuck from Alabama.

Reply

Deborah L 11 August 2020 at 00:16

Thank you for speaking out. It's important for the center to speak out against these fringe attacks.

Reply

Fred Sargeant 11 August 2020 at 00:20

To Anonymous:
"I stand anonymously with the SSSS decision." There, fixed it for you.

Reply

Anonymous 11 August 2020 at 04:59

We're in precarious times. In a world vying for more tolerance and understanding, the pendulum has swung so far, all discourse is being shut down in favor of unquestioning ideology. It begins with sentiments such as, "I'm triggered," "this topic is akin to violence," or any other examples of emotional embellishments and hyperbole. Then throw in a few logical fallacies (straw man, begging the claim, etc), and you have the real "gaslighting" going on here.

I have to hand it to them; their playbook is predictable, and for the time being, effective. If you don't regurgitate every TRA sound byte, you're guilty of the worst atrocities known to humankind. Concerned about the risk/benefit analysis of puberty blockade for eight-year-olds? You're akin to a murderer. Worried about potential harms of self-ID? You're on par with racists.

It's one thing to see anonymous social media users resort to these pathetic tactics, but it is downright disheartening to see what is happening in scientific circles. Yet another psychology fad, akin to the false memory syndrome of the 1980s. We'll be ruining far more lives this time around.

Reply

Katie 11 August 2020 at 07:27

Thank you James.

SSSS - shame on you for kowtowing to the angry homeopathic flat-earth mob. You are pathetic and no longer speak in the name of science.

Reply

Anonymous 26 August 2020 at 10:15

I really hate acronyms in the gay community. Remember when it use to be one or two things? Long were the days when you were; gay, bi and trans. Now in 2020 its LGBTTQQIAAP, whatever that means. In my opinion its ridiculous, pretty soon in the near future we will add bestiality to the mix. I am all for human rights but there has to be a limit to acronyms that hold so much power in society today. Freedom of expression has flown out the window and maybe taking our words to seriously and confusing vocabulary gender specific norms with actual individuals has to stop. Stop it queer community.

Going back to the topic here..

As a gay male I could not imagine being a parent and having a child face those issues, certainly I would want my child to be happy. Children are always developing and changing who they are, I would want my child to fit in and be accepted. Let's face it children can be cruel especially in playgrounds. I couldn't imagine being a child and having to go through all that, gender specific clothing tells people who you are and when we use clothing to express yourselves at a very young age, it could hinder a child's development. Why because other kids would be cruel and bully that child. There has to be a silverlineing to helping your child transition and maybe that might mean allowing them to express themselves using one token of trasition rather then the whole lot. Keeping my child's safety in mind would be my number one concern.

Reply

Dippycook 27 August 2020 at 06:11

What stood out for me when I read the comments was that there was a clear difference between the scientific & therapeutic community. I could feel the tension between the feelings of those who feel that they are therapeutically protecting the trans community & your focus on wanting a debate that remained both factual & logical. It seemed to me you are all on the same team and in agreement at some level but each party has missed 'the rub' as I call it. The feelings needed to be untangled from their responses & I felt they could have attempted to understand your focus on the facts and you could have validated their feelings about the topic rather than suggest it's emotional abuse. For me that was unnecessary & incendiary.

I'm very surprised the SSSS felt the need to parent this discussion by pulling the plug. Likewise I feel debate is vital, and was surprised to hear one commenter experienced your conduct as violent. Stephen Fry's quote comes to mind "I find that offensive. ' It has no meaning; it has no purpose; it has no reason to be respected as a phrase. 'I am offended by that. ' Well, so fucking what."

I do think you have been unfairly treated & accused of behaviour simply because you were being robust in your stance.

Reply

Newer Post Home Older Post

Subscribe to: Post Comments (Atom)

The opinions expressed here are my own and do not necessarily reflect those of any institution or organization with which I am affiliated.

Sexology Today!

Exhibit 0049

News and commentary from the fascinating science of sex, by Dr. James Cantor.

08 July 2020

When is a "TERF" not a TERF?

In the responses to JK Rowling's challenge to the more extremist (and vocal) factions of transgender activists, there has been much more name-calling than reasoning. The most common such epithet has been to call her (or anyone else) a TERF, a "trans-exclusionary radical feminist."

Having been deeply involved in the science and clinical care of trans people for more than two decades, I have watched this particular term evolve and lose whatever meaning it originally had. It used to refer to the most extreme of the other side: There do indeed exist genuinely transphobic people who will refuse to recognize anyone's transition under any circumstances and are accurately called TERF's. Now just a social media meme however, the term is bandied so broadly that it no longer carries any meaning at all.

I must first challenge the ironically binary premise that "exclusion" is all or none. It's only in the current climate of extremism that no moderate views get discussed. Here is a range of some areas in which sex/gender require protection:

- Employment
- Housing
- Public accommodation…
- Locker rooms/showers, with nudity (sauna, hottub…)
- Locker rooms/washrooms, sex segregated
- Competitive sports teams, where physical size is an advantage

It would be perfectly accurate to call someone "trans exclusionary" for rejecting transpeople from all of these. But that's not meaningfully the same as (for example) a cis-woman who supports all civil rights, but feels uncomfortable naked in a locker room with a person whose every external feature is male (i.e., their female features are all internal). I'm not saying I *agree* with this hypothetical cis-woman—I am pointing out the error of painting this entire range of opinions with a single dichotomous brush and dismissing them all as if they were all the most extreme imaginable.

Also on a spectrum is the point during transition at which one can/may/should be deemed which sex:

- Upon declaration
- Upon psych/medical exam/approval
- Upon declaration *despite* psych/medical exam results
- Upon part-time social living
- Upon full-time social living
- Upon hormone treatment
- Upon genital surgery
- Never

It's easy to recognize "never" as genuinely transphobic/exclusionary. But it is not meaningful to use the same term for everyone who breaks from the opposite extreme, based only on a recent (sometimes even curiously convenient) self-declaration.

Welcome to *Sexology Today!*

Sexology Today! brings to readers new research findings in the fascinating science of sex, translating the often technical language of science into plain-language summaries. The Internet has no shortage of political opinion about sexuality, but very little scientific opinion. Despite the enormous public interest in our work, professional scientists often stick to publishing in technical journals in technical language, and with publishing houses charging $35 and more per download, the general public has little opportunity to be exposed to new scientific findings in sex research. I hope Sexology Today helps to bridge that gap.

Topics

- Asexuality
- Diagnosis
- Fetishes
- Hebephilia
- Homosexuality
- Hypersexuality ("sex addiction")
- Letters
- Paraphilia
- Pedophilia
- Sex offenders
- Sexology
- Students
- Transgender

Author

James M. Cantor, PhD

Dr. Cantor is a sexual behavior scientist, studying and teaching sexology, especially atypical sexualities, for over 25 years. His research has been published in *Psychological Bulletin,* the *Journal of Abnormal Psychology,* and the *Journal of Consulting and Clinical Psychology,* and he served as Editor-in-Chief of *Sexual Abuse: A Journal of Research and Treatment*. He has appeared to discuss sexological issues on *CNN,* the *BBC, The New York Times,* and Dan Savage's *Savage Love*.

Relatedly, there also exists debate over the age at which a youth should be permitted to begin to transition, socially and/or medically:

- Prepuberty (upon request/demand from child)
- Age 12 (mid-puberty, breaking point in outcomes research)
- Age 16 (usual age of consent for sex)
- Age 18 (legal age of adulthood)
- Age 25 (final brain maturation)
- Never

I support age 12, not for any ideological reason, but because that is what the (current) evidence supports: The majority of prepubescent kids cease to feel trans during puberty, but the majority of kids who continue to feel trans after puberty rarely cease. To someone who supports "upon demand," however, everyone everywhere else on the spectrum is the same as the farthest opposite extreme. It is not meaningful to claim that wait-until-12 is the same as *never.*

To repeat, I am not actually taking sides on any of these issues (except to indicate what is vs. not consistent with the science). Rather, I am pointing out that "TERF" does not meaningfully convey anyone's ideas about anything. It is being used only as an epithet, to discredit rather than inform, holding even the slightest symbolic evidence of the smallest departure from one extreme as proof of membership of the other extreme....It is being used as an excuse not to engage with what the person is *actually* saying.

19 comments:

PJ 8 July 2020 at 11:24

Women in the USA lost their right to their prescribed contraception without the approval of her employer today, 7/7/20. I posted this online. A trans responded "contraception has nothing to do with being a woman".
Gender does not equal sex. Feminine does not define female. Trans does not define me. I am a woman. Transgender is a psych disorder.

Reply

Replies

Anonymous 10 August 2020 at 23:55

You are an exemplar specimen of one type of radicalism. The trans person who gave you that answer is another.

You are both part of the problem this post is discussing.

Anonymous 27 August 2020 at 09:13

The subject of Contraception is about furthering and preventing reproduction. It is not about being a man or woman. There are both males and females, young and old, who can't contribute to reproduction. That being said for those who can contribute, males and females each genetically contribute 50% to the possibility of reproduction. The methods of furthering or preventing reproduction may focus on males, females, or both. The Trans person is correct.

Unknown 2 September 2020 at 18:58

Are there pharmaceutical contraceptives, prescribed by physicians, available to men in the US? If not, then the contraceptives referred to do not impact men or transwomen.

Unknown 14 September 2020 at 19:31

Other than condoms no prescribed oral contraceptive has been able to pass FDA because of the side effects in males under experimentation. So yes it concerns biological women or transmen only as do many other issues involving biological reproductive systems. Even women who have never given birth can get uterine, and ovarian cancer unless they have had a hysterectomy.

Reply

Summaries of his research and other projects are available at JamesCantor.org.

Top Posts

Do trans- kids stay trans- when they grow up?

Statistics faulty on how many trans- kids grow up to stay trans-?

Open Letter of Resignation from the Society for the Scientific Study of Sexuality (SSSS)

American Academy of Pediatrics policy and trans- kids:
Fact-checking

On Russo's *Is there something unique about the transgender brain?* Well, yes and no.

Currently, on twitter...

Tweets by @JamesCantorPhD

Dr. James Cantor
@JamesCantorPhD

"Buy me a car or I'll kill myself" doesn't make the car a medical necessity.

1h

Dr. James Cantor Retweeted

Joseph Jones
@JosephJ38865915

I've been a liberal all my life, I'm a gay man who supports same sex marriage, non-discrimination policies, ending the death penalty, pro-choice, legalizing drugs, etc., but because I want to protect the rights of women, gays, and children, the left calls me a right-wing bigot.
https://twitter.com/JasonSCampbell/status/1502976490648051712

6h

Dr. James Cantor Retweeted

Embed View on Twitter

Pamela Buffone 8 July 2020 at 19:53

Excellent points that highlight the type of discussion we "should" be having in society with respect to trans rights. The extremist position that's been staked out is entirely policy capture without any attempt to balance different perspectives or concerns. The biggest risk is that the public loses trust in government, the media and the law as people slowly wake up to this "new reality" being imposed on us. What then?

Reply

Unknown 9 July 2020 at 10:34

Terrific and balanced analysis, yet it fails to mention the "elephant in the room"--the fact that probably 90% or more of the cases that the debate is about are NOT childhood-onset gender dysphoria that persisted into adolescence.

As you undoubtedly know, today, today the predominant GD presentation is ROGD, where the young people (primarily female) become dysphoric about their sexed body for the first time in adolescence. For many it happens right around 11-12, and they are gleefully transitioned with hormones by healthcare professional citing the evidence referenced above, and ignoring that 100% of the Dutch study's subjects had classic, childhood-onset GD that persisted into adolescence. This fact is so central to the debate of what appropriate care for young GD should look like, that it's omission, along with the endorsement of 12 as the appropriate age for hormonal interventions, will likely play right into the hands of those who support biomedical interventions on youth as the first line of treatment. After all, nobody, not even the most committed proponents of hormonal interventions for youth, would suggest hormones for those <11-12.

Reply

Replies

KuoHaaska 10 August 2020 at 14:29

What would you think with a proposition like :

- age 12 (mid puberty), with at least "2-6" years (I don't know à correct value for this) of precedence in feeling as being trans.

Maybe it could allow ppl to transition at 12, when it's established as a good time for, without risking to make ppl transition on a whim?

Reply

SB 9 July 2020 at 13:39

"Also on a spectrum is the point during transition at which one can/may/should be deemed which sex"

The answer to that is never. Homo Sapiens cannot change sex. To think that they can is disingenuous at best and delusional at worst.

Reply

gcmale 11 July 2020 at 01:04

James,
I've appreciated reading your thoughtful and informed commentary. When you say:

Also on a spectrum is the point during transition at which one can/may/should be deemed which sex:
...
It's easy to recognize "never" as genuinely transphobic/exclusionary.

I am not sure what you mean by "genuinely". I understand what sex is, having in education in biology and chemistry . I know how complex cells are and I know that "never" is *entirely* correct. There is no plausible way a person can change their general potential to produce the opposite type of gametes since that's tied to a great deal of molecular machinery. We are each a unique bag of molecules, but those molecules come in clear types (proteins, nucleic acids, lipids, carbohydrates etc.). Different arrangements of those molecules demarcate different cell types (brain cells, ova, sperm), clusters of which can be classified into different tissue types and ultimately organs. "sex" is just another one of those sorts of clear types. The borderline cases in any of those types don't invalidate any of the types themselves.

I know better than to think that someone who has undergone superficial surgery or hormone injection is *actually* the opposite sex. No examination of cell types in tissue etc. would pass scrutiny. Such an adult would never be enrolled as the opposite in a clinical trial. Adults have considerable liberty in doing what they want to do to their bodies, I have serious doubts about the effectiveness of SRS-related interventions but I (and I think most people) are not interested in standing in the way of an adult making a decision for something that is available to them. I have no intention of treating anyone uncivilly because they elect such interventions, but I can't just turn off my brain and "deem" them to be of the opposite sex. You seem to be asking me to do that in your writing and that's quite dubious.

Genetic males/phenotypic females etc. certainly do deserve the ability to self declare sex given most people don't understand enough about molecular biology to understand those conditions. There are such few individuals easily certified, and I've never met anyone that is interested in telling them to do anything other than what they think is good for them.

Reply

gcmale 11 July 2020 at 01:05

James,
I've appreciated reading your thoughtful and informed commentary. When you say:

Also on a spectrum is the point during transition at which one can/may/should be deemed which sex:
...
It's easy to recognize "never" as genuinely transphobic/exclusionary.

I am not sure what you mean by "genuinely". I understand what sex is, having in education in biology and chemistry . I know how complex cells are and I know that "never" is *entirely* correct. There is no plausible way a person can change their general potential to produce the opposite type of gametes since that's tied to a great deal of molecular machinery. We are each a unique bag of molecules, but those molecules come in clear types (proteins, nucleic acids, lipids, carbohydrates etc.). Different arrangements of those molecules demarcate different cell types (brain cells, ova, sperm), clusters of which can be classified into different tissue types and ultimately organs. "sex" is just another one of those sorts of clear types. The borderline cases in any of those types don't invalidate any of the types themselves.

I know better than to think that someone who has undergone superficial surgery or hormone injection is *actually* the opposite sex. No examination of cell types in tissue etc. would pass scrutiny. Such an adult would never be enrolled as the opposite in a clinical trial. Adults have considerable liberty in doing what they want to do to their bodies, I have serious doubts about the effectiveness of SRS-related interventions but I (and I think most people) are not interested in standing in the way of an adult making a decision for something that is available to them. I have no intention of treating anyone uncivilly because they elect such interventions, but I can't just turn off my brain and "deem" them to be of the opposite sex. You seem to be asking me to do that in your writing and that's quite dubious.

Genetic males/phenotypic females etc. certainly do deserve the ability to self declare sex given most people don't understand enough about molecular biology to understand those conditions. There are such few individuals easily certified, and I've never met anyone that is interested in telling them to do anything other than what they think is good for them.

Reply

Anonymous 13 July 2020 at 16:05

I think the key word there is "deemed" - perhaps as with the concept of "legal fiction", or "treated as". There are cases where a person who is male in the sense of having proceeded some way down the developmental pathway for producing small gametes is rightly treated socially as a woman and it would be transphobic or some sort of meanness to refuse.

Reply

Replies

gcmale 13 July 2020 at 19:38

If we presume the contexts already listed in the post (under "require protection") and other obvious ones (medical treatment, clinical trials etc.)...what does it mean to "rightly" treat someone as a woman?

Reply

Stopthief 19 July 2020 at 18:45

Honestly, a lot of casual onlookers to the trans debate ought to be shown this article, whichever "side" they feel inclined to take. It's such a simple and straightforward summary of how complex the debate actually is.

Reply

Anonymous 21 July 2020 at 19:08

It's probably no coincidence that the sex drive usually kicks in around age 12.

Reply

Replies

Anonymous 2 September 2020 at 22:20

But is that true for people on the autistic spectrum, which many trans identifying people are? If this is a disorder characterized by people who poorly understand social norms feeling like they do not fit social norms, perhaps the binary is the problem, not the person?

Reply

Unknown 25 August 2020 at 19:53

"Here is a list of the Human Rights of Women that transactivism is eliminating. No-one is saying trans-identified individuals should not have rights civil protections such as housing, employment etc. But what IS wrong is that the transactivists are erasing females as a protected sex class and erasing the sex-based protections of women and girls."

• Removing the right of women to assemble outside the presence of men.

• Removing the legal right of women to organize politically against sex-based oppression by males.

• Elimination of the patient right of dependent females to hospital/facility/bed assignments separate from males.

• Removing the legal right of women to be free from the presence of men in areas of public accommodation where nudity occurs.

• Elimination of the right of dependent females to prefer female providers for their intimate personal care requirements.

• Removing the legal right of women to educational programs created for women outside the presence of men.

• Elimination of sex-based crime statistics.

• Elimination of the human right of female prisoners under state confinement to housed separately from male prisoners.

• Eliminating athletic programs and sports competition for women and girls.

• Eliminating data collection of sex-based inequalities in areas where females are underrepresented.

• Elimination of grants, scholarships, board and trustee designations, representative positions and affirmative programs for women.

• Removing the legal right of women to create reproductive clinics, rape crisis services, support groups or any organisation for females.

• Eliminating media and all public discourse specific to females.

• Removal of the right of journalists to report the sex, and history, of subjects.

• Eliminating the legal right of lesbians to congregate publicly.

• Elimination of lesbian-specific organisations and advocacy groups.

• Removing the legal right of women to free speech related to sex roles and gender.

• Elimination of the legal right of women to protection from state-enforced sex-roles (appearance/behaviour/thought).

• Elimination of the legal right of girls to protection from state-enforced sex-roles in public education

Reply

Replies

Anonymous 2 September 2020 at 22:30

I would also say that in terms of social norm, it reinforces a rigid gender binary instead of allowing people to do as they like

Unknown 14 September 2020 at 19:34

Thank you so much for this. I get in frequent debates over the loss of biological women's rights. I want Trans people as human beings to have rights but no ones should be stripped of their to comply. Women have been fighting since the dawn of time to be considered equal to men.

Reply

Newer Post Home Older Post

Subscribe to: Post Comments (Atom)

(2.2/1000AE 95% CI: 0.9–3.4), and cheerleading (1.1/1000AE 95% CI: 0.4–1.8). Most concussions occurred during practice (47.3%) compared with games (34.5%).

Conclusions Consistent with previous research a greater proportion of MMS injuries occurred during practices than games. However, our finding that girls' basketball had the highest injury rate differs from prior research. Overall, our rates were higher than those previously reported, yet more precise given our sample size and detailed documentation AE. Our findings highlight the need for medical care at MMS practices and games.

056 EPIDEMIOLOGY OF SPORTS INJURIES AMONG MIDDLE SCHOOL STUDENTS

Shane Caswell,[1,2] Matthew Prebble,[1,2] Kaitlyn Romm,[1,2] Jatin Ambegaonkar,[1,2] Amanda Caswell,[1,2] Nelson Cortes[1,2]. [1]*George Mason University, Fairfax, USA;* [2]*Sports Medicine Assessment, Research and Testing Laboratory, Manassas, USA*

10.1136/bjsports-2016-097372.56

Background Participation in youth sports is increasing. Yet, limited research on the incidence and severity of injuries within middle school sports(MSS) exists.

Objective Describe the epidemiology of injuries in MSS within a large metropolitan school system located in the US.

Design Prospective descriptive epidemiology.

Setting Injury data collected in 9-MSS programs during 2015–2016.

Patients (or Participants) Boy(n=1194) and girl(n=1008) athletes(age=11–14 years) participating in 12 school sponsored sports.

Interventions (or Assessment of Risk Factors) Each of the 9 schools was assigned an athletic trainer(AT) who attended all practices and games during the fall, winter, and spring sports seasons of 2015–2016.

Main Outcome Measurements ATs collected injury and athlete-exposure(AE) data. Injury frequencies and rates with 95% confidence intervals(95%CI) for practices and games were described by sport for all injuries (no-time-loss and time-loss). Injury proportions were described by body region.

Results 1643 injuries(182.6±64.8 injuries/school) were reported. Highest rates of injury were: Girls' basketball (39.4/1000AE 95% CI: 34.7–44.1), American football (30.7/1000AE 95% CI: 27.3–34.1), wrestling (28.7/1000AE 95% CI: 24.4–32.9), and girls' soccer (26.3/1000AE 95% CI: 22.1–30.6). Lowest rates of injury were: Girls' volleyball (9.1/1000AE 95% CI: 6.7–11.6), boys' baseball (10.2/1000AE 95% CI: 6.8–13.5), cheerleading (10.6/1000AE 95% CI: 8.5–12.8), and boys' soccer (14.7/1000AE 95% CI: 11.8–17.7). Most injuries occurred during practice (n=889, 54%) compared with games (n=390, 24%). The top three injured body regions injured were: lower extremity (n=744, 45%), upper extremity (n=409, 25%), and head (n=297, 18%). The highest rates of concussion were: football (2.9/1000AE 95% CI: 1.9–4.0), girls' soccer

Normative health-related fitness values for children: analysis of 85347 test results on 9–17-year-old Australians since 1985

Mark J Catley, Grant R Tomkinson

► The appendix to this paper is published online only. To view this file please visit the journal online (http://bjsm.bmj.com/content/early/recent)

Health and Use of Time (HUT) Group, Sansom Institute for Health Research, University of South Australia, Adelaide, Australia

Correspondence to
Dr Grant R Tomkinson, Health and Use of Time (HUT) Group, Sansom Institute for Health Research, University of South Australia, GPO Box 2471, Adelaide, SA 5001, Australia; grant.tomkinson@unisa.edu.au

Received 13 May 2011
Accepted 22 September 2011

ABSTRACT

Objectives To provide sex- and age-specific normative values for health-related fitness of 9–17-year-old Australians.

Methods A systematic literature search was undertaken to identify peer-reviewed studies reporting health-related fitness data on Australian children since 1985—the year of the last national fitness survey. Only data on reasonably representative samples of apparently healthy (free from known disease or injury) 9–17-year-old Australians, who were tested using field tests of health-related fitness, were included. Both raw and pseudo data (generated using Monte Carlo simulation) were combined with sex- and age-specific normative centile values generated using the Lambda Mu and Sigma (LMS) method. Sex- and age-related differences were expressed as standardised effect sizes.

Results Normative values were displayed as tabulated percentiles and as smoothed centile curves for nine health-related fitness tests based on a dataset comprising 85347 test performances. Boys typically scored higher than girls on cardiovascular endurance, muscular strength, muscular endurance, speed and power tests, but lower on the flexibility test. The magnitude of the age-related changes was generally larger for boys than for girls, especially during the teenage years.

Conclusion This study provides the most up-to-date sex- and age-specific normative centile values for the health-related fitness of Australian children that can be used as benchmark values for health and fitness screening and surveillance systems.

BACKGROUND

Physical fitness is considered to be an important marker of current and future health in children and adults.[1] In children, cardiovascular fitness is a weak-to-strong predictor of total and abdominal adiposity, cardiovascular disease risk factors, cancer and mental health.[1 2] Certain muscular fitness components (eg, strength and endurance) are moderate predictors of cardiovascular disease risk factors, skeletal health and mental health.[1] Meaningful relationships have also been reported between running speed (another muscular fitness component) and skeletal health.[3] In adults, cardiovascular fitness is a strong and independent predictor of all-cause mortality and cardiovascular disease mortality and morbidity,[4] stroke,[5] cancer, mental health,[6] health-related quality of life[7] and many other cardiometabolic risk factors and comorbidities.[8 9] Moreover, physical fitness tracks moderately well from childhood through to adulthood.[10–13] This evidence highlights the need to include health-related fitness testing (ie, the testing of fitness components such as cardiovascular and muscular fitness that have the strongest links with health outcomes) as part of existing health and fitness screening and surveillance systems.

Although the most valid assessments of fitness require sophisticated laboratory equipment and a high level of tester expertise, they unfortunately are not suitable for mass testing. On the other hand, properly conducted field tests offer simple, feasible, and practical alternatives, which typically demonstrate good reliability and validity.[2 14–17] In Australia, unlike in Europe and North America where standardised test batteries such as the Eurofit[18] or FITNESSGRAM[19] are widely administered, a number of different field-based fitness tests and testing protocols have been used over time. For example, the most popular field test of cardiovascular fitness in Australia in the 1960s and 1970s was the 549-m (600 yd) run; in the 1980s and 1990s, it was the 1600-m run; and over the past decade or so, it has been the 20-m shuttle run.[20] Many physical educators and sports coaches in Australia continue to administer tests that are no longer in favour, largely because normative data (which are now several decades old) are available. This makes it difficult to assess the current status of health-related fitness in Australian children. Further compounding the problem is that the last national fitness survey of Australian children was conducted in 1985,[21] and with convincing evidence of recent temporal changes in several components of fitness,[22–24] the usefulness of such data seems to be limited.

Because there has never been a follow-up to the 1985 national survey, this study aimed to locate large and reasonably representative datasets of Australian children to generate normative centile values for health-related fitness. This study also aimed to quantify sex- and age-related differences in health-related fitness. These normative data will facilitate the identification of children with (a) low fitness in order to set appropriate goals and to promote positive health behaviours, and (b) specific fitness characteristics that may be considered important for sporting success.

METHODS

Data sources

A systematic review of the peer-reviewed scientific literature was undertaken to locate studies

Table 1 Summary of the included studies that have been used to assess the health-related fitness of 9–17-year-old Australians since 1985

								Tests reported in included studies								
Study	Year	Age (years)	*N*	Raw data	Sampling method	Sample base	Protocol	Push-ups	Sit-ups	Standing broad jump	Basketball throw	50 m sprint	Sit-and-reach	Hand-grip	1.6 km run	20 m shuttle run
ACHPER[50]	1994	9–18	39–104	yes	School-based; stratified, proportional	State (VIC)	ACHPER[50]		•		•		•		•	•
Barnett *et al*[51]	2007	15–17	21–69	no	School-based; stratified, random	State (NSW)	ACHPER[50]									•
Birchall[52]	1990	5–12	6–184	yes	School-based; convenience	State (VIC)	Pyke[21]	•	•	•		•	•		•	
Booth *et al*[53]	1997	9, 11, 13,15	399–634	no	School-based; stratified, proportional	State (NSW)	ACHPER[50]		•		•					•
Booth *et al*[54]	2004	9–15	357–466	no	School-based; stratified, proportional	State (NSW)	ACHPER[50]									•
Burke *et al*[55]	2004	10–13	38–117	yes	School-based; stratified, proportional	Capital city (WA)	ACHPER[50]									•
Cooley and McNaughton[56]	1998	11–16	339–636	no	School-based; stratified, proportional	State (TAS)	ACHPER[50]									•
Dollman *et al*[57]	1997	10–12	118–450	yes	School-based; stratified, proportional	State (SA)	Pyke[21]			•		•			•	
Dollman pers. comm.	2002	11–12	19–154	yes	School-based; stratified, random	State (SA)	Pyke[21]						•	•		
Dollman, pers. comm.	2002	8–12	8–389	yes	School-based; stratified, proportional	State (SA)	ACHPER[50] Pyke[21]			•						•
Hands[58]	2000	6–12	14–37	yes	School-based; stratified, random	Capital city (WA)	ACHPER[50] Pyke[21]		•	•	•	•	•	•		•
McIntyre, pers. comm.	2009	10–11	23–44	yes	School-based; stratified, random	Capital city (WA)	ACHPER[50]									•
McNaughton *et al*[59]	1995	7–10	30–83	no	School-based; stratified, random	State (TAS)	Pyke[21]					•			•	
Pyke[21]	1985	7–15	405–497	yes	School-based; stratified, proportional	National	Pyke[21]	•	•	•		•	•	•	•	
Vandongen *et al*[60]	1990	11	485–486	no	School-based; stratified, random	Capital city (WA)	ACHPER[50]								•	•

identifies test data that are available.
ACHPER, Australian Council for Health, Physical Education and Recreation; year, year of testing; *n*, sample size range per sex by age by test group; VIC, TAS, SA, WA, NSW

reporting descriptive summary data on Australian children tested for health-related fitness using field tests. Candidate studies were searched for in November 2009 using a computer search of online bibliographic databases (Ausport, CINAHL, Medline, PubMed, Scopus and SPORTDiscus). The search string used for the computer search was: ((((((((((((((((fitness) OR aerobic) OR anaerobic) OR cardio*) OR endurance) OR agility) OR flexibility) OR speed) OR power) OR strength) OR sprint*) OR jump*) OR push-up*) OR sit-up*) OR grip strength) OR sit and reach) AND (((((((child*) OR paediatric*) OR adolesc*) OR boy*) OR girl*) OR youth*) OR teen*) AND (Australia*). All titles and abstracts (when available) were assessed to identify eligible articles, with full-text articles retrieved if there was doubt in an article's eligibility. A number of Australian researchers were contacted through email to ask whether they knew of any appropriate studies or unpublished datasets.

Inclusion/exclusion criteria

Studies were included if they explicitly reported descriptive health-related fitness test data for apparently healthy (free from known disease or injury) 9–17-year-old Australians who were tested from 1985 onwards and if they reported data at the sex by age by test level, on children directly measured using field-based fitness tests for which explicit testing protocols were available. Studies were excluded if they reported descriptive data that were published in another identified study. The reference lists of all included studies were examined and cross-referenced to identify additional studies. Attempts were made to contact the corresponding

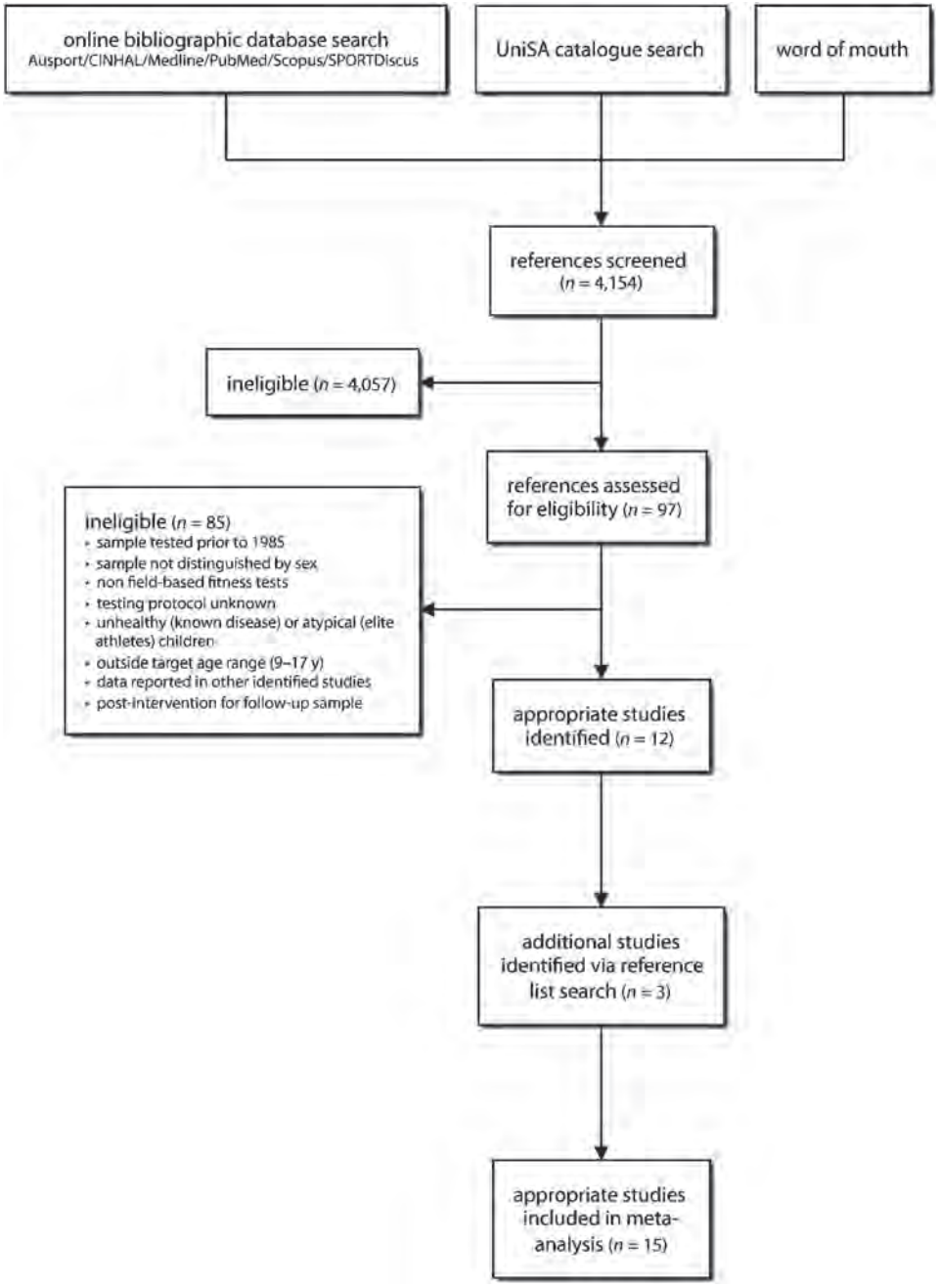


Figure 1 Flowchart outlining the identification of the included studies.

author of each study to request raw data and/or to clarify study details.

Initial data analysis

The following descriptive data were extracted from each included study: sex, age, year of testing, sample size, mean, SD, fitness test type and test protocol. Only data for commonly used fitness tests that were collected using protocols that were originally described in national or state-based fitness surveys of Australian children were retained for further analysis. Tests were considered to be 'common' if they were used to measure fitness in children across a broad range of ages and in at least two separate studies. Data for each fitness test were expressed in a common metric, and protocol differences were corrected where possible (eg, 20 m shuttle run data were expressed as the number of completed stages using the correction procedures described by Tomkinson *et al*).[25] However, if protocol correction was not possible, then only fitness data collected using the most common test protocol were retained. All available raw data were checked for anomalies by running range checks with data ±3 SDs away from the respective study by sex by age by test mean excluded. Age was expressed in whole years as the age at last birthday.

Statistical analysis

Sex- and age-specific normative centile values were calculated using a dataset comprising raw data and pseudo data that were generated using the method described by Tomkinson.[24] Normative centile values were generated using

Table 2 1.6 km run (s) centile values and LMS summary statistics by sex and age in 9- to 17-year-old Australians

Age (year)	P_5	P_{10}	P_{20}	P_{30}	P_{40}	P_{50}	P_{60}	P_{70}	P_{80}	P_{90}	P_{95}	L	M	S
Boys														
9	750	684	618	578	547	522	499	476	452	423	401	−1.042	521.963	0.183
10	732	666	602	564	535	511	489	469	447	420	400	−1.284	511.053	0.175
11	710	646	585	549	523	500	480	461	441	416	397	−1.466	500.394	0.166
12	682	621	563	530	505	485	467	449	430	408	392	−1.721	484.819	0.157
13	643	587	535	505	483	465	448	432	415	395	380	−1.895	464.529	0.148
14	605	556	509	482	462	446	431	416	401	382	369	−1.987	445.569	0.140
15	575	531	490	465	447	432	418	404	390	373	360	−1.979	431.504	0.133
16	552	514	477	454	437	423	410	397	383	366	354	−1.865	422.693	0.128
17	534	500	467	446	430	417	404	392	379	362	350	−1.707	416.545	0.123
Girls														
9	829	769	706	666	635	609	584	559	533	499	475	−0.779	608.674	0.167
10	820	759	697	657	626	600	576	552	526	494	470	−0.878	600.149	0.166
11	801	741	680	641	611	586	562	539	514	483	460	−0.929	585.820	0.165
12	784	726	666	629	600	575	552	529	505	474	452	−0.921	574.682	0.164
13	771	716	658	621	593	569	546	524	500	469	447	−0.852	568.706	0.163
14	763	711	655	620	592	567	545	523	498	468	445	−0.737	567.486	0.162
15	760	710	656	621	594	570	547	525	500	469	446	−0.591	569.809	0.161
16	757	710	658	624	597	573	550	527	502	471	446	−0.428	572.723	0.160
17	753	708	658	625	598	575	552	529	504	471	446	−0.263	574.536	0.159

Note, percentile data were calculated from 11 423 1.6 km run performances collected between 1985 and 1997.
L, skew; M, median; P, percentile; S, coefficient of variation.

Table 3 20 m shuttle run (completed stages) centile values and LMS summary statistics by sex and age in 9- to 17-year-old Australians

Age (year)	P_5	P_{10}	P_{20}	P_{30}	P_{40}	P_{50}	P_{60}	P_{70}	P_{80}	P_{90}	P_{95}	L	M	S
Boys														
9	1	1	2	2	2	3	3	3	4	5	6	0.213	2.573	0.568
10	1	2	2	3	3	4	4	5	5	7	8	0.373	3.537	0.543
11	1	2	3	3	4	4	5	5	6	7	8	0.520	4.131	0.517
12	1	2	3	3	4	4	5	6	6	8	9	0.643	4.460	0.486
13	2	2	3	4	4	5	5	6	7	8	9	0.744	4.888	0.453
14	2	3	4	4	5	6	6	7	8	9	10	0.835	5.664	0.418
15	3	3	4	5	6	7	7	8	9	10	11	0.926	6.527	0.380
16	3	4	5	6	7	7	8	8	9	10	11	1.031	7.159	0.343
17	4	5	6	6	7	8	8	9	10	11	11	1.143	7.690	0.306
Girls														
9	1	1	1	1	2	2	2	2	3	4	5	−0.065	1.842	0.535
10	1	1	2	2	2	2	3	3	4	5	6	0.086	2.468	0.557
11	1	1	2	2	2	3	3	4	4	6	7	0.220	2.844	0.573
12	1	1	2	2	3	3	3	4	5	6	7	0.324	3.016	0.577
13	1	1	2	2	3	3	4	4	5	6	7	0.400	3.138	0.569
14	1	1	2	2	3	3	4	4	5	6	7	0.457	3.225	0.554
15	1	1	2	3	3	3	4	4	5	6	7	0.505	3.412	0.536
16	1	2	2	3	3	4	4	5	5	6	7	0.554	3.672	0.518
17	1	2	2	3	4	4	5	5	6	7	8	0.603	4.032	0.499

Percentile data were calculated from 18 075 20 m shuttle run performances collected between 1990 and 2009.
The 20 m shuttle run can be scored in different metrics other than as the number of completed stages, such as the number of completed laps, the speed at the last completed stage and as mass-specific peak oxygen uptake estimates (see Tomkinson *et al*[25] for details on how to correct 20 m shuttle run performances to different metrics).
L, skew; M, median; P, percentile; S, coefficient of variation.

LMSChartmaker Light (v2.3, The Institute of Child Health, London) software, which analyses data using the LMS method.[26] The LMS method fits smooth centile curves to reference data by summarising the changing distribution of three sex- and age-specific curves representing the skewness (L: expressed as a Box-Cox power), the median (M) and the coefficient of variation (S). Using penalised likelihood, the curves can be fitted as cubic splines using non-linear regression, and the extent of smoothing required can be expressed in terms of smoothing parameters or equivalent degrees of freedom.[27]

For each fitness test, differences in means between: (a) age-matched Australian boys and girls (eg, 10-year-old boys vs 10-year-old girls); (b) sex-matched Australian children of different ages (eg, 10-year-old boys vs 11-year-old boys); and (c) sex- and age-matched Australian and international children[18 28–30] were expressed as standardised effect sizes.[31] Positive effect

Table 4 50 m sprint (s) centile values and LMS summary statistics by sex and age in 9- to 15-year-old Australians

Age (year)	P_5	P_{10}	P_{20}	P_{30}	P_{40}	P_{50}	P_{60}	P_{70}	P_{80}	P_{90}	P_{95}	L	M	S
Boys														
9	10.6	10.2	9.8	9.5	9.3	9.1	9.0	8.8	8.6	8.3	8.1	−1.837	9.136	0.078
10	10.5	10.1	9.7	9.4	9.2	9.0	8.8	8.7	8.5	8.2	8.0	−2.185	9.009	0.080
11	10.4	10.0	9.6	9.3	9.1	8.9	8.7	8.5	8.3	8.1	7.9	−2.405	8.877	0.081
12	10.2	9.8	9.3	9.1	8.9	8.7	8.5	8.3	8.1	7.9	7.7	−2.446	8.673	0.081
13	9.8	9.4	9.0	8.8	8.6	8.4	8.2	8.1	7.9	7.7	7.5	−2.489	8.377	0.079
14	9.4	9.0	8.7	8.4	8.2	8.1	7.9	7.8	7.6	7.4	7.2	−2.701	8.063	0.076
15	9.0	8.6	8.3	8.1	7.9	7.7	7.6	7.5	7.3	7.1	7.0	−3.021	7.738	0.073
Girls														
9	11.7	11.3	10.8	10.5	10.3	10.0	9.8	9.6	9.3	9.0	8.8	−0.981	10.033	0.088
10	11.1	10.7	10.3	10.0	9.8	9.5	9.3	9.1	8.9	8.6	8.4	−1.453	9.542	0.084
11	10.7	10.3	9.9	9.6	9.4	9.2	9.0	8.8	8.6	8.3	8.1	−1.803	9.161	0.082
12	10.4	10.0	9.6	9.3	9.1	8.9	8.7	8.6	8.4	8.1	7.9	−1.977	8.919	0.080
13	10.2	9.8	9.4	9.2	9.0	8.8	8.6	8.4	8.3	8.0	7.8	−1.991	8.787	0.078
14	10.0	9.7	9.3	9.1	8.9	8.7	8.5	8.4	8.2	7.9	7.8	−1.884	8.686	0.076
15	9.9	9.6	9.2	9.0	8.8	8.6	8.5	8.3	8.1	7.9	7.7	−1.724	8.638	0.075

Note, percentile data were calculated from 10 104 50 m sprint performances collected between 1985 and 1999.
L, skew; M, median; P, percentile; S, coefficient of variation.

Table 5 Basketball throw (m) centile values and LMS summary statistics by sex and age in 9- to 17-year-old Australians

Age (year)	P_5	P_{10}	P_{20}	P_{30}	P_{40}	P_{50}	P_{60}	P_{70}	P_{80}	P_{90}	P_{95}	L	M	S
Boys														
9	2.3	2.5	2.7	2.9	3.1	3.3	3.4	3.6	3.8	4.1	4.4	0.623	3.260	0.198
10	2.5	2.8	3.0	3.3	3.4	3.6	3.8	4.0	4.2	4.5	4.8	0.675	3.608	0.192
11	2.8	3.1	3.4	3.6	3.8	4.0	4.2	4.4	4.7	5.0	5.3	0.733	4.026	0.188
12	3.1	3.4	3.8	4.0	4.3	4.5	4.7	4.9	5.2	5.6	5.9	0.792	4.471	0.188
13	3.5	3.8	4.2	4.5	4.8	5.0	5.3	5.5	5.8	6.2	6.6	0.843	5.012	0.187
14	3.9	4.2	4.7	5.0	5.3	5.5	5.8	6.1	6.4	6.9	7.2	0.898	5.522	0.186
15	4.2	4.6	5.0	5.4	5.7	6.0	6.3	6.6	6.9	7.4	7.8	0.943	5.975	0.185
16	4.4	4.8	5.3	5.6	6.0	6.3	6.5	6.9	7.2	7.7	8.2	0.964	6.254	0.185
17	4.5	4.9	5.5	5.8	6.2	6.5	6.8	7.1	7.5	8.0	8.5	0.966	6.467	0.187
Girls														
9	2.1	2.3	2.5	2.7	2.9	3.0	3.2	3.3	3.5	3.7	3.9	1.116	3.015	0.182
10	2.3	2.6	2.8	3.0	3.2	3.3	3.5	3.7	3.8	4.1	4.3	1.024	3.336	0.181
11	2.6	2.8	3.1	3.3	3.5	3.6	3.8	4.0	4.2	4.5	4.7	0.942	3.646	0.180
12	2.8	3.1	3.4	3.6	3.8	4.0	4.2	4.3	4.6	4.9	5.2	0.873	3.970	0.180
13	3.0	3.3	3.6	3.9	4.1	4.3	4.5	4.7	4.9	5.3	5.6	0.816	4.265	0.179
14	3.2	3.4	3.8	4.0	4.2	4.4	4.6	4.8	5.1	5.4	5.7	0.739	4.410	0.175
15	3.3	3.6	3.9	4.1	4.3	4.5	4.7	4.9	5.1	5.5	5.8	0.606	4.486	0.169
16	3.4	3.7	4.0	4.2	4.4	4.6	4.7	5.0	5.2	5.6	5.9	0.394	4.557	0.162
17	3.6	3.8	4.1	4.3	4.5	4.6	4.8	5.0	5.3	5.6	5.9	0.140	4.634	0.154

Note, percentile data were calculated from 5,541 basketball throw performances collected between 1994 and 1999; L, skew; M, median; P, percentile; S, coefficient of variation.

sizes indicated that mean fitness test performances for boys (age-matched analysis), older children (sex-matched analysis) or Australian children (sex- and age-matched analysis) were higher than those for girls, younger children or international children, respectively. Effect sizes of 0.2, 0.5 and 0.8 were used as thresholds for small, moderate and large.[31]

RESULTS

Table 1 summarises the 15 included studies. Of these, 12 were identified through bibliographic database searching and word of mouth, and three were identified through reference list searching. Corresponding authors of all the studies were contacted through email to clarify study details and/or to request raw data. All authors satisfactorily clarified study details, and seven of them supplied raw data (figure 1).

The final dataset comprised 85347 individual test results and 142 sex by age by test groups with a median sample size of 537 (range: 54–2612). Data were available for six fitness components and nine fitness tests: cardiovascular endurance (20 m shuttle run, 1.6 km run), muscular strength (hand-grip), muscular endurance (push-ups and sit-ups), muscular power (standing broad jump and basketball throw), muscular speed (50 m sprint) and flexibility (sit-and-reach). Raw data were available for 74% of all data points.

Table 6 Standing broad jump (cm) centile values and LMS summary statistics by sex and age in 9- to 15-year-old Australians

Age (year)	P_5	P_{10}	P_{20}	P_{30}	P_{40}	P_{50}	P_{60}	P_{70}	P_{80}	P_{90}	P_{95}	L	M	S
Boys														
9	105	113	121	127	133	138	142	147	153	161	168	1.244	137.506	0.138
10	109	117	126	133	138	143	148	154	160	168	174	1.490	143.430	0.138
11	112	121	131	138	144	149	154	160	166	174	181	1.654	149.322	0.138
12	117	126	137	144	150	156	161	167	173	182	189	1.704	155.838	0.137
13	126	136	147	154	161	166	172	178	185	194	201	1.629	166.340	0.135
14	137	146	157	165	172	178	184	190	197	206	214	1.526	177.688	0.131
15	148	157	169	177	183	189	196	202	209	219	228	1.446	189.485	0.127
Girls														
9	95	102	110	116	122	126	131	136	142	150	157	1.098	126.379	0.149
10	100	108	117	123	128	133	138	143	149	158	165	1.152	133.177	0.147
11	106	114	123	129	135	140	145	151	157	166	173	1.197	140.142	0.145
12	111	118	128	135	140	145	151	156	163	171	179	1.211	145.432	0.142
13	115	123	132	139	145	150	155	161	167	176	183	1.183	150.080	0.138
14	119	127	136	143	148	154	159	164	171	180	187	1.158	153.551	0.134
15	122	129	139	145	151	156	161	166	173	181	188	1.148	155.661	0.130

Percentile data were calculated from 11 194 standing broad jump performances collected between 1985 and 2002.
L, skew; M, median; P, percentile; S, coefficient of variation.

Table 7 Push-ups (no. in 30 s) centile values and LMS summary statistics by sex and age in 9- to 15-year-old Australians

Age (year)	P_5	P_{10}	P_{20}	P_{30}	P_{40}	P_{50}	P_{60}	P_{70}	P_{80}	P_{90}	P_{95}	L	M	S
Boys														
9	4	6	8	9	11	12	14	15	17	20	22	0.846	12.310	0.452
10	4	6	8	10	11	13	14	16	18	21	23	0.894	12.943	0.447
11	4	6	8	10	12	13	14	16	18	20	22	0.940	12.942	0.438
12	4	6	9	10	12	13	15	16	18	20	22	0.980	13.200	0.422
13	5	7	9	11	13	14	16	17	19	22	24	1.020	14.255	0.399
14	6	8	11	13	14	16	17	19	21	23	25	1.070	15.954	0.370
15	7	10	13	15	16	18	19	21	23	25	27	1.126	17.697	0.337
Girls														
9	2	3	5	7	8	9	10	12	13	16	18	0.719	8.989	0.550
10	2	3	5	6	7	9	10	11	13	16	18	0.652	8.655	0.583
11	2	3	4	6	7	8	9	11	13	16	18	0.584	8.142	0.624
12	1	2	4	5	6	7	9	10	12	15	18	0.518	7.395	0.672
13	1	2	3	4	6	7	8	10	12	15	18	0.453	6.792	0.720
14	1	2	3	4	5	6	8	9	11	15	18	0.390	6.384	0.765
15	1	2	3	4	5	6	7	9	11	14	18	0.329	5.818	0.812

Percentile data were calculated from 7,342 push-up test performances collected between 1985 and 1991.
L, skew; M, median; P, percentile; S, coefficient of variation.

Normative fitness data for 9–17-year-old Australians are presented as tabulated percentiles from 5 to 95 (P_5, P_{10}, P_{20}, P_{30}, P_{40}, P_{50}, P_{60}, P_{70}, P_{80}, P_{90}, P_{95}) in tables 2–10. The sex- and age-specific LMS values for all fitness tests are also shown. The LMS values depict the nature of the age-related distributions for boys and girls and can be used to calculate *z*-scores and hence percentile values by looking up a *z*-table, using the following formula:

$$z = \frac{\left(\frac{x}{M}\right)^{L} - 1}{L \times S}$$

where *z* is *z* score, *x* is performance, *L* is sex- and age-specific *L*-value, *M* is the sex- and age-specific *M*-value and *S* is the sex- and age-specific *S*-value.

Figures 2 and 3 show the smoothed centile curves (P_{10}, P_{50}, P_{90}).

Figure 4 shows the sex-related differences in mean fitness. Boys consistently scored higher than girls on health-related fitness tests, except on the sit-and-reach test, with the magnitude of the differences typically increasing with age and often accelerating from about 12 years of age. Overall, the magnitude of differences between boys and girls was large for the 1.6 km run, 20 m shuttle run, basketball throw and push-ups; moderate for the 50-m sprint, standing broad jump and sit-and-reach; and small for sit-ups and hand-grip strength. Figure 5 shows the age-related changes in mean fitness. The age-related changes were typically larger for boys than for girls, especially during the teenage years, and for muscular fitness tests than for cardiovascular fitness tests. Fitness also tended to peak from about the age of 15 years. Figure 6 shows that the differences in health-related fitness between Australian and international children

Table 8 Sit-ups (no. in 180 s) centile values and LMS summary statistics by sex and age in 9- to 17-year-old Australians

Age (year)	P_5	P_{10}	P_{20}	P_{30}	P_{40}	P_{50}	P_{60}	P_{70}	P_{80}	P_{90}	P_{95}	L	M	S
Boys														
9	3	5	8	11	14	17	21	25	30	40	48	0.321	17.046	0.755
10	5	8	13	17	20	24	29	34	40	50	59	0.466	24.459	0.669
11	6	10	16	21	25	29	34	39	45	55	60	0.629	29.422	0.594
12	8	14	21	26	31	36	40	45	51	60	60	0.841	35.561	0.514
13	10	17	25	31	36	40	45	50	55	60	60	1.056	40.288	0.443
14	12	20	29	34	39	43	48	52	57	60	60	1.232	43.454	0.389
15	14	22	31	36	41	45	49	53	58	60	60	1.335	44.942	0.359
16	16	24	32	38	42	46	50	54	58	60	60	1.426	46.209	0.332
17	18	26	34	40	44	47	51	55	59	60	60	1.517	47.466	0.306
Girls														
9	5	8	12	15	18	21	25	29	35	43	51	0.394	21.258	0.642
10	7	10	14	18	22	26	30	34	40	50	58	0.485	25.666	0.605
11	8	11	17	21	25	29	34	39	45	54	60	0.571	29.444	0.569
12	9	13	19	24	28	32	37	42	48	57	60	0.646	32.123	0.534
13	10	15	21	26	30	34	39	44	50	59	60	0.705	34.408	0.504
14	11	15	22	27	31	35	40	45	50	59	60	0.741	35.334	0.482
15	11	16	22	27	31	35	40	44	50	58	60	0.757	35.327	0.464
16	12	17	23	28	32	36	40	44	50	57	60	0.761	35.690	0.447
17	13	18	24	28	32	36	40	45	50	58	60	0.761	36.333	0.431

Percentile data were calculated from 8 837 sit-up test performances collected between 1985 and 1999.
L, skew; M, median; P, percentile; S, coefficient of variation.

Table 9 Hand-grip strength (kg) centile values and LMS summary statistics by sex and age in 9- to 15-year-old Australians (taken as the mean of both hands)

Age (year)	P_5	P_{10}	P_{20}	P_{30}	P_{40}	P_{50}	P_{60}	P_{70}	P_{80}	P_{90}	P_{95}	L	M	S
Boys														
9	11.5	12.5	13.8	14.8	15.6	16.4	17.2	18.1	19.2	20.8	22.1	0.600	16.415	0.197
10	13.1	14.3	15.9	17.0	18.0	19.0	19.9	21.0	22.2	23.9	25.4	0.728	18.967	0.198
11	14.5	15.9	17.7	19.0	20.1	21.2	22.3	23.5	24.9	26.8	28.5	0.764	21.217	0.200
12	15.4	17.0	18.9	20.3	21.5	22.7	23.8	25.1	26.6	28.7	30.5	0.747	22.655	0.203
13	17.5	19.3	21.5	23.1	24.5	25.8	27.2	28.6	30.4	32.8	34.9	0.738	25.819	0.205
14	20.8	22.9	25.5	27.4	29.1	30.7	32.4	34.1	36.2	39.1	41.6	0.742	30.731	0.207
15	24.6	27.1	30.3	32.6	34.6	36.5	38.4	40.5	43.0	46.5	49.5	0.752	36.517	0.207
Girls														
9	9.8	10.8	12.0	12.9	13.7	14.4	15.1	16.0	17.0	18.4	19.5	0.639	14.396	0.205
10	11.4	12.6	14.1	15.2	16.2	17.1	18.0	19.0	20.1	21.8	23.1	0.842	17.072	0.210
11	12.5	13.9	15.5	16.8	17.8	18.8	19.8	20.9	22.1	23.9	25.3	0.932	18.816	0.208
12	14.4	16.0	17.8	19.1	20.3	21.4	22.5	23.6	25.0	26.9	28.5	0.922	21.374	0.200
13	16.4	18.0	19.9	21.3	22.5	23.6	24.8	26.0	27.4	29.5	31.1	0.880	23.641	0.190
14	18.2	19.7	21.6	23.0	24.3	25.4	26.5	27.8	29.2	31.3	33.0	0.828	25.390	0.178
15	19.8	21.3	23.2	24.6	25.8	26.9	28.0	29.2	30.7	32.7	34.4	0.770	26.881	0.165

Percentile data were calculated from the 3 707 hand-grip strength performances collected between 1985 and 1999.
L, skew; M, median; P, percentile; S, coefficient of variation.

were generally small, with Australian children scoring slightly higher on hand-grip strength (mean ±95% CI: 0.20±0.03 SDs) and 50 m sprint tests (0.24±0.02 SDs), and slightly lower on sit-and-reach (–0.36±0.02 SDs), standing broad jump (–0.25±0.02 SDs) and 20 m shuttle run tests (–0.49±0.01 SDs).

DISCUSSION

This study provides the most up-to-date sex- and age-specific normative centile values for 9–17-year-old Australians across a range of health-related fitness tests, which can be used as benchmark values for health and fitness screening and surveillance of children. These data complement a growing literature reporting growth percentiles across a range of different health measures, such as body mass index,[32] waist girth[33] and blood pressure,[28] and a range of other health-related fitness measures.[29 30] It also quantifies the magnitude and direction of sex- and age-related differences in children's health-related fitness and shows that boys consistently scored higher than girls on fitness tests (except on the sit-and-reach test of flexibility) and that boys experience larger age-related changes in fitness. The developmental patterns of children's fitness have been well studied and extensively reviewed (eg, for cardiovascular fitness, refer to Armstrong

Table 10 Sit-and-reach (cm) centile values and LMS summary statistics by sex and age in 9- to 17-year-old Australians.

Age (y)	P_5	P_{10}	P_{20}	P_{30}	P_{40}	P_{50}	P_{60}	P_{70}	P_{80}	P_{90}	P_{95}	L	M	S
boys														
9	10.4	12.9	15.7	17.7	19.4	20.9	22.4	23.9	25.8	28.2	30.3	1.211	20.877	0.285
10	10.0	12.5	15.3	17.3	19.0	20.5	22.1	23.7	25.5	28.0	30.1	1.190	20.537	0.294
11	9.6	12.1	15.0	17.0	18.7	20.3	21.9	23.5	25.4	28.0	30.1	1.167	20.313	0.305
12	9.3	11.8	14.8	16.9	18.7	20.3	21.9	23.6	25.6	28.3	30.5	1.133	20.292	0.315
13	9.4	12.0	15.1	17.2	19.1	20.8	22.5	24.3	26.4	29.2	31.6	1.091	20.785	0.322
14	9.8	12.5	15.7	18.0	20.0	21.8	23.6	25.5	27.8	30.9	33.4	1.054	21.804	0.328
15	10.4	13.2	16.6	19.1	21.2	23.1	25.1	27.1	29.5	32.9	35.7	1.017	23.112	0.332
16	11.1	14.0	17.6	20.1	22.3	24.4	26.5	28.7	31.3	34.9	37.8	0.984	24.392	0.334
17	11.7	14.8	18.5	21.2	23.5	25.7	27.9	30.2	33.0	36.8	40.0	0.953	25.686	0.335
girls														
9	13.0	15.8	18.9	21.1	22.9	24.6	26.2	28.0	29.9	32.6	34.8	1.285	24.614	0.264
10	13.0	15.7	18.8	20.9	22.7	24.4	26.0	27.7	29.7	32.4	34.6	1.259	24.402	0.265
11	13.2	15.9	19.0	21.2	23.0	24.7	26.4	28.1	30.1	32.8	35.0	1.235	24.705	0.265
12	14.0	16.7	19.9	22.2	24.1	25.8	27.5	29.3	31.3	34.2	36.4	1.230	25.790	0.262
13	15.3	18.2	21.6	24.0	25.9	27.7	29.5	31.4	33.6	36.5	38.9	1.250	27.740	0.256
14	16.5	19.5	23.1	25.5	27.6	29.4	31.3	33.2	35.4	38.4	40.9	1.293	29.440	0.248
15	17.0	20.1	23.7	26.1	28.1	30.0	31.8	33.7	35.9	38.8	41.2	1.350	29.997	0.241
16	17.0	20.0	23.5	25.9	27.9	29.6	31.4	33.2	35.3	38.1	40.3	1.412	29.647	0.235
17	16.8	19.8	23.2	25.5	27.4	29.1	30.7	32.5	34.4	37.1	39.2	1.472	29.074	0.229

Note, percentile data were calculated from 9,124 sit-and-reach performances collected between 1985 and 2000; L = skew; M = median; S = coefficient of variation. Note, a score of "20 cm" corresponds to the participant reaching their toes.

et al,[34] Krahenbuhl *et al*[35] and Rowland[36]; for muscular fitness, refer to Blimkie and Sale,[37] Froberg and Lammert[38] and De Ste Croix[39]). Although the underlying causes of sex- and age-related differences are clear for some fitness test performances, such as those for muscular strength, power and speed, which are largely explained by physical differences (eg, differences in muscle mass or height), they are less clear for others, such as for cardiovascular endurance, which may be explained by physiological differences (eg, differences in mechanical efficiency and/or fractional utilisation).[15 36] It is, nonetheless, beyond the scope of this article to discuss the causes that underscore the sex- and age-related changes in fitness test performance.

International comparisons

Although several studies have previously compared the health-related fitness of Australian children with their sex- and age-matched international peers,[20 40] comparisons have only been made for cardiovascular fitness. Figure 6 compares the 20-m shuttle run, 50 m sprint, standing broad jump, handgrip strength and sit-and-reach performance of 9–17-year-old Australians with 1 894 971 test results from sex-, age- and test-matched international children from 48 countries who have been measured using the same test protocols as those referenced in table 1 and described in Appendix 1. Figure 6 also shows typically small differences in health-related fitness between Australian and international children. Furthermore, the sex- and age-related differences in fitness of Australian children are strikingly similar to those observed in international children. Given that the differences are generally small, the normative centile data presented in this study could be used as approximate benchmark values for health-related fitness of international children.

Fitness thresholds for cardiometabolic risk

Fitness is widely recognised as a powerful marker of current and future cardiovascular, skeletal and mental health. Unfortunately, there are no universally accepted recommendations for health-related levels of fitness. In recent years however, sex- and age-specific threshold values for cardiovascular fitness (operationalised as mass-specific peak oxygen uptake in ml/kg/min) have been established for European and US children using linked cardiometabolic risk-based values from receiver operator characteristic curve analyses.[41–44] To estimate the prevalence of Australian children with 'healthy' cardiovascular fitness (ie, those above the thresholds), 'international' sex- and age-specific thresholds for 9–17-year-old children were estimated by determining best-fitting polynomial regression model (quadratic or cubic) relating age (predictor variable) to previously reported threshold values (response variable) in Adegboye *et al*,[41] Lobelo *et al*,[42] Ruiz *et al*[43] and Welk *et al*.[44] Separate models were generated for boys and girls. Peak oxygen uptake values in Australian children were estimated using 1.6 km run and 20 m shuttle run data and the Cureton *et al*[45] and Léger *et al*[46] regression equations, respectively.

Using these thresholds, about 71% of Australian boys (median ±95% CI: 71%±8%) and 77% of Australian girls (median ±95% CI: 77%±10%) apparently have 'healthy' cardiovascular fitness. Although in light of recent secular declines in cardiovascular fitness,[20 22 23 25] and with a median testing year of 1993 in this study's cardiovascular fitness dataset, it is likely that these prevalence rates somewhat overestimate those of today. These prevalence rates are better than (for girls), or similar to (for boys), those observed in European (61% of boys and 57% of girls)[29] and US (71% of boys and 69% of girls)[42] children. Geographical differences in prevalence rates may reflect differences in (a) threshold levels, (b) the year(s) of testing, (c) sampling methodology, (d) test methodology and (e) the way in which peak oxygen uptake was measured or estimated.[47]

Ultimately, it is important to remember that the normative data presented in this study show how well Australian children perform on health-related fitness tests relative to their sex- and age-matched peers. For example, using a percentile

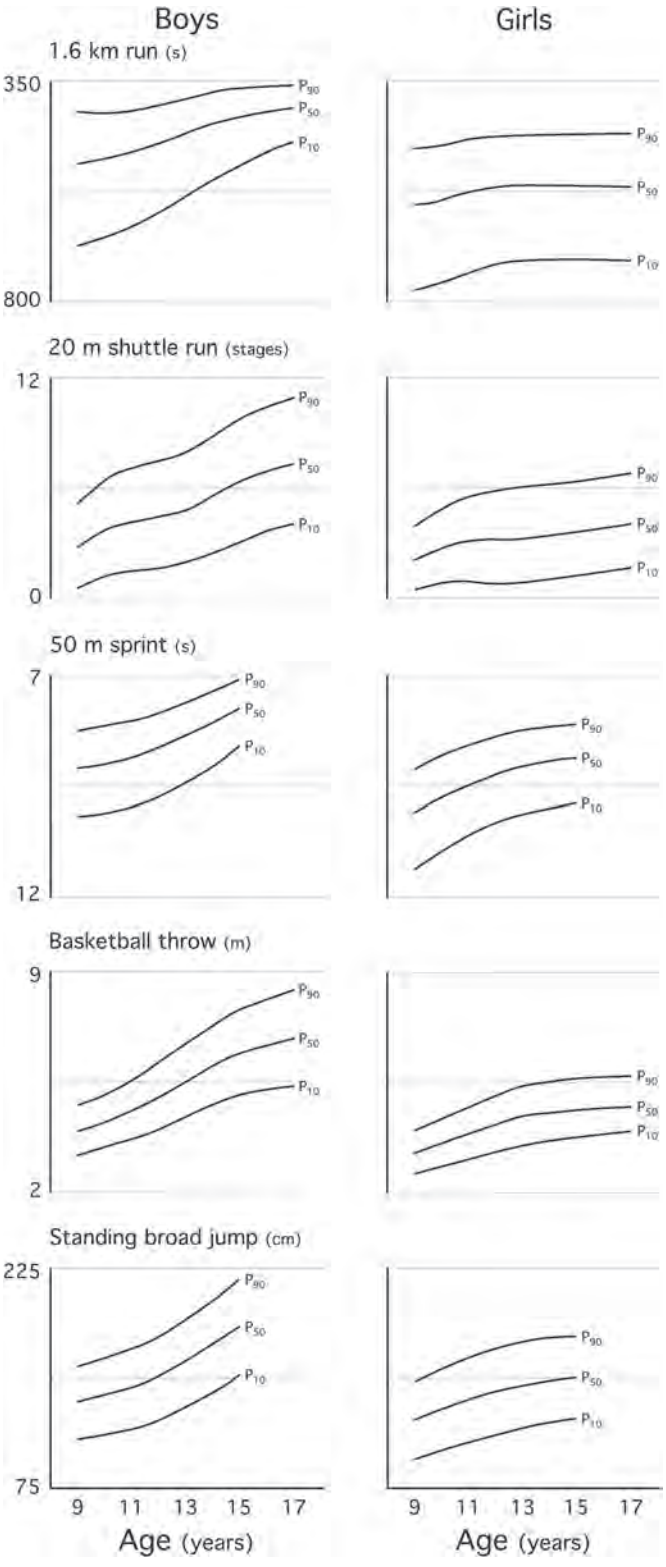


Figure 2 Smoothed centile curves (P_{10}, P_{50} and P_{90}) for (A) 1.6 km run (s), (B) 20 m shuttle run (number of completed stages), (C) 50 m sprint (s), (D) basketball throw (m) and (E) standing broad jump (cm).

classification, children with fitness in the bottom 20% can be classified as having 'very low' fitness; those between the 20th and 40th percentiles as having 'low' fitness; those between the 40th and 60th percentiles as having 'average' fitness; those between the 60th and 80th percentiles as having 'high' fitness; and those above the 80th percentile as having 'very high'

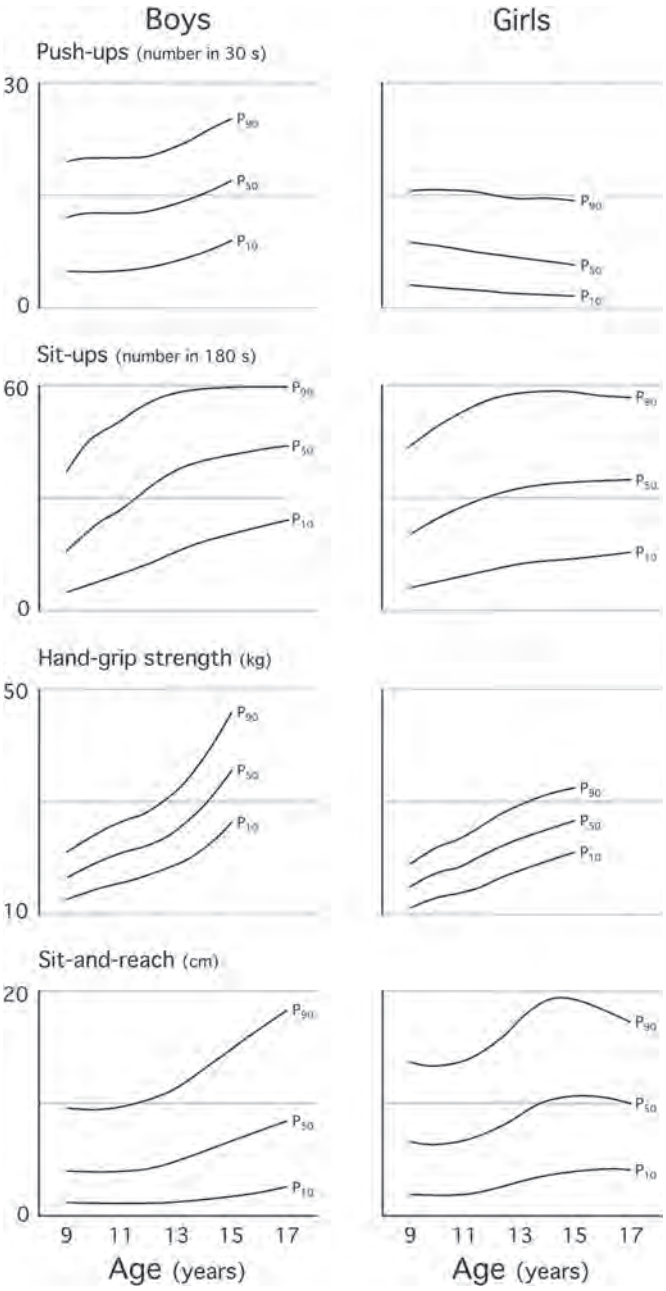


Figure 3 Smoothed centile curves (P_{10}, P_{50} and P_{90}) for (A) push-ups (number in 30 s), (B) sit-ups (number in 180 s), (C) hand-grip strength (kg) and (D) sit-and-reach (cm) tests.

fitness. These data are not criterion-referenced in that they do not indicate whether children with 'very low' or 'low' (or any other classification for that matter) have 'unhealthy' cardiovascular fitness or increased cardiometabolic risk. Despite the fact that previous Australian evidence has linked low childhood cardiovascular fitness with increased cardiometabolic risk in adulthood,[48] future Australian studies are required to examine whether childhood thresholds for cardiovascular fitness (or other health-related fitness components) are significantly associated with clustered cardiometabolic risk (or other health outcomes, such as mental or skeletal health outcomes).

Strengths and limitations

Despite the fact that the last national fitness survey of Australian children was in 1985, this study provides the most

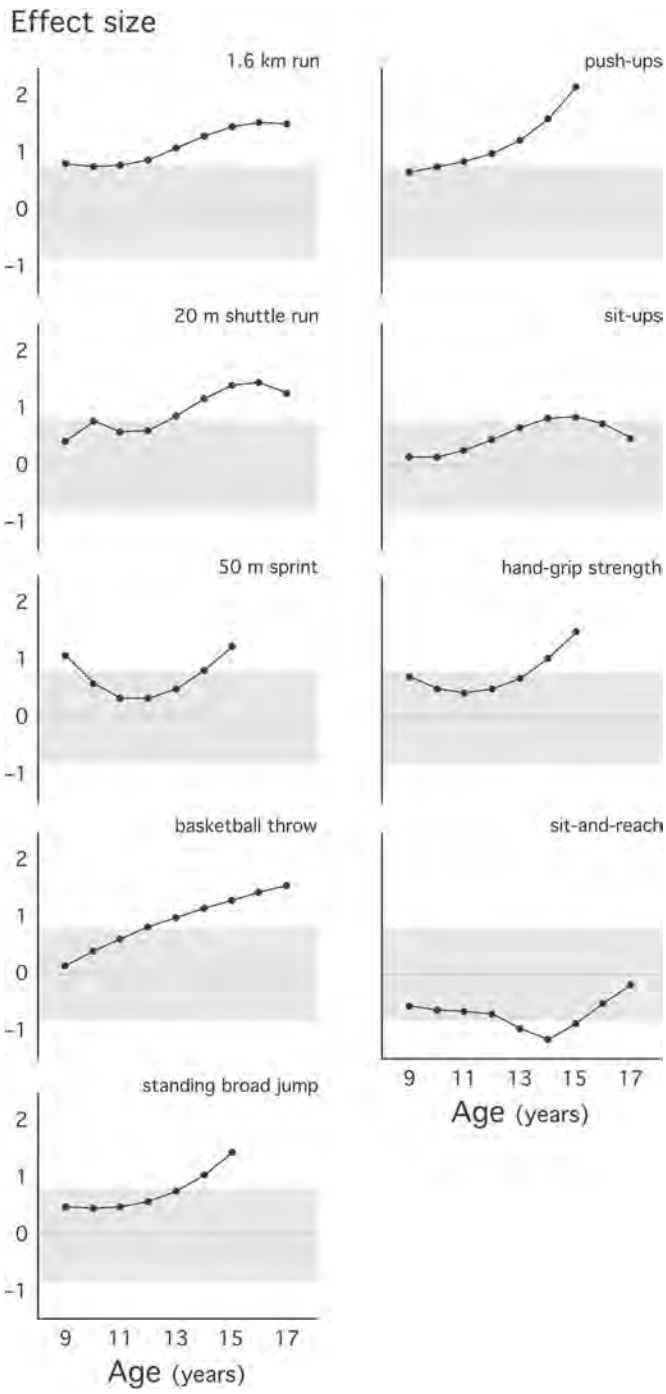


Figure 4 Sex-related differences in mean fitness expressed as effect sizes. Data are shown for 9–17-year-old children tested on the (A) 1.6 km run, (B) 20 m shuttle run, (c) 50 m sprint, (D) basketball throw, (E) standing broad jump, (F) push-ups, (G) sit-ups, (H) hand-grip strength and (I) sit-and-reach tests. The limits of the grey zone represent effects sizes of 0.8 and –0.8, beyond which large differences are observed.

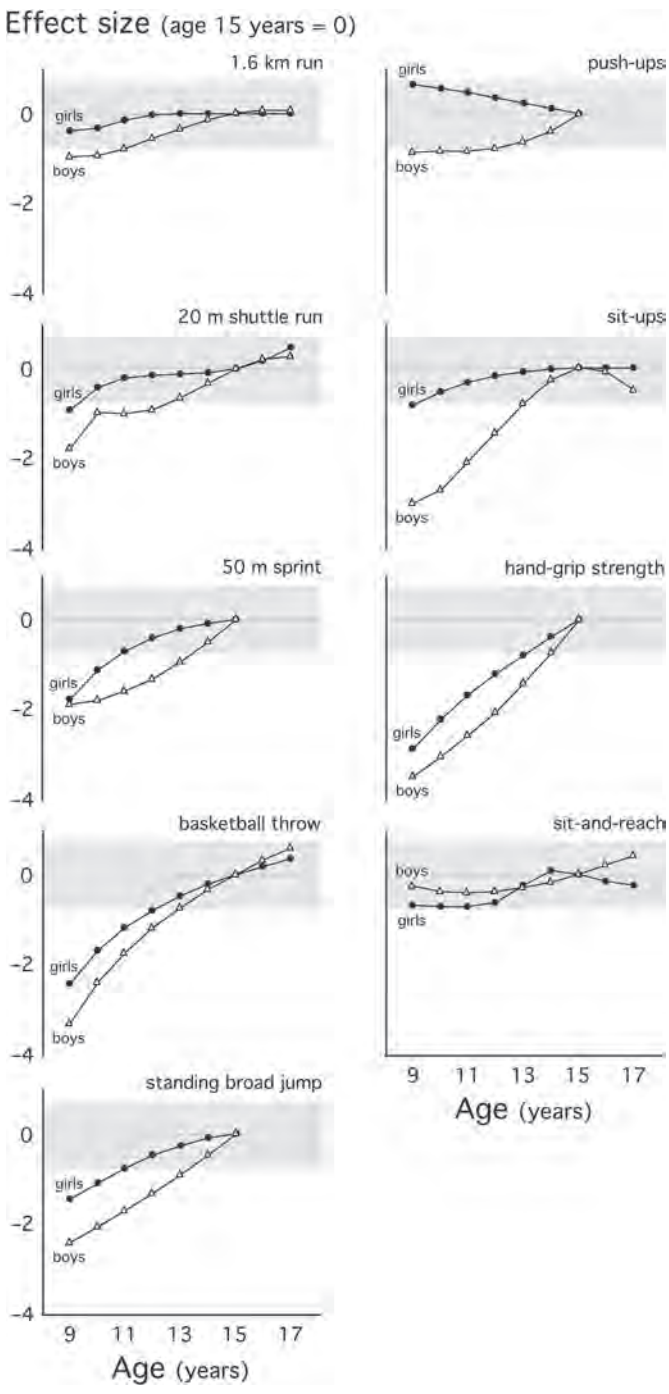


Figure 5 Age-related changes in mean fitness expressed as effect sizes standardised to an effect size of age 15 years=0. Data are shown for 9–17-year-old boys (triangles) and girls (circles) separately tested on the (A) 1.6 km run, (B) 20 m shuttle run, (C) 50 m sprint, (D) basketball throw, (E) standing broad jump, (F) push-ups, (G) sit-ups, (H) hand-grip strength and (I) sit-and-reach tests. The limits of the grey zone represent effects sizes of 0.8 and –0.8, beyond which large differences are observed.

up-to-date normative dataset for nine widely administered health-related fitness tests, using cumulated data from 85347 Australian children aged 9–17 years collected between 1985 and 2009. This study used a strict set of inclusion and exclusion criteria and rigorous initial data analysis procedures to systematically control for any factors (eg, differences in test methodology) that might have biased the normative values or the estimates of the sex- and age-related differences. It used a novel pseudo-data method to allow both descriptive and raw data to be merged before using the LMS method to create sex- and age-specific smoothed percentiles. It also quantified sex- and age-related differences as standardised effects sizes, allowing for comparison between sexes, among different ages, and with sex, age and test-matched international children.

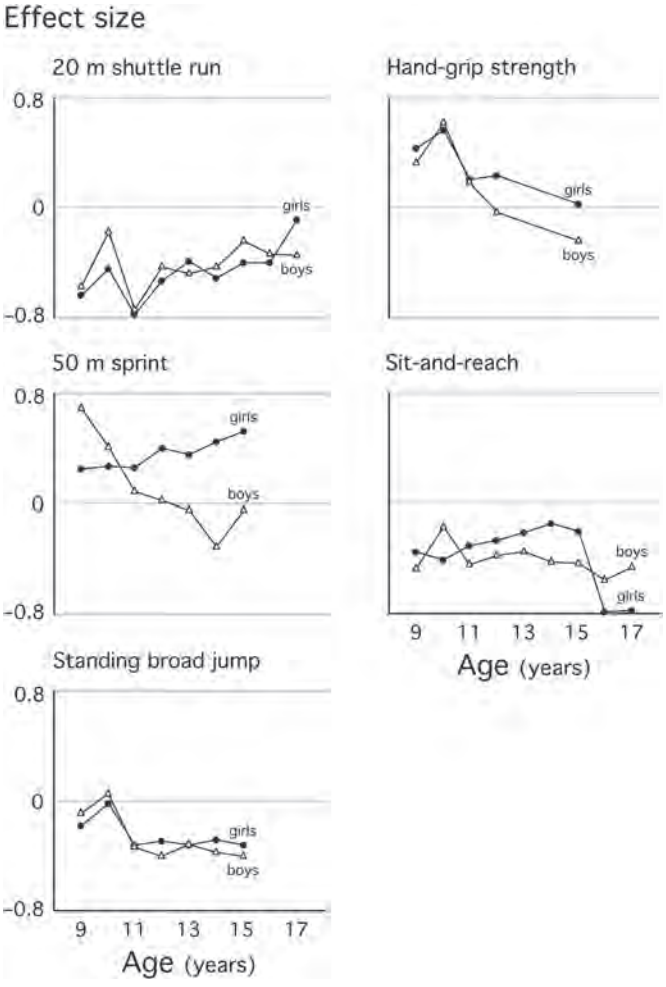


Figure 6 Sex- and age-specific effect sizes for (A) 20 m shuttle run, (B) 50 m sprint, (C) hand-grip strength, (D) sit-and-reach and (E) standing broad jump for 9–17-year-old Australian boys (triangles) and girls (circles) relative to their international peers. Positive effects indicate higher fitness scores for Australian children and negative effects indicate lower fitness scores. Comparative data represent n=284 508 20 m shuttle run performances,[28–30] n=1 216 452 50 m sprint performances,[18] n=126 361 hand-grip strength performances,[29] n=102 664 sit-and-reach performances,[29] and n=164 986 standing broad jump performances[29] of 9–17-year-old children from 48 international countries.

However, this study is not without limitations. Only one of the 15 included studies was based on a nationally representative sample, which obviously raises the issue of representativeness. Most of the included studies used similar sampling frames (table 1). Schools with a greater interest in sport and fitness may have been more willing to participate, and because participation at the individual level was voluntary, it is possible that children with low fitness levels chose not to participate. This might have resulted in fitness test performances unrepresentative of the population, but it should not have affected the sex- and age-related differences. Fitness data were also collected at different times during the 1985–2009 period, and given convincing evidence of recent temporal declines in some (but not all) components of Australian children's fitness,[23 49] it is possible that the normative data presented in this study represent a better 'health-related picture' than what would be observed today. A temporal analysis of the data accumulated in this study suggests that these normative data would probably overestimate the fitness of Australian children in 2009 by an average of 0.3 SDs or 13 percentile points, assuming of course that the observed temporal changes remained consistent across the entire 1985–2009 period. Nonetheless, these data represent the best available and most up-to-date health-related fitness data on Australian children. It must also be remembered that despite being simple, cheap, easy, reliable, reasonably valid and widely used alternatives of laboratory-based criterion measures, field tests are affected by factors other than underlying construct fitness. For example, validity data for field tests of cardiovascular fitness suggest that (at best) only 50–60% of the variance in field test performance is explained by the variance in underlying peak oxygen uptake, indicating that other physiological, physical, biomechanical, psychosocial and environmental factors also play a part.[15] In addition, although criterion-related validity has not been established for all of the included tests, face validity is generally accepted.[17] Most of the included tests are also considered to demonstrate good reliability, although tests requiring a reasonable degree of subjective judgement (eg, the subjective scoring of a properly performed sit-up or push-up) typically demonstrate poorer reliability.[14]

CONCLUSION

Physical fitness is considered to be an excellent marker of current and future health. In anticipation of a follow-up national fitness survey, this study provides the most up-to-date and most comprehensive set of sex- and age-specific normative centile values of health-related fitness of Australian children, which can be used as benchmark values for health and fitness screening and surveillance systems. These normative centile values will facilitate the identification of children with low fitness to set appropriate fitness goals, monitor individual changes in fitness and promote positive health behaviours. They will also facilitate the identification of children who possess specific fitness characteristics that may be considered important for sporting success, in the hope of recruiting the high achievers into elite sporting development programs.

Correction notice This article has been corrected since it was published Online First. The authors have noticed that the normative data in Table 10 are incorrect. The correct table has been inserted.

Acknowledgements The authors thank the authors of the included studies for generously clarifying details of their studies and/or for providing raw data. The University of South Australia Divisional Development Research Scheme supported this study.

Competing interests None.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1. **Ortega FB**, Ruiz JR, Castillo MJ, *et al*. Physical fitness in childhood and adolescence: a powerful marker of health. *Int J Obes (Lond)* 2008;**32**:1–11.
2. **Ruiz JR**, Castro-Piñero J, Artero EG, *et al*. Predictive validity of health-related fitness in youth: a systematic review. *Br J Sports Med* 2009;**43**:909–23.
3. **Vicente-Rodriguez G**, Ara I, Perez-Gomez J, *et al*. High femoral bone mineral density accretion in prepubertal soccer players. *Med Sci Sports Exerc* 2004;**36**:1789–95.
4. **Kodama S**, Saito K, Tanaka S, *et al*. Cardiorespiratory fitness as a quantitative predictor of all-cause mortality and cardiovascular events in healthy men and women: a meta-analysis. *JAMA* 2009;**301**:2024–35.
5. **Lee CD**, Blair SN. Cardiorespiratory fitness and stroke mortality in men. *Med Sci Sports Exerc* 2002;**34**:592–5.
6 **Dishman RK**, Washburn RA, Heath GW. *Physical Activity Epidemiology*. Champaign, IL: Human Kinetics 2004:358–9.
7. **Gillespie LD**, Robertson MC, Gillespie WJ, *et al*. Interventions for preventing falls in older people living in the community. *Cochrane Database Syst Rev* 2009;**2**:CD007146.
8. **Katzmarzyk PT**, Church TS, Janssen I, *et al*. Metabolic syndrome, obesity, and mortality: impact of cardiorespiratory fitness. *Diabetes Care* 2005;**28**:391–7.

9. **Katzmarzyk PT**, Church TS, Blair SN. Cardiorespiratory fitness attenuates the effects of the metabolic syndrome on all-cause and cardiovascular disease mortality in men. *Arch Intern Med* 2004;**164**:1092–7.
10. **Andersen LB**, Hasselstrøm H, Grønfeldt V, *et al*. The relationship between physical fitness and clustered risk, and tracking of clustered risk from adolescence to young adulthood: eight years follow-up in the Danish Youth and Sport Study. *Int J Behav Nutr Phys Act* 2004;**1**:6.
11. **Cleland VJ**, Ball K, Magnussen C, *et al*. Socioeconomic position and the tracking of physical activity and cardiorespiratory fitness from childhood to adulthood. *Am J Epidemiol* 2009;**170**:1069–77.
12. **Twisk JW**, Kemper HC, van Mechelen W. Tracking of activity and fitness and the relationship with cardiovascular disease risk factors. *Med Sci Sports Exerc* 2000;**32**:1455–61.
13. **Kristensen PL**, Wedderkopp N, Møller NC, *et al*. Tracking and prevalence of cardiovascular disease risk factors across socio-economic classes: a longitudinal substudy of the European Youth Heart Study. *BMC Public Health* 2006;**6**:20.
14. **Artero EG**, España-Romero V, Castro-Piñero J, *et al*. Reliability of field-based fitness tests in youth. *Int J Sports Med* 2011;**32**:159–69.
15 **Tomkinson GR**, Olds TS. Field tests of fitness. In: Armstrong N, Van Mechelen W, eds. *Pediatric Exercise Science and Medicine*. Oxford: Oxford University Press 2008:109–28.
16. **Castro-Piñero J**, Artero EG, España-Romero V, *et al*. Criterion-related validity of field-based fitness tests in youth: a systematic review. *Br J Sports Med* 2010;**44**:934–43.
17. **Ruiz JR**, Castro-Piñero J, España-Romero V, *et al*. Field-based fitness assessment in young people: the ALPHA health-related fitness test battery for children and adolescents. *Br J Sports Med* 2011;**45**:518–24.
18 Council of Europe. *Eurofit: Handbook for the Eurofit Tests Of Physical Fitness*. Rome: Council of Europe 1988.
19 The Cooper Institute. *FITNESSGRAM/ACTIVITYGRAM Test Administration Manual*. Fourth edition. Champaign, IL: Human Kinetics 2007.
20 **Olds TS**, Tomkinson GR. The aerobic fitness of 9–15 year old South Australian children: norms, trends and international comparisons. *ACHPER Healthy Lifestyles Journal* 2003;**50**:25–30.
21 **Pyke JE**. *Australian Health And Fitness Survey 1985: The Fitness, Health And Physical Performance Of Australian School Students Aged 7–15 Years*. Parkside, SA: The Australian Council for Health, Physical Education and Recreation 1987.
22. **Tomkinson GR**, Olds TS. Secular changes in pediatric aerobic fitness test performance: the global picture. *Med Sport Sci* 2007;**50**:46–66.
23. **Tomkinson GR**, Olds TS. Secular changes in aerobic fitness test performance of Australasian children and adolescents. *Med Sport Sci* 2007;**50**:168–82.
24. **Tomkinson GR**. Global changes in anaerobic fitness test performance of children and adolescents (1958-2003). *Scand J Med Sci Sports* 2007;**17**:497–507.
25. **Tomkinson GR**, Léger LA, Olds TS, *et al*. Secular trends in the performance of children and adolescents (1980-2000): an analysis of 55 studies of the 20m shuttle run test in 11 countries. *Sports Med* 2003;**33**:285–300.
26. **Cole TJ**, Green PJ. Smoothing reference centile curves: the LMS method and penalized likelihood. *Stat Med* 1992;**11**:1305–19.
27 **Pan H**, Cole T. *User's Guide to Lms Chart Maker*. Council MR, ed. 1997–2005 Medical Research Council. UK: Medical Research Council 2005.
28. **Jackson LV**, Thalange NK, Cole TJ. Blood pressure centiles for Great Britain. *Arch Dis Child* 2007;**92**:298–303.
29. **Ortega FB**, Artero EG, Ruiz JR, *et al*. Physical fitness levels among European adolescents: the HELENA study. *Br J Sports Med* 2011;**45**:20–9.
30. **Castro-Piñero J**, González-Montesinos JL, Keating XD, *et al*. Percentile values for running sprint field tests in children ages 6-17 years: influence of weight status. *Res Q Exerc Sport* 2010;**81**:143–51.
31 **Cohen J**. *Statistical Power Analysis For The Behavioral Sciences*. Second edition. New Jersey: Lawrence Erlbaum 1988.
32. **Cole TJ**, Freeman JV, Preece MA. Body mass index reference curves for the UK, 1990. *Arch Dis Child* 1995;**73**:25–9.
33. **Eisenmann JC**. Waist circumference percentiles for 7- to 15-year-old Australian children. *Acta Paediatr* 2005;**94**:1182–5.
34 **Armstrong N**, McManus AM, Welsman JR. Aerobic fitness. In: Armstrong N, Van Mechelen W, eds. *Paediatric Exercise Science and Medicine*. Oxford: Oxford University Press 2008:269–82.
35. **Krahenbuhl GS**, Skinner JS, Kohrt WM. Developmental aspects of maximal aerobic power in children. *Exerc Sport Sci Rev* 1985;**13**:503–38.
36. **Rowland TW**. Evolution of maximal oxygen uptake in children. *Med Sport Sci* 2007;**50**:200–9.
37 **Blimkie CJR**, Sale DG. Strength development and trainability during childhood. In: Van Praagh E, eds. *Pediatric Anaerobic Performance*. Champaign, IL: Human Kinetics 1998:193–224.
38 **Froberg K**, Lammert O. Development of muscle strength during childhood. In: Bar-Or O, ed. *The Child and Adolescent Athlete*. Oxford: Blackwell Science 1996:42–53.
39 **De Ste Croix MDA**. Muscle strength. In: Armstrong N, Van Mechelen, W, eds. *Paediatric Exercise Science and Medicine*. Second edition. Oxford: Oxford University Press 2008:199–212.
40. **Olds T**, Tomkinson G, Léger L, *et al*. Worldwide variation in the performance of children and adolescents: an analysis of 109 studies of the 20-m shuttle run test in 37 countries. *J Sports Sci* 2006;**24**:1025–38.
41. **Adegboye AR**, Anderssen SA, Froberg K, *et al*. Recommended aerobic fitness level for metabolic health in children and adolescents: a study of diagnostic accuracy. *Br J Sports Med* 2011;**45**:722–8.
42. **Lobelo F**, Pate RR, Dowda M, *et al*. Validity of cardiorespiratory fitness criterion-referenced standards for adolescents. *Med Sci Sports Exerc* 2009;**41**:1222–9.
43. **Ruiz JR**, Ortega FB, Rizzo NS, *et al*. High cardiovascular fitness is associated with low metabolic risk score in children: the European Youth Heart Study. *Pediatr Res* 2007;**61**:350–5.
44 **Welk GJ**, Laurson KR, Eisenmann JC, *et al*. Development of youth aerobic capacity standards using receiver operator characteristic curves. *Am J Prev Med* 2011;**41**:S111–6.
45. **Cureton KJ**, Sloniger MA, O'Bannon JP, *et al*. A generalized equation for prediction of VO2peak from 1-mile run/walk performance. *Med Sci Sports Exerc* 1995;**27**:445–51.
46. **Léger L**, Lambert J, Goulet A, *et al*. [Aerobic capacity of 6 to 17-year-old Quebecois–20 meter shuttle run test with 1 minute stages]. *Can J Appl Sport Sci* 1984;**9**:64–9.
47. **Tomkinson G**. Aerobic fitness thresholds for cardio metabolic health in children and adolescents. *Br J Sports Med* 2011;**45**:686–7.
48. **Dwyer T**, Magnussen CG, Schmidt MD, *et al*. Decline in physical fitness from childhood to adulthood associated with increased obesity and insulin resistance in adults. *Diabetes Care* 2009;**32**:683–7.
49 **Tomkinson GR**, Hamlin MJ, Olds TS. Secular changes in anaerobic test performance in Australasian children and adolescents. *Pediatr Exerc Sci* 2006;**18**:314–28.
50 Australian Council for Health, Physical Education and Recreation. *Australian Fitness Education Award: User's Manual And Curriculum Ideas*. Adelaide, SA: Australian Council for Health, Physical Education and Recreation 1996.
51. **Barnett LM**, Van Beurden E, Morgan PJ, *et al*. Does childhood motor skill proficiency predict adolescent fitness? *Med Sci Sports Exerc* 2008;**40**:2137–44.
52 **Birchall J**. Health related fitness testing. In: Went S, ed. *A Healthy Start: Holistic Approaches to Health Promotion in School Communities*. Second edition. Melbourne: Monash University 1992:251–65.
53 **Booth M**, Macaskill P, McLellan L, *et al*. NSW Schools Fitness and Physical Activity Survey 1997. Sydney: NSW Department of Education and Training, 1997.
54 **Booth M**, Okely AD, Denney-Wilson E, *et al*. NSW Schools Physical Activity and Nutrition Survey (SPANS) 2004: Full Report. Sydney: NSW Department of Health, 2006.
55. **Burke V**, Beilin LJ, Durkin K, *et al*. Television, computer use, physical activity, diet and fatness in Australian adolescents. *Int J Pediatr Obes* 2006;**1**:248–55.
56. **Cooley D**, McNaughton L. Aerobic fitness of Tasmanian secondary school children using the 20-m shuttle run test. *Percept Mot Skills* 1999;**88**:188–98.
57 **Dollman J**, Olds T, Norton K, *et al*. The evolution of fitness and fatness in 10–11-year-old Australian schoolchildren: changes in distributional characteristics between 1985 and 1997. *Pediatr Exerc Sci* 1999;**11**:106–21.
58 **Hands B**. *Fitness and motor skill levels of Western Australian Primary School Children*. Perth, WA: University of Western Australia 2000.
59 **McNaughton L**, Morgan R, Smith P, *et al*. An investigation into the fitness levels of Tasmanian primary schoolchildren. *ACHPER Healthy Lifestyles Journal* 1996;**43**:4–10.
60. **Vandongen R**, Jenner DA, Thompson C, *et al*. A controlled evaluation of a fitness and nutrition intervention program on cardiovascular health in 10- to 12-year-old children. *Prev Med* 1995;**24**:9–22.

Normative health-related fitness values for children: analysis of 85347 test results on 9–17-year-old Australians since 1985

Mark J Catley and Grant R Tomkinson

Br J Sports Med published online October 21, 2011

Notes

Re-Affirming the Value of the Sports Exception to Title IX's General Non-Discrimination Rule

DORIANE LAMBELET COLEMAN* MICHAEL J. JOYNER** DONNA LOPIANO***

It all grew from the power of an idea.[1]

INTRODUCTION

In just two years, we will celebrate Title IX's fiftieth anniversary. The statute was designed to address pervasive sex inequality in educational settings, including in admissions and programming, and in benefits and treatment. Although sex equality in education-based sport was not an original focus of Title IX's proponents, it became an integral part of the project from the date of its enactment in 1972. Indeed, by the time the statute was in effect re-enacted in 1988, Title IX had become synonymous with sport.[2]

Title IX's structure reflects a hybrid approach to sex equality. That is, the statute consists of a sex-blind non-discrimination rule, and its regulations contain a set of limited, sex-affirmative exceptions. Thus, the statutory text provides in relevant part that

> [n]o person in the United States shall, on the basis of sex, be excluded from participation in, be denied the benefits of, or be subjected to discrimination under any education program or activity receiving Federal financial assistance.[3]

And the exceptions permit schools to take sex into account to address imbalances in admissions, academic programming, and sport.[4] Because these

* Professor of Law, Duke Law School.

** Caywood Professor of Anesthesiology and Perioperative Medicine, Mayo Clinic School of Medicine.

*** Adjunct Professor of Sports Management, Southern Connecticut State University. President, Sports Management Resources. Former Chief Executive Officer of the Women's Sports Foundation and Director of Women's Athletics at the University of Texas at Austin.

1. Cary McTighe Musil, *The Triumph of Title IX*, MS. MAG., Fall 2007, at 42, https://www.feminist.org/education/TriumphsOfTitleIX.pdf (quoting Professor David Sadker, whose groundbreaking work with his wife Professor Myra Sadker on gender bias and educational equity helped to support the development and eventual success of Title IX).

2. *See generally* EQUAL PLAY: TITLE IX AND SOCIAL CHANGE (Nancy Hogshead-Makar & Andrew Zimbalist eds., 2007). *See also infra* notes 38–78 and accompanying text (setting out the history of Title IX).

3. 20 U.S.C. § 1681 (2018).

4. 34 C.F.R. § 106 (2019); 34 C.F.R. §§ 106.41(a)–(b) (2019) (containing the general non-discrimination provision and the sports exception to that provision).

original regulations were specifically required by Congress,[5] they have traditionally been accorded heightened deference by the courts and are tightly woven into Title IX's legal fabric.[6]

This regulatory approach was designed to and has yielded extraordinary results for women and girls, and for society more generally. The pre- and post-Title IX narrative is well worn at this point: "Title IX has successfully changed the lives of girls and of women educators, protecting their rights, broadening their horizons and setting them up for success in later stages of their education and careers."[7] Still, coming almost two centuries after Mary Wollstonecraft's original argument for educating women, its recency still surprises.[8]

There is this, for example, from *The Harvard Crimson* on November 15, 1968, three years before Title IX became law:

> Yale President Kingman Brewster announced yesterday that Yale will become coeducational in September 1969. The announcement came shortly after the Yale faculty approved with only one dissenting vote a plan to admit 250 freshman women plus 250 upperclass women by transfer. Eventually 1500 women will be admitted in addition to the 4000 male students . . . The faculty first approved under-graduate coeducation at Yale in 1962, after women graduate students had been admitted for several years. The administration considered establishing an independent coordinate college for women, similar to Radcliffe, two years ago. Later, Vassar was invited to consider affiliating with Yale, but its trustees declined to abandon Poughkeepsie for New Haven.[9]

This last bit of regional disrespect was righted in *The New York Times* which, a month later, reported that "[a] Radcliffe dormitory at Harvard is applying as a unit for admission to Yale next year . . . to end the frustrations of semi-

---

5. Pursuant to a now-defunct process, the regulations were presented by Secretary Weinberger to President Ford for his signature in May 1975, and then to Congress on June 4, 1975. From that date, Congress had 45 days to review and approve or disapprove them. Letter from Caspar Weinberger, Sec'y of the Dep't of Health, Education, & Welfare, to the President (Feb. 28, 1975), at A-1–2 [hereinafter Letter from Casper Weinberger] (delivering and explaining HEW's final regulation), (on file with the Gerald R. Ford Presidential Library). Had Congress failed to act, the regulations would have become valid by default. *Id.* In its review period, senators and congressmen opposed to Title IX in its entirety or as applied to sport presented a series of resolutions rejecting the regulations, but ultimately, on July 21, the regulations were approved. *See History of Title IX*, WOMEN'S SPORTS FOUND. (Aug. 13, 2019), https://www.womenssportsfoundation.org/advocate/title-ix-issues/history-title-ix/history-title-ix/. We detail this history, including the still-active practice of according the regulations heightened deference, *infra* at notes 40–69 and accompanying text.,

6. *See, e.g.*, McCormick ex rel. McCormick v. Sch. Dist. of Mamaroneck, 370 F.3d 275, 286 (2d Cir. 2004) (summarizing the legislative history of Title IX and noting this heightened deference point). *See also infra* notes 65–69 and accompanying text (further discussing the modern legal basis for heightened deference in this context).

7. *See* Jennifer Hahn, *Schoolgirl Dreams,* MS. MAG., Fall 2007, at 46.

8. MARY WOLLSTONECRAFT, A VINDICATION OF THE RIGHTS OF WOMAN (1792).

9. *Yale Will Admit Women in 1969; May Have Coeducational Housing,* HARV. CRIMSON (Nov. 15, 1968), https://www.thecrimson.com/article/1968/11/15/yale-will-admit-women-in-1969/ [hereinafter *Yale Will Admit Women in 1969*].

coeducation."[10] The utter lack of seriousness with which the matter was discussed among otherwise respectable people and institutions belied its significance for women. But it appears to have been in line with the administration's rationale for the policy change, which was reported more-or-less consistently by the rival schools' newspapers. According to *The Crimson*, President Brewster denied that a "student-sponsored Coed Week [which] brought women students from throughout the northeast for academic and social activity . . . had a direct effect on his decision, but his proposal to the faculty praised the organizers of the week and their guests for providing Yale with 'uncommon excitement.'"[11] The *Yale Daily News* noted that "the decision to have women in 1969 was based not on Yale's seeing her mission as the education of both sexes, but on fears that Yale cannot continue to attract the nation's top males to a non-coeducated campus."[12]

In this context, Title IX was very much an "idea revolution."[13] It sought to force colleges and universities to matriculate, educate, and graduate female students not to service men and their institutional needs, but on an equal basis with their male students. It required them to see women as they saw men: as students and as alumni without regard to their sex; and to own sex equality in all respects as one of their own institutional goals.

The transition wasn't always smooth, but with some notable exceptions, mostly today colleges and universities have fully embraced these ideas. Indeed, as we write this almost fifty years later, although their numbers are still disproportionately low in STEM fields and in leadership positions in the professional ranks, women have become the majority of college graduates.[14] As a result, they are an increasingly important part of the work force and the economy.[15] Title IX was not singularly responsible for these developments, of course, but the statute's role in the larger societal project that is sex equality, and specifically the empowerment of women and girls, is widely recognized and celebrated.[16]

This narrative is generally mirrored in the context of education-based sport. High schools, colleges, and universities understand that the education of women

---

10. *A Timeline of Women at Yale*, YALE U., https://celebratewomen.yale.edu/sites/default/files/files/Timeline-of-Women-at-Yale.pdf (last visited Dec. 1, 2019) (quoting N.Y. TIMES). This is not an isolated example of the sexism on display in the paper in that period. *See also, e.g*, Faith A. Seidenberg, *The Federal Bar v. The Ale House Bar: Women and Public Accommodations*, 5 VAL. U. L. REV. 318, 324–25 (1971) (noting that "The New York Times, in reporting the . . . story [of *Seidenberg v. McSorleys' Old Ale House, Inc*., 317 F. Supp. 593 (S.D.N.Y. 1970) (holding that bar violated Equal Protection Clause when it refused to admit women as patrons)], said, 'There was, perhaps a trace of wistfulness in the ruling [that] the sawdust-floored haven was just another "public place" that must admit any customer who comes in, even a woman.'" N.Y. TIMES, June 26, 1970, at 1.).

11. *Yale Will Admit Women in 1969*, *supra* note 9.

12. *A Timeline of Women at Yale*, *supra* note 10.

13. Musil, *supra* note 1, at 42 (quoting Professor Sadker).

14. Dani Matias, *New Report Says Women Will Soon Be Majority of College-Educated U.S. Workers*, NPR (June 20, 2019), https://www.npr.org/2019/06/20/734408574/new-report-says-college-educated-women-will-soon-make-up-majority-of-u-s-labor-f.

15. *Id.*

16. *See, e.g.*, Hogshead-Makar & Zimbalist, *supra* note 2.

and girls includes providing them with opportunities both to participate and to compete that are on par with those that are provided to boys and men. As is the case in STEM, the numbers are not equal, and schools often struggle to achieve their regulatory obligations.[17] Still, "[s]ince the enactment of Title IX, women's participation in sport has grown exponentially."[18] In high school, girls' participation numbers have grown from 294,000 in 1971–72 to more than 3.4 million in 2018-19.[19] In college, women's numbers have grown from 30,000[20] in 1972 to more than 288,000[21] in 2017–18. Women and girls today have the opportunity only boys and men had in the previous period to reap the widely recognized and highly valued benefits of being physically strong, of being on teams and developing the myriad skills associated with competitive sport, of attending college on athletic scholarships, and of high-end competitive experiences.[22] Again, although Title IX is not singularly responsible for these developments, it is generally credited with a central role in this aspect of the empowerment project.[23]

There is an important difference between the two stories, however. Where there has been a clear and steady upward trajectory for women and girls in academics, especially outside of STEM, in athletics the momentum has always been mixed, and it has been a constant battle to gain and to retain ground.[24] Among other things, the "fate" of women's sports "has always been tied to the larger political climate."[25] Because of this, and because of the recency in historical time of the broader commitment to sex equality, it is important for those who are devoted to the idea to remain attuned to shifts in that climate.

The earliest and most persistent political opponents of women's sports came from the broader football community and from men's sport more generally.[26] Whatever they thought about women's equality in theory, they were clear that it

17. *See* Gerald Gurney, Donna Lopiano & Andrew Zimbalist, *Chapter 6: A Continuing Disgrace: Discrimination Based on Gender, Race, Ethnicity, and Disability*, *in* UNWINDING MADNESS: WHAT WENT WRONG WITH COLLEGE SPORTS AND HOW TO FIX IT (2017).

18. *This Day in History: Title IX Enacted*, HISTORY (July 28, 2019), https://www.history.com/this-day-in-history/title-ix-enacted [hereinafter *This Day in History*].

19. *2018-19 High School Athletics Participation Survey*, NAT'L FED'N OF STATE HIGH SCH. ASS'NS 50, 54 (2019), https://www.nfhs.org/sports-resource-content/high-school-participation-survey-archive/.

20. *This Day in History*, *supra* note 18.

21. *The Equity in Athletics Data Analysis Cutting Tool*, OFFICE OF POSTSECONDARY EDUC., U.S. DEP'T OF EDUC., http://ope.ed.gov/athletics (last visited Feb. 29, 2020).

22. See *infra* notes 128–147 and accompanying text (detailing this point).

23. *See, e.g.*, *This Day in History*, *supra* note 18.

24. This aspect of the story is routinely emphasized in Title IX retrospectives. *See, e.g.*, Hogshead-Makar & Zimbalist, *supra* note 2, at 5–6 (summarizing "the struggle for gender equity in athletics").

25. SUSAN WARE, TITLE IX: A BRIEF HISTORY WITH DOCUMENTS, 13 (2007). Interestingly, the explanations for the relative weakness of sport and STEM may be similar or the same, i.e., a lack of commitment on the part of original stakeholders to seeing females succeeding in those areas in particular, together with related sex-linked stereotypes and cultural norms. The numbers and experiences are better in medical school and medicine, which our co-author Michael Joyner suggests may be related to the higher likelihood of predictability, promotion, and success in that STEM field.

26. *See* Hogshead-Makar & Zimbalist, *supra* note 2; *infra* note 41 and accompanying text.

should not come at the expense of existing men's programs. In addition to resisting the threshold assumption that Title IX should apply to sport, they have specifically resisted the integration of teams, the inclusion of men's revenue producing sports in spending comparisons, and the reallocation of spending from men's programs to fund women's programs. Throughout, the underlying premise has been that men's sport produces higher value social goods and thus should not be diminished in an effort to achieve different goods that are of lesser or questionable value.

The most recent political challenge has come from the identity movement and affiliated advocacy groups whose goal has been to secure much needed protections for people who are transgender. To that end, movement advocates have pushed for policy reforms that would grow the circumstances in which law is sex-blind. Where sex remains a basis for classification, they have worked to ensure that people who are transgender are included in spaces and programming consistent with their gender identity.[27]

The merits of both approaches are clear to us in contexts where sex does not actually matter. But in sport, where sex and the sex-linked physical traits associated with the male and female body are outcome determinative, the effects of the proposed reforms would be revolutionary: they would require either the dismantling of Title IX's existing sex-segregated architecture and thus of the female category, or the unconditional inclusion of males who identify as females in girls and women's sport.[28] While the latter is less obviously existential than the former, both would signal that policymakers were abandoning the original commitment to sex equality in this setting.

There is no question about this, as the goal is expressed, unambiguously, in the public statements of movement advocates; for example, in this one from a set of prominent civil rights organizations dedicated to ensuring, among other things, that Title IX evolves to disallow any distinctions on the basis of sex as "sex" is normally defined:

> [W]e support laws and policies that protect transgender people from discrimination, including in participation in sports, and reject the suggestion that cisgender [sex typical] women and girls benefit from the exclusion of women and girls who happen to be transgender.[29]

---

27. Doriane Lambelet Coleman, *Sex in Sport*, 80 LAW & CONTEMP. PROBS 63, 102–111 (2017).

28. We understand that language and word choices are fraught in this discussion. We explain our approach and vocabulary immediately below, at and around notes 31–37. The bottom line is that our goal is to communicate to a broad audience in a highly contested space, including about what sex is and how it is relevant. To do this well, we can't adopt an unfamiliar or unclear lexicon, or one that assumes a particular political outcome since this paper is in part about what that outcome should be. We know this will make some readers uncomfortable, but we hope that this explanation will help. We intend no disrespect.

29. *Statement of Women's Rights and Gender Justice Organizations in Support of Full and Equal Access to Participation in Athletics for Transgender People*, NAT. WOMEN'S LAW CTR. (Apr. 10, 2019), https://nwlc.org/wp-content/uploads/2019/04/Womens-Groups-Sign-on-Letter-Trans-Sports-4.9.19.pdf [hereinafter *Statement of Women's Rights and Gender Justice Organizations*].

The underlying premise of those who support this move to elide the relevant physical differences between females on the one hand, and males who identify as women and girls on the other, is that reconceiving of sex as gender identity—or privileging gender identity over sex—will produce the highest value overall.[30]

The goals of this paper are to provide the legal, factual, and normative background necessary to evaluate the merits of this most recent challenge to the sports exception to Title IX's general nondiscrimination rule, and then to present the case for re-affirming the exception in a form that is appropriate for this next period of its history. It proceeds in three parts as follows: Part I describes the legal history of Title IX's sports exception, its goals, and the current state of the legal doctrine. Part II explains its scientific basis and rationale. Part III sets out the best case for and against affirming the commitment to sex equality in education-based sport, and then presents our argument for resolving the collision of interests at issue. The paper concludes that the original "idea revolution" continues to do important work and should not be abandoned, including in the sports space where equality requires not only recognizing but also celebrating physical sex differences. Including trans people within this design is difficult by definition, but because they are also entitled to dignity and respect, policymakers should accept the challenge.

* * *

Standardizing vocabulary is critical to communication among the different groups concerned with this topic. Many do not use the same words and phrases, and even if they do, they often use them differently. Most difficult are those instances when a word or term that is important to one group for descriptive or political reasons is politically anathema or even painful to another. Because capturing and controlling language is part of movement strategy, solving the latter impasse is especially complicated. We have attempted to standardize our use of the language in a way that avoids unnecessary harm or discomfort, but to the extent we cannot always do this, we intend no disrespect. Our goal is to communicate to a broad audience using standard terms, not to demean.[31]

Most importantly, given current debates, we do not work from the assumption that sex is or includes gender identity. Whether "sex" in law includes or is distinct from "gender identity" is at issue in both the debates about H.R. 5, The Equality Act (2019),[32] and in the Title VII case currently pending in the United

30. *See, e.g., id.* (arguing that it is good for all women that transgender women and girls are included in the category "women" including in athletics competition on the basis that there are no cognizable differences among transwomen and females that are not "driven by stereotypes and fear . . . nondiscrimination protections for transgender people—including women and girls who are transgender—are not at odds with women's equality or well-being, but advance them"). For an analysis of this claim from our perspective, i.e., from a different feminism, see *infra* notes 184–193 and accompanying text.

31. For a set of up-to-date working definitions from medicine and endocrinology, see Joshua D. Safer & Vin Tangpricha, *Care of Transgender Persons*, 381 NEW ENG. J. MED. 2451, 2451–60 (2019).

32. Earlier versions of H.R. 5 distinguished "sex" from "gender identity" where the current bill defines "sex" to include "gender identity." *See* Equality Act, H.R. 5, 116th Cong. (2019). This drafting move is based in evolving views in the advocacy and biosciences communities about how gender

States Supreme Court.[33] The related question whether gender identity is biologically based and, if so, whether it should be considered an aspect or characteristic of biological sex is the subject of ongoing consideration by relevant experts in the scientific community.[34] For now, however, both remain contested claims in a context in which "sex" is otherwise understood to be the word we use to denote the individual's biological and reproductive classification as male or female.[35] Because this paper is precisely about whether policy should be reformed so that sex in this standard sense comes to be replaced by or to include gender identity in the context of education-based sport, it is necessary for us clearly to distinguish the operative terms. Again, we intend no disrespect.

Sex – Biological sex. "Either of the two divisions, designated female and male, by which most organisms are classified on the basis of their reproductive organs and functions."[36] The cluster of sex-linked traits—i.e., chromosomal, gonadal, endocrinological (hormonal), and phenotypic characteristics—commonly used to establish and denote sex. Primary and secondary sex characteristics. Although they are terms of art and commonly used in science and medicine, "biological sex" and "biological (fe)male" may be hurtful to those who are triggered by references to sexed bodies. Because sport relies on the biological distinction between males and females to justify separate sex sport, however, we need to use the terms in their scientific sense and to consider their substance in this discussion.[37] Again, we intend no disrespect.

---

identity is appropriately characterized and also how best to craft a political path to full equality and inclusion.

33. Equal Emp't Opportunity Comm'n v. R.G. & G.R. Harris Funeral Homes Inc., 884 F.3d 560 (6th Cir. 2018).

34. *See* Safer & Tangpricha, *supra* note 31, at 2451–52 (addressing this point). For a summary description of some of the ongoing work on "brain sex" more generally, see also Coleman, *Sex in Sport*, *supra* note 27, at 75–77.

35. *See, e.g.*, R (on the Application of Miller) v. College of Policing & Chief Constable of Humberside, [2020] EWHC 225 (Admin) [225], ¶ 267 (Eng.) (decision out of the United Kingdom's High Court of Justice, addressing this contested issue and the expert witness statement of Professor Kathleen Stock on the point that "For many English speakers, 'woman' is strictly synonymous with 'biologically female[']and 'man' with 'biologically male'.").

36. *Sex*, AM. HERITAGE DICTIONARY OF THE ENGLISH LANGUAGE (5th ed. 2020).

37. One of our co-authors, Michael Joyner, is a biomedical researcher who, in the 2000s, was admonished in the peer review process to use the phrase "sex differences" and not "gender differences" when describing biological phenomenon. This insistence is increasingly routine in the biomedical setting where the study of sex differences is now well established and producing important value both in the basic sciences and in applications in personalized medicine. For example, there is a journal dedicated to the subject, see BIOLOGY OF SEX DIFFERENCES, https://bsd.biomedcentral.com/about/ (last visited Feb. 15, 2020); and the federal government routinely emphasizes this focus in its own work and in the distribution of research funding. *See, e.g.*, U.S. FOOD & DRUG ADMIN., *Understanding Sex Differences at FDA* (Apr. 12, 2019), https://www.fda.gov/science-research/womens-health-research/understanding-sex-differences-fda. This focus is in part the result of the report from the NAT'L. ACAD. OF SCI. INST. OF MED., EXPLORING THE BIOLOGICAL CONTRIBUTIONS TO HUMAN HEALTH: DOES SEX MATTER? (2001) [hereinafter EXPLORING THE BIOLOGICAL CONTRIBUTIONS], http://www.nationalacademies.org/hmd/~/ media/Files/Report%20Files/2003/Exploring-theBiological-

Female—An individual whose sex is female, i.e., who has ovaries not testes and a natural estrogenic not androgenic endocrine system. A person's designation as "female" may not correspond with their gender identity (in the case of a trans person) or their sex recorded at birth (in the case of some intersex persons).

Male—A person whose sex is male, i.e., who has testes not ovaries and a natural androgenic not estrogenic endocrine system. A person's designation as "male" may not correspond with their gender identity (in the case of a trans person) or their sex recorded at birth (in the case of some intersex persons).

Sex stereotype—An assumption (which may be evidence-based or not) or a generalization (which may have a factual basis) about the aptitudes, preferences, and capacities of males and females based on their biological sex.

Gender—Social and cultural expression of masculine or feminine behavior. Often used differently as a synonym for biological sex.

Gender identity—A person's deeply held internal sense of themselves as male, female, both, neither, or fluid, and which can be different from their biological sex or their sex recorded at birth.

Trans—Transgender—Gender incongruent—The word or term used to describe a person whose gender identity is different from their biological sex or their sex recorded at birth.

Trans(gender) woman/girl or man/boy—A person who is trans(gender) who identifies as a woman/girl or man/boy.

I. THE LEGAL HISTORY, MISSION, AND CURRENT DOCTRINE

Title IX was developed to secure equality for women and girls in federally funded educational settings. It filled the gap that was left by Title VII of the 1964 Civil Rights Act, which protects against sex discrimination in employment but excludes educational settings otherwise, and by Title VI, which prohibits federally funded programs from discriminating on the basis of race, color, and national origin, but not on the basis of sex. Because of this gap, there was no federal statutory remedy to address the educational disparities women and girls experienced in relation to boys and men before Title IX. There was also no effective constitutional remedy to address laws that supported sex discrimination, as the

Contributions-to-Human-Health-Does-Sex-Matter/DoesSexMatter8pager.pdf (explaining the difference between sex and gender (identity), reporting on the extensive physiological processes and medical contexts in which "sex matters", and because of this indicating, among other things, the need for researchers to address "the inconsistent and often confusing use of the terms "sex" and "gender" in the scientific literature and popular press").

United States Supreme Court had yet to subject them to more than rational basis scrutiny.[38]

As described in the Introduction, the architects of Title IX settled on a hybrid approach to achieving sex equality in education. They paired a general, sex-blind non-discrimination rule with a set of limited, sex-affirmative exceptions, which allow educational institutions to take sex into account where doing so is necessary to address particular imbalances, i.e., in admissions, in programming, and in sport. In sport, at least, if an institution can meet its sex equality obligations using a sex-blind approach—without taking sex into account—it need not use these sex-affirmative tools; they are formally permissive not mandatory. But they become mandatory in effect if these obligations are not or cannot be met otherwise.[39]

Assurances that sports teams would be sex segregated were material to Title IX's passage and to congressional approval of its implementing regulations.[40] For

---

38. *See* Craig v. Boren, 429 U.S. 190 (1976) (holding for the first time that sex discrimination was subject to heightened, i.e., intermediate, scrutiny under the Constitution's Equal Protection Clause).

39. As it concerns sports, this requirement is clear in the statute's legislative history, in the original 1975 regulations, and in the original 1979 Policy Interpretation. *See* Regulations of the Department of Education, Nondiscrimination on the Basis of Sex in Education Programs or Activities Receiving Federal Financial Assistance, 34 C.F.R. §106.12 (regulations governing athletics); Title IX of the Education Amendments of 1972; The Policy Interpretation: Title IX and Intercollegiate Athletics, 44 Fed. Reg. 71,419 (Dec. 11, 1979) (again making clear the requirement of parity of competitive opportunities). For example, the 1979 Policy Interpretation provides, *inter alia*, as follows:

> 4. Application of the Policy—Selection of Sports.
>
> In the selection of sports, the regulation does not require institutions to integrate their teams nor to provide exactly the same choice of sports to men and women. However, where an institution sponsors a team in a particular sport for members of one sex, it may be required either to permit the excluded sex to try out for the team or to sponsor a separate team for the previously excluded sex.
>
> a. Contact Sports—Effective accommodation means that if an institution sponsors a team for members of one sex in a contact sport, it must do so for members of the other sex under the following circumstances:
>
> (1) The opportunities for members of the excluded sex have historically been limited; and
>
> (2) There is sufficient interest and ability among the members of the excluded sex to sustain a viable team and a reasonable expectation of intercollegiate competition for that team.
>
> b. Non-Contact Sports—Effective accommodation means that if an institution sponsors a team for members of one sex in a non-contact sport, it must do so for members of the other sex under the following circumstances:
>
> (1) The opportunities for members of the excluded sex have historically been limited;
>
> (2) There is sufficient interest and ability among the members of the excluded sex to sustain a viable team and a reasonable expectation of intercollegiate competition for that team; and
>
> (3) Members of the excluded sex do not possess sufficient skill to be selected for a single integrated team, or to compete actively on such a team if selected.

In this context "skill" is understood to include physical not just learned capacity. *See infra* notes 40–54 and accompanying text. Together with the original 1975 regulations, the 1979 Policy Interpretation is woven into the fabric of Title IX, including in statutory law. *See infra* notes 51–69 and accompanying text (elaborating on the Title IX scheme).

40. As the bill made its way through Congress, "[a] few people (very few) noticed that athletics might be affected . . . and so there was a discussion on the floor of the Senate about whether [it] required

different reasons, both men's and women's groups supported and/or insisted on this. Men's teams simply did not want to have to include females on their rosters or to be made to subsidize the equality project.[41] Women's groups wanted separate opportunities because they were keen to secure equality in education-based sports, and they understood that it couldn't be achieved without this separation.

In particular, women and women's groups fell into one of two camps. Both accepted that there was a performance gap between male and female athletes that necessitated sex segregation and thus a sports exception to Title IX's general non-discrimination rule. But they disagreed about the source of the gap and thus about the terms of the exception:

One group took the position that sex segregation and thus the sports exception in the regulations would be necessary only for a period, until females were afforded the (equal) training and competition opportunities that would be required eventually to close the performance gap; after that, sport could be co-ed.[42]

---

educational institutions to allow women to play on football teams." WARE, *supra* note 25, at 41. The answer from its floor manager Senator Birch Bayh was no. *Id.* "Having inserted that notion into the legislative history, higher education retreated." *Id.* That brief discussion on the Senate floor was significant in two respects: it presaged the decades-long resistance to Title IX that would come from college football and men's sport in general. *See supra* note 26 and accompanying text (summarizing this resistance), and *infra* note 41 and accompanying text (further describing the legislative and regulatory history). And it laid the necessary legal foundation for sex segregation in this setting. The "sports exception" or "carve out" to Title IX's prohibition on sex discrimination in federally funded educational settings powered girls' and women's sport and continues to secure its success today.

41. *See supra* note 26 and accompanying text. The sports question exploded in the aftermath of the bill's passage through Congress. From the time President Richard Nixon signed the Education Amendments into law in the summer of 1972 to the summer of 1975 when Congress formally approved the implementing regulations, the institutional powerhouse that is men's college football went into overdrive to ensure that, in fact men's teams would not have to accommodate female athletes, and that equality for female athletes would not come at the expense of men's revenue producing sports. This activity was particularly heated in 1974 and 1975. In that period, the Senate heard but ultimately declined to pass a bill sponsored by Senator John Tower (R-TX) to amend Title IX to exclude revenue producing sports from compliance tabulation. Had it succeeded, this would have meant that spending on female athletes and women's sports programs could be only be compared with spending on men's non-revenue producing sports, simultaneously insulating (primarily) men's football and basketball programs from Title IX's effects, and reducing schools' obligations to women's programs. Instead, Congress passed an alternative bill sponsored by Senator Jacob Javitz (R-NY), formally the Educational Amendments of 1974, which required the Secretary of the Department of Health, Education, and Welfare (HEW) to develop "with respect to intercollegiate athletic activities reasonable [regulations] concerning the nature of particular sports." The Education Amendments of 1974, Pub. L. 93–380, 88 Stat. 612. Although some advocates for men's sport continued to press the point that men's revenue producing sports should not be made to pay (either directly or indirectly) for women's programming, the Department's focus during this rulemaking process was mainly on the question whether men's and women's sport and teams would be sex segregated. Comments and lobbying from men's groups were consistently against integration. The NCAA, for example, had no interest in having Title IX cover sport at all, and it was opposed to including women in any of its programming.

42. The National Organization for Women (NOW) originally disagreed with the AIAW and the NCAA that the goal should be separate sex sport. In a letter to President Ford, HEW Secretary Casper Weinberger explained NOW's position that "the 'separate but equal' concept is inappropriate for any civil rights regulation and that open access should be required for all athletic teams with one exception.

In general, one might characterize advocates in this first group as holding the view that sex differences were entirely the result of disparate treatment and sex stereotypes, both of which could be eradicated over time. In their view, like the classroom, eventually sport could also be sex-blind. The language they used in this context is reminiscent of that which appears in cases involving race-based affirmative action measures.[43]

The other group took the position that the performance gap was the result of a combination of disparate treatment, sex stereotypes, and biological differences. For those in this second group, even if that part of the performance gap that was the result of disparate treatment and sex stereotype could be eradicated, sex segregation would always be necessary because immutable biological differences would remain.[44] This view is consistent with the Supreme Court's current substantive equality jurisprudence, which distinguishes sex from stereotype, but also race from sex on the ground that the latter but not the former involve inherent differences—these differences are properly considered when doing so serves to empower rather than to subordinate.[45]

Importantly, although textualists reject the role of legislative history in statutory interpretation, to the extent that it remains important to others, and that there is an ongoing debate about what policymakers meant when they used the word "sex" in the drafting period,[46] it is easily resolved in the sports setting. At least in this context, the legislative history is clear that "sex" meant biological sex, which was distinguished from sex stereotype. Specifically, the biological

---

When women are effectively excluded from open teams (where skill in the given sport is the criteria, it is still conceded by all that open competition for a tackle football team would result in an all-male team), separate teams should be provided for them on the basis that the training and sports traditionally available to women have been limited and the provision of separate teams until such time as the training gap is filled would best fulfill the purposes of the Act." *See* Letter from Caspar Weinberger, *supra* note 5, at A-6, A-7.

43. *See, e.g.*, Grutter v. Bollinger, 539 U.S. 306 (2003) ("We expect that 25 years from now, the use of racial preferences will no longer be necessary to further the interest [in student body diversity] approved today.").

44. *See* Kathleen Megan, *Transgender Sports Debate Polarizes Women's Advocates*, CT MIRROR (July 22, 2019), https://ctmirror.org/2019/07/22/transgender-issues-polarizes-womens-advocates-a-conundrum/ (quoting our co-author Donna Lopiano: "Title IX was passed 47 years ago to ensure an equal education for girls, but included a 'carve out' allowing separate programs for girls because of the clear biological advantage that males have over females in athletics. 'It was the notion that there are distinct biological differences in sex that are immutable.'"); Memorandum from Patricia Sullivan Lindh, the President's Special Assistant for Women's Programs, to James Cannon, White House Domestic Policy Advisor (May 1, 1975) (on file with the Gerald R. Ford Presidential Library) (noting that allowing schools to field only one "open" team would let them off the hook in terms of providing equal opportunities for women, and that to assure sex equality, schools should have to take into account differences between men and women in "competitive skill and physical ability").

45. *See, e.g.*, United States v. Virginia, 518 U.S. 515, 532–33 (1996).

46. *See, e.g.*, Brief of Walter Dellinger, et al. as Amici Curiae in Support of the Employees, Bostock v. Clayton County, GA, Altitude Express, Inc. v. Zarda, and Harris Funeral Homes, Inc. v. Stephens, 888 F.3d 100 (Nos.17-1618, 17-1623,18-107), 2019 WL 3027045; Brief for the Federal Respondent Supporting Reversal, Harris Funeral Homes, Inc. v. EEOC, et al., 884 F.3d 560 (No.19-107) (debating this question in the context of Title VII of the 1964 Civil Rights Act).

differences between males and females that account for the performance gap, as well as those sex traits and related customs that raised safety and privacy concerns, were key to the discussions and decisions around inclusion and segregation.[47]

Similarly, to the extent that there is debate today about whether Title IX was designed to ensure that girls and women were able not only to participate but also to compete on an equal basis with boys and men, the legislative history also confirms this commitment. While some early proponents of the statute suggested that females were or should be interested only in ("cooperative and inclusive") participation not ("patriarchal capitalist") competition,[48] the stereotypes and

47. The hearings throughout the month of June 1975, before the House of Representatives Subcommittee on Post-Secondary Education of the Committee on Education and Labor, are illustrative, including in that members and witnesses distinguished race from sex, and focused on the extent to which the performance gap was the result of inherent differences between the sexes rather than historical disparate treatment. *See Sex Discrimination Regulations: Hearings Before the Subcomm. on Post-Secondary Educ. of the Comm. of Educ. & Labor*, 94th Cong. 54 (1975) (statement of Bob Blackman, Head Football Coach, Univ. of Illinois) ("HEW has already [taken sex differences into consideration], they have already stated . . . that because of physiological differences between men and women, the women are not expected to compete in the so-called contact sports . . . . So they have already stated the fact that there are differences."); *id.* at 130 (statement of Joan Holt, President, Eastern District, Ass'n for Intercollegiate Athletics for Women) ("[W]e have been discriminated against in the past due to physiological limitations that women do have, we are not capable of getting a place on the men's team, and they then have an obligation, both because of the discrimination of the past and because of our competitive interests, that they would have to provide a separate team for the women in this case."); *id.* at 197 (statement of Rep. McKinney) ("We know that until puberty, girl and boy children have roughly the same athletic capacity. After this point there is a significant difference in their ability in most sports. However, until we stop punishing girl children for being tomboys and allow their full participation in scholastic athletics, we will never know their true capacity as sportspersons."); *id.* at 390 (statement of Bernice Sandler, Dir., Project of the Status and Educ. of Women & Exec. Assoc., Ass'n of Am. Colls.) ("In almost all other areas of discrimination, the precedents and principles developed by the courts in race discrimination cases can readily and easily be applied to sex discrimination problems. Because of the general physical differences between men and women as a whole, the principles developed in other discrimination areas do not easily fit athletic issues, particularly in the area of competitive sports, where the issue of single sex and integrated teams is a difficult one to solve. 'Separate but equal,' which is a discredited legal principle in terms of civil rights, may have some validity when applied to some areas of competitive athletics . . . ."); *id.* at 339 ("Before puberty, males and females are nearly identical in their physical abilities. Tests of strength, muscular endurance, cardiovascular endurance and motor performance show few differences between the sexes up to this age. Beyond that age, however, the male becomes considerably stronger, possesses greater muscular and cardiovascular endurance and is more proficient in almost all motor skills."); *id.* at 343 ("To some, complete integration of the sexes in all sports would appear to be both the simplest and the least discriminatory solution. Upon closer examination, however, it becomes clear that because of the differences in training and physiology, such an arrangement would effectively eliminate opportunities for women to play in organized competitive athletics. For these reasons, this alternative would not appear to be in line with the principle of equal opportunity."); *id.* ("[T]he 'separate-but-equal' principal in competitive athletics can be justified for sex discrimination (but not for race discrimination) because there are general physical differences between [women] and men (but not between blacks and whites).").

48. *See* Interview with our co-author Donna Lopiano, Adjunct Professor of Sports Mgmt., S. Conn. State Univ. (Oct. 13, 2019) (describing the views of some within the "old" NOW and the AIAW who rejected NCAA-style sports administration and competition). This approach to women's sport is

norms they advanced were rejected by other advocates and ultimately by lawmakers.[49] To use a currently topical distinction, education-based sport was not designed to be like selection for the military's special forces, where women are entitled to participate in selection rounds but are rarely, if ever, competitive for full status because of their physical disadvantages relative to men;[50] rather, sport was sex segregated because the goal was parity across all categories of opportunity.

The regulations that were formally presented to, reviewed, and passed over by Congress in 1975 mimic Title IX's hybrid approach, and reflect the general consensus at the time regarding sex segregation in sport. Specifically, they begin with this general nondiscrimination provision:

> 86.41(a) *General.* No person shall, on the basis of sex, be excluded from participation in, be denied the benefits of, be treated differently from another person or otherwise discriminated against in any interscholastic … athletics offered by a recipient, and no recipient shall provide any such athletics separately on such basis.[51]

The provision is followed by the exception, or "carve out," for sex-segregation in sport. The exception emphasizes the two factors that make sex segregation

---

described in Ann Travers, *The Sport Nexus and Gender Injustice*, 2 STUD. IN SOC. JUST. 79, 86 (2008) (describing "radical and cultural feminist . . . scholars [who] indict sport in its current patriarchal capitalist iteration and seek to replace it with cooperative and nonhierarchical celebrations of physicality and play based on feminist principles of cooperation and inclusion").

49. DEP'T OF EDUC., OFFICE FOR CIVIL RIGHTS, A POLICY INTERPRETATION: TITLE IX AND INTERCOLLEGIATE ATHLETICS (1979) (making clear the requirement of parity of competitive opportunities); McCormick ex rel. McCormick v. Sch. Dist. of Mamaroneck, 370 F.3d 275, 282 (2d Cir. 2004) (recognizing the difference between the opportunity to participate and the opportunity to compete for the win, including for championships, and holding that Title IX requires schools to provide females with opportunities in both categories that are on par with those provided to males). Notably, when the subject is not trans athletes or (intersex) athletes with differences of sex development, the premise that sex segregated sport exists in part to ensure that there are the same numbers of spots in finals and on podiums for females as for males is generally not controversial. Indeed, even their advocates tend to take this as a given, i.e., they appear to appreciate the benefits, including for their clients, of sex segregation. This has them arguing not for co-ed sport, but rather for what is in effect an exception for those whose sex is male but who identify legally and/or personally as women and girls. *See, e.g., Statement of Women's Rights and Gender Justice Organizations*, *supra* note 29 (arguing that "transgender women and girls" should benefit fully and equally from participation "in women's sports"); Doriane Lambelet Coleman, *Semenya and ASA v. IAAF: Affirming the Lawfulness of a Sex-Based Eligibility Rule for the Women's Category in Elite Sport*, 19 SWEET & MAXWELL'S INT'L SPORTS L. REV. 83 (detailing how a version of this approach was presented in Ms. Semenya's case at CAS) [hereinafter Coleman, *Semenya and ASA v. IAAF*] .

50. Meghann Myers, *A Female Soldier Has Made It Through the Army's Special Forces Selection*, ARMY TIMES (Nov. 14, 2018), https://www.armytimes.com/news/your-army/2018/11/14/a-female-soldier-has-made-it-through-the-armys-special-forces-selection/.For more information on the integration of women into special operations career fields and concerns about sex equality and sex specific or neutral standards in that highly competitive context, see KRISTY M. KAMARCK, CONG. RESEARCH SERV., R42075, WOMEN IN COMBAT: ISSUES FOR CONGRESS (2016), https://fas.org/sgp/crs/natsec/R42075.pdf.

51. 34 C.F.R. § 106.41 (a) (2019).

necessary for the attainment of equality in sport, that is, concerns about competitive fairness and physical safety:

> 86.41(b) *Separate teams*. A recipient may operate or sponsor separate teams for members of each sex where selection for such teams is based upon competitive skill or the activity involved is a contact sport. However, where a recipient operates or sponsors a team in a particular sport for members of one sex but operates or sponsors no such team for members of the other sex, and athletic opportunities for members of that sex have previously been limited, members of the excluded sex must be allowed to try-out for the team offered unless the sport involved is a contact sport. For purposes of this part, contact sports include boxing, wrestling, rugby, ice hockey, football, basketball and other sports the purpose or major activity of which involved bodily contact.[52]

As is the case with Title IX more generally, the affirmative approach is permissive, not mandatory, in the first instance, meaning that if a school can find a way to provide equal training and competition opportunities for females without taking sex into account, they can proceed in a sex-blind way; but if proceeding in a sex-blind way perpetuates disparities, the affirmative approach becomes mandatory.[53]

According to Susan Ware, "[i]n the early days of the law, much discussion centered on whether teams should be coeducational based on skill (the model adopted in elementary and high school physical education classes) and whether women should be eligible to play on men's teams. On the high school and intercollegiate level, a consensus soon emerged that sex-segregated but comparable sports teams were a better model."[54] As our co-author Donna Lopiano has explained, "[i]t was the notion that there are distinct biological differences in sex that are immutable . . . Everybody agreed that . . . if you have boys and girls competing after puberty, who would be more likely to get on a team? Who would win? It would be men. There would be very few women."[55]

From 1975 through 1988, proponents of girls' and women's sport continued to face resistance from boys and men and the male sports establishment, including with respect to funding, facilities, coaching staff, and competition opportunities.[56]

---

52. *Id.* § 106.41 (b).

53. *See, e.g.*, Yellow Springs Exempted Village School Dist. Bd. Of Educ. v. Ohio High School Athletic Ass'n, 647 F.2d 651, 656 (6th Cir. 1981) (Title IX "grant[s] flexibility to the recipient of federal funds to organize its athletic program as it wishes [one or separate teams] so long as the goal of equal athletic opportunity is met.").

54. WARE, *supra* note 25, at 5. This model—co-ed "prior to puberty"— is represented in the still-current position of the Women's Sports Foundation (WSF). *See* WOMEN'S SPORTS FOUND., ISSUES RELATED TO GIRLS AND BOYS COMPETING WITH AND AGAINST EACH OTHER IN SPORTS AND PHYSICAL ACTIVITY SETTINGS, https://www.womenssportsfoundation.org/wp-content/uploads/2019/08/issues-related-to-girls-and-boys-competing-with-and-against-each-other-in-sports-and-physical-activity-settings-the-foundation-position.pdf [hereinafter WOMEN SPORTS FOUND., ISSUES RELATED TO GIRLS AND BOYS COMPETING]. Although the original position paper was developed and published several years ago, the WSF re-affirmed it on August 14, 2019. *See id*.

55. Megan, *supra* note 44 (quoting Donna Lopiano).

56. *See* Hogshead-Makar & Zimbalist, *supra* note 2, at Part III (describing this period as "The Initial Backlash"); Mary C. Curtis & Christine Grant, *Landmark Title IX Cases in History*, GENDER EQUITY IN SPORT, http://bailiwick.lib.uiowa.edu/ge/historyRE.html (listing key dates in the resistance).

This resistance was particularly fierce in circumstances that involved cuts to boys' and men's programs that were considered—or described as—necessary to meet Title IX requirements.[57] It culminated in the 1984 decision, *Grove City College v. Bell*, in which the United States Supreme Court sided with the Reagan Administration's position that Title IX and its sex equality requirements applied only to the particular programs that received federal funds, not more broadly to the institutions of which they were a part.[58] Because the federal government did not contribute directly to education-based sports programs, *Grove City* in effect "gutted" Title IX.[59] As a result, the merits of the "idea revolution"—that there should be sex equality across educational settings including in education-based sport—were once again put to Congress.[60]

The Education Amendments of the Civil Rights Restoration Act of 1987 were finally passed in 1988, over President Reagan's veto, extending Title IX's sex equality requirements to all programs within institutions receiving federal funds.[61] As Title IX expert and three-time Olympic Gold Medalist Nancy Hogshead-Makar explains, although sport was not prominent in the legislative history prior to the statute's passage in 1972, "sports for girls and women were the driving narrative behind the imperative to pass the law again in 1988. Sports for women swung Republicans and average families."[62] Since then, although resistance has been ongoing,[63] the legislative, executive, and judicial branches of the federal government have consistently reaffirmed at least the essential aspects of the statutory scheme, including that parity of competitive opportunities matter and that the original regulations remain an integral part of the law.[64]

---

57. Cases involving cuts to boys' and men's wrestling were particularly prevalent. *See generally* Bradley David Ridpath et al., *Changing Sides: The Failure of the Wrestling Community's Challenges to Title IX and New Strategies for Saving NCAA Sport Teams*, 1 J. INTERCOLLEGIATE SPORT 255 (2008). *See, e.g.*, Nat'l Wrestling Coaches Ass'n v. Dep't of Educ., 366 F.3d 930 (D.C. Cir. 2004) (affirming district court's finding that the decision to drop wrestling is a matter of institutional preference not a requirement in fact or in effect of Title IX).

58. Grove City College v. Bell, 465 U.S. 555 (1984) (rejecting Association's challenge to Title IX Policy Interpretation).

59. *See* E-mail from Nancy Hogshead-Makar, Chief Exec. Officer, Champion Women, to Doriane Lambelet Coleman, Professor of Law, Duke Law Sch. (Feb. 19, 2020, 3:17 PM) (on file with authors).

60. *See supra* note 13 and accompanying text (introducing "the idea revolution").

61. *See S. 557 (100th): Civil Rights Restoration Act of 1987*, 100th Cong. (1987), https://www.govtrack.us/congress/bills/100/s557 (providing timeline of the history and text of the legislation).

62. *See* E-mail from Hogshead-Makar, *supra* note 59. *See also* WARE, *supra* note 25, at 36–43 (describing this history).

63. *See* Hogshead-Makar & Zimbalist, s*upra* note 2, at Part V (describing the period from 2001 to 2008 as "The Second Backlash").

64. For example, the 1994 Equity in Athletics Disclosure Act requires "co-educational institutions of postsecondary education that participate in a Title IV, federal student financial assistance program, and have an intercollegiate athletic program, to prepare an annual report to the Department of Education on athletic participation, staffing, and revenues and expenses, by men's and women's teams. The Department . . . use[s] this information in preparing its required report to the Congress on gender equity in intercollegiate athletics." *Equity in Athletics Disclosure Act*, U.S. DEPT. OF EDUC. (Jan. 24, 2017), https://www2.ed.gov/finaid/prof/resources/athletics/eada.html. The data in detail are available from *The Equity in Athletics Data Analysis Cutting Tool*, OFFICE OF POSTSECONDARY EDUC., U.S. DEP'T OF

Although the legislative veto was declared unconstitutional in 1983,[65] in 1984, the United States Supreme Court affirmed that "where Congress has specifically delegated to an agency the responsibility to articulate standards governing a particular area"—as it did in 1972 with respect to Title IX's standards governing athletics—"we must accord the ensuing regulation considerable deference."[66] The standards in the 1975 Regulations as well as their 1979 Policy Interpretation have continued to benefit from such protection, even as the Court has increasingly rejected the use of legislative history as a tool for statutory interpretation.[67] In part,

---

EDUC., http://ope.ed.gov/athletics. In *McCormick ex rel. McCormick v. Sch. Dist. Of Mamaroneck*, 370 F.3d 275, 282 (2d Cir. 2004), the Second Circuit held that a school district was out of compliance with Title IX when it established separate boys' and girls' teams but provided boys with more and more important competitive opportunities. And in 2016, the Obama Administration issued guidance for the inclusion of transgender student-athletes in education-based sports that made clear its commitment to sex segregation when this remains necessary to secure competitive fairness and physical safety. *See infra* note 73 and accompanying text (providing the details of this guidance).

65. INS v. Chadha, 462 U.S. 919, 959 (1983).

66. Kelley v. Bd. of Trs., 35 F.3d 265, 270 (1994) (citing Chevron, U.S.A., Inc. v. Nat. Res. Def. Council, Inc., 467 U.S. 837 (1984)).

67. In 2018, the Eastern District of Michigan explained that heightened deference under *Chevron* is accorded to both the 1975 Regulations and the 1979 Policy Interpretation of the 1975 Regulations "because Congress explicitly delegated to the agency the task of prescribing standards for athletic programs under Title IX." *See* Mayerova v. E. Mich. Univ., 346 F. Supp. 3d 983, 989 (E.D. Mich. 2018) (citing the First Circuit's decision in Cohen v. Brown Univ., 991 F.2d 888, 895 (1st Cir. 1993) and the Sixth Circuit's decision in Miami Univ. Wrestling Club v. Miami Univ., 302 F.3d 608, 615 (6th Cir. 2002)). *See also* Ollier v. Sweetwater Union High Sch. Dist., 768 F.3d 843, 855 (9th Cir. 2014) (affirming its practice of giving *Chevron* deference to the 1979 Policy Interpretation and applying it to the high school setting, citing the Second Circuit's decision in *Mamaroneck*, 370 F.3d at 300, and the Sixth Circuit's decision in Horner v. Ky. High Sch. Athletic Ass'n, 43 F.3d 265, 272–75 (6th Cir. 1994)); Biediger v. Quinnipiac Univ., 691 F.3d 85, 96–97 (2d Cir. 2012) (the Second Circuit reaffirming its decision in *Mamaroneck*); Mansourian v. Regents of Univ. of Cal., 602 F.3d 957, 965 (9th Cir. 2010) (the Ninth Circuit reaffirming that "both the Policy Interpretation and the Clarification are entitled to deference under *Chevron*").

Chevron was reaffirmed by the Court itself in 2013, in *City of Arlington v. FCC*, 569 U.S. 290 (2013) (Justice Scalia writing for the Court and noting that courts cannot substitute themselves as policymakers when this is precisely the job of the federal agencies). *City of Arlington* was a 6/3 decision, with Justices Roberts, Kennedy, and Alito in dissent. Their concern was the extension of what they describe as essentially legislative authority to the executive—and to the consequent creation of an ever-growing administrative state—even in circumstances where Congress did not clearly delegate this authority. *Id.* at 312 (Roberts, C.J., dissenting). In their view, "before a court may grant such deference, it must on its own decide whether Congress—the branch vested with lawmaking authority under the Constitution—has in fact delegated to the agency lawmaking power over the ambiguity at issue." *Id.* at 317. Unless *Chevron* itself is repealed, because Congress "explicitly delegated to the agency the task of prescribing standards for athletic programs under Title IX," both the 1975 Regulations and the 1979 Policy Interpretation should continue to be accorded heightened deference by the courts. *Mayerova*, 346 F. Supp. 3d at 989. Of course, this would not be the case should Title IX, the Regulations, and/or the Policy Interpretation be repealed.

The final deference point relates to the Department of Education's own ongoing interpretation of the 1975 Regulations and the 1979 Policy Interpretation. Such interpretations are afforded regular—not heightened—deference only when the standards themselves are "genuinely ambiguous" and: the agency's interpretation is (1) its own "'authoritative' or 'official position,'" (2) "reasonable"; (3)

this is because the Regulations and Policy Interpretation establish the architecture of and rationale for sex segregated education-based sport; and, they are embedded in an inextricably linked web of related law, including in statutory law. For example, the 1994 Equity in Athletics Disclosure Act (EADA) requires federally funded colleges and universities to produce annual reports regarding their athletic programs so that the Department of Education can monitor compliance with Title IX.[68] The fact that there is deep bipartisan support for girls and women's sport surely influences ongoing deference to the regulations as well.[69]

To date, the movement to include trans people in spaces and opportunities based on their gender identity rather than on their sex has not altered this legal state of affairs at the federal level. That is, as of this writing, there are no new regulations that require recipients of federal education dollars to include transgender people in sport on the basis of their gender identity rather than their biological sex; and there are no federal cases that expand the meaning of "sex" to include or to be replaced by "gender identity" in a Title IX sports context.[70] As they have done under Title VII, a few federal circuits have expanded the meaning of "sex" under Title IX, but so far only in the contexts of restrooms and locker room access.[71]

Supporters of transgender student-athletes have argued that these precedents are applicable to sport: that just as transgender girls must be permitted to use girls' restrooms they must also be permitted to be on girls' sports teams and included without condition in girls' competitions.[72] But as the Obama

---

"implicate[s] its substantive expertise"; and (4) "reflect[s] 'fair and considered judgment.'" Kisor v. Wilkie, 139 S. Ct. 2400, 2412, 2415–17 (2019) (explaining that deference to agencies under the Court's decision in *Auer v. Robbins*, 519 U.S. 452 (1997), are based in the presumption that "when granting rulemaking power to agencies, Congress usually intends to give them, too, considerable latitude to interpret the ambiguous rules they issue" but that courts need not defer to those interpretations when they are not justified according to these requirements).

68. *See supra* note 64 (discussing the EADA).

69. This bipartisan support for girls and women's sport, in a climate where such issues are often difficult to find, is presumably one of the reasons Republicans have seized on trans inclusion in girls and women's sport as an election issue for the 2020 cycle. *See, e.g.*, James Freeman, Opinion, *Did Democrats Just Create a Problem with Soccer Moms and Dads? A Friday House Vote Could Be the Sleeper Issue of 2020,* WALL ST. J. (May 20, 2019, 4:56 PM), https://www.wsj.com/articles/did-democrats-just-create-a-problem-with-soccer-moms-and-dads-11558385818.

70. As of this writing, the first federal case to address this issue has just been filed in the United States District Court in the District of Connecticut. *See* Soule et al. v. Conn. Interscholastic Athletic Conference et al., No. 3:20-cv-00201, (D. Conn., filed Feb. 12, 2020). Otherwise, a Westlaw search of the All Federal database for "Title w/1 IX & transgender & sport" yields only thirteen cases, none of which apply to selection for sex segregated sports teams. Almost all are bathroom and/or locker room privacy cases. The others are not even indirectly on point.

71. *See, e.g.*, Adams ex rel. Kasper v. Sch. Bd. of St. Johns Cty., 318 F. Supp. 3d 1293 (M.D. Fla. 2018) (restrooms); G.G. ex rel. Grimm v. Gloucester Cty. Sch. Bd., 822 F.3d 709 (4th Cir. 2016), *vacated and remanded*, 137 S. Ct. 1239 (2017) (restrooms); Johnston v. Univ. of Pittsburgh, 97 F. Supp. 3d 657 (W.D. Penn. 2015) (restrooms and locker rooms).

72. *See, e.g.*, Shayna Medley & Galen Sherwin, *Banning Trans Girls from School Sports Is Neither Feminist Nor Legal*, ACLU (Mar. 12, 2019, 5:45 PM), https://www.aclu.org/blog/lgbt-rights/transgender-rights/banning-trans-girls-school-sports-neither-feminist-nor-legal (arguing that "[w]hen

Administration apparently recognized in 2016 when it was interpreting Title IX in its Transgender Guidance to schools, sport is different from restrooms not only in its policy objectives but also in the extent to which sex actually matters:[73] Where sport is designed to develop and showcase the capacities of the physical body—including mental control of the physical body, and the girls' and women's categories are designed to secure sex equality with respect to the benefits that flow from sports, restrooms are designed to provide a space for people to relieve themselves, and girls' and women's restrooms are designed to secure safety and privacy as they do.[74] In law, at least, institutional design and objectives matter, as do the facts about whether individuals are similarly or dissimilarly situated with respect to the characteristics that are relevant to their attainment. In any event, the Trump Administration has withdrawn the Obama Guidance, restoring the original regulatory status quo;[75] and the Department of Education's Office of Civil Rights is investigating a complaint alleging that the Connecticut Interscholastic

misinformation about biology and gender is used to bar transgender girls from sports in schools receiving federal funds, it amounts to the same form of sex discrimination that has long been prohibited under Title IX"); Dave Zirin, *Transphobia's New Target Is the World of Sports: First It Was Bathrooms, Now It's Athletics*, NATION (Mar. 5, 2019), https://www.thenation.com/article/archive/trans-runner-daily-caller-terry-miller-andraya-yearwood-martina-navratilova (analogizing the two in general).

73. Dear Colleague Letter on Transgender Students from Catherine E. Lhamon, Assist. Sec'y for Civil Rights, U.S. Dep't of Educ. & Vanita Gupta, Principal Deputy Assist. Att'y Gen. for Civil Rights, U.S. Dep't of Justice (May 13, 2016), https://www2.ed.gov/about/offices/list/ocr/letters/colleague-201605-title-ix-transgender.pdf. *Compare*

> Restrooms and Locker Rooms. A school may provide separate facilities on the basis of sex, but must allow transgender students access to such facilities consistent with their gender identity. A school may not require transgender students to use facilities inconsistent with their gender identity or to use individual-user facilities when other students are not required to do so. A school may, however, make individual-user options available to all students who voluntarily seek additional privacy.

with

> Athletics. Title IX regulations permit a school to operate or sponsor sex-segregated athletics teams when selection for such teams is based upon competitive skill or when the activity involved is a contact sport. *A school may not, however, adopt or adhere to requirements that rely on overly broad generalizations or stereotypes about the differences between transgender students and other students of the same sex (i.e., the same gender identity) or others' discomfort with transgender students*. Title IX does not prohibit age-appropriate, tailored requirements based on sound, current, and research-based medical knowledge about the impact of the students' participation on the competitive fairness or physical safety of the sport.

Id. at 3 (emphasis added). The first and third sentences in the "Athletics" paragraph are original to the regulations. The second or middle sentence is guidance developed by the administration concerning the application of the traditional rule to the transgender context. The administration's position at least on the first and third points, but arguably also the second, was well-grounded in the legislative history as recognized over the years by the courts interpreting the statute and its regulations. *See, e.g.*, Kelley v. Bd. of Trs., 35 F.3d 265, 270 (7th Cir. 1994) (noting that "Congress itself recognized that addressing discrimination in athletics presented a unique set of problems not raised in areas such as employment and academics").

74. The difference between sport and restrooms is further detailed in Coleman, *Sex in Sports*, *supra* note 27, at nn.316–317 and accompanying text.

75. Andrew Mytelka, *Trump Administration Rescinds Obama-Era Guidance on Transgender Students*, CHRON. OF HIGHER EDUC. (Feb. 22, 2017), https://www.chronicle.com/blogs/ticker/trump-administration-rescinds-obama-era-guidance-on-transgender-students/117025.

Athletic Conference (CIAC) policy allowing unconditional inclusion of trans girls in girls competition violates Title IX.[76]

Regardless of the Trump Administration's motivation for taking up this complaint,[77] as a doctrinal matter it is on sound footing. Not only is the legal history of the sports exception to Title IX's general nondiscrimination rule clear that it is focused on equality for females in relation to males, but it is also generally accepted that sex segregated sport is constitutional because of its grounding in the "inherent [biological] differences between men and women" and because its purposes are to "compensate" women (and girls) for past and ongoing sex-related discriminations, to "promote equal [sports] opportunity [between men and women]," and to "advance full development of [women's and girls'] talent and capacities."[78] As we explore further below, if both "sex" and "gender identity" became the basis for eligibility for girls' and women's sport—or, as applied, if both girls' and boys' sport included both males and females—inherent differences would no longer be the rationale for separate sex sport, and the girls and women's categories would no longer serve their equality and empowerment goals. They would lose their constitutional grounding.

II. THE SCIENTIFIC EVIDENCE SUPPORTING THE SPORTS EXCEPTION

The disagreement among women and women's groups in the 1970s about whether sex differences in athletic ability were merely stereotype or in fact inherent has long since been resolved. In 2008, Gina Kolata of *The New York Times* reported that "even though some scientists once predicted that women would eventually close the gender gap in elite performances—it was proposed that all they needed was more experience, better training and stronger coaching—that idea is . . . largely discredited, at least for Olympic events."[79] As we write this paper in 2020, it is clear that Kolata's point is accurate across the board, at both the elite and non-elite levels of almost all standard sports and events. To say, as some advocacy groups do, that there is "no evidence"—or that it is "myth" and

76. Dan Brechlin, *Federal Office of Civil Rights Agrees to Investigate Connecticut's High School Transgender Athlete Policy*, HARTFORD COURANT (Aug. 8, 2019), https://www.courant.com/sports/high-schools/hc-sp-high-school-connecticut-transgender-policy-20190808-20190808-j5yfbovklvf4fjrir4ouybssj4-story.html.

77. Mark Joseph Stern, *Betsy DeVos May Force High Schools to Discriminate Against Trans Athletes*, SLATE (Aug. 9, 2019), https://slate.com/news-and-politics/2019/08/trump-education-department-title-ix-trans-athletes-discrimination.html (noting that this decision may be part of a broader anti-trans agenda).

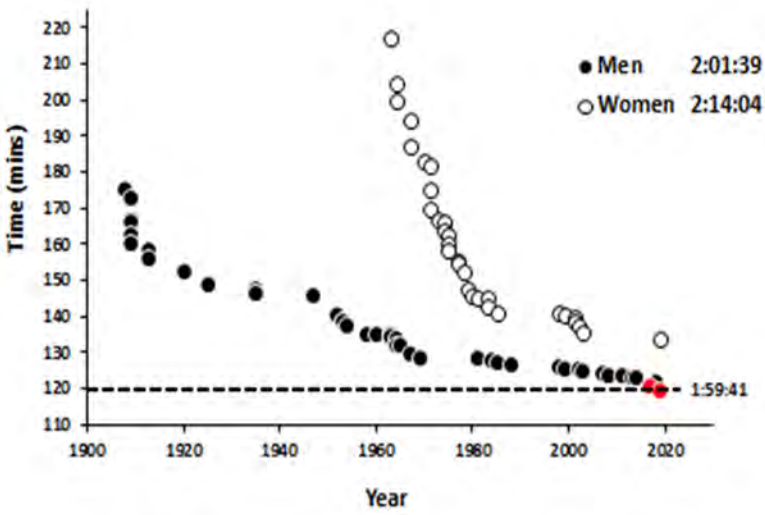
78. Coleman, *Sex in Sport*, *supra* note 27, at 67–70 (quoting United States v. Virginia, 518 U.S. 515, 532 (1996)); *see also* Kelley v. Bd. of Trs., 35 F.3d 265, 272 (7th Cir. 1994) (upholding the constitutionality of the statute and regulations against an equal protection challenge on these grounds); Cohen v. Brown Univ., 991 F.2d 888, 900–01 (same).

79. Gina Kolata, *Men, Women and Speed. 2 Words: Got Testosterone?* N.Y. TIMES (Aug. 22, 2008), https://www.nytimes.com/2008/08/22/news/22iht-22testosterone.15533354.html; *see also* Robinson Meyer, *We Thought Female Athletes Were Catching Up to Men, but They're Not*, ATLANTIC (Aug. 9, 2012), https://www.theatlantic.com/technology/archive/2012/08/we-thought-female-athletes-were-catching-up-to-men-but-theyre-not/260927/.

"outdated stereotype"—that males, including trans women and girls not on gender affirming hormones, are "better" in sport than females is simply to deny science.[80]

Sporting opportunities are not always identical, but there is now substantial parity in training and competition, especially at the elite level, and this has resulted in important performance gains for female athletes. But as the following figure by our co-author Mike Joyner tracking marathon performances illustrates, better training, better races, and more competitive opportunities throughout the world have resulted in gains for both sexes with a compressed time frame for women.

Marathon World Record Progression



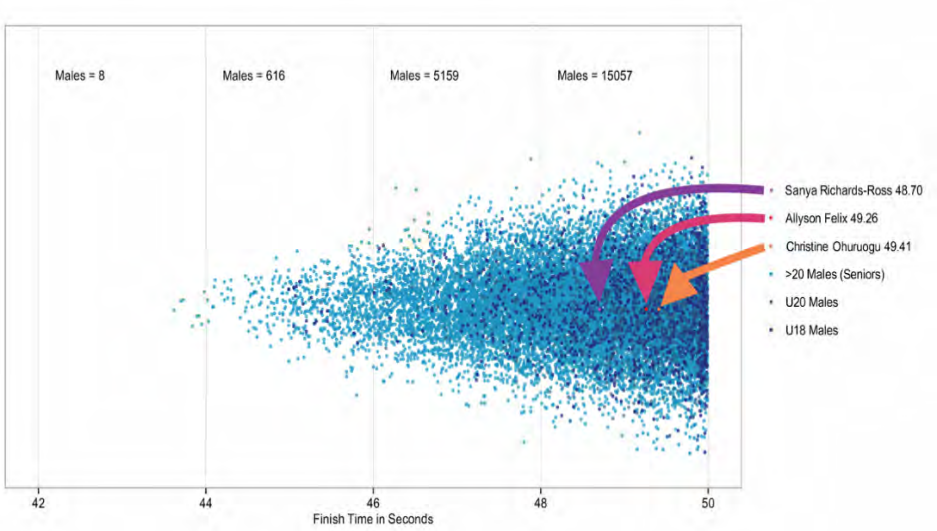
The upshot is that today, depending on the sport and event, the gap between the best male and female performances remains somewhere between 7 to 25 percent; and even the best female is consistently surpassed by many elite and non-elite males, including both boys and men.[81] If elite sport were co-ed or competition were open, even the best female would be rendered invisible by the sea of men and boys

80. *See, e.g.,* Medley & Sherwin, *supra* note 72 (using the terms "myth", "stereotype" and the claim of "no research" to describe the biological evidence in this context); Emilie Kao, *How Pelosi's "Equality Act" Would Ruin Women's Sport*, HERITAGE FOUND. (Apr. 24, 2019), https://www.heritage.org/gender/commentary/how-pelosis-equality-act-would-ruin-womens-sports (quoting Sunu Chandy from the National Women's Law Center using these ACLU talking points).

81. For data comparing male and female performances in a number of events on the track, including the number of males (boys and men) who surpass the very best females, see, e.g., Doriane Lambelet Coleman &Wickliffe Shreve, *Comparing Athletic Performances: The Best Elite Women to Boys and Men,* DUKE CTR. FOR SPORTS LAW & POL'Y (2018), https://law.duke.edu/sites/default/files/centers/sportslaw/comparingathleticperformances.pdf.

who would surpass her. As this next visual using the 400 meters on the track reflects, in percentage terms, the best female is bettered by relatively non-elite boys and men starting at 0.01 percent.[82] Simulating the final 100 meters of that event, it shows three of the fastest ever females on their very best day, against the thousands of boys and men whose performances—just in the single year 2017—would be competitive with or better than them in that final stretch.

Comparing the Best Elite Females to Boys & Men:
Personal Bests For 3 Female Gold Medalists vs 2017 Performances by Boys & Men



The same is true outside of the professional ranks in education-based sport, including in high school regular and post-season competition. For example, in 2016, Vashti Cunningham—the daughter of former NFL quarterback Randall Cunningham—set the high school American record in the high jump outdoors at 6 feet, 4½ inches. Since she joined the professional ranks, she has jumped 6 feet, 6¾ inches, and is ranked in the top ten in the world.[83] Still, in just one year—2018—and just in the state of California, 50 high school boys jumped higher than her high school best. Nationwide, in 2019, 760 boys jumped higher.[84] As the following figure simulating a high jump competition demonstrates, if high school sport were co-ed or competition were open, Cunningham would not have made it to her state

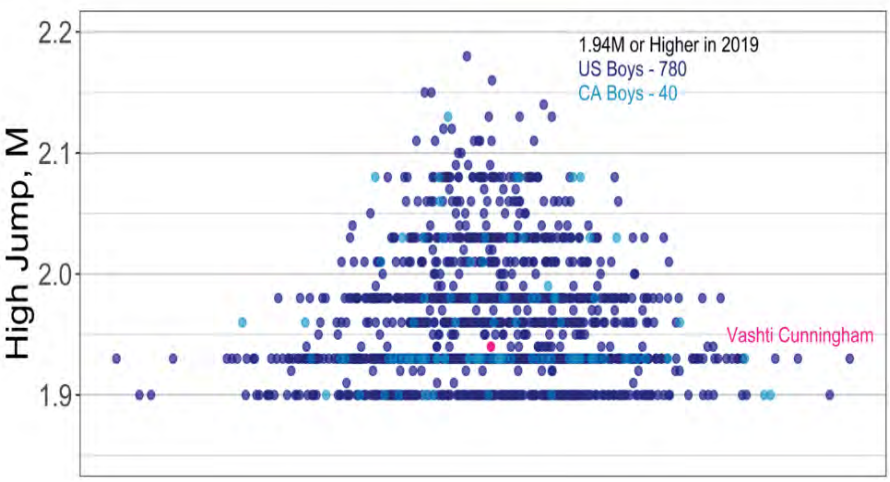
82. Jeff Wald, Doriane Lambelet Coleman, Wickliffe Shreve & Richard Clark, *Comparing the Best Elite Females to Boys and Men: Personal Bests for 3 Female Gold Medalists Versus 2017 Performances by Boys and Men*, DUKE CTR. FOR SPORTS LAW & POL'Y (2018).

83. *Athlete Profile: Vashti Cunningham*, WORLD ATHLETICS, https://www.iaaf.org/athletes/united-states/vashti-cunningham-280887 (last visited Jan. 15, 2020).

84. *See 2019 High School Men's High Jump Rankings*, ATHLETIC.NET, https://www.athletic.net/TrackAndField/Division/Event.aspx?DivID=97967&Event=9&page=7.

meet, she would not be on the national team, and we would not know her name other than as a footnote on her father's Wikipedia page.[85]

High Jump: Best American Boys in 2019
Compared to the Girls' American Record



It is perhaps more important for policy purposes that those girls who are only average high school athletes—for example, those who might just or occasionally win an invitational event or regional competition—would fare even worse. Indeed, a review of age-group performance data confirms what policymakers understood already in the 1970s: if sport were not sex segregated, most school-aged females would be eliminated from competition in the earliest rounds.[86] The following chart illustrates this point, using California intra-state regional results from the 2019 outdoor season, where the best boy in the state jumped 7 feet, 0 inches, and the best girl jumped 5 feet, 10 ½ inches. The average differential was

85. This figure by Jeff Wald is based on data from Athletics.net. According to that database and that of the NAT'L FED'N OF STATE HIGH SCH. ATHLETIC ASS'NS, https://www.nfhs.org/RecordBook/Record-book-result.aspx?CategoryId=1712 (last visited Jan. 28, 2020), only five other females residing and competing in the U.S. have jumped in this range: Alyxandria Treasure (1.92 meters in 2017), Jeannelle Scheper (1.91 meters in 2016), Toni Young (1.93 meters in 2009), Amy Acuff (1.91 meters in 1992), and Latrese Johnson (1.90 meters in 1985).

86. *See supra* notes 44, 54–55 and accompanying text (discussing this concern as it arose in the original policymaking process). *See also* E-mail from Michael J. Joyner, Caywood Professor of Anesthesiology and Perioperative Med., Mayo Clinic Sch. of Med., to Doriane Lambelet Coleman, Professor of Law, Duke Law Sch. (Oct. 6, 2019, 9:08 AM) (on file with authors) ("My younger boys 9 and 7 are good swimmers and they are in mixed races and it is hit or miss boy or girl who wins. The best swimmer in the club is a tall skinny 16-year-old girl who is getting recruited by good schools. She has real ability. She gets crushed by guys who will swim D3 if they choose to."); Coleman, *Sex in Sport*, *supra* note 27, at n.173 (describing the experience in Massachusetts with boys at the girls' state swimming championships).

approximately 12 inches. In percentage terms, across the state the performance gap ranged from 11.88 percent to 20.73 percent.

2019 California Regional High Jump Results[87]

REGION	BEST BOY	BEST GIRL	% DIFFERENCE
Central	2.0828	1.778	14.63%
Central Coast	1.9812	1.6764	15.38%
Los Angeles	1.8796	1.5748	16.22%
North Coast	2.0828	1.651	20.73%
Northern	1.9558	1.6764	14.29%
Oakland	1.8034	1.4732	18.31%
Sac-Joaquin	2.032	1.73355	14.69%
San Diego	2.032	1.7907	11.88%
San Francisco	1.8288	1.4732	19.44%
Southern	2.1336	1.7399	18.45%

The point that from puberty on, co-ed competition relegates most, if not all, females to being only participants in the game is easiest to prove in the case of sports with objective metrics, but "it is well-understood that a version of this story can be told across the board, almost no matter the event."[88] Indeed, the performance gap is so well-understood, and so abundantly documented in easily searchable databases, that it's difficult to take seriously the claim that it is merely

87. We developed this chart using data from the query "California High Jump Results," in ATHLETIC.NET , https://www.athletic.net/ (last visited September 25, 2019).

88. Coleman, *Sex in Sport*, *supra* note 27, at 91; Robinson Meyer, *We Thought Female Athletes Were Catching Up to Men, but They're Not*, *supra* note 79.

"myth" and "false stereotype."[89] Indeed, many on the sport and science side of the discussion have not bothered to try.[90]

Beyond the data, the sex-specific biology underlying the performance gap is also well-studied and well documented. Like other scientific fields that have focused on biological sex differences and that have come to recognize the extensive (beyond reproductive) reach of sex in the human body,[91] sports science and related disciplines—e.g., cardiology, hematology, endocrinology—have advanced tremendously in their understanding of the bases for the sex differences in athletic performance. What is clear from the evidence is that "the differences aren't the result of boys and men having a male gender identity, more resources, better training or superior discipline. It's because they have androgenized bodies."[92] Specifically, scientists agree that "males and females are materially different with respect to the main physical attributes that contribute to athletic performance," and that "the *primary* reason for sex differences in these attributes is exposure in gonadal males to much higher levels of testosterone (T) during growth and development (puberty), and throughout the athletic career."[93]

Before the onset of puberty, males and females produce similar, low levels of T, that is, on the order of 0.25 milligrams (mg) per day. But starting at puberty, male testes begin to produce much more than female ovaries and adrenal glands combined. On average, males (including elite male athletes) produce about 7 mg per day, and females (including elite female athletes) continue to produce about 0.25 mg per day. The normal male range is from 7.7 to 29.4 nmol/L. The normal

89. For example, domestic databases like Athletics.net provide not only national coverage but also regional, state, and local coverage that goes deep into college, high school, junior high school, and age group results. And international databases run by the governing bodies do a version of the same on a global level. *See supra* notes 81–83 (providing data from the IAAF's interactive database). Nevertheless, in this period the ACLU regularly insists that there is "no evidence" that males are better than females in sport. And even the NWLC publicly repeats this "no evidence" claim and adds that all sex differences are "unfounded stereotype." *See supra* note 80 and accompanying text (citing to these talking points). It is only if one accepts their predicate that the category "women" includes males who identify as female—or, as they put it, "women who happen to be transgender"—that stereotype theory works in the sports space. *See* Coleman, *Sex in Sport*, *supra* note 27, at 105–106, 109–11 (describing and responding to these rhetorical claims as they relate to sport). Otherwise, their argument is either uneducated or convenient science denial. *See* Doriane Coleman, Martina Navratilova & Sanya Richards-Ross, *Pass the Equality Act but Don't Abandon Title IX*, WASH. POST (Apr. 29, 2019), https://www.washingtonpost.com/opinions/pass-the-equality-act-but-dont-abandontitleix/2019/04/29/2dae7e58-65ed-11e9-a1b6-b29b90efa879_story.html?noredirect=on.

90. From statistics, for example, see, e.g., *For Crying Out Loud 2019, Biology in Sports Matters*, STATHOLE SPORTS (Apr. 18, 2019), http://statholesports.com/for-crying-out-loud-2019-biology-in-sports-matters/.

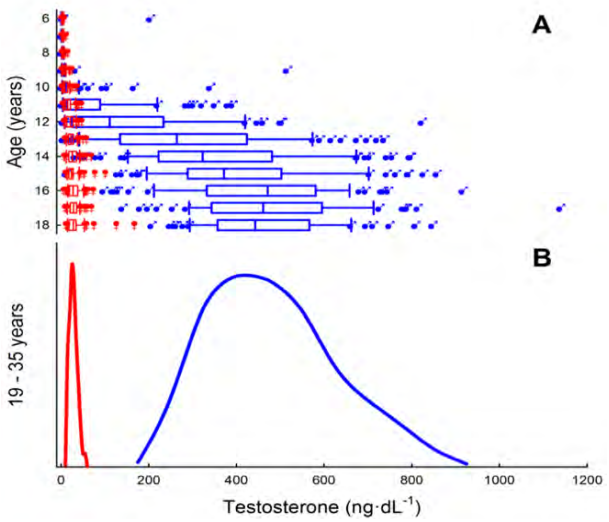
91. *See* EXPLORING THE BIOLOGICAL CONTRIBUTIONS, *supra* note 37 (IOM report examining the evolving study of sex differences and making the case that barriers to knowledge about sex differences must be eliminated).

92. Coleman & Shreve, *supra* note 81.

93. Various experts have recognized the impact of testosterone in athletic performance. *The Role of Testosterone in Athletic Performance*, DUKE CTR. FOR SPORTS LAW & POLICY (Jan. 2019) [hereinafter *Testosterone in Athletic Performance*] (emphasis added), https://law.duke.edu/sites/default/files/centers/sportslaw/Experts_T_Statement_2019.pdf.

female range is from 0.06 to 1.68 nmol/L.[94] In other words, as the following figure from Jonathon Senefeld and our co-author Michael Joyner shows, beginning at puberty, testosterone distributes bi-modally and males (whether they are trans or not) generally produce four to fifteen times more testosterone than females (whether they are trans or not). Female T readings are represented in red, male T readings in blue/purple.

Nationally Representative Data for Total Testosterone
for the U.S. Population Ages 6–35 Years[95]



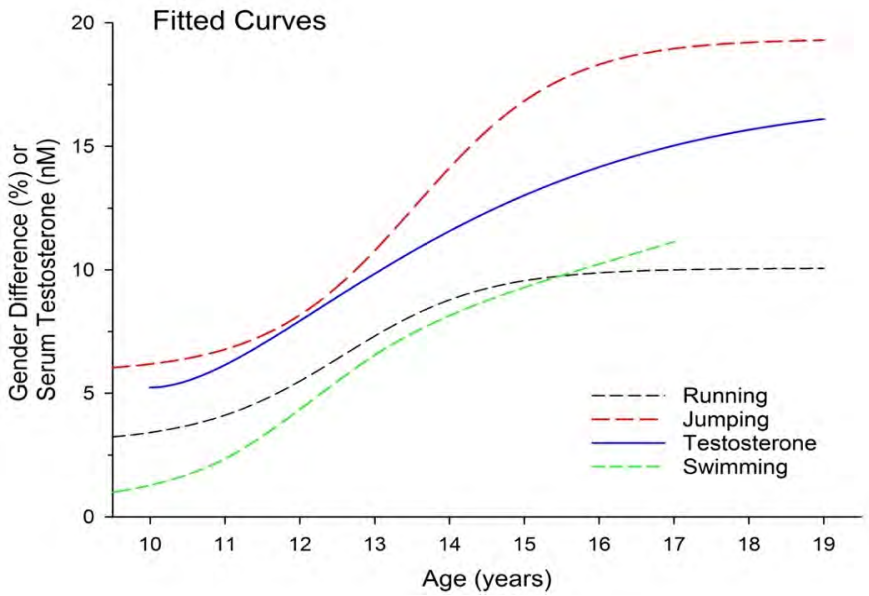
94. Females with polycystic ovaries (PCOS) may have levels upward of 4.8 nmol/L, and those with with untreated congenital adrenal hyperplasia (CAH) may have levels that are higher than that. But no healthy female, e.g., no elite athlete, has a natural T level above 5 nmol/L.

95. Jonathon W. Senefeld & Michael J. Joyner (2019). The data in the figure are from the National Health and Nutrition Examination Survey (NHANES), a study of health of random sample of children and adults in the United States. Using standard, validated clinical laboratory measurement tools (isotope dilution liquid chromatography tandem mass spectrometry) the NHANES laboratory precisely determined total testosterone of 4,229 girls/women and boys/men from ages 6 –35 years. These data are collected biennially to create a longstanding, National database of normative data. Panel A demonstrates the distribution of total testosterone concentration in children age 6–18 for each year using a box-plot to show the 25th, 50th and 75th percentile scores of testosterone. The error bars represent 3 standard deviations (SD) from the mean testosterone, and all outliers (greater or lesser than 3 SD from the mean) are marked using symbols. The girls (red colored box-plots and symbols) have a ~10-fold increase in testosterone (~3 ng·dL^{-1} to ~30 ng·dL^{-1}) that plateaus at 14 years. The boys (blue colored box-plots and symbols) have a substantially greater increase in testosterone than girls, an increase of over 100-fold (~4 ng·dL^{-1} to ~450 ng·dL^{-1}) that begins to plateau at 16 years. Testosterone concentrations are high and steady after age 18 (during years of peak endurance), and the distribution of testosterone for adults in this age range (19-35 years) are displayed in Panel B. This data set of over 1,400 samples from men and women shows the distribution curves from 99 percent of the data, with upper and lower outliers (0.05 percent above and below the mean) removed. The narrow range of the distribution for normative values for women (10-60 ng·dL^{-1}) is much smaller than the large range of normative values for men (175-925 ng·dL^{-1}).

As the next two figures demonstrate, this different exposure literally builds the male body in the respects that matter for sport. Specifically, "compared to biological females, biological males have greater lean body mass (more skeletal muscle and less fat), larger hearts (both in absolute terms and scaled to lean body mass), higher cardiac outputs, larger hemoglobin mass, (also both in absolute terms and scaled to lean body mass), larger $V\dot{O}_{2max}$ (higher aerobic capacity) (also both in absolute terms and scaled to lean body mass), greater glycogen utilization, higher anaerobic capacity, and different economy of motion."[96]

The figure immediately below, from David Handelsman, shows that the emergence of the performance gap in running, jumping, and swimming tracks the rise in male T levels at puberty:

Sex Differences in Athletic Performance Coinciding with the Onset of Male Puberty: Running, Jumping, and Swimming[97]

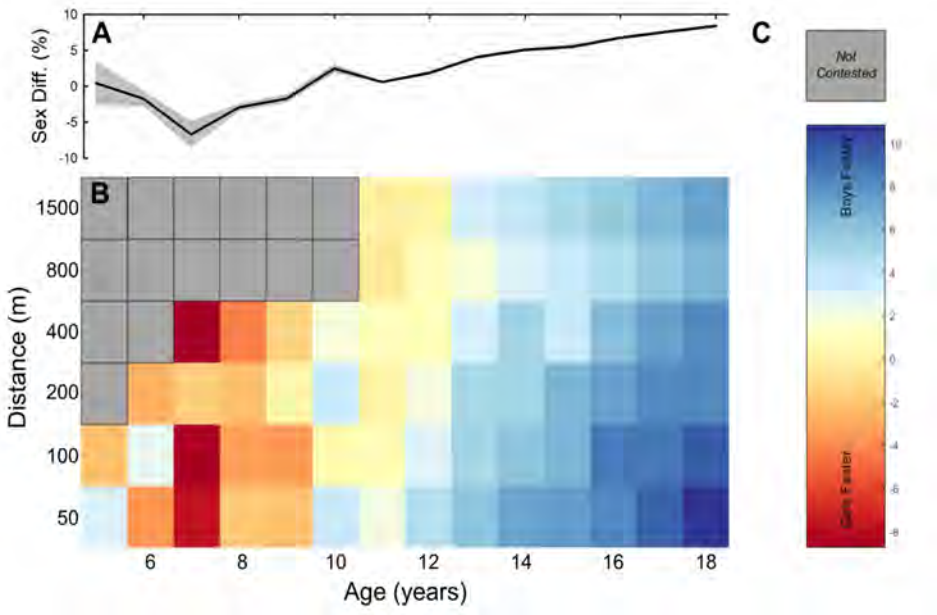


96. *Testosterone in Athletic Performance*, *supra* note 93.

97. This figure is part of a series that was published by David J. Handelsman et al., *Circulating Testosterone as the Hormonal Basis of Sex Differences in Athletic Performance*, 39 ENDOCRINE REVS. 803–29 (2018), https://academic.oup.com/edrv/article/39/5/803/5052770/.

This final figure, from our co-author Michael Joyner and his colleague Jonathon Senefeld, homes in on the swimming line. It confirms that pre-pubertal children of both sexes are competitive for the win in co-ed events, with females having some advantage in the six to eight-year-old age brackets. This—together with the requirement of sex-blind approaches when these do not undermine equality goals—is the basis for co-educational programming in elementary school and some age-group competitions.[98] The figure also confirms that from the onset of physical puberty to late adolescence, the competitiveness of females decreases essentially to zero. This—together with the preference for sex conscious approaches when these are necessary to meet equality goals—is the basis for policies that separate males and females in athletic competition starting in middle school and beyond.[99]

Sex Differences in Swimming Performance by Age[100]



98. *See generally* WOMEN'S SPORT FOUND., ISSUES RELATED TO GIRLS AND BOYS COMPETING, *supra* note 54 (describing the circumstances in which it is appropriate legally and scientifically to provide for co-ed and sex segregated sports and physical activities).

99. *See id.* (preferring co-ed sport until puberty). Swimming provides a particularly powerful example of the need for sex segregation in this context because the socio-cultural explanations for the performance gap are neutralized if not eliminated. Women and girls have had significant competitive opportunities in swimming that preceded Title IX, more girls participate in USA swimming than boys, training is systematic, rigorous and mixed from an early age, and records are typically set under standardized conditions. Additionally, swimmers typically come from resource rich homes where sex differences in nutrition or access to medical care are unlikely.

100. Michael J. Joyner & Jonathon W. Senefeld, *Sex Differences in Youth Elite Swimming*, 14 PLOS ONE 5 (2019). The data in the figure are from the top five swimming performances of all-time by US

The State of Connecticut is illustrative. It explains the applicability of these science facts in the FAQs that accompany its public school health, fitness, and performance standards:[101]

> *Why do some standards for boys and girls differ?*
>
> Two factors must be taken into account when determining criterion-referenced health standards: inherent physiologic differences between genders, and differences in health risks between genders. Due to physiologic and anatomic differences between the genders, there may be inherent performance differences between boys and girls for a specific fitness component. For example, differences in cardiac function and body composition between adolescent boys and adolescent girls result in adolescent boys, as a general rule, having a higher aerobic capacity than adolescent girls. Specifically, if the minimum $\dot{V}O_{2max}$ for healthy girls is 28 ml. kg-1. min-1 and for healthy boys, 32 ml. kg-1. min-1, setting the same standard for both on the One-Mile Run Test would not be appropriate. In the case of aerobic capacity, gender differences are taken into account, along with existing data on health risks, in order to determine the standards. In addition to physiologic differences, the two genders do not face the same health risks during their growth. To reflect these differentiated health risks, the standards are adjusted.
>
> *Why are some standards for boys and girls the same?*
>
> When there is no valid reason for expecting a difference in the performance between boys and girls, the standards are the same. For example, young children, particularly in Grades 1-6, do not always possess the physical and physiological differences that appear as children approach puberty (Falls & Pate, 1993). When this is true, the same standards may be used for both genders.
>
> *Why are standards for aerobic endurance lower for girls than for boys?*
>
> Inherent gender-related differences in body composition and in hemoglobin concentration cause aerobic capacity, referred to as VO2 max, for boys and girls who have the same level of physical activity, to be different. The differences prior to puberty are very small or nonexistent (for hemoglobin

---

swimmers for each 1-year age group from age 5 to 18 years, a database maintained and verified by USA Swimming. These data are using long course, freestyle swimming performances. Panel A demonstrates the sex difference in swimming performance across age (5–18 years). The sex difference in negative (indicating faster performances by girls) until age 10 (no sex differences) and then the sex difference markedly increases (faster boys) with a plateau at age 17. This plateau at ~8.5 percent maintains steady until ~age 50. The black line is the mean sex difference, and the grey area is the 95 percent confidence interval. Panel B demonstrates that similar trends were observed for each major freestyle swimming distance. Notably, the sex difference is largest in 'sprint events' (50, 100 and 200 m) and smallest in 'endurance events' (400, 800 and 1,500 m). Panel C is the legend for panel B. The y-axis for panel C is sex difference (%).

101. CONN. STATE DEP'T OF EDUC., CONNECTICUT PHYSICAL FITNESS ASSESSMENT: THIRD GENERATION 12–13 (2019–2020), https://portal.ct.gov/-/media/SDE/Phys-Ed/CPFA----Test-Administrators-Manual---2019.pdf?la=en.

> concentration), but they increase during puberty and adolescence. These differences are linked in part to differences in the reproductive hormones. The lower VO2 max in girls compared to boys with the same physical activity level are not thought to be associated with increased health risk. The standards for boys and girls reflect the different levels of VO2 max that are associated with increased health risk in adults.[102]

Notably, based on its interpretation of Connecticut anti-discrimination law, the Connecticut Interscholastic Athletics Conference (CIAC) has taken a position that is at odds with the state's physical standards, that is, despite the science and related policy, it permits trans girls—whether or not they were on gender affirming hormones—to compete in girls' events.[103] As a result, in the three year period 2016 to 2019, two trans girls who would not have been successful, had they competed in the boys' division, won fifteen individual state championships in the girls division.[104] It is this set of facts that is the basis for the federal Title IX complaint that the United States Department of Education's Office of Civil Rights is currently investigating.[105]

Apart from testosterone, different sex-linked factors and processes also contribute to sex differences in athletic performance. What this means is that even when trans women and girls use blockers and/or gender affirming hormones, male legacy advantages remain if their therapy begins only after the onset of puberty.[106] These include—among others—different muscle mass, bone density, and airway size.[107] Legacy advantages are more or less pronounced, depending on the age at

---

102. *Id*.

103. *See infra* note 104 and accompanying text (summarizing the results of CIAC policy).

104. In spring 2018, Terry Miller and Andraya Yearwood placed first and second at the 2018 Class M State Championship in the girls' 100 meters with times of 11.77 and 12.22, respectively. Miller broke the meet record of 12.16 that was set in 1995. *2018 CIAC Spring Championships: Class M Outdoor Track*, CIAC TOURNAMENT CENT., https://content.ciacsports.com/ot18m.shtml (last visited Jan. 28, 2020). The top five finishers from the Class M meet go on to the State Open, which includes the top twenty-five athletes in the state. Miller and Yearwood took 1st and 2nd place there too, with times of 11.72 and 12.29 respectively. *See 2018 CIAC Spring Championships: Open Outdoor Track*, CIAC TOURNAMENT CENT., https://content.ciacsports.com/ot18o.shtml (last visited Jan. 28, 2020). Because they had only run 11.87 and 12.28 going into the State Championships, and the boys' qualifying time was 11.84, neither would have qualified for post-season competition had they been competing in the boys' division. The top six from the State Open represent Connecticut in the New England Championship, which Miller went on to win as well. *See* 2018 New England High School Outdoor Track and Field Championships, RUNNERSPACE.COM, https://new-england-interscholastic-outdoor-championships.runnerspace.com/eprofile.php?do=info&event_id=5773&year=2018 (last visited Jan. 28, 2020).

105. S*ee supra* notes 76–77 and accompanying text (discussing this complaint).

106. Among other related questions, whether male legacy advantages are offset by disadvantages associated with transition is the subject of ongoing inquiry, including by Joanna Harper and colleagues at the University of Loughborough.

107. For an easy to follow description of sex differences in this category, see Yvonne van Dongen's interview with physiologist Alison Heather, *Even Before Birth, Genetic Building Blocks Are Giving Athletes Born Male a Massive Advantage Over Females*, STUFF (July 21, 2019), https://www.stuff.co.nz/sport/other-sports/114327152/even-before-birth-genetic-building-blocks-are-giving-athletes-born-male-a-massive-advantage-over-females; *see also, e.g.,* Paolo B. Dominelli et al., *Sex Differences in Large Conducting Airway Anatomy*, J. APPLIED PHYSIOLOGY (1985).

which the person physically transitions, and Joanna Harper has cautioned that they may be offset by the disadvantages associated with taking exogenous hormones and performing in a male frame with female T levels.[108] Finally, legacy advantages are more or less relevant to performance depending on the sport and event. The latter is why, for example, track and field has taken the position that winding down testosterone levels is an acceptable compromise, but power lifting has not.[109]

In contrast, despite the frequent contrary assertion from advocates for unconditional transgender inclusion, there is no evidence that other physical traits—such as height, wingspan, muscle fiber type, or lactic acid processing ability, among others—are similarly determinative of outcomes in sport.[110] As our co-author Doriane Coleman has written elsewhere,

> [t]here is no characteristic that matters more than testes and testosterone. Pick your body part, your geography, and your socioeconomic status and do your comparative homework. Starting in puberty there will always be boys who can beat the best girls and men who can beat the best women.
>
> Because of this, without a women's category based on sex, or at least these sex-linked traits, girls and women would not have the chance they have now to develop their athletic talents and reap the many benefits of participating and winning in sports and competition.[111]

Indeed, even scientists Ross Tucker and Eric Vilain, who are on record in support of gender inclusive policies, have acknowledged that this is why "we separate men and women into categories . . . we want women to be able to win some competitions;"[112] and, "[w]ithout a women's category" based on sex-linked traits, "elite sport would be exclusively male."[113] Because the traits that account for the performance gap are in play starting at the onset of male puberty, the same would be true of non-elite, education-based sport. To paraphrase them both, without a girls' category based on sex, or at least on sex-linked traits, education-based

---

108. Katherine Kornei, *This Scientist is Racing to Discover How Gender Transitions Alter Athletic Performance—Including Her Own*, SCI. MAG. (July 25, 2018, 9:00 AM), https://www.sciencemag.org/news/2018/07/scientist-racing-discover-how-gender-transitions-alterathletic-performance-including.

109. *Compare* WORLD ATHLETICS ELIGIBILITY REGULATIONS FOR TRANSGENDER ATHLETES, WORLD ATHLETICS, http://www.athletics.org.tw/Upload/Web_Page/WA/Eligibility%20Regulations%20for%20Transgender%20Athletes,%20.pdf, *with* USA POWERLIFTING TUE COMMITTEE REPORT 2019, USA POWERLIFTING,https://www.usapowerlifting.com/wp-content/uploads/2019/05/USA-Powerllifting-TUE-Committee-Report-2019.pdf.

110. *See Testosterone in Athletic Performance*, *supra* note 93 ("No other endogenous physical or physiological factors have been identified as contributing substantially and predominantly to [sex] differences [in athletic performance].").

111. Doriane Lambelet Coleman, *Sex, Sport, and Why Track and Field's New Rules on Intersex Athletes are Essential*, N.Y. TIMES (Apr. 30, 2018), https://www.nytimes.com/2018/04/30/sports/track-gender-rules.html.

112. Coleman, *Sex in Sport*, *supra* note 27, at 91 (quoting Eric Vilain).

113. *Id.* at 86 (quoting Ross Tucker).

competitive sport would be mostly, if not exclusively, male. And so, if we want females to win some competitions, we need to separate them.[114]

III. REAFFIRMING THE MISSION IN LIGHT OF THE CHALLENGE FROM THE IDENTITY MOVEMENT

As we show in Parts I and II, the question presented by advocates for unconditional inclusion of transgender athletes in girls and women's sport is not whether there is a difference between the male body and the female body that justifies current Title IX policy; nor is it whether that policy supports sex segregation in this arena. Rather, the question is the normative one whether, in light of new attention to and growing knowledge about differences of sex development and gender incongruence, we are and should remain committed to equality for females in relation to males in the education-based sports space. Specifically, is that aspect of the idea revolution that was equality for females still viable, or should we be moving on to a new one that rejects the original focus on sex so that we can be inclusive of individuals who are gender diverse?

The voices with the megaphone at the moment prioritize this new and different revolution;[115] but in fact, feminists are split on this issue, as are transgender people whether or not they are also feminists.[116] In this respect, the dispute is as much "in the family"[117] as it is among traditional political opponents.[118] Biological sex, sex equality, and sport are matters that mean a lot to

---

114. None of this is new to gym teachers and secondary school coaches, who have long used sex specific performance charts both to assess physical health and development and to sort students for competitive games and teams. *See supra* notes 101–102 and accompanying text (setting out the State of Connecticut's standards). Nor is it new to the military, which similarly uses sex specific standards for initial inclusion into the armed services and for fitness reviews, and sex neutral standards in circumstances where operational needs outweigh inclusion and equality concerns. *See generally* KAMARCK, *supra* note 50.

115. *See infra* text and accompanying notes 177–190 (discussing the public positions taken by, among other organizations, the ACLU, the National Women's Law Center, and the Women's Sports Foundation).

116. *Compare* Shasta Darlington, *Transgender Volleyball Star in Brazil Eyes Olympics and Stirs Debate*, N. Y. TIMES (Mar. 17, 2018), https://www.nytimes.com/2018/03/17/world/americas/brazil-transgender-volleyball-tifanny-abreu.html (including the views of Joanna Harper and Tifany Abreu), *with* Scott Gleeson & Erik Brady, *These Transgender Cyclists Have Olympian Disagreements on How to Define Fairness*, USA TODAY (Jan. 11, 2011), https://www.usatoday.com/story/sports/olympics/2018/01/11/these-transgender-cyclists-have-olympian-disagreement-how-define-fairness/995434001/ (comparing the views of Rachel McKinnon with those of some of her female competitors).

117. *See, e.g.*, Coleman, Navratilova, & Richards-Ross, *supra* note 89; Doriane Coleman, *Who Is a "Woman" in Sport*, VOLOKH CONSPIRACY (Mar. 11, 2019, 8:02 AM), https://reason.com/2019/03/11/who-is-a-woman-in-sport/; Dave Zirin, *Martina Navratilova is Expelled From an LGBTQ Advocacy Group Over Transphobia Accusations*, NATION (Feb. 25, 2019), https://www.thenation.com/article/martina-navratilova-athlete-ally-transphobia/.

118. *See, e.g.*, Andrea Jones & Clare Hepler, *Males Don't Belong in Women's Sports—Even If They Don't Always Win*, HERITAGE FOUND. (Nov. 27, 2019), https://www.heritage.org/gender/commentary/males-dont-belong-womens-sports-even-if-they-dont-always-win (one of a series of commentaries on the subject by the foundation); Megan, *supra* note 44 (reporting on filing of a Title IX complaint by the Alliance Defending Freedom).

many people regardless of politics, and so many people are interested in the conversation.

In this final part of the paper, we set out what we see as the most persuasive arguments on both sides of this debate, and we offer a proposal for policy reform that would simultaneously secure Title IX's existing commitment to sex equality and make clear policymakers' authority to develop approaches to the inclusion of gender diverse students in education-based sport that do not undermine that commitment.

A. The Affirmative Case for Sex Equality in Sport

As we see it, the most persuasive argument for sex equality in sport, including in competitive education-based sport, is the following:

Equality for females is a broadly-held commitment and a high value social good. Extending this commitment to education-based sport is part and parcel of securing this value for the individual females who are directly involved, for related stakeholders, and for society more generally. Because of biological ("inherent") differences between males and females, this value cannot be achieved if teams and events are not separated by sex. And, because of the way equal protection doctrine has evolved, there is likely no different—other than sex-based—rationale for segregated sport that is likely to survive constitutional scrutiny: if the classifications "girls' sport" and "women's sport" were based on gender identity rather than sex, they would be difficult if not impossible to sustain. Thus, for both biological and legal reasons, unless society is prepared to forego the benefits that flow from girls' and women's sport, the classification must continue to be based on sex, or at least on reproductive sex-linked traits.

1. Sex Equality is a High-Value Social Good

Both sex equality in general and empowering females in particular are societal priorities. For example, here using the word "gender" as a synonym for "sex" the United Nations Population Fund explains that:

> Gender equality is intrinsically linked to sustainable development and is vital to the realization of human rights for all. The overall objective of gender equality is a society in which women and men enjoy the same opportunities, rights and obligations in all spheres of life. Equality between men and women exists when both sexes are able to share equally in the distribution of power and influence; have equal opportunities for financial independence through work or through setting up businesses; enjoy equal access to education and the opportunity to develop personal ambitions, interests and talents; share responsibility for the home and children and are completely free from coercion, intimidation and gender-based violence both at work and at home.
>
> Within the context of population and development programmes, gender equality is critical because it will enable women and men to make decisions that impact more positively on their own sexual and reproductive health as well as that of their spouses and families. Decision-making with regard to such issues as age at marriage, timing of births, use of contraception, and recourse to harmful practices (such as female genital cutting) stands to be improved with the achievement of gender equality.

> However it is important to acknowledge that where gender inequality exists, it is generally women who are excluded or disadvantaged in relation to decision-making and access to economic and social resources. Therefore a critical aspect of promoting gender equality is the empowerment of women, with a focus on identifying and redressing power imbalances and giving women more autonomy to manage their own lives. This would enable them to make decisions and take actions to achieve and maintain their own reproductive and sexual health. Gender equality and women's empowerment do not mean that men and women become the same; only that access to opportunities and life changes is neither dependent on, nor constrained by, their sex.[119]

That males and females are physically different in relevant ways, and that females are routinely subject to discrimination because of their distinct physical characteristics, explains the UN's commitment to an equality toolbox that includes sex affirmative measures, that is, measures that "see" sex or that are not "sex blind." It is understood that such measures can be effective in ways that sex neutral measures are not: [120]

> Taking gender concerns into account when designing and implementing population and development programmes therefore is important for two reasons. First, there are differences between the roles of men and women, differences that demand different approaches. Second, there is systemic inequality between men and women. Universally, there are clear patterns of women's inferior access to resources and opportunities. Moreover, women are systematically under-represented in decision-making processes that shape their societies and their own lives. This pattern of inequality is a constraint to the progress of any society because it limits the opportunities of one-half of its population. When women are constrained from reaching their full potential, that potential is lost to society as a whole. Programme design and implementation should endeavour to address either or both of these factors.[121]

These principles are codified at the international level in, for example, the UN Convention on the Elimination of Discrimination Against Women, and in European Union law.[122]

The goal to "[*de*]limit the opportunities of one-half of [the] population," as well as the utility of sex conscious approaches to achieving it, have been embraced in the domestic context as well. In the United States, both the means and the ends are supported by a constitutional standard that recognizes the fact of sex differences and the distinction between sex and sex stereotype. The now-classic

---

119. *Frequently Asked Questions About Gender Equality* (FAQs), UNITED NATIONS POPULATION FUND, https://www.unfpa.org/resources/frequently-asked-questions-about-gender-equality (last visited Jan. 28, 2020) [hereinafter *Gender Equality FAQs*] .

120. Mona Lena Krook & Diana Z. O'Brien, *The Politics of Group Representation: Quotas for Women and Minorities Worldwide*, 42 COMP. POL. 253 (2010); Stephanie M. Wildman, *Affirmative Action: Necessary for Equality for All Women*, 12 BERKELEY LA RAZA L.J. 429 (2000); Laurel Wamsley, *California Becomes 1st State To Require Women On Corporate Boards*, NPR (Oct. 1, 2018), https://www.npr.org/2018/10/01/653318005/california-becomes-1st-state-to-require-women-on-corporate-boards.

121. *Gender Equality FAQs*, *supra* note 119.

122. *See* Coleman, *Sex in Sport*, *supra* note 27, at 68, n.17.

articulation of this standard comes from Justice Ruth Bader Ginsburg's majority opinion in *United States v. Virginia*:[123]

While the law cannot "rely on overbroad generalizations about the different talents, capacities, or preferences of males and females," it "does *not* make sex a proscribed classification."[124] Indeed, it recognizes that "[p]hysical differences between men and women . . . are enduring" and that these "'[i]nherent differences' . . . remain a cause for celebration, but not for denigration of the members of either sex or for artificial constraints on an individual's opportunity."[125] As applied, this means is that "[s]ex classifications *may* be used to compensate women 'for particular economic disabilities [they have] suffered,' to 'promot[e] equal employment opportunity,' [and] to advance full development of the talent and capacities of our Nation's people";[126] but, they *may not* be used to "den[y] to women, simply because they are women, full citizenship stature—equal opportunity to aspire, achieve, participate in and contribute to society based on their individual talents and capacities."[127]

2. Sex Equality in Sport is a High-Value Social Good

Sport is one of the areas in which the sex equality project has particularly thrived. The goals of competitive sport are "to showcase the best athletes, to produce related goods for stakeholders, and to use sport as a means to spread certain values throughout society. In all three respects, sport seek specifically to reverse society's traditional subordination of women by providing them with opportunities for equal treatment and empowerment."[128]

The Olympic Movement supports sex equality in international sport by setting aside separate competitive opportunities for males and for females. For example:

FIFA sponsors a football (soccer) World Cup for men and a World Cup for women so that both sexes have the chance to field teams, to experience high level international competition, and to be regional and world champions in the sport. Through the production of its various events, FIFA is able to promote the sport and to express its emerging commitment to sex equality—or at least to the value of showcasing empowered females.[129] FIFA President Gianni Infantino explained the impact of the 2019 World Cup on the organization this way:

---

123. U.S. v. Virginia, 518 U.S. 533 (1996).

124. *Id*. at 533 (emphasis added).

125. *Id.* (distinguishing the "[s]upposed 'inherent differences'" between the races and national orgins and adding that "'the two sexes are not fungible; a community made up exclusively of one [sex] is different from a community composed of both'"). *See also id*. at 532 ("Without equating gender classifications, for all purposes, to classifications based on race or national origin, the Court . . . has carefully inspected official action that closes the door or denies opportunity to women (or to men).").

126. *Id*. at 533 (emphasis added).

127. *Id.* at 516 (emphasis added).

128. Coleman, *Sex in Sport*, *supra* note 27, at 85.

129. *FIFA Women's World Cup France 2019 Watched by More Than 1 Billion*, FIFA.COM (Oct. 18, 2019), https://www.fifa.com/womensworldcup/news/fifa-women-s-world-cup-2019tm-watched-by-morethan-1-billion [hereinafter *FIFA Women's World Cup France 2019*]. *See* Andreas Themistokleous, *The Need*

> More than a sporting event, the FIFA Women's World Cup 2019 was a cultural phenomenon attracting more media attention than ever before and providing a platform for women's football to flourish in the spotlight. The fact that we broke the 1 billion target just shows the pulling power of the women's game and the fact that, if we promote and broadcast world-class football widely, whether it's played by men or women, the fans will always want to watch[.][130]

The official blog of the U.S. Department of State added that:

> Every four years, the FIFA Women's World Cup brings the participation and empowerment of women through sports to the international stage and reminds us of the essential contributions of women to societies around the world. From the first tournament held in China twenty eight years ago to France today, the arena of the Women's World Cup not only continues to inspire, but also demonstrates the progress that has been made through the leadership of female athletes, role models and their supporters on gender equality. A key priority for the State Department is to support women and girls' empowerment across economic, political, and social spheres. One of the ways we achieve this is through sports diplomacy . . . as the U.S. Women's National Team holds the Women's World Cup Champion title, the positive impacts of Title IX continue to be felt at home and abroad.[131]

Toward these same combined ends, the sport of athletics (track and field) has men and women competing separately but at the same event in all of the same disciplines and arenas.[132] From the international federation—formerly the IAAF now World Athletics—pay is also equal. For example, a world record at the World Championships is compensated at the rate of $100,000, whether it is set by a male or a female.[133] At the 2019 World Championships, Dalilah Muhammad was the

---

for Female Role Models in Sport, MONEY SMART ATHLETE BLOG (Mar. 13, 2019), http://moneysmartathlete.com/2019/03/13/the-need-for-female-role-models-in-sports/ (discussing research on the broader effects of seeing strong female athletes and of the related Irish campaign, "If she can't see it, she can't be it").

130. *FIFA Women's World Cup France 2019*, *supra* note 129.

131. Erin Brown, *The 2019 FIFA Women's World Cup: A Women's Team That Dares To Shine*, DIPNOTE: U.S. DEP'T OF STATE OFFICIAL BLOG (July 8, 2019), https://blogs.state.gov/stories/2019/07/08/en/2019-fifa-women-s-world-cup-women-s-team-dares-shine.

132. Eddie Pells & Pat Graham, *Felix, Other Top Stars, Fight Track's Pregnancy Penalty*, ASSOCIATED PRESS (Sept. 28, 2019), https://www.apnews.com/80b9e2db9a614cef99fadc5f7f0f8902 (noting that some sponsors have subjected female track and field athletes to a "pregnancy penalty", but that exposure in competition is equal because "Diamond League meets have just as many female events as male events" and because "[t]he women's side of the sport has long produced as much talent and star power as the men"). *See also* Peter Bodo, *Follow the Money: How the Pay Gap in Grand Slam Tennis Finally Closed*, ESPN.COM (Sept. 6, 2018), https://www.espn.com/tennis/story/_/id/24599816/us-open-follow-money-how-pay-gap-grand-slam-tennis-closed (noting that while pay is not equal across all tennis events, "the equivalent prize money that men and women receive at Grand Slam events still puts tennis ahead of other leagues and associations in terms of equality").

133. *See, e.g.*, Press Release, IAAF, TDK and QNB to Support World Record Program in Doha (Sept. 17, 2019), https://www.iaaf.org/news/press-release/world-championships-doha-2019-record-program m. In 2019 in Doha, Dalilah Muhammad of Team USA and Team USA's mixed 4x400m relay were paid under this program, which has been the IAAF standard since the 2009. *See IAAF $100,000 IAAF*

only individual athlete to earn the bonus, for her world record in the women's 400 meters hurdles;[134] and for her extraordinary achievements she was subsequently celebrated, alongside Eliud Kipchoge, who broke the two-hour barrier in the marathon, as the sport's Athlete of the Year.[135]

Sex equality in international sport remains an aspiration, not a perfected goal. There is no doubt, however, that we have come a long way toward parity in this setting even as it remains elusive in different institutional contexts. In important part, this is because the sex affirmative measures in use in sport—including sex segregation and a quota system which ensure an equal number of spots on teams, in finals, and on podiums—have not been embraced elsewhere.

Females have most of the same matching opportunities domestically that exist internationally, not only to participate but also to make those teams, finals, and podiums; but here, as the State Department's statement following Team USA's victory in the FIFA Women's World Cup suggests, this is primarily the result of Title IX's equality mandate.[136] In the United States, most sport and athlete development, and most competitions, take place under the auspices of secondary and post-secondary educational institutions and affiliated sports organizations. This includes state interscholastic athletic associations and the NCAA. These institutions and organizations embrace sex equality not only because they have to as recipients of federal funds and actors in interstate commerce, but also because they understand the important social value that is created when opportunities for participation and competition are distributed not only to boys and men but also to women and girls. The set of synergistic goods that flow from sport to individual, institutional, and community stakeholders is believed to be worth the investment.

Thus, starting in high school if not already in middle school, educational institutions and affiliated organizations support separate local, state, regional, and national competitions for males and females which are designed—like elite sport—to isolate and celebrate the champions. They also support a combination of co-ed and sex-segregated opportunities for participation, the latter as pathways to developing competitive athletes and to inculcating the values of fitness and athleticism for lifelong health and wellness. The opportunity to be engaged in competitive sport in particular—regardless of the level at which the competition occurs—is understood to impart additional socially valuable traits including teamwork, sportsmanship, and leadership, as well as individually valuable traits including goal setting, time management, perseverance, discipline, and grit.[137]

World Record Programme Supported by TDK and Toyota – Berlin 2009, WORLD ATHLETICS (Aug. 12, 2009), https://www.iaaf.org/news/news/100000-iaaf-world-record-programme-supported.

134. Scott Cacciola, *Dalilah Muhammad Breaks Her Own World Record in the 400-Meter Hurdles*, N.Y. TIMES (Oct. 4, 2019), https://www.nytimes.com/2019/10/04/sports/dalilah-muhammad-world-record-hurdles.html.

135. OlympicTalk, *Dalilah Muhammad, Eliud Kipchoge Named World Athletes of the Year*, NBC SPORTS (Nov. 23, 2019, 3:43 PM), https://olympics.nbcsports.com/2019/11/23/eliud-kipchoge-dalilah-muhammad-world-athletics-athlete-year/.

136. *See supra* note 131 and accompanying text (quoting that statement).

137. Coleman, *Sex in Sport*, *supra* note 27, at 95–96.

As in the international context, opportunities for participation and competition are still not equal;[138] and compared to boys, girls have "disproportionate drop-out rates . . . which [are] heightened as girls transition from childhood to early adolescence."[139] Nonetheless, because of Title IX and its sports exception, it is no longer just boys and men who have the opportunity to experience the sense of strength and power—both physical and mental—that comes from consistent training and competition; the proverbial "thrill of victory" and "agony of defeat"; the notion of failure as opportunity; and those "fourth and goal" high intensity, high impact moments when the team is counting on the individual to be at their best not only for themselves but also for the collective endeavor. Because of Title IX and its sports exception, it is no longer just boys and men who experience being celebrated as champions in their communities, who are recruited to compete in college, and who provide the optics necessary for those who look like them and come from their circumstances realistically to dream of following in their footsteps. The latter point about optics is controversial in some circles because it is focused on the female phenotype.[140] But as the pervasive

138. Congress continues to require institutions to produce an annual accounting of their efforts toward Title IX's equality mandate. *See supra* notes 64 and 68 and accompanying text (discussing the Equity in Athletics Disclosure Act). And, litigation is ongoing to ensure that they are held accountable also at the local level. *See, e.g.*, Portz v. St. Cloud State Univ., No. CV 16-1115 (JRT/LIB), 2019 WL 6727122 (D. Minn. Dec. 11, 2019) (denying defendants' motion to stay injunction requiring schools to "take immediate steps to provide its female students with an equitable opportunity to participate in varsity intercollegiate athletics and with an equitable athletic-related treatment and benefits at every tier of its athletic department"; injunction was based on a finding that the schools "did not comply with Title IX in its allocation of athletic participation opportunities and treatment and benefits and had not since at least 2014); Robb v. Lock Haven Univ. of Pa., No. 4:17-CV-00964, 2019 WL 2005636 (M.D. Pa. May 7, 2019) (denying summary judgment because facts are in dispute on the issue whether the university's plans to eliminate its women's varsity swim team and demote its women's varsity field hockey team from Division I to Division II discriminated against female student athletes in violation of Title IX's requirement that schools provide "equivalently advanced competitive opportunities"); D.M. by Bao Xiong v. Minn. State High Sch. League, 335 F. Supp. 3d 1136, 1139–40 (D. Minn. 2018*), rev'd and remanded*, 917 F.3d 994 (8th Cir. 2019) (rejecting boy's challenge to girls'-only competitive dance team on the grounds that the exclusion is designed to delimit competitive opportunities for girls in state's high school sports space: "[I]t is not unfair discriminatory practice to restrict membership on an athletic team to participants of one sex whose overall athletic opportunities have previously been limited").

139. NICOLE ZARRETT, ET AL., WOMEN'S SPORTS FOUND., COACHING THROUGH A GENDER LENS: MAXIMIZING GIRLS' PLAY AND POTENTIAL, EXECUTIVE SUMMARY 1 (2019), http://www.womensportsfoundation.org/wp-content/uploads/2019/04/coaching-through-a-gender-lens-executive-summary-web-1.pdf ; Laura Mallonee, *The Importance of Photographing Women in Sports*, WIRED (June 26, 2019, 4:56 PM), https://www.wired.com/story/female-hockey-players-photo-gallery/ ("By age 14, girls drop out of sports at a rate nearly twice that of boys, due to lack of access, social stigma, and other inequities.").

140. *See* Coleman, *Sex in Sport*, *supra* note 27, at 91–93 (discussing the controversy over the optics of the female body).

mantra "If she can't see it, she can't be it" suggests,[141] it is both well-studied and widely-embraced.[142]

Finally, it is because of Title IX and its sports exception—together with the fight that Title IX advocates have put to those who would impede the project—that we have now experienced four generations of empowered little girls becoming empowered women. According to Donna de Varona and Beth Brooke-Marciniak, "Girls who play sport stay in school longer, suffer fewer health problems, enter the labor force at higher rates, and are more likely to land better jobs. They are also more likely to lead. [Ernst & Young] research shows stunningly that 94% percent of women C-Suite executives today played sport, and over half played at a university level."[143] Although we have not been able to verify their particular numbers,[144] and education-based sport is certainly not the only path to

141. *See, e.g.*, Themistokleous, *supra* note 129 (discussing the use of this phrase by the Federation of Irish Sport); Melody Glenn, *If She Can't See It, She Can't Be It: Part 1*, FEMINEM (May 25, 2017), https://feminem.org/2017/05/25/cant-see-cant-part-1/ (using this phrase in support of female role models in emergency medicine). The phrase is a variant of others such as "you can't be what you can't see" that have also been applied to the cause of women's equality and the use of female role models in that context. *See, e.g.*, Tasnuva Bindi, *If You Can't See It, You Can't Be It: Female Founders Crushing Stereotypes*, STARTUP DAILY (May 13, 2014), https://www.startupdaily.net/2014/05/cant-see-cant/ (using this phrase in support of female role models in business, tech, and politics).

142. It is beyond legitimate dispute that role models are effective when their observable characteristics and trajectories are relevant to the aspirant. This is why, for example, we say that girls need to see women in positions of authority and that children of color need to see people of color in those same positions. What we can see—the optics—matter. And the point of reference is the viewer or aspirant, not the individual who would self-identify as a role model. *See* Coleman, *Sex in Sport*, *supra* note 27, at n.256; ZARRETT, ET AL., *supra* note 139, at 3 (listing "female coaches" as one of the things that can encourage girls to stay engaged in sport once they have chosen to participate: "Girls more readily identify with and see a female coach as a mentor and role model, which in turn, can help counter stereotypes and boost girls' confidence, self-efficacy, and sense of belonging."). *See also* Coleman, *Semenya and ASA v. IAAF*, *supra* note 49 ("It is well understood that the empowerment effects of [sex segregated female sport] are different from those that result from seeing men compete together, and also different from seeing open competition among men and women.").

143. Beth A. Brooke-Marciniak & Donna de Varona, *Amazing Things Happen When You Give Female Athletes the Same Funding as Men*, WORLD ECON. FORUM (Aug. 25, 2016), https://www.weforum.org/agenda/2016/08/sustaining-the-olympic-legacy-women-sports-and-public-policy/.

144. *See also* Rebecca Hinds, *The 1 Trait 94 Percent of C-Suite Women Share (And How to Get It)*, INC. MAG. (Feb. 8, 2018), https://www.inc.com/rebecca-hinds/the-1-trait-94-percent-of-c-suite-women-share-and-how-to-get-it.html (citing these EY statistics and noting that being "former or current athletes" is "a trait that Meg Whitman, Indra Noovi, Marissa Mayer, and many other top female executives possess"); Monica Miller, *4 Female C-Suite Executives Who Played College Sports*, NCAA AFTER THE GAME (Mar. 8, 2018), http://www.ncaa.org/student-athletes/former-student-athlete/4-female-c-suite-executives-who-played-college-sports (noting that being a former athlete is a characteristic top female executives share); Valentina Zarya, *What Do 65% of the Most Powerful Women Have in Common? Sports*, FORTUNE (Sept. 22, 2017), https://fortune.com/2017/09/22/powerful-women-business-sports/ (same). We were especially interested in the study design that resulted in the 94 percent figure and so sought to establish how the company that ran the survey—Longitude—identified its recipients. A representative from the company explained that "[a]t the time of recruiting for a particular study, our vendors reach out to the general survey audience and screen respondents according to the survey requirements. Due to the nature of our work, we therefore deploy purposive sampling methods where we purposely target a specific type of audience to eventually qualify the right people." E-mail from

the C-Suite,[145] its multiple health, welfare, competence, and confidence effects have been well-understood for decades.[146] Indeed, as the award-winning sports reporter Christine Brennan offered in the wake of Team USA's victory at the 2019 FIFA Women's World Cup:

> This is a watershed moment. The 1999 U.S. Women's World Cup victory was a revelation. Back then, the nation fell in love with what it created with Title IX. But this, the 2019 U.S. Women's World Cup victory—this is an affirmation. This is the nation saying, 'We want to see more of this.' We've been watching these little girls running to practice every day in our neighborhoods for a couple of decades. They grow up and this is what happens. They become strong, powerful, fearless women who can do anything. The success of the 2019 U.S. soccer team is set against the backdrop of more than 100 women in Congress and 25 women in the Senate. This is that conversation, #MeToo, women speaking out, equal pay, it's all wrapped in one.[147]

The women on this and other teams stand on the shoulders of their predecessors and all of them started in school.[148]

3. Sex Segregation is Necessary to Protect This Good

If schools could not "carve out an exclusive [girls' and] women's category defined by sex, females and their associates would be excluded from the most important of the[] benefits" that flow from participation in competitive sport.[149] As we explain in Part II, this is because females as a group are not competitive with males as a group beginning from the onset of male puberty. Starting then,

---

Ali Syed, Research Operations Manager, Longitude to Eugene Volokh, Gary T. Shchwartz, Distinguished Professor of Law, UCLA Law Sch. (May 20, 2019, 8:40 AM) (on file with authors). We were not provided with further detail about how the company's purposive approach may have influenced who received and responded to the survey. The approach may have caused the number to be higher than it would have been otherwise.

145. Sport is one way that individuals can gain the set of traits that are commonly associated with success. It is ultimately that set of traits that is valuable to employers. *See, e.g.*, Christina DesMarais, *7 Reasons Athletes Make the Best Employees,* INC. MAG. (Nov. 22, 2017), https://www.inc.com/christina-desmarais/heres-why-kids-who-play-sports-do-better-in-life.html?cid=search (describing the traits are commonly associated with athleticism); Coleman, *Sex in Sport*, *supra* note 27, at 96 (noting that the traits athletes develop "are socially valuable in part because they are highly transferrable . . . which is why 'executives like to hire athletes'"); Nanette Fondas, *Research: More Than Half of Top Female Execs Were College Athletes*, HARV. BUS. REV. (Oct. 9, 2014), https://hbr.org/2014/10/research-more-than-half-of-female-execs-were-college-athletes (same).

146. *See* Coleman, *Sex in Sport*, *supra* note 27, at 95–96; Donna Lopiano, *Modern History of Women in Sports: Twenty-five Years of Title IX*, 19 CLIN. SPORTS MED. 163, 163–73 (2000).

147. E-mail from Christine Brennan, Sports Columnist, CNN, to Doriane Lambelet Coleman, Professor of Law, Duke Law Sch. (Aug. 17, 2019, 10:14 AM) (on file with authors) (discussing CNN broadcast on day of Team USAs return from France to the parade in New York City).

148. *Women's Interest in Sport Continues to Grow,* IBERDROLA, https://www.iberdrola.com/about-us/womens-sport/other-sports/women-sport-today (last visited Jan. 28, 2020) (detailing "the influence of participation in sports at school [beginning in the 1970s] on women's sustained interest in sports now").

149. Coleman, *Sex in Sport*, *supra* note 27, at 96.

even the very best females are surpassed by second-tier males, and second-tier females have no realistic chance to be anything but early participants in the game.

The case for re-affirming the sports exception is based in the goods produced by girls' and women's sport and in the causal link between sex segregation and those goods. The more specific case for not including—or for conditioning the inclusion of—transgender women and girls in girls' and women's sport is related: If they haven't been on feminizing hormones for a relevant period of time,[150] trans women and girls remain fully male-bodied in the respects that matter for sport; because of this, their inclusion effectively de-segregates the teams and events they join.[151] Beyond this basic structural point is the fact that if they are just decent athletes, they will displace females who are the classification's raison d'être,[152] including in championship positions.[153] This matters for the individual females who are displaced, for those who would aspire to be champions, and for the broader expressive effects we expect from the classification. Even an exception risks swallowing the rule and defeating the category.

Finally, the position that there is no legally cognizable difference between females and trans women and girls destroys the legal basis for separate sex sport.[154] This position—encapsulated in the movement mantra, "Girls who are transgender are girls. Period."[155]—is presumably designed to erase sex-linked traits from consideration in the analysis whether the two groups are similarly situated for purposes of equal protection doctrine. If we are not permitted legally to notice that girls who are female and girls who are transgender are dissimilarly situated with respect to their anatomy and physiology, and if we are not permitted to distinguish among them in circumstances where sex actually matters, we will have dismantled the legal scaffolding that supports separate sex sport. Unlike restrooms, which are segregated for safety and privacy, sport does not have an argument that it needs to separate girls from boys, men from women, for any reason other than sex.[156] And, sex discrimination, including sex segregation, is only lawful if it is necessary.

---

150. *Feminizing Hormone Therapy*, MAYO CLINIC, https://www.mayoclinic.org/tests-procedures/ mt f-hormone-therapy/about/pac-20385096 (last visited Jan. 28, 2020).

151. *See supra* notes 91–102 and accompanying text (summarizing the effects of male puberty on the body).

152. *See supra* notes 81–88 and accompanying text (explaining that even second-tier males routinely surpass not only second-tier females but also the very best elite females). We don't separate men from women, girls from boys, in competitive sport because they have a different gender identity; we separate them because they have different sex-linked anatomy and physiology. *See supra* and *infra* notes 111–114 and 218 and accompanying text (elaborating on this point).

153. *See infra* notes 181–182 and accompanying text (noting recent victories by trans women athletes on and not on hormones).

154. *See supra* notes 77–78 and accompanying text (describing this legal point).

155. *See, e.g., Support Trans Student Athletes*, ACLU, https://action.aclu.org/petition/support-trans-student-athletes?ms_aff=NAT&initms_aff=NAT&ms=190726_lgbtrights_transathletepledge&initms=190726_lgbtrights_transathletepledge&ms_chan=tw&initms_chan=tw&redirect=transathletesbelong (last visited Jan. 28, 2020).

156. For additional analysis of the difference between sport and restrooms, see *supra* notes 72–74 and accompanying text.

B. The Affirmative Case for Inclusion on the Basis of Gender Identity

There are important arguments on the other side. From our perspective, this is the most persuasive case:

A just and ethical society aspires to be inclusive of and to secure equal protection for all of its citizens, especially for its most vulnerable. Because schools are one of the settings in which children learn societal values, a just and ethical society inculcates inclusivity through its educational programming, and then ensures that all students have equal access to high value spaces and opportunities. Education-based sports are among the high value spaces and opportunities schools provide. Eligibility rules for teams and events that sort students based on biological sex may have the effect of excluding those who are transgender. (This effect exists when individual transgender students are sufficiently uncomfortable with the barrier to entry that they choose to exclude themselves.) When this happens, they are denied equal access to sports. It also denies their schools and athletic associations the ability to pursue their pedagogical goals and to perfect themselves as just and ethical institutions. Finally, where a particular student is especially vulnerable, schools and organizations acting in quasi loco parentis are denied the means to secure their health and welfare.

1. Equality and Inclusivity as Hallmarks of a Just and Ethical Society

A just and ethical society aspires to be inclusive of and to secure equal protection for its most vulnerable citizens. Policies that exclude or deny them equal access are flawed as a matter of principle because they have these effects, and because they impede the perfection of a virtuous society. Where the policies are otherwise valuable, a just and ethical society should provide for exceptions. Where their value is in doubt and cannot be established, they should be dismantled. The goal should be to acknowledge everyone's humanity, to practice generosity, and to make room at the table for everyone. It is often difficult but especially important to do so in the face of incomprehensible or inexperienced difference.

In the international context, "social integration to create an inclusive society … [is] one of the key goals of social development."[157] For example, the United Nations Department of Economic and Social Affairs understands social integration to be "a dynamic and principled process of promoting the values, relations and institutions that enable all people to participate in social, economic, cultural, and political life on the basis of equality of rights, equity and dignity."[158] Social inclusion "is understood as a process by which efforts are made to ensure equal opportunities for all, regardless of their background, so that they can achieve their full potential in life."[159]

157. UNITED NATIONS DEP'T OF ECON. & SOC. AFFAIRS, CREATING AN INCLUSIVE SOCIETY: PRACTICAL STRATEGIES TO PROMOTE SOCIAL INTEGRATION 4 (2009) (draft), https://www.un.org/esa/socdev/egms/docs/2009/Ghana/inclusive-society.pdf.

158. *Id.* at 3.

159. *Id.*

In the domestic context, these same inclusion and equal access goals motivated the creation of the Civil Rights Division of the Department of Justice in 1957, the passage of the 1964 Civil Rights Act, and, among other subsequent civil rights legislation, the Americans with Disabilities Act. Throughout, the idea has been to re-define "We the people" in our Constitution to include all of us within its protections. And then, through the law and the social movements that constantly re-work its application, to evolve the culture and social norms also to be inclusive of and even generous toward those who were previously excluded. As we do, we not only make room at the table but we also perfect our society.

In both the international and domestic contexts, civil and human rights advocates have called on states and institutions to make room at the table for trans people specifically, by protecting them from discrimination and securing their full social integration according to these principles. This advocacy

> [s]eeks to persuade institutional decisionmakers to develop policies designed to recognize, normalize, include, and empower . . . trans people who throughout history have been erased or severely marginalized and often subject to violence . . . [T]he trans community . . . is . . . particularly "associated with high levels of stigmatization, discrimination and victimization, contributing to negative self-image and increased rates of other mental disorder." For example, in the United States, "[t]ransgender individuals are at a higher risk of victimization and hate crimes than the general public" and "[a]dolescents and adults with gender dysphoria are at increased risk for suicide." In less tolerant parts of the world, trans people are at even greater risk of violence, social isolation, and reduced life span.[160]

We are in the midst of this particular social movement, which has garnered important support. In the United States, the 116th Congress passed H.R. 5 - The Equality Act in 2019, which re-defines "sex" in federal civil rights law to include "gender identity."[161] This move was designed to make it unlawful to discriminate against individuals based on their gender identity; among other things, it would disallow distinctions among people on the basis of sex. Because it faces significant opposition in the Republican-controlled Senate and President Trump likely would not sign it, it probably will not become law; but the vote in the House of Representative expresses an important social and political viewpoint. Also in 2019, a consortium of the most important human rights organizations within the United Nations signed a joint statement "call[ing] on States [and other stakeholders] to act urgently to end [among other things] . . . discrimination against . . . transgender and intersex . . . adults, adolescents and children."[162] Like H.R. 5, this statement is

---

160. Coleman, *Sex in Sport*, *supra* note 27, at 102. For example, in the United States, the Human Rights Campaign "envisions a world where LGBTQ people are ensured of their basic rights, and can be open, honest and safe at home, at work and in the community." *HRC's Mission Statement*, HUMAN RIGHTS CAMPAIGN, https://www.hrc.org/hrc-story/mission-statement (last visited Jan. 28, 2020). Its Transgender Equality Council advocates specifically for full inclusion and equality for "transgender children and gender expansive youth." *Explore: Transgender Children & Youth*, HUMAN RIGHTS CAMPAIGN, https://www.hrc.org/explore/topic/transgender-children-youth (last visited Jan. 28, 2020).

161. Equality Act, H.R. 5, 116th Cong. (2019).

162. *Ending Violence and Discrimination Against Lesbian, Gay, Bisexual, Transgender and Intersex People*, UNITED NATIONS (Sept. 2015), https://www.ohchr.org/Document/Issues/Discrimination/

not itself formal law, but it nevertheless demonstrates important support for the cause.

2. The Mission of Schools in a Just and Ethical Society Includes Providing Equal Access and Inculcating and Expressing Inclusivity

Part of the mission of educational institutions in any society is to inculcate and express community values. The extent to which they do depends on whether the institutions are public or private, and also on the degree of consensus within the community about what those values are. Where there is ideological homogeneity, values are most likely to be inculcated through the schools. Where there is ideological heterogeneity, this is less likely.

In a just and ethical society, there is or should be a high degree of consensus that inclusivity and equality are among the most important societal values. Educational institutions within such a society are likely to be permitted or even required by the government and citizenry to inculcate and express both. Doing this successfully means ensuring that school programming is accessible to all students. This is especially important with respect to high value spaces and opportunities which are most likely to be arbitrarily exclusive if they are not carefully monitored.

Education-based sport, including competitive sport, is understood to be a high value space and opportunity. This is because of the direct physical and health benefits it yields, and because in elementary and secondary school sports are also—if not mostly—a co-curricular social space where students learn to interact successfully with their peers. Because of this, all students should have access to school sports and related structures should be designed to make this possible.

It is noteworthy that school sports programming has traditionally provided the basis for inculcating inclusivity. While this is especially evident in the context of noncompetitive games, it is also apparent in lower level competitive ones. "Everyone can play" policies, rotation and substitution practices, and the selection of teams balanced by ability are all ways in which the existing practices of competitive education-based sport express and inculcate inclusivity.

Providing equality of opportunity for transgender students is consistent with this approach and the values that drive it. Because they may be excluded in effect by programming that sorts students according to their sex, it is arguably incumbent on schools and athletic organizations either to sort students differently or to accommodate them within existing structures according to their gender identity. Like other students, transgender students should have the benefit of the positive social, health, and empowerment effects of school sports. Rashaan

---

Joint_LGBTI_Statement_ENG.PDF (adopted by the International Labour Organization, Office of the High Commissioner of the United Nations Human Rights Council, UNDP, UNESCO, UNFPA, UNHCR, UNICEF, UNODC, UN Women, World Food Programme, World Health Organization, and UNAIDS); *see also* UNITED NATIONS, THE ROLE OF THE UNITED NATIONS IN COMBATTING DISCRIMINATION AND VIOLENCE AGAINST LESBIAN, GAY, BISEXUAL, TRANSGENDER AND INTERSEX PEOPLE: A PROGRAMMATIC OVERVIEW (2019), https://www.ohchr.org/Documents/Issues/Discrimination/LGBT/UN_LGBTI_summary_2019.pdf.

Yearwood, an educator who is also the father of a transgender girl, expresses the point this way:

> My daughter is a trans-female. That means she needs to compete on girls' teams in order to feel most comfortable. [This isn't about competition for her.] She's running because she wants to be part of a team at this age. You know, comradery. Perseverance, grit, teamwork. My job is to raise a healthy child. And we all know that being part of groups and being included allows students to develop in a healthier way than when you're excluded."[163]

Because all students gain from exposure to important social values, including transgender children in school activities like sport according to their gender identity separately supports the institutions' broader pedagogical goals.

3. Schools Need to be Able to Take Care of the Especially Vulnerable Child

Schools stand in quasi loco parentis. Although they are not formally in the shoes of parents, they do have physical custody of the children during the school day while they are on campus, and their charge is not only to ensure that the children are safe but also that they are prepared to engage and to learn. Where an individual child is especially vulnerable—where they might not be safe or when the environment is such that they might be unable successfully to participate—schools should have the tools to address their circumstances.

This is a commonplace. When a child has a peanut allergy, we say that other children cannot bring peanut butter to school. When a child becomes ill, it is or should be "all hands on deck." The interests of the at-risk child are understood to outweigh the interests of others in their teachers' focused attention or in their lunch preferences or even their nutritional needs. This balancing analysis does not always come out in the vulnerable child's favor—for example when they are disruptive of the educational process, when they risk the health and welfare of other children, or when the school is not and cannot be equipped to handle their special needs. But it should come out in their favor when the interests on the other side are not so significant.

Not all transgender children are struggling and fragile. Being trans does not mean having dysphoria.[164] But many are and do. *Individual* transgender students may be struggling and fragile because of the discordance between their sex and their gender identity, because they are in the process of transitioning socially and maybe also physically, and because of the ways in which others perceive and treat them:

> The disconnect between their experienced gender and their assigned gender can result in acute stress called gender dysphoria. Gender dysphoria can be a source of profound suffering. A recent study of transgender teens found that more than 50 percent of transgender males and almost 30 percent of transgender females reported attempting suicide. Transgender adolescents are often vulnerable to

163. Nia Hamm, *Father of Transgender Student Athlete Pushes Back Against Petition to Change Competition Policy*, FOX 61 (June 26, 2018), https://fox61.com/2018/06/26/father-of-transgender-student-athlete-pushes-back-against-petition-to-change-competition-policy/.

164. *See* Safer & Tangpricha, *supra* note 31, at 2451 (explaining this point).

> bullying and family rejection. And even when families are supportive, it can be a very difficult transition for both the teen and the parents.[165]

The medical standard of care for trans and gender diverse children includes living and having others treat them in accordance with their gender identity.[166] Although medicine can't prescribe other-than-medical policy, schools should take that standard under advisement as they develop their policies; and they should follow it in individual instances when the benefits of doing so outweigh the costs. As applied, taking care of transgender children means including them in sex segregated spaces and opportunities like sport on the basis of their gender identity. Where the *individual* child is especially vulnerable, supporting their successful social transition is arguably more important than the integrity of the sex-segregated competitions they would enter. If the student is a transgender girl, supporting her is arguably more important than the interests of all female students in winning.

This last argument is often coupled with evidence that transgender teens are at an especially high risk of suicide. For example, Helen Carroll of the National Center for Lesbian Rights explains that "[s]port can be life-saver for transgender people, who are at high risk of suicide . . . 'They've been fighting themselves and feeling like they're in the wrong body, and sport gives them a place to be happy about their body and what it can do.'"[167] Carroll is right on the facts; an extraordinarily disturbing 30 to 50 percent of transgender teens report attempting suicide.[168] Carroll herself continues to be invaluable to all stakeholders in her support for trans athletes within existing structures.[169] But of course her point is

---

165. Caroline Miller, *Transgender Kids and Gender Dysphoria*, CHILD MIND INST., https://childmind.org/article/transgender-teens-gender-dysphoria/ (last visited Jan. 28, 2020). *See also* Doriane Lambelet Coleman, *Transgender Children, Puberty Blockers, and the Law: Solutions to the Problem of Dissenting Parents*, 19 AM. J. BIOETHICS 82 (2019) (addressing the question whether and how the law is able to assist transgender children whose families are not supportive).

166. News Release, Endocrine Soc'y, Endocrine Society Urges Policymakers to Follow Science on Transgender Health (Oct. 29, 2019), https://www.eurekalert.org/pub_releases/2019-10/tes-esu102919.php ("As noted in our evidence-based guideline, transgender individuals, both children and adults, should be encouraged to experience living in the new gender role and assess whether this improves their quality of life."). *See also* Jason Rafferty, *Ensuring Comprehensive Care and Support for Transgender and Gender-Diverse Children and Adolescents*, 142 PEDIATRICS 1 (2018); https://pediatrics.aappublications.org/content/pediatrics/142/4/e20182162.full.pdf (American Academy of Pediatrics does the same.); Joshua Safer & Vin Tangpricha, *Care of the Transgender Patient*, ANNALS OF INTERNAL MED. (July 2, 2019), https://annals.org/aim/article-abstract/2737401/care-transgender-patient (American College of Physicians guidelines support the same.).

167. Christie Aschwanden, *Trans Athletes Are Posting Victories and Shaking Up Sports*, WIRED (Oct. 29, 2019, 12:00 PM), https://www.wired.com/story/the-glorious-victories-of-trans-athletes-are-shaking-up-sports/.

168. Rokia Hassanein, *New Study Reveals Shocking Rates of Attempted Suicide Among Trans Adolescents*, HUMAN RIGHTS CAMPAIGN BLOG (Sept. 12, 2018), https://www.hrc.org/blog/new-study-reveals-shocking-rates-of-attempted-suicide-among-trans-adolescen.

169. *See generally* Cyd Zeigler, *LGBTQ Sports Advocate Helen Carroll Retires From NCLR*, OUTSPORTS (June 1, 2017, 10:27 PM), https://www.outsports.com/2017/6/1/15723406/helen-carroll-nclr-lgbtq-sports-retire and *infra* note 225 and accompanying text (discussing her work with Pat Griffin on the NCAA transgender guidelines).

an application of more general ones, about the empowerment effects of sport for females and its therapeutic effects for those who are suffering as a result of difficult personal and mental health issues whatever their source.[170] It is an unfortunate fact that major depression and associated feelings of hopelessness are on the rise among children in the United States, with girls being especially affected.[171] Suicide is now the second leading cause of death among adolescents in general.[172] Risk factors include "psychiatric disorders and comorbidities, family history of depression or suicide, loss of a parent to death or divorce, physical and/or sexual abuse, lack of a support network, feelings of social isolation, and bullying"[173] as well as "barriers to access treatment, homosexual orientation" and being an "early or late developing girl[]."[174] Educational institutions that have the necessary resources should follow the evidence and do what they can to mitigate these risks regardless of the child at issue.[175]

There is a related, overlapping claim from vulnerability. It is that *as a group*, transgender children are struggling and fragile so that their interests in being included in sex-segregated spaces and opportunities on the basis of their gender identity always trump the interests of classmates who, *as a group*, cannot be described as similarly vulnerable. This is a standard move in advocacy circles which has, in turn, influenced sports policymakers on the ground. For example, the Executive Director of the National Federation of State High School Associations Karissa Niehoff argues that:

> [This] is not about the winning and losing. It's about the successful development of these [transgender] kids. [I]f we don't treat them respectfully, their development is going to lose. That's a much bigger issue than someone not getting a medal or a place in a race. Much bigger issue.[176]

---

170. *See generally* Emily Pluhar et al., *Team Sport Athletes May Be Less Likely To Suffer Anxiety and Depression Than Individual Sport Athletes*, 18 J. SPORTS, SCI., & MED. 490 (2019) (noting positive mental health effects of sports in general while distinguishing results in team and individual sports); Nick Pearce et al., *The Role of Physical Activity and Sport in Mental Health*, FACULTY OF SPORT & EXERCISE MED. UK (May 2018), https://www.fsem.ac.uk/position_statement/the-role-of-physical-activity-and-sport-in-mental-health/.

171. Patti Neighmond, *A Rise In Depression Among Teens and Young Adults Could Be Linked to Social Media Use*, NPR (Mar. 14, 2019, 11:01 AM), https://www.npr.org/sections/health-shots/2019/03/14/703170892/a-rise-in-depression-among-teens-and-young-adults-could-be-linked-to-social-medi.

172. Melonie Heron, *Deaths: Leading Causes for 2017*, 68 NAT'L VITAL STAT. REP. 1 (2019).

173. *Teen Suicide*, AMERICA'S HEALTH RANKINGS, https://ww.americashealthrankings.org/explore/health-of-women-andchildren/measure/teen_suicide/ state/ALL (last visited Jan. 28. 2020).

174. Stephanie Secord Fredrick et al., *Can Social Support Buffer the Association Between Depression and Suicidal Ideation in Adolescent Boys and Girls?*, 55 PSYCHOL. IN THE SCHOOLS 490, 491 (2018).

175. *See generally* CTRS. FOR DISEASE CONTROL & PREVENTION, NAT'L CTR. FOR INJURY PREVENTION & CONTROL, DIV. OF VIOLENCE PREVENTION, THE RELATIONSHIP BETWEEN BULLYING AND SUICIDE: WHAT WE KNOW AND WHAT IT MEANS FOR SCHOOLS (2014), https://www.cdc.gov/violenceprevention/pdf/bullying-suicide-translation-final-a.pdf.

176. Mary Albl, CIAC's *Transgender Policy Faces Test With New Lawsuit*, DYESTAT (June 24, 2019, 1:20 PM), https://www.runnerspace.com/gprofile.php?mgroup_id=44531&do=news&news_id=580238.

Niehoff's statement reflects the sense that the successful development of different kids is not similarly dependent on equality, inclusion, and success in school and sport; and that to the extent it might be, this will either be an only occasional conflict, or else the opportunity to participate will be enough for most or all kids who are not transgender. It also reflects the view that although education-based sport sometimes promotes competition and winning, this is ultimately not its primary institutional focus.

* * *

We close out this section with brief reactions to four arguments that are not persuasive from our perspective. They include some that are particularly prominent in the public relations strategies of the advocacy groups that currently control the megaphone.

The first of these is the argument, grounded in science denial, that we have already addressed in Part II.[177] It is fact, not myth or stereotype, that beginning at the onset of male puberty, an insurmountable performance gap between males and females emerges such that even the very best females are not competitive for the win against males, including against second-tier males. If we care about sex equality in sport, that is, if we care about seeing females in finals and on the podium however they might happen to identify, competitive sport has to be segregated on the basis of sex.

The second is the suggestion that because there are few transgender women and girls relative to the numbers of females, any disruption of the competitive hierarchy is unlikely to be substantial enough to jeopardize sex equality goals.[178] That some transgender women and girls may be on feminizing hormones is said to reduce their potential impact even further. As we explain in Part III(C) below, we agree that a consistent course of hormone replacement therapy (HRT) can wind down the male advantage to the point that a policy exception at the development elite and collegiate levels is justified. But we don't think that transition hormones should be a requirement for participation in secondary school competition; and in any event, many transgender teens do not want or have access to hormones. As students in the latter category are increasingly comfortable coming out at school—which from our perspective is a good thing—the number of "out" trans kids is growing beyond the small percentages described in earlier population surveys. This increase is not yet well understood, but it appears to be an upward trajectory.[179] Because it is well-established that athletic but not necessarily elite

177. *See supra* Part II. *See, e.g.,* Medley & Sherwin, *supra* note 72 (describing as "myth" and "impermissible sex stereotype" the fact of the sex-linked performance gap, and as "arbitrary" the rules of sports governing bodies that use sex-linked traits as the basis for classification into and out of the women's category, and claiming that there is "ample evidence that girls can compete and win against boys").

178. *See, e.g., id.* ("The truth is, transgender women and girls have been competing in sports at all levels for years, and there is no research supporting that they maintain a competitive advantage.").

179. *See* Rafferty, *supra* note 166 (discussing recent statistics); Michelle M. Johns et al., *Transgender Identity and Experiences of Violence Victimization, Substance Use, Suicide Risk, and Sexual Risk Behaviors Among High School Students—19 States and Large Urban School Districts, 2017*, 68 CDC MORBIDITY & MORTALITY WKLY. REP. 67 (reporting a 1.8 percent incidence rate as of 2017); Jesse Singal, *When Children*

males dominate females in almost every sport and event, which is true without regard to how individuals identify, it is reasonable to expect that trans girls not on hormones will affect results in important ways—as athletes with 46,XY differences of sex development have in the international arena.[180] It is not a fluke that, in the last three years, we have seen the first trans girls as state champions,[181] the first trans woman as DII national champion,[182] and the first potentially consequential trans woman in a DI sport.[183]

The third is the political and sociological claim[184] that transgender girls are girls and, because of this, their reproductive sex-linked traits are irrelevant to the conversation about their classification into spaces and opportunities designed specifically to empower females.[185] This claim is weak not only because it has its

---

Say They're Trans, ATLANTIC (July/Aug. 2018), https://www.theatlantic.com/magazine/archive/2018/07/when-a-child-says-shes-trans/561749/ (reporting on rise in numbers of people identifying as trans).

180. *See* Coleman, *Sex in Sport*, *supra* note 27, at 106–108 (responding to the numbers argument in that context).

181. *See 2018 CIAC Spring Championships: Class M Outdoor Track*, *supra* note 104 and accompanying text (discussing Connecticut state championships in track and field).

182. Press Release, Franklin Pierce Univ., NATIONAL CHAMPION! Telfer Claims Women's Track & Field's First NCAA Title (May 26, 2019), https://www.franklinpierce.edu/about/news/National_Champion_CeCe_Telfer.htm.

183. Kyle Hansen, *Montana Cross Country Runner, Belgrade Native to Make History as Transgender Athlete*, MISSOULIAN (Aug. 30, 2019), https://missoulian.com/sports/college/big-sky-conference/university-of-montana/montana-cross-country-runner-belgrade-native-to-make-history-as/article_2a37bd80-9eea-519d-8636-040473b84cc8.html. As of this writing, June Eastwood's performances appear to be consistent with Joanna Harper's hypothesis that, at least for distance runners, a year of consistent use of gender affirming hormones winds down the male advantage to the point that trans women return to their place in the hierarchy, e.g., roughly speaking if they were the tenth best man they will be approximately the tenth best woman. Eastwood's announcement that she was transitioning from male to female, and from the Montana men's team to the Montana women's team, sounded alarm bells within the sport, however, because her pre-transition times in high school and early college, especially in the middle distances, would have immediately put her at or close to the professional women's world records. In this respect she was, and depending on how she runs going forward may still be, the most significant trans woman athlete to date.

184. We characterize this as a political and sociological claim because it is based in an effort to expand the standard definitions of female, girl, and woman beyond their basis in female reproductive sex to include a subset of individuals whose reproductive sex is male. It is one of a number of strategies that might be employed to secure equality for trans people. There is a counterargument that this is a factual claim since gender identity likely has a neurobiological basis and may be related to our natural reproductive inclinations. Even if this is eventually established, however, individuals would still have either male or female reproductive sex, which is what the words and dichotomies male/female, girl/boy, and man/woman are generally understood to connote; and we would still need words that did this descriptive work. The arguments that we wouldn't or shouldn't, or that it shouldn't or couldn't be the ones we use now because they exclude trans people, are undoubtedly political.

185. *See, e.g.*, Medley & Sherwin, *supra* note 72 (rejecting the "policing of gender [that] has been used to justify subjecting transgender student athletes to numerous additional barriers to participating in sports, from onerous medical requirements to segregation in locker rooms to outright bans on their participation"); *Statement of Women's Rights and Gender Justice Organizations*, *supra* note 29 (rejecting any distinctions among cis and trans women and girls on the argument that "Transgender girls are girls

advocates making easily dismissed arguments about testes and T levels being no different than—for example—height and wingspan, but also because it ignores the reasons these spaces and opportunities exist and censors legitimate discussions about the implications of erasing biological sex even where it is undoubtedly relevant. This includes contexts in which sex is outcome determinative, like competitive sport. It also includes contexts in which sex differences are an affirmative individual and societal good. Discussions about sex and sex-linked anatomy and physiology must be had kindly, but it is wrong to censor them.[186]

The fourth is the argument from intersectional feminism that it is good for all women that we accept that transgender women are women, and that transgender girls are girls.[187] Accepting trans women and girls is the good and right thing to do for a lot of reasons, including that our default should be inclusion unless there are persuasive reasons to make distinctions. But the suggestion that women should understand that eliding relevant sex differences is good for them particularly—even if they don't know it—is patronizing and otherwise deeply problematic.[188] From what we have gathered, it appears to be based in a number of different assumptions all of which we reject: that sex classifications are always

---

and transgender women are women. They are not and should not be referred to as boys or men, biological or otherwise").

186. For an analysis of the harm caused by the censorship of females talking about the female body see Doriane Lambelet Coleman, *A Victory for Female Athletes Everywhere*, QUILLETTE (May 3, 2019), https://quillette.com/2019/05/03/a-victory-for-female-athletes-everywhere/.

187. *See, e.g, Statement of Women's Rights and Gender Justice Organizations*, *supra* note 29 ("[W]e . . . reject the suggestion that cisgender women and girls benefit from the exclusion of women and girls who happen to be transgender"; "we recognize the harm to all women and girls that will flow from allowing some women and girls to be denied opportunities to participate and cast out of the category of 'woman' for failing to meet standards driven by stereotypes and fear"; "we speak from experience and expertise when we say that nondiscrimination protections for transgender people—including women and girls who are transgender—are not at odds with women's equality or well-being, but advance them."); *Support Trans Student Athletes*, *supra* note 155 ("The marginalization of trans student-athletes is rooted in the same kind of gender discrimination and stereotyping that has held back cisgender women athletes. Transgender girls are often told that they are not girls (and conversely transgender boys are told they are not really boys) based on inaccurate stereotypes about biology, athleticism, and gender. . . . Girls who are transgender are girls. Period."). *See also, e.g.*, Carol Hay, *Who Counts as a Woman*, N.Y. TIMES (Apr. 1, 2019), https://www.nytimes.com/2019/04/01/opinion/trans-women-feminism.html (noting the origins of this claim in intersectional feminism); Jack Guy, *Women or 'Womxn'? Students Adopt Inclusive Language*, CNN (Nov. 27, 2018, 11:38 AM), https://www.cnn.com/2018/11/27/uk/womxn-inclusive-language-gbr-scli-intl/index.html (describing the "growing use of inclusive language, designed to avoid excluding particular groups of people" including that "Womxn is a more inclusive term which promotes intersectionality" and is thus "more inclusive of all kinds of women, including trans women" and that "Womxn is used to demonstrate our commitment to inclusiveness").

188. Especially patronizing is the suggestion that these organizations have particular "experience and expertise" about women—including about how they should define themselves and their wellbeing—that women themselves don't have. *See, e.g., Statement of Women's Rights and Gender Justice Organizations*, *supra* note 29 ("[W]e speak from experience and expertise when we say that nondiscrimination protections for transgender people—including women and girls who are transgender—are not at odds with women's equality or well-being, but advance them."). It is this kind of talk that has alienated many, including many women, who might otherwise be natural allies.

harmful or at least a net harm to women and girls; that all sex classifications are based in false and damaging sex stereotypes; that women all want or should want to be freed from the yoke that is their secondary sex characteristics and cultural expectations around femininity; that women and girls are by nature or necessity inclusive and self-sacrificing, so that the category that describes them itself should be generously so; and that women and girls—and their allies—don't or shouldn't care as much as boys and men do about competition and being competitive for the win. We reject each of these because they are themselves false and damaging sex stereotypes.

The last of these is especially problematic as we discuss the future of Title IX. It has led a group of prominent civil rights organizations to return to the posture of some of their 1970s counterparts who argued that high-end NCAA-style competitions were not for women;[189] today, they have determined to limit their advocacy to protecting opportunities for women and girls to participate—not necessarily to win—in sport.[190] It is presumably just helpful coincidence and not coordinated strategy that their position aligns with the position of some in the trans advocacy community that because trans women and girls are women and girls, their victories in events in which females retain opportunities to participate should be celebrated, not discredited.[191] This move is particularly insidious as applied to the high school sports space which can be described as relatively unimportant to protect if it's really just about participation.

---

189. *See supra* notes 42 and 48 and accompanying text (describing that earlier position); and Interview with our co-author Donna Lopiano, President & Founder, Sports Mgmt. Res. (Oct. 13, 2019), (noting that in the development of Title IX, "the focus on participation opportunities came first, followed by the focus on competition, because we had to build the ranks of those participating before we could think about competition, but also because early advocates preferred a health-focused, student-led, physical education model that would concentrate on intramural or junior varsity style events, as distinguished from the varsity and NCAA commercial model for competition they rejected").

190. *Id.* (noting that women's organizations that fought for both participation and competition opportunities in earlier periods have decided as a strategic matter to focus their efforts going forward on participation numbers and not to use organizational resources to continue to secure and protect the right of women and girls also to win, i.e., also to spots in finals and on podiums); E-mail from N.F. to M.R. (Mar. 23, 2019, 3:17 PM) (on file with authors) (sharing that "the Title IX advocacy community is in lockstep" in its commitment to the unconditional inclusion of trans kids in high school sports on the basis of their gender identity so that the right of cis-girls to more than just participation in this—as opposed to the college sports—space "is not the battle of [the organizations] at this time"). This position appears to be reflected in the approach of the Connecticut CIAC, among other state athletic associations. *See New Hampshire House Bill 1251*, SAVE WOMEN'S SPORTS (Jan. 15, 2020), https://savewomenssports.com/original-articles/f/new-hampshire-house-bill-1251-hearing (quoting a testimonial from Connecticut mother, Christy Mitchell, "I was astounded to hear from state officials that 'girls have the right to participate not to win.'").

191. *See, e.g.*, Rachel McKinnon, *I Won a World Championship. Some People Aren't Happy.*, N.Y. TIMES (Dec. 5, 2019), https://www.nytimes.com/2019/12/05/opinion/i-won-a-world-championship-some-people-arent-happy.html ("Trans women are women. We are female . . . It is a human right to be able to compete. I will continue to show up. I hope you'll consider cheering."); Dave Zirin, *Transphobia's New Target Is The World of Sports*, *supra* note 72 ("Terry Miller and Andraya Yearwood finished first and second place respectively in the state open indoor-track championships last month. Instead of celebrating one of the great moments in their lives, they were immediately put on the defensive about their right to compete in the first place.").

The proposition that females don't need to be competitive for the win is no longer viable; see the FIFA Women's World Cup.[192] The proposition that showcasing strong female-bodied champions is not a high value social good is no longer viable; see Serena Williams and Allyson Felix.[193] We do not doubt that they are well-intentioned, but coming in 2020, from organizations that otherwise decry sex stereotypes and that previously championed the right of women and girls also to equality of competitive opportunity, they are nothing short of extraordinary. We need to take care of transgender women and girls, but the path should not involve weakening the commitment to females. And here we should be clear: It is weakening the commitment to females and to sex equality to accept that the boys' state championship will probably *always* be won by a male where the girls' state championship will *not always* be won by a female.

C. Updating Title IX for Its Next Half Century

Biological sex matters. It is real, not socially constructed; and it affects peoples' lives, opportunities, and experiences in even the most benign or egalitarian situations and societies. In particular, because of their different reproductive biology and secondary sex characteristics, females in general have different physical capacities and experiences than males. This is true regardless of how they identify.

Sex matters in ways that are empowering and celebrated, and also in ways that are damaging and censured. For females, physical—including sexual—violence perpetrated primarily by males is a destructive commonplace, as are the routine exclusions and subordinations they experience based on their phenotype and their reproductive biology. Sometimes these exclusions and subordinations are driven by false stereotypes. But they may also be driven by a reluctance to re-imagine structures to accommodate real and otherwise appreciated sex-linked physical differences. Regardless, sex is relevant if not defining.

Sex equality also matters. Although we can imagine a world in which sex is not relevant or defining, in which we classify people on entirely different terms or else not at all, this isn't ours. As ours exists, because sex matters, so does sex equality. The United Nations is not wrong to think about the world's population in male and female halves, because this is how we tend in the first instance to sort ourselves, and then how our experiences and opportunities line up. They generally line up this way, again at least in the first instance, without regard to race, class, or gender identity. Because of this, it is also not wrong to make anti-subordination commitments specifically to the female half of the world's population. This should not be our only anti-subordination commitment, but is it an entirely rational and important one.

---

192. *See also* ZARRETT, ET AL., *supra* note 139, at 3 (noting that one of the things that can work to encourage girls to stay engaged in sport once they have chosen to participate is "[a]n emphasis on winning . . . when combined with an emphasis on fun and skill development . . . . Healthy forms of competition are ideal for fostering girls' engagement").

193. Mallonee, *supra* note 139 ("One of the biggest things researchers are finding that keeps girls engaged in sports is access to their heroes and mentors—even if it's just seeing them . . . [t]o see a banner or a poster or an ad featuring someone like you is monumental.").

Sex equality can sometimes be achieved without affirmative consideration of sex, and in this period in the United States, sex-blind policies are preferred. But when sex blindness would be counterproductive, ineffective, or insufficiently remedial, sex conscious approaches should be tools in the equality toolbox. These are especially useful in circumstances where, to fix inequities, it is important to see, not to erase, the female body. This includes at least aspects of the workplace; the military; medicine and bio-medical research; and competitive sport. If we were to have to ignore sex differences in these circumstances, sex equality would remain elusive.

Efforts to secure inclusion and equality for transgender people that are premised on erasing sex and sex differences—conceptually and from the discourse—are fundamentally incompatible with these facts, priorities, and approaches. Where sex, including being able to name it, is central for many if not most females, both are anathema for many in the transgender community. For this reason, but also because it fits their legal and political strategy, some trans rights advocates seek to redefine sex to be or to include gender identity; and further to ensure that no distinctions can be made on the basis of sex on the ground that such distinctions are a rejection of transgender people.[194] Those who do not play ball are labeled "transphobic" and, if they are liberal and feminist, also as "trans exclusionary radical feminists" (TERFs), which is intended as an insult.[195]

The competitive sports question is hard, maybe even impossible to resolve with a win on all sides, because what females need to have recognized is precisely what advocates for trans students suggest should be erased. For example, if we continue to be committed to equality for females in relation to males in the high school sports space—equality not only for its own sake but also for the myriad individual, institutional, and societal benefits that flow from its terms—we need structures that recognize sex differences in athletic performance and that sort athletes on the basis of sex.[196] At the same time, because they are properly focused on the individual children in their care—and not on the implications for others of their approach—erasure of these same sex differences has been built into the therapeutic model developed by pediatricians working with transgender children, and impressed on their educational custodians.[197]

As we have already discussed, efforts to will this collision away do not work because they rest on a series of ultimately unacceptable fictions: That all sex is socially constructed stereotype that inures to the detriment of women. That there are no cognizable differences between women who are transgender and women who are not. And that "women" is a concept that, in its best iteration, describes

194. *See supra* notes 29–30 and accompanying text.

195. Colleen Flaherty, *'TERF' War*, INSIDE HIGHER ED (Aug. 29, 2018), https://www.insidehighered.com/news/2018/08/29/philosophers-object-journals-publication-terf-reference-some-feminists-it-really. *See* R (on the Application of Miller) v. College of Policing & Chief Constable of Humberside, [2020] EWHC 225 (Admin) [225], ¶ 241–46 (Eng.) (describing and quoting the expert witness statement of Professor Kathleen Stock on the use of the term 'TERF' in the context of the gender-related divisions among feminist academics).

196. *See supra* notes 112–114, 149–156 and accompanying text.

197. *See supra* note 166 and accompanying text.

people who are relatively anti-competitive and selfless and so do not mind being relegated to the bench if this is necessary to secure the health and welfare of others who may be more vulnerable.[198]

The competitive sports question is also hard because it involves claims for inclusion and equality from both sides; and, because of how sport and sex-linked biology work together, including one group necessarily means excluding or limiting the opportunities of the other. The carve-out that is the sports exception to Title IX's general, sex-blind nondiscrimination rule recognizes this. And so, to secure the inclusion of and equality for females, it formally permits—but also sometimes requires—schools to exclude males from their sports and events.[199] Carving out what would be, in effect, an exception to this exception for the subset of males who identify as women and girls would result in their inclusion; but it would have exclusive effects in the other direction. Longer term, it would inevitably signal that we don't actually need sex segregation, which would result in the full re-integration of sport and thus the re-exclusion of females. Shorter term, because competition itself is exclusionary in the sense that making teams requires try-outs and cuts, and making it through to the finals and championships involves a version of the same—there can be only one state, regional, or national champion—carving out an exception to the sex segregation rule for males who identify as women and girls will result in the exclusion of females from teams, finals, and podiums.[200]

The problem is hard but, in general, there are four possible approaches: affirming the carve-out as a biological classification tied to natal sex; re-imagining the carve-out as an identity classification; re-imagining the carve-out as a biological classification tied to the onset of male puberty; and formalizing an accommodations approach.

1. Affirming the Carve-Out as a Biological Classification Tied to Natal Sex

The first option is to affirm the traditional approach. This approach ties eligibility to natal sex, which would generally be based on the sex recorded at birth on the individual's birth certificate. It is mostly efficient and effective because natal sex and sex recorded at birth are typically the same, and both typically correspond to the relevant primary and secondary sex characteristics that matter for sport.

Nevertheless, the traditional approach is both over and underinclusive: On the one hand, it includes trans boys and trans men who go on puberty blockers and gender affirming hormones beginning at the onset of female puberty; as a result, these athletes develop the male secondary sex traits that girls' and women's sport exist to exclude. On the other hand, it excludes trans girls and trans women who do the same at the onset of male puberty, and who, as a result, never develop those traits.

198. *See supra* notes 185–193 and accompanying text.

199. *See supra* notes 52–53 and accompanying text.

200. *See supra* notes 104–105 and accompanying text (detailing how this has already happened in the State of Connecticut and how the related data are being used in a Title IX OCR complaint).

In addition, many jurisdictions now permit people who are transgender to change the sex that is recorded on their birth certificates so that it accords with their gender identity; in some places, parents can do the same for their children.[201] Depending on the jurisdiction, the individual need not have gone on hormones before making this switch; that is, the switch may be permitted on some form of self-declaration.[202] As a result, at least for some trans people, the sex recorded on the birth certificate is no longer a reliable proxy either for natal sex or for the relevant sex traits.

Because of these administrative and policy concerns, but also because the law requires sex-related criteria to be closely tailored to institutional ends, we do not support this approach. In our view, transgender women and girls should not be excluded from girls' and women's sport if they have not gone through any part of male puberty. Moreover, to include transgender men and boys who are on gender affirming androgens is, in effect, no different from condoning the use of performance enhancing drugs. Still, whether it is as a matter of choice or inertia, many jurisdictions continue to provide that natal sex as recorded on the individual's birth certificate is the standard for eligibility for sex segregated sport.[203]

2. Re-imagining the Carve-out as an Identity Classification

The second option is to abandon the carve-out for females only. This would necessarily entail a rejection of the legitimacy and value of the claim from females for equality in relation to males in the education-based competitive sports space. It would involve either the full integration of sports and events, that is, all would be co-ed; or it would involve a challenge to re-imagine girls' and women's sport as a category that includes anyone who identifies as a girl or woman most broadly defined, even if they retain their full male-linked performance advantages. The best analog would be to an all-women's college that admits and retains students who identify as women regardless of their sex.[204]

Advocates for transgender rights have persuaded many secondary school athletic associations to adopt this approach. The unfounded claim that it is required by Title IX may have been at play in some cases.[205] Nevertheless, it is

201. *See generally ID Documents Center*, NAT'L CTR. FOR TRANSGENDER EQUAL., https://transequality.org/documents (last updated Jan. 2020) (providing state-by-state information and details on New York City rules permitting parents to change their child's birth certificate).

202. *See id.*

203. *See High School Policies*, TRANSATHLETE, https://www.transathlete.com/k-12 (last visited Jan. 28, 2020).

204. *See, e.g.*, Jeremy Bauer-Wolf, *At Women's Colleges, Rules Vary Widely for Trans and Nonbinary Students*, EDUC. DIVE (Nov. 18, 2019), https://www.educationdive.com/news/at-womens-colleges-trans-and-nonbinary-applicants-face-inconsistent-rules/567537/; Rebecca Brenner Graham, *Women's Colleges Should Admit Trans Students. It's Wholly Consistent with Their Mission*, WASH. POST (Jan. 10, 2019), https://www.washingtonpost.com/outlook/2019/01/10/womens-colleges-should-admit-trans-women-its-wholly-consistent-with-their-mission/.

205. *See supra* notes 70–76 and accompanying text (discussing the state of the law); *supra* note 203 (providing an up-to-date description of state high school athletic association policies). *See e.g., Complaint Targets Transgender HS Track Athletes*, ESPN (June 20, 2019), https://www.espn.com/high-school/story/_/id/27015115/complaint-targets-transgender-hs-track-athletes (discussing the claim that

viable in the long run only if these advocates can convince law and policy makers at the national level that Title IX should be revised to these ends. Specifically, they will need to convince the federal government that the original Title IX carve-out for females in sport is not a commitment it wants to keep; but that we still need—and need to support with federal funds—two classifications, both of which would be comprised of a combination of males and females, and both of which would see males in championship positions.

These are important hurdles, particularly in an adverse political climate. Not only have the last two administrations held firm to the premise that distinctions on the basis of sex in Title IX sport are necessary and appropriate, but also, this policy choice appears to be based in or at least consistent with a clear, bipartisan, nationwide consensus in favor of the traditional approach.[206] Nevertheless, the option is attractive to those who prefer that education-based sport focus on opportunities for participation not competition, including whenever possible in co-ed settings.[207] It is also attractive to those whose focus is specifically on the health and welfare of transgender people, rather than on females, on the view that at least in this period, the former need support more than the latter.[208] Finally, it is attractive to those who prefer a wait-and-see to a precautionary approach to restrictions on eligibility; they predict that trans women and girls won't have an important impact on girls' and women's sport, but that we can re-evaluate if they do.

The best argument in favor of this approach today is based in the view that almost everyone has a gender identity that aligns with their natal sex, in the still small numbers of trans girls and women who seek to be classified according to their gender identity, and in the fact that, until recently, we had not known of any in championship positions. Until recently, including them was—in effect—a non-category defeating accommodation within the existing sex segregated structure. As Rachel McKinnon has argued,

---

Title IX requires the inclusion of trans girls in girls sport); *Reference Guide for Transgender Policy*, CONN. INTERSCHOLASTIC ATHLETIC CONF., https://www.casciac.org/pdfs/Principal_Transgender_Discussion_Quick_Reference_Guide.pdf (last visited on Feb. 19, 2020) (explaining that "[t]he CIAC has concluded that it would be fundamentally unjust and contrary to applicable state and federal law to preclude a student from participation on a gender specific sports team that is consistent with the public gender identity of that student for all other purposes").

206. *See supra* notes 69–76 and accompanying text. *See, e.g., Most Oppose Transgender Athletes on Opposite Sex Teams*, RASMUSSEN REP. (June 4, 2019), https://www.rasmussenreports.com/public_content/lifestyle/social_issues/most_oppose_transgender_athletes_on_opposite_sex_teams (finding that "just 28% of American Adults favor allowing transgender students to participate on the sports team of the gender they identify with"). Given that this is a breakout conversation, see *supra* note 69, this number is likely to shift in one or the other direction, as is the number representing undecideds, i.e., 18 percent in the same survey.

207. *See infra* note 208 and accompanying text (describing this position).

208. *See* Navratilova, Coleman & Richards-Ross, *supra* note 89 (noting that "Advocates of the Equality Act who know sports or aren't science deniers . . . [argue] that it's time to shift our focus from supporting female-bodied athletes for whom Title IX has already done a lot of work, to supporting transgender women and girls who [now] need our help more").

> Since the 2004 Athens Olympics, there have been over 54,000 Olympians. Not one of them has been openly trans. There also aren't any cases of men pretending to be (trans) women. Next year, there are a few athletes who have the potential to be the first openly trans athlete to compete in the Games. None are a medal favorite. This is not the beginning of the end of women's sports.[209]

This moment is passing quickly, however, as more children identify as transgender; as we are learning to embrace them as they are; as they are increasingly comfortable coming out in high school even if they are not on hormones; and as advocacy to establish trans rights is increasingly successful in other contexts.[210] This includes recent legislation in the United States and in many countries around the world that permits trans people more easily to reform their identity documents to provide that their legal sex is their gender identity.[211] As we have already noted, it is not a coincidence that it is in the last three years we have seen the first male-to-female transgender state, national, and international champions in girls' and women's Olympic events,[212] and that international regulators are working to ensure that policies are in place to address the expected increase in numbers.[213]

McKinnon herself made history when she won gold in the women's sprint event at the 2018 and 2019 Masters Track Cycling World Championships and in the process set a women's world record.[214] She is a trans woman who has met the hormonal conditions required for inclusion in women's cycling events; she is playing by the rules established by her governing body.[215] Those rules are in play, however. On the one hand, many have expressed concern that they insufficiently account for the legacy advantages retained by trans women who physically transition after puberty.[216] On the other, some, including McKinnon, have argued that they should be revised to permit unconditional inclusion on the view that trans women are women based on their gender identity, without respect to the decisions they might take privately about gender affirming hormones or

209. McKinnon, *supra* note 191.

210. *See supra* note 179.

211. *See, e.g., ID Documents Center*, *supra* note 201 (providing up-to-date information on relevant state laws and practices).

212. *See supra* notes 181–182 and 209 and accompanying text.

213. Press Release, World Athletics, International Federations Discuss Consensus on Establishing Rules for Transgender Athletes (Oct. 31, 2019), https://www.worldathletics.org/news/press-release/international-federations-rules-transgender-a.

214. Karleigh Webb, *Trans Cyclist Rachel McKinnon Keeps Winning Championships and Her Detractors Don't Like It*, OUTSPORTS (Oct. 23, 2019), https://www.outsports.com/2019/10/23/20928252/rachel-mckinnon-trump-cycling-trans-athletes-transphobia-world-championships.

215. McKinnon, *supra* note 191.

216. *See, e.g.*, Sean Ingle, *Sports Stars Weigh in on Row Over Transgender Athletes*, GUARDIAN (Mar. 3, 2019), https://www.theguardian.com/society/2019/mar/03/sports-stars-weigh-in-on-row-over-transgender-athletes (summarizing this argument). For a good description and evaluation of the debate around legacy advantages, see generally Ross Tucker, *On Transgender Athletes and Performance Advantages*, SCI. SPORT (Mar. 24, 2019), https://sportsscientists.com/2019/03/on-transgender-athletes-and-performance-advantages/?doing_wp_cron=1576790435.8559970855712890625000.

surgery.[217] Although the call for unconditional inclusion has not yet succeeded in the elite sport space, related efforts directed at state high school athletic associations across the United States have altered that landscape.

3. Re-Imagining the Carve-Out as a Biological Classification Tied to Puberty

In 2019, in the challenge brought by Caster Semenya to the eligibility criteria for the women's category in the sport of track and field, the Court of Arbitration for Sport (CAS) explained that:

> the purpose of the male-female divide in competitive athletics is not to protect athletes with a female legal sex from having to compete against athletes with a male legal sex. Nor is it to protect athletes with a female gender identity from having to compete with athletes with a male gender identity. Rather, it is to protect individuals whose bodies have developed in a certain way following puberty from having to compete against individuals who, by virtue of their bodies having developed in a different way following puberty, possess certain physical traits that create such a significant performance advantage that fair competition between the two groups is not possible.[218]

Consistent with this rationale, the third approach affirms the carve-out for those who have not experienced male puberty. We peg this to male puberty to include in the classification all those who cannot be said to have developed the male sex-linked advantages that justify sex-segregated sport. It would include all females—however they identify—so long as they are not on masculinizing hormones; and all trans girls who, because they were on blockers and then feminizing hormones, have not experienced male puberty. This option is a challenge to imagine education-based sport as analogous to the medical setting, where sex and associated physical characteristics remain relevant; and where gender identity and expression are for the individual to resolve and others to respect as they would any central aspect of an individual's personhood.

This option is attractive for several reasons:

It is fully consistent with the carve-out's raison d'être and with its legal grounding.[219] In the language of the law, an approach that hews closely to this rationale is, as required, narrowly tailored and neither over- nor under-inclusive.

Moreover, as in other modern contexts where physical facts remain relevant to the enterprise, it distinguishes only on those objective grounds and otherwise respects the inherently personal nature of gender identity and expression. That is, it does not seek to establish, judge, or sort anyone on those different grounds that—in contrast with the physical—are ultimately unrelated to the institution's goals. But it does commit to respecting them without challenge. As it should be

217. *See, e.g.*, Charlie Ashworth, *Women's Sports: Rachel McKinnon Believes Trans Women Have a 'Human Right' to Compete*, GIVE ME SPORT (Oct. 21, 2019), https://www.givemesport.com/1514948womens-sports-rachel-mckinnon-believes-trans-women-have-a-human-right-t-ocompete ("By preventing trans women from competing or requiring them to take medication, you're denying their human rights.").

218. Mokgadi Caster Semenya & Athletics S. Afr. v. Int'l Ass'n of Athletics Fed'ns, CAS 2018/O/5794 ¶ 559 (2019).

219. *Id.*

developmentally, there is no questioning by a school or athletic association of a child's credibility about or commitment to a particular gender identity or expression; nor is there a requirement that they agree to be fixed for a season or an academic year to their expressed preferences. They can be who they are, including in flux, throughout the relevant period.

Consistent with Title IX's original design, all females who are not on transition hormones—however they might identify—have a space in which they can not only play but also compete for the win. Except in the rare case of an exceptionally precocious child star, the policy cannot be said to incentivize a student's choice to go on gender affirming hormones. And because they have benefitted from either endogenous or exogenous male testosterone levels, the subset of trans kids who start on hormones only after the onset of puberty is not be eligible to compete in the female category; but they are eligible to participate and welcome in the male category, and—importantly—their private needs and choices will not dictate outcomes for others. In this respect, this option is the least disruptive of the existing model that does a lot of good work for the vast majority of stakeholders.

It would be most easily implemented in circumstances where school sports teams don't have to train separately, that is, where only competition itself needs to be sex segregated. In such contexts, the social aspects of participation in education-based sport would not be linked to sex or gender. Swimming, cross country, and track and field are among the sports where—at least as practiced in some places—a version of this model already exists.[220]

The principal costs associated with this approach are as follows:

Because it is not simply a re-naming of the existing sex-segregated structure, imagining the classifications as we suggest would take work, even if everyone were on board. This work would range from the relatively simple and technical to the more difficult and conceptual. For example, we would have to decide what to call the classifications and how to establish where students belong. We have labeled the classifications "male" and "female" here, but they could be labeled differently. Uniform rules would need to be changed to allow students to select the style that makes them most comfortable regardless of sex or gender. And the notion of school sports as sex or gender-specific social spaces would need to be wound down. The latter would be resisted by traditionalists but also by some trans kids and their advocates who embrace binary sex classifications and see inclusion within them as having the potential to contribute to their successful transition.[221]

Finally, regardless of how respectful and welcoming the environment is made for gender diverse students within the categories, those who may be passing do not want to be outed by sex classifications. Others who are not passing but who suffer from dysphoria and are deeply (not just politically) hurt by references to their sex-linked traits may continue to be effectively excluded from participation

220. *See, e.g.*, Josh Weinreb, *Thetford Academy Transgender Runner Embraces Her Identity and Gains Freedom*, VALLEY NEWS (May 25, 2019, 9:25 AM), https://www.vnews.com/Running-career-has-helped-Thetford-Academy-senior-Bel-Spelman-navigate-her-gender-identity-24764677.

221. *See supra* note 166 and accompanying text (noting this strategy).

and competition. Still others who do not suffer from dysphoria but who prefer to occupy spaces where sex is irrelevant may reject school sport because it is still so focused. To the extent the institution could and should produce important health and welfare benefits for these individuals, this approach would not be effective. As is the case today for kids who for various reasons eschew school sports, some will remain left out or will choose to exclude themselves.

4. Formalizing an Accommodations Approach

The fourth option is to affirm the commitment to the Title IX carve-out for females but also formally to grant authority to policymakers to accommodate gender diverse students in ways that are not category defeating. Accommodations are often preferred in circumstances that involve competing rights claims.[222] To date, some form of this approach has been preferred by those working most thoughtfully on inclusion in the elite sports space, for example within the Olympic Movement, the NCAA, and the National Scholastic Athletics Foundation (NSAF).[223]

Although it has not been formally tested to date, the NCAA policy is especially relevant as it operates within institutions governed by Title IX. It permits transgender student-athletes to compete either according to their natal sex or their gender identity. In the latter case, if the athlete is a trans man, they are required to have a therapeutic use exemption (TUE) if they are on gender affirming hormones (testosterone); and if they are a trans woman, they are required to have been on testosterone suppressants for at least a year before they are eligible for women's teams and competitions.[224] We do not have enough studied experience with trans women and girls in sport to know that suppression to a certain level for a given period of time is sufficient to wind down male-linked advantages to the point where they are not category defeating in particular sports and events. Nevertheless, the rule fits the model because it presumes an ongoing primary commitment to female athletes, which is the raison d'être for the sports exception to Title IX's non-discrimination rule; and it includes trans women when they meet

222. *See, e.g., Notice, Enforcement Guidance: Reasonable Accommodation and Undue Hardship Under the Americans with Disabilities Act*, U.S. Equal Emp. Opportunity Comm'n (Oct. 17, 2002), https://www.eeoc.gov/policy/docs/accommodation.html (applying the model in the context of conflicts between employers and employees with disabilities). Accommodation is related to compromise. *See, e.g.*, Dale Eilerman, *Agree to Disagree: The Use of Compromise in Conflict Management*, MEDIATE (Oct. 2006), https://www.mediate.com/articles/eilermanD7.cfm (describing when compromise is useful in mediation).

223. Almost all elite sports institutions have adopted a rule of conditional inclusion based on testosterone levels. *See, e.g., IOC Consensus Meeting on Sex Reassignment and Hyperandrogenism*, INT'L OLYMPIC COMM. (2015), https://stillmed.olympic.org/Documents/Commissions_PDFfiles/Medical _commission/201511_ioc_consensus_meeting_on_sex_reassignment_and_hyperandrogenism-en.pdf; *NCAA Inclusion of Transgender Student-Athletes*, NAT'L COLLEGIATE ATHLETIC ASS'N (2011), https://13248aea-16f8-fc0a-cf26-a9339dd2a3f0.filesusr.com/ugd/2bc3fc_4a135824fabc462183c71357c93a99b4.pdf; *National Scholastic Athletic Foundation Transgender Participation Policy and Procedure*, NAT'L SCHOLASTIC ATHLETICS FOUND. (2019), https://www.nationalscholastic.org/nbin/transgender/.

224. *NCAA Inclusion of Transgender Student-Athletes*, *supra* note 223, at 13.

relevant physical conditions.[225] At least conceptually, because testosterone is the primarily driver of the performance gap, the accommodation is viable as category affirming not defeating. As an evidentiary matter, and thus legally, the rule would be especially defensible if the NCAA were to establish a maximum allowable T level that is within the female range and then to develop a protocol for monitoring compliance.[226]

A different example of the accommodations approach outside of elite sports can be found in the policies of state athletic associations that have experience integrating male students who are not transgender into girls' sports and events. This happens where there is no boys' team and the male students can show—consistent with Title IX requirements—that they are the excluded sex. Where particular males threaten to disrupt the championship experience and hierarchy, officials have sought solutions to their inclusion that are consistent with the goals of the carve-out, such as adding lanes or running separate male and female sections of a final, and featuring separate podiums for male and female finishers. Consistent with the goals of the sports exception to Title IX's general non-discrimination rule, this has ensured that male student-athletes are not precluded from participating in their chosen sports, but also that there cannot be a male winner of the girls state championship.[227] These strategies aren't perfect fits, given that trans girls and women identify as girls and women, not as boys and men. But since sport is segregated on the basis of sex, not identity, and identity is ultimately irrelevant to sports performance, they can be useful as examples of solutions that might resolve certain impasses.

Other models that could be adapted depending on the sport and event are quotas and adjusted scores and start lines. Quotas might be especially useful in team sports situations. Joanna Harper and Tiffany Abreu have suggested they could work in volleyball and basketball, for example.[228] The basic concept of adjusted scores and start lines comes from golf, which designate different Tee boxes for males and females and—to level the playing field for golfers of different abilities—use adjusted scores (handicaps) to compare relative performances.

---

225. According to the Justice Department, the NCAA rule was developed in "consult[ation] with medical experts, athletics officials, affected students, and [with advocates for transgender student-athletes]." Dear Colleague Letter, *supra* note 73 (citing NCAA Inclusion of Transgender Student-Athletes 2, 30–31, and Pat Griffin & Helen J. Carroll, On the Team: Equal Opportunity for Transgender Student Athletes (2010)). On the Team itself notes that "policies that may be appropriate at the college level may be unfair and too complicated for [the high school] level of competition." It thus encourages the development at the high school level of age-appropriate policies. *Id.* at 26. It was co-authored by two of the leading players in this area, Helen Carroll of the National Center for Lesbian Rights (NCLR) and Pat Griffin of the University of Massachusetts Amherst, in consultation with Shannon Minter, also of the NCLR, and Eric Vilain, who also consulted on the IOC policy. Both Minter and Vilain are leading experts in their respective fields, i.e., civil rights litigation and biological sex, respectively.

226. As of this writing, although we expect that most trans girls and women on hormones follow the medical standard of care which has a target of well under 5 nmol/L, the NCAA has not set a maximum allowable level, nor does it monitor compliance.

227. *See, e.g.*, Coleman, *Sex in Sport*, *supra* note 27, at 173–77 (describing the State of Massachusetts' approach to the inclusion of boys in girls' swimming events).

228. Darlington, *supra* note 116.

Different start lines could be adjusted based on the sport's average performance gap between males and females, in the manner that the distance between Tee boxes is adjusted to reflect the relative power of female and male golfers, and scores could be adjusted at the outset on the same group—rather than individual—basis. More complicated iterations of golf's handicapping system have been described elsewhere.[229] Ultimately, the key to such approaches would be assuring their efficacy and their administrative feasibility.

Like accommodations in general, accommodations in sport have the benefit of being adaptable over time based on new knowledge. For example, as we learn more about the nature and extent of the legacy advantages of going through male puberty, as well as about their particular effects in different sports and events, the specific requirements within the model could be adjusted without altering the commitment to the model itself. And as we develop a better sense of the political community's relative commitments to female sport on the one hand and to trans inclusion on the other, the reasonableness of specific conditions within models will also evolve.

The merits of and problems inherent in accommodations models are that they tend mostly, but not entirely, to satisfy principal policy goals while reducing, but not eliminating, the concerns of affected individuals. In other words, like all compromises, accommodations generally mean that no one gets everything they wanted; and, depending on the specifics, one or the other side still faces a complete loss. As it considers the question of transgender inclusion, sport is no different. Purists on both sides of the debate decry all proposed concessions: Those who want girls' and women's sport to remain exclusively for females say they cannot abide a solution that would ever see a transgender athlete in a championship position, even if she is following all of the rules. Those who want girls and women's sport to be unconditionally inclusive of transgender athletes say they cannot abide a solution that recognizes that there is a difference between females and transgender women and girls: "Transgender women are women. Period. Transgender girls are girls. Period." Those who argue from "the messy middle" can struggle to get a foothold. But given the stakes on both sides, it is surely worthwhile also to consider solutions in this space.

D. Our Recommendations

Because sex equality in education-based sport produces enormous value, and because the development of inclusive policies is separately consistent with educational institutions' goals, policymakers should affirm Title IX's original design and work to include gender diverse students within that design. Because institutional goals are different in non-elite and elite settings, approaches to inclusion should track those different goals. Throughout, policymakers should endeavor to develop strategies that will encourage as many students as possible to remain engaged in school sports as participants and as competitors.

Where it can be effective to all ends to combine teams or at least team practices and only to segregate competition itself on the basis of sex, the approach

229. *See, e.g.,* Aschwanden, *supra* note 167 (discussing the concept in general and the work of Alison Heather and colleagues in particular).

we detail in Part IIIC3—re-imagining the carve-out as a biological classification pegged to puberty—should be preferred. So long as it doesn't result in diminished coaching opportunities for females who remain underrepresented in those ranks or deter female students from staying engaged with sport, it is the most inclusive, least intrusive, and simplest to administer. Existing co-ed arrangements can be an ongoing model for this purpose.

Where combined teams or practices coupled with sex segregated competition cannot accomplish institutional goals, the accommodations approach detailed in Part IIIC4 should be adopted. This will be the case in circumstances where sex segregated teams and events remain necessary to secure parity of opportunity for females. Where the accommodations approach is adopted, trans students will train and compete consistent with their gender identity so long as their inclusion can be relevantly conditioned. The NCAA transgender policy is illustrative of a hormonal condition in this category; others that do not require medicalization— such as handicaps, offsets, and quotas—exist as more appropriate models for the high school sports space.

In high school intramural, junior varsity, and regular season play, where institutional goals are primarily related to health and fitness and to the development of social skills, unconditional inclusion of gender diverse students according to their gender identity rather than their sex will usually be category affirming. Exceptions will arise where this is not the case, for example in contact sports situations where physical safety is tied to sex-linked differences, and where regular season play determines invitational and post-season opportunities. But to the extent that including trans students according to their gender identity merely makes others uncomfortable, educators should be encouraged to educate, including to inculcate empathy and inclusivity, rather than to exclude.

Once the focus shifts to competition and to the establishment of hierarchy and the isolation and celebration of champions, unconditional inclusion of trans girls and women who have benefited from male puberty becomes category defeating. Conditional inclusion in this context is therefore appropriate. This position will not satisfy purists on either side of the issue and both have strong arguments in support of their views. Most immediately, it won't satisfy those who believe the category is inevitably defeated by the inclusion of students whose natal sex is male regardless of how their participation is conditioned. And it won't satisfy either medical providers who have built girls sport or invariable inclusion on the basis of gender identity into their treatment design or trans advocates whose movement strategy is to elide the differences between sex and identity. Ultimately, however, the standard that prevails should be one that provides for reasonable accommodations given institutional goals.

Finally, because the legal landscape has become muddied in this period, to the point that there are questions about what Title IX does or should require, the re-commitment to its original design should be codified by statute along with an allowance for reasonable, non-category defeating accommodations.[230] To the

230. The Obama and Trump Administrations were both put to this question in the context of claims for transgender inclusion and both have re-affirmed the federal government's commitment to sex equality, albeit in different forms. *See supra* notes 73–75 and accompanying text. But because the

extent possible the legislation should be based in existing language so as not to disrupt the well-established body of accompanying law, with definitions and clarifications as appropriate given the current context. Consistent with this prescription, we propose the following draft language:

> No person shall, on the basis of sex, be excluded from participation in, be denied the benefits of, be treated differently from another person or otherwise discriminated against in any interscholastic athletics offered by a recipient, and no recipient shall provide any such athletics separately on such basis.
>
> However, to secure Title IX's commitment to sex equality, a recipient may operate or sponsor separate teams and events based on sex where selection and advancement are affected by sex-linked competitive advantages or the activity involved is a contact sport in which physical safety is implicated.
>
> So long as they do not imperil female students' physical safety or diminish their competitive opportunities, a recipient that operates or sponsors separate sex teams and events may include persons of the excluded sex when their gender identity is concordant. On the same conditions, a recipient that sponsors a team for only one sex may include persons of the excluded sex. Reasonable accommodations consistent with these conditions are encouraged.
>
> For purposes of this statute, sex retains its dictionary definition as "either of the two divisions, designated female and male, by which most organisms are classified on the basis of their reproductive organs and functions."[231] It does not include sex stereotypes or legal or gender identity.

CONCLUSION

Title IX expresses society's commitment to sex equality in educational settings. At the time of the statute's enactment in 1972, this commitment was revolutionary. Today, in no small part because Americans across the political spectrum are invested in the goal, Title IX's value is mostly a given.[232] From the focus on increasing the numbers of women in STEM to the effort to eradicate the conditions that enable sexual assault, the idea that women belong as equals on campus persists. Notably, the commitment to this idea is not merely normative. As Nicholas Kristof wrote in his year-end column for the *New York Times* in 2019, "few forces change the world so much as education and the empowerment of women."[233]

---

executive branch has discretion in the interpretation of federal regulations, and because administrations come and go, they can foster unnecessary confusion and ensure that the matter remains unsettled.

231. *Sex*, AM. HERITAGE DICTIONARY OF THE ENGLISH LANGUAGE (5th ed. 2020).

232. Sandra Guy, *Title IX at 45*, SOC'Y OF WOMEN ENG'RS MAG. (Mar. 20, 2017), https://alltogether.swe.org/2017/03/title-ix-45/ ("Title IX is hugely popular, and it's a bipartisan issue. We don't expect that to change.").

233. Nicholas Kristof, *This Year Has Been The Best Year Ever*, N.Y. TIMES (Dec. 31, 2019), https://www.nytimes.com/2019/12/28/opinion/sunday/2019-best-year-poverty.html. *See also* Ana Revenga & Sudhir Shetty, *Empowering Women is Smart Economics*, 49 IMF FIN. & DEV. (2012), https://www.imf.org/external/pubs/ft/fandd/2012/03/revenga.htm.

The structure of the Title IX regulatory scheme makes clear that the goal is sex equality, not sex neutrality. Consistent with American equal protection jurisprudence and our general political inclinations, the latter is merely the preferred means to the former end. Like other sex equality measures, Title IX recognizes that females often remain disadvantaged in relation to males because of their reproductive biology and because of stereotypes about them based on that biology. Sex affirmative approaches are appropriate when sex neutrality cannot effectively address that disadvantage. Thus, such approaches may be used to overcome entrenched discriminatory patterns that are not explained by inherent differences; see special provisions for women and girls in fields in which they remain underrepresented. And they may be used to ensure that such differences are not unnecessary obstacles to important opportunities; see separate sex sport.

Notwithstanding our general preference for sex neutral measures, the sports exception to Title IX's general nondiscrimination rule has long been one of the statute's most popular features.[234] This affirmative approach is understood to be necessary to ensure that the sex-linked differences that emerge from the onset of male puberty do not stand as obstacles to sex equality in the athletic arena. From the beginning, it was understood that any different, sex neutral measure would ensure precisely the opposite—that spaces on selective teams and spots in finals and on podiums would all go to boys and men. The sports exception makes it possible for women and girls also to benefit from the multiple positive effects of these experiences, and for their communities and the broader society to reap the benefits of their empowerment.

The challenge in the beginning of the Title IX era was to conceive of and equally to support females as athletes, coaches, and sports administrators. We continue to fight for equal support as important institutions still stumble—see, for example, the dearth of female coaches in NCAA programs;[235] Nike's recently-revealed failure to keep its brightest female stars under contract when they become pregnant;[236] and USA Soccer's refusal to provide equal pay to the members of its male and female teams.[237] But as Title IX concludes its first semi-centennial, we no longer struggle as we did in the beginning with the basic concept of females as athletes. It is no longer commonplace for an athletic department to assume that a female is on the field to land a husband rather than a medal. Female puberty, pregnancy, and motherhood remain visible indicia of difference, but because of

234. *See supra* note 62 and accompanying text (quoting Nancy Hogshead-Makar on this point).

235. Jeré Longman, *Number of Women Coaching in College Has Plummeted in Title IX Era*, N. Y. TIMES (Mar. 30, 2017), https://www.nytimes.com/2017/03/30/sports/ncaabasketball/coaches-women-titleix.html.

236. Alisia Montaño, *Nike Told Me to Dream Crazy, Until I Wanted a Baby*, N. Y. TIMES (May 12, 2019), https://www.nytimes.com/2019/05/12/opinion/nike-maternity-leave.html. The company was smart enough to continue to pay Serena Williams through her pregnancy, but other stars—including most notably Allyson Felix—were not similarly treated. Scott Davis, *Serena Williams Supports Nike After Its Maternity Pay Controversy, Saying the Company Is 'Learning from Mistakes and Doing Better'*, BUS. INSIDER (May 28, 2019, 12:51 PM), https://www.businessinsider.com/serena-williams-backs-nike-maternity-pay-controversy-2019-5.

237. Andrew Das, *U.S. Women's Soccer Team Granted Class Status in Equal Pay Lawsuit*, N. Y. TIMES (Nov. 8, 2019), https://www.nytimes.com/2019/11/08/sports/uswnt-equal-pay-lawsuit.html.

the sports exception, they are no longer disqualifying. Indeed, when the promotion is done right, these are affirmatively empowering and celebrated.[238]

The challenge as we move into Title IX's second semi-centennial is to persuade institutions finally to address the remaining disparities in their support of female athletes and female sport at the same time that we enter a new revolutionary period in which we are being asked to imagine that "female" includes individuals of both biological sexes so long as they identify as women and girls. This ask reflects the intellectual choice to conceive of sex as a social construct rather than as a fact of biology tied to reproduction, and also the strategic choice of trans rights advocates to work toward law reform that would disallow any distinctions on the basis of reproductive sex. A popular manifestation of this strategy is their insistence that we accept as threshold truth rather than as political claim the proposition that "Trans women are women, period."

The problem is that female sport is by design and for good reasons, a reproductive sex classification. These reasons have nothing to do with transphobia and everything to do with the performance gap that emerges from the onset of male puberty. Whether one is trans or not, if one is in sport and cares about sex equality, this physical phenomenon is undeniably relevant. Changing how we define "female" so that it includes individuals of both sexes, and then disallowing any distinctions among them on the basis of sex, is by definition and in effect a rejection of Title IX's equality goals. Whatever their earlier allegiances, and however they would seek to re-tool the relevant vocabulary to obscure this point, we should be clear that those push for these changes today are committed to sex neutrality, not to sex equality.

We need to find a path to equality also for trans people. And we need to be thoughtful about how they are included within an institution whose design is at odds with who they are. But given the enormous social utility and popularity of that design, as well as the work that still needs to be done to fulfill its promise, a path that involves a rejection of its principal terms is a non-starter.

In this paper, we have provided the legal history and the science that make sense of the sports exception to Title IX's general nondiscrimination rule. We have also developed the policy arguments for affirming the commitment to sex equality in the education-based sports space, and for including trans kids in that space in ways that support their healthy development without undermining either the statute's sex equality goals or its allowance for sex affirmative measures to achieve them. Finally, we have described and evaluated the options that are and ought to be on the table as civil rights advocates and policymakers work through this challenge. We do not expect that we have thought of everything; indeed, because the science and social norms are evolving as we write, we assume that regular

238. This video from the Olympic Channel, *Aiming for the Olympics After Child-Birth ft. Allyson Felix/ Top Performer*, YOUTUBE (Oct. 21, 2019), https://www.youtube.com/watch?v=TNuL38NRppg is illustrative. *See also, e.g.*, Annabelle Timsit, *Serena Williams's New Ad Gives Working Moms the Nuanced Representation They Deserve*, QUARTZ AT WORK (Aug. 29, 2018), https://qz.com/work/1372139/serena-williams-thismama-ad-offers-a-powerful-vision-of-working-moms/. For a summary of the development of the market for female sport, see Ross Andrews, *Women's Sports Popularity is Growing, According to Nielsen Study*, GLOB. SPORT MATTERS (Nov. 13, 2018), https://globalsportmatters.com/business/2018/11/13/womens-sports-popularity-is-growing-according-to-nielsen/.

updating will be necessary. But we hope that the structure, background, and arguments we've set out will be useful in the process.

ORIGINAL RESEARCH

Psychological Support, Puberty Suppression, and Psychosocial Functioning in Adolescents with Gender Dysphoria

Rosalia Costa, MD,*† Michael Dunsford, PsyD,* Elin Skagerberg, PhD,* Victoria Holt, MRCPsych,* Polly Carmichael, PhD,*[1] and Marco Colizzi, MD†‡[1]

*Gender Identity Development Service, Tavistock and Portman NHS Foundation Trust, Tavistock Centre, London, UK;
†Department of Medical Basic Sciences, Neuroscience and Sense Organs, University of Bari "A. Moro," Bari, Italy;
‡Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK

DOI: 10.1111/jsm.13034

ABSTRACT

Introduction. Puberty suppression by gonadotropin-releasing hormone analogs (GnRHa) is prescribed to relieve the distress associated with pubertal development in adolescents with gender dysphoria (GD) and thereby to provide space for further exploration. However, there are limited longitudinal studies on puberty suppression outcome in GD. Also, studies on the effects of psychological support on its own on GD adolescents' well-being have not been reported.

Aim. This study aimed to assess GD adolescents' global functioning after psychological support and puberty suppression.

Methods. Two hundred one GD adolescents were included in this study. In a longitudinal design we evaluated adolescents' global functioning every 6 months from the first visit.

Main Outcome Measures. All adolescents completed the Utrecht Gender Dysphoria Scale (UGDS), a self-report measure of GD-related discomfort. We used the Children's Global Assessment Scale (CGAS) to assess the psychosocial functioning of adolescents.

Results. At baseline, GD adolescents showed poor functioning with a CGAS mean score of 57.7 ± 12.3. GD adolescents' global functioning improved significantly after 6 months of psychological support (CGAS mean score: 60.7 ± 12.5; $P < 0.001$). Moreover, GD adolescents receiving also puberty suppression had significantly better psychosocial functioning after 12 months of GnRHa (67.4 ± 13.9) compared with when they had received only psychological support (60.9 ± 12.2, $P = 0.001$).

Conclusion. Psychological support and puberty suppression were both associated with an improved global psychosocial functioning in GD adolescents. Both these interventions may be considered effective in the clinical management of psychosocial functioning difficulties in GD adolescents. **Costa R, Dunsford M, Skagerberg E, Holt V, Carmichael P, Colizzi M. Psychological support, puberty suppression, and psychosocial functioning in adolescents with gender dysphoria. J Sex Med 2015;12:2206–2214.**

Key Words. Gender Dysphoria; Adolescents; Psychosocial Functioning; Puberty Suppression

Introduction

Gender dysphoria (GD) individuals experience a marked incongruence between their assigned gender and their experienced gender [1]. GD refers to this stressful condition resulting in clinically significant distress or impairment in

[1]Joint last authors.
The study was conducted in the Gender Identity Development Service, Tavistock and Portman NHS Foundation Trust, Tavistock Centre, 120 Belsize Lane, London NW3 5BA.

important areas of functioning [2,3]. When supporting and treating children and adolescents with GD, health professionals should broadly conform to the Standards of Care of the World Professional Association for Transgender Health (WPATH) [4]. These guidelines indicate that psychological support should focus on exploring gender identity, role, and expression; addressing the negative impact of GD and stigma on mental health; alleviating internalized transphobia; enhancing social and peer support; improving body image; promoting resilience. Psychological interventions such as individual, couple, family, or group therapy should be provided within a multidisciplinary gender identity specialty service [4].

Studies indicate that cross-sex hormonal treatment (CSHT) improves well-being in GD adults [5,6]. However, it has been observed that despite many years of psychotherapy the GD of most adolescents does not often abate. Rather, once these young persons, who are already experiencing considerable distress over their gender identity, undergo the pubertal development of their biological sex, their psychological well-being deteriorates significantly [7]. Because this risk can be so great, the need for an early intervention has become paramount.

Delemarre-van de Waal and Cohen-Kettenis have proposed an early intervention approach, the Dutch model [8], which aims to eliminate the exposure to unwanted pubertal hormones, limit GD, and improve the ability to "pass" as the desired gender in adulthood. It considers adolescents, after a comprehensive psychological evaluation with many sessions over a longer period of time, eligible for puberty suppression, cross-sex hormonal treatment (CSHT), and gender reassignment surgery (GRS) at the respective ages of 12, 16, and 18 years when there is a history of GD; no psychosocial problems interfering with assessment or treatment; adequate family or other support; and good comprehension of the impact of medical interventions. According to this protocol, suppressing puberty and allowing young individuals the opportunity to explore their gender identity would provide some relief from the distress associated with the development of secondary characteristics [8]. Consistently, some studies indicate that puberty suppression leads to a better psychosocial outcome [2,9].

Since the release of the Dutch model, there has been disagreement about the appropriateness of treatment in minors. Some practitioners have questioned the ethics and safety of this intervention. Conversely, other health care professionals have argued they have an obligation to alleviate suffering and it would be unethical to allow a patient to suffer through the distress of pubertal development when there is a way of preventing it [10]. Anyway, puberty suppression by gonadotropin-releasing hormone analogs (GnRHa) has increasingly become accepted in clinical management of adolescents with GD. Even if further studies are needed, GnRHa are considered a safe and putatively reversible intervention which should be provided to people in need of it, especially if allowing puberty to progress appears likely to harm the young person [7].

There are limited longitudinal studies on the psychosocial functioning of GD adolescents after puberty suppression [2,9]. Also, studies on the effects of psychological support on its own on GD adolescents' psychosocial functioning have not been reported.

Aims

The aim of this study was to assess GD adolescents' psychosocial functioning in follow-up evaluations. Based on previous literature [2,9] and our clinical experience, we hypothesized a poor general functioning at baseline, an improvement after psychological support, and a further improvement after the beginning of the GnRHa.

Methods

Study Design and Participants

This longitudinal study was conducted at the Gender Identity Development Service (GIDS) in London. The health care pathway provided at the GIDS is described in Figure 1. A consecutive series of 436 adolescents (mean age = 15.74 ± 1.38 years; natal male/natal female ratio = 1:1.7) were referred between 2010 and 2014 to the GIDS. 201 adolescents (mean age = 15.52 ± 1.41 years; natal male/natal female ratio = 1:1.6) completed the diagnostic procedure (about 6 months) and were invited to take part in the follow-up evaluations. No GD adolescent refused to participate and all participants and their parents gave informed consent. By clinical interview, all adolescents fulfilled DSM-IV-TR criteria in use at the time for Gender Identity Disorder. The GIDS has adopted the WPATH Standards of Care [4]. There were no significant differences in socio-demographic characteristics as well as baseline CGAS scores

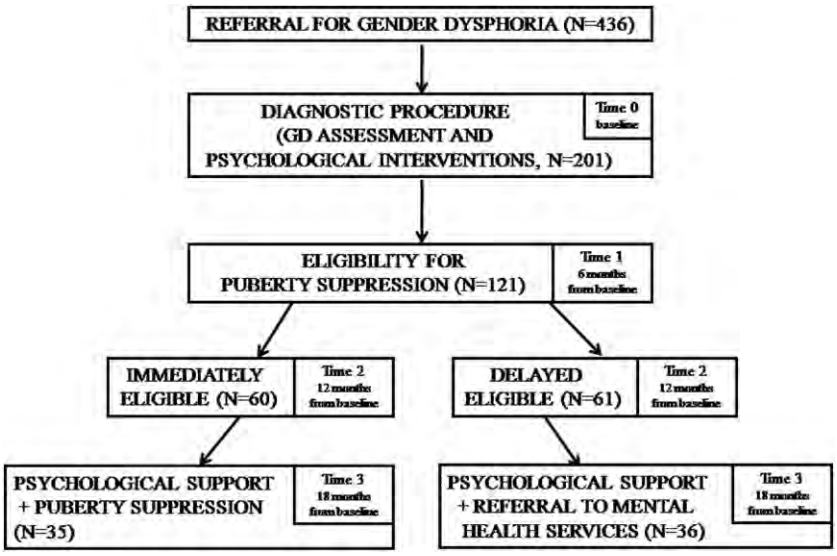


Figure 1 Health care pathway at the Gender Identity Development Service (GIDS)

between adolescents with a GD diagnosis enrolled in this study (N = 201) and adolescents who did not complete the diagnostic procedure (N = 235; all $P > 0.1$).

Psychological Support

The GIDS has developed a standardized psychological assessment which is part of the diagnostic procedure, in accordance with the WPATH guidelines [4]. This model emphasizes the early recognition and non-judgmental acceptance of gender identity problems as well as the importance of ameliorating associated behavioral, emotional and relationship difficulties [11]. Ample room is given to adolescents to explore different options for gender expression. Together with their families GD adolescents are supported in making difficult decisions regarding the extent to which they are allowed to express a gender role that is consistent with their gender identity. Also the timing of changes in gender role and possible social transition are extensively explored. This ensures that decisions about gender expression and the treatment of GD are thoughtfully and recurrently considered. Health care professionals help families to make decisions regarding the timing and process of any gender role changes for their young children. Information is provided to parents to weigh the potential benefits and challenges of choices.

The aims outlined are achieved through various psychotherapeutic interventions, ranging from individual to family and group therapy, which are carried out on a regular basis (at least once a month). Social and educational interventions are also provided if necessary. All these interventions are well coordinated and integrated in a comprehensive management plan agreed with local services (The Network Model). Moreover, the care pathway provides continuous psychological support to the patients' emotional and behavioral changes that may occur during the puberty suppression treatment. All adolescents received psychological support for the entire duration of the study.

Eligibility for Puberty Suppression

In accordance with the WPATH Standards of Care [4], adolescents were able to commence puberty suppression with GnRHa if they met the following criteria: (i) a presence of GD from early childhood on; (ii) an increase of the GD after the first pubertal changes; (iii) an absence of psychiatric comorbidity that interferes with the diagnostic work-up or treatment; (iv) adequate psychological and social support during treatment; and (v) a demonstration of knowledge and understanding of the effects of GnRHa, cross-sex hormone treatment, surgery, and the social consequences of sex reassignment. All GD adolescents were considered eligible for puberty suppression. Eligible adolescents were divided into two groups: immediately eligible and delayed eligible adolescents, consistently with Cohen-Kettenis and colleagues [12]. Immediately eligible adolescents started GnRHa at the end of the diagnostic procedure (0.75 ± 0.59 years from baseline). On the contrary, some adolescents were considered delayed eligible and continued to receive psychological support without

any type of physical intervention until they felt ready to make a decision in collaboration with their families and the clinicians. In those specific cases clinicians needed more time to make the decision of starting GnRHa because of possible comorbid psychiatric problems and/or psychological difficulties. If concomitant problems were observed (e.g., psychiatric problems, substantial problems with peers, or conflicts with parents or siblings), the young person was referred to a local mental health service. All possible medical and/or psychosocial interventions were well coordinated, integrated in a comprehensive management plan agreed with local services, and tended to be individualized in relation to the psychopathology/difficulty. The primary aim was for the child and the family to function better. After being assessed and, if necessary, treated for a psychiatric comorbidity, all delayed eligible GD individuals received puberty suppression. The interval from the start of the diagnostic procedure to the start of puberty suppression took about 1.5 years (1.5 ± 0.63 years from baseline). None of the delayed eligible individuals received puberty suppression at the time of this study.

Main Outcome Measures

Socio-Demographic Information

The data collected included: natal gender (male–female ratio), age (at assessment, at start of GnRHa), education level (yes/no), living arrangement (both parents, one parents, other), living in the chosen gender (partly, i.e., by wearing clothing and having a hairstyle that reflects gender identity/completely, i.e., by also using a name and pronouns congruent with gender identity/no), and change of name (yes/no).

GD-Related Discomfort

The Utrecht GD Scale (UGDS) was used to measure adolescents' GD-related discomfort. This is a 12-item questionnaire specifically developed to measure GD in a dimensional way. In particular, the UGDS focuses on core aspects of GD and gender identity. The adolescents are asked to rate their agreement on a 5-point scale. The total score ranges from 12 to 60. Higher UGDS total scores indicate high level of GD [13]. The scale has shown a high reliability (a Cronbach's alpha of 0.66–0.80 in one sample, and 0.78–0.92 in another); as reported by the authors, the lower alphas on the scale were only found among control subjects, which may be related to the lower variability of GD in these groups [13]. Cronbach's alpha for UGDS in our sample was 0.76–0.88. The UGDS has also shown a good discriminant validity, when adolescents and adults with and without a GD diagnosis were compared.

Measure of Global Psychosocial Functioning

The Children's Global Assessment Scale (CGAS) was used to assess adolescents' psychosocial functioning. The CGAS is one of the most widely used rating scales designed to measure how children and adolescents function psychosocially in daily life [14]. This clinical-rated instrument is divided into 10-point intervals and ranges from 1 to 100, with higher scores indicating better psychosocial functioning. The CGAS is useful to assess psychosocial/psychiatric outcomes, socio-cognitive competence and changes because of treatment [15]. In particular, it has been used in several longitudinal and epidemiological studies in clinical and non-clinical populations, naturalistic cohorts [16], and young GD individuals [9]. The inter-rater reliability was tested by Shaffer and his colleagues [14] before publication of CGAS, in order to minimize variation because of clinician background. Test–retest has been described in different studies with raters' consistence over time [16].

All CGAS were administered by qualified psychologists, psychotherapists, and psychiatrists who attended training and intra-class correlation assessment ($0.76 \leq$ Cronbach's $\alpha \leq 0.94$). Participants were assessed at baseline (Time 0) and every following 6 months, for a total of four evaluations over an 18-month period. Follow-up evaluations were performed 6 months from the baseline (Time 1: after 6 months of psychological support); 12 months from the baseline (Time 2: after 12 months of psychological support for delayed eligible GD adolescents, and after 12 months of psychological support + 6 months of puberty suppression for immediately eligible GD adolescents); 18 months from the baseline (Time 3: after 18 months of psychological support for delayed eligible GD adolescents, and after 18 months of psychological support + 12 months of puberty suppression for immediately eligible GD adolescents).

Participants were compared with a sample of young individuals without observed psychological/psychiatric symptoms ($N = 169$), using the same methodology of this study, the CGAS scale [16]. This sample was part of a large naturalistic cohort

of children/adolescents who attended child and adolescent mental health services (CAMHS; N = 12,613) in Stockholm in order to be evaluated for their psychosocial functioning.

Statistical Analysis

Chi-squared and independent *t*-tests were used to test for possible differences in socio-demographic characteristics and CGAS scores between natal men and natal women; adolescents who did not complete the diagnostic procedure and adolescents who received a GD diagnosis; immediately eligible and delayed eligible individuals. Dependent and independent *t*-tests were used to test for possible differences in CGAS scores between baseline and follow-up evaluations, in both immediately eligible and delayed eligible individuals.

Finally, independent *t*-tests were used to compare GD adolescents' CGAS scores with CGAS scores from a sample of children/adolescents without observed psychological/psychiatric symptoms [16].

Ethics

The study received ethical approval from the National Research Ethics Service (NRES) Committee London-Camden and Islington.

Results

Socio-Demographic Characteristics of the Sample

Socio-demographic characteristics of the sample (N = 201) are reported in Table 1. The majority of GD adolescents were living with one parent, were in education, were living as a member of the desired gender, and had changed their names. However, compared with natal women, a higher proportion of natal men did not live with their biological parents, had left school, were not living as a member of the desired gender, and had not changed their names. Moreover, natal women reported a significantly higher GD-related discomfort than natal men. Natal men and women did not differ in their age, both at assessment and when GnRHa was started (Table 1).

Table 1 General characteristics of 201 adolescents with gender dysphoria

Age in years, M (SD)	All participants	Natal men	Natal women	Statistical comparisons *t*-test; *P* value
Baseline	15.52 (1.41)	15.61 (1.70)	15.46 (1.22)	0.73; 0.47
Range	12–17	12–17	12–17	
At start of GnRHa	16.48 (1.26)	16.64 (1.22)	16.39 (1.28)	0.74; 0.46
Range	13–17	13–17	13–17	
Living arrangement, N (%)				χ^2; *P*
Both parents	78 (41.5)	25 (33.7)	53 (44.2)	8.95; 0.01
One parent	100 (53.2)	35 (51.5)	65 (54.2)	
Other*	10 (5.3)	8 (11.8)	2 (1.6)	
No details	13	8	5	3.47; 0.06
Education				
Yes	168 (89.8)	56 (83.6)	112 (93.3)	20.52; <0.001
No	19 (10.2)	11 (16.4)	8 (6.7)	
No details	14	9	5	
Living in role				
Completely	117 (62.6)	29 (42.6)	88 (73.9)	23.14; <0.001
Partly	27 (14.4)	12 (17.7)	15 (12.6)	
No	43 (23.0)	27 (39.7)	16 (13.5)	
No details	14	8	6	
Change name				
Yes	107 (57.5)	23 (33.8)	84 (71.2)	
No	79 (42.5)	45 (66.2)	34 (28.8)	
No details	15	8	7	
	Mean (SD)	Mean (SD)	Mean (SD)	*t*-test; *P* value
UGDS†	54.7. (6.8)	51.6 (9.7)	56.1 (4.3)	4.07; <0.001
CGAS at baseline	57.7 (12.3)	55.4 (12.7)	59.2 (11.8)	2.15; 0.03

*Living in children's home, living with other family's members
†Data available in 160 individuals, 50 natal men (31.25%), 110 natal women (68.75%)
M (SD) = mean (standard deviation); UGDS = Utrecht Gender Dysphoria Scale; CGAS = Children's Global Assessment Scale; GnRHa = gonadotropin-releasing hormone analogs

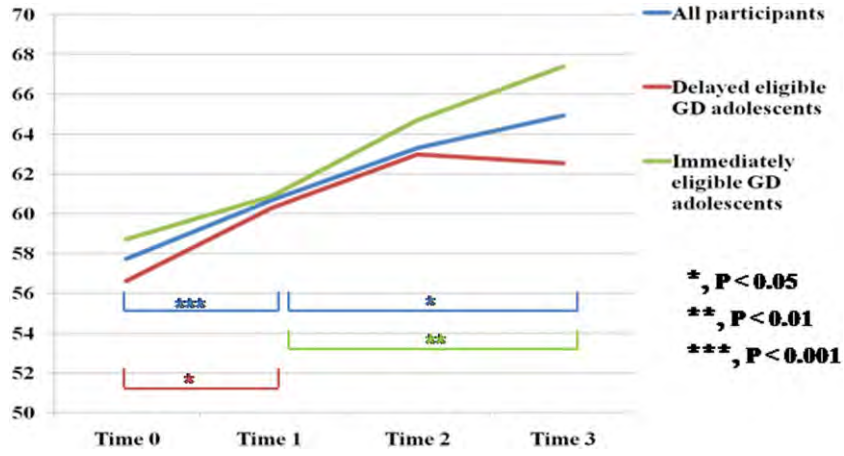


Figure 2 Gender dysphoria adolescents' psychosocial functioning (CGAS) at baseline, after psychological support, and after puberty suppression

CGAS, Children's Global Assessment Scale; Time 0, baseline; Time 1, 6 months from baseline (after 6 months of psychological support); Time 2, 12 months from baseline (delayed eligible gender dysphoria [GD] adolescents, after 12 months of psychological support; immediately eligible GD adolescents, after 12 months of psychological support + 6 months of puberty suppression); Time 3, 18 months from baseline (delayed eligible GD adolescents, after 18 months of psychological support; immediately eligible GD adolescents, after 18 months of psychological support + 12 months of puberty suppression)

CGAS at Baseline

GD adolescents' CGAS at baseline (Time 0, $M = 57.7 \pm 12.3$) revealed a score suggestive of "variable functioning with sporadic difficulties or symptoms in several but not all social areas" (range 50–59). Natal men had a significantly lower functioning than natal women at baseline ($P = 0.03$; Table 1). CGAS scores were not associated with any demographic variable, in both natal men and women (all $P > 0.1$). GD adolescents' CGAS scores at baseline were significantly lower ($t = 7.4$, $P < 0.001$) than that found in a sample of children/adolescents without observed psychological/psychiatric symptoms ($N = 169$, 67.1 ± 12) [16].

CGAS at Follow-Up

Compared with baseline, GD adolescents' psychosocial functioning was increasingly higher at each of the following evaluations (Figure 2). In particular, CGAS scores were significantly higher after 6 months of psychological support (Time 0 vs. Time 1, $P < 0.001$). Also there was a further significant improvement 18 months from baseline (Time 1 vs. Time 3, $P = 0.02$; Table 2).

Delayed eligible GD adolescents, who received only psychological support for the entire duration of the study, had a significantly better psychosocial functioning after six months of psychological support (Time 0 vs. Time 1, $P = 0.05$). However, despite scoring better at the following evaluations they did not show any further significant improvement in their psychosocial functioning (Table 2). Also, the delayed eligible group continued to score lower than a sample of children/adolescents without observed psychological/psychiatric symptoms [16], even after 18 months of psychological support (Time 3, $t = 2.0$, $P = 0.04$).

On the contrary, the immediately eligible group, who at baseline had a higher, but not significantly different psychosocial functioning than the delayed eligible group, did not show any significant improvement after 6 months of psychological support. However, immediately eligible adolescents had a significantly higher psychosocial functioning after 12 months of puberty suppression compared with when they had received only psychological support (Time 1 vs. Time 3 $P = 0.001$; Table 2). Also, their CGAS scores after 12 months of puberty suppression (Time 3) coincided almost perfectly with those found in a sample of children/adolescents without observed psychological/psychiatric symptoms ($t = 0.01$, $P = 0.99$) [16].

There were no significant differences in CGAS scores between GD natal men and women in all the follow-up evaluations (all $P > 0.1$). Also delayed eligible and immediately eligible GD adolescents did not differ in their demographic variables (all $P > 0.1$). Finally, even if at the end of the

Table 2 Gender dysphoria adolescents' psychosocial functioning (CGAS) at baseline, after psychological support, and after puberty suppression

	Time 0	Time 1	Time 2	Time 3	Statistical comparisons
	N	N	N	N	
	M/F ratio	M/F ratio	M/F ratio	M/F ratio	
	M (SD)	M (SD)	M (SD)	M (SD)	*t*-test; *P* value
All participants	N = 201 1:1.6 57.73 (12.27)	N = 201 1:1.6 60.68 (12.47)	N = 121 1:1.6 63.31 (14.41)	N = 71 1:1.6 64.93 (13.85)	4.87*; <0.001 3.70†; <0.001 4.11‡; <0.001 1.73§; 0.08 2.40¶; 0.02 0.76**; 0.45
Delayed eligible GD adolescents	N = 100 1:1.6 56.63 (13.14)	N = 100 1:1.6 60.29 (12.81)	N = 61 1:1.6 62.97 (14.10)	N = 36 1:1.6 62.53 (13.54)	1.99*; 0.05 2.89†; 0.005 2.29‡; 0.02 1.24§; 0.22 0.89¶; 0.37 0.15**; 0.88
Immediately eligible GD adolescents	N = 101 1:1.7 58.72 (11.38)	N = 101 1:1.7 60.89 (12.17)	N = 60 1:1.7 64.70 (13.34)	N = 35 1:1.7 67.40 (13.93)	1.31*; 0.19 3.02†; 0.003 3.66‡; <0.001 1.85§; 0.07 2.63¶; 0.001 0.94**; 0.35
Statistical comparisons *t*-test; *P* value	1.21††; 0.23	0.34††; 0.73	0.69†; 0.49	1.49†; 0.14	

*Comparison between baseline and Time 1
†Comparison between baseline and Time 2
‡Comparison between baseline and Time 3
§Comparison between Time 1 and Time 2
¶Comparison between Time 1 and Time 3
**Comparison between Time 2 and Time 3
††Comparison between delayed eligible GD adolescents and immediately eligible GD adolescents
CGAS = Children's Global Assessment Scale; M/F = natal male/natal female; M (SD) = mean (standard deviation)

follow-up study (Time 3) the immediately eligible group had a 5-point higher CGAS score than the delayed eligible group, this difference failed to reach significance, possible because of sample size (Table 2).

Discussion

Results from this study indicate that psychological support is associated with a better psychosocial functioning in GD adolescents, especially if presenting psychological/psychiatric problems. Moreover, puberty suppression was associated with a further improvement in global functioning. Finally, global functioning improved steadily over time in GD adolescents receiving both psychological support and GnRHa.

Medical and surgical interventions are considered to be necessary components of effective management in GD adults. These partially reversible/irreversible treatments aim to align the individuals' physical appearance with their internal gender identity and have been shown to improve the patients' psychosocial well-being [3,5,6]. GD adolescents may experience psychosocial problems at puberty onset because of an intensification of feelings of incongruence between self-perception and their natal gender [2,9]. Therefore, in the prepubertal population, the suppression of puberty using continuous GnRHa is a fully reversible treatment which has the fundamental benefit for children of gaining time to reflect over their gender identity, have a real-life experience living as the other gender (i.e., in dress and behavior) and determine whether or not they desire the transition [12,13]. Preventing the development of a body contrary to the experienced gender, puberty suppression allows GD adolescents to experience a smooth transition into their desired gender role. This translates into an improvement in many aspects of their psychosocial functioning, such as mood improvement and school integration [2,9]. Consistently, these results underline the importance of puberty suppression for GD adolescents' well-being.

The GD adolescents' improved global functioning after only 6 months of psychological support may have different explanations. First, it

could indicate that the timely addressing of psychosocial problems contributes to enhanced psychological well-being. Second, as also reported in previous studies among both GD adults and adolescents [2,3,5,9], our clinical experience suggests that patients attending a gender unit are pleased in the knowledge that the puberty suppression will be performed within a reasonable time and refer a distress reduction because of their accepted and understood requirements. Moreover, the initiation of the puberty suppression may have a psychological meaning which *per se* could be fundamental in reducing distress. In any case, data are too limited to express conclusively.

Both natal men and women benefited from the clinical approach, although natal men had a significantly worse functioning than natal women at baseline. It is even more important if we consider that natal men reported more social difficulties than natal women (higher dropout from school and more frequently not living with their parents). Interestingly, natal women reported significantly more GD-related discomfort than natal men. As already suggested [2], with a mean of 15 years most natal women had developed their breasts and had their menarche, which are likely to be associated with higher levels of distress. Therefore, natal men and women may need to be thought about separately and may require different interventions. Also, as the revised Dutch model [8] encourages considering GD individuals eligible for puberty suppression when they are 12 years old, studies are ongoing at our service to explore the possible benefit of further reducing the age for being eligible for puberty suppression. Even if the absence of a control group in our study does not allow us to pronounce conclusively on these comparisons, GD adolescents undergoing puberty suppression in addition to the psychological support result in psychosocial functioning levels that are impossible to differentiate from a sample of peers. These additional findings further indicate the effectiveness of both psychological support and puberty suppression in enabling young GD individuals to reach a satisfactory psychosocial functioning.

In the present study, there are some limitations. Even if psychosocial functioning is of crucial importance to identify clinical or sociocognitive difficulties [17], we focused only on a measure of psychosocial well-being. Also, the study sample was relatively small and came from only one clinic. Most importantly, despite the findings seem to suggest a cumulative and increasing over time positive effect of psychological support and GnRHa on young GD patients' well-being, results could have also different explanations because of the study design. For instance, getting older has been positively associated with maturity and well-being [18]. Ideally, a blinded randomized controlled trial design should have been performed. However, it is highly unlikely that adolescents would be motivated to participate. Also, disallowing puberty suppression, resulting in irreversible development of secondary sex characteristics, may be considered unethical [2]. Moreover, we cannot be conclusive on the higher GD-related distress in natal women compared with natal men. There are different versions of the UGDS scale for men and women, with specific items reversely coded because of gender. These differences do not allow drawing strong conclusions from the gender difference analysis.

Conclusions

In conclusion, this study confirms the effectiveness of puberty suppression for GD adolescents. Recently, a long-term follow-up evaluation of puberty suppression among GD adolescents after CSHT and GRS has demonstrated that GD adolescents are able to maintain a good functioning into their adult years [2]. The present study, together with this previous research [2], indicate that both psychological support and puberty suppression enable young GD individuals to reach a psychosocial functioning comparable with peers.

Corresponding Author: Rosalia Costa, MD, Gender Identity Development Service, Tavistock and Portman NHS Foundation Trust, Tavistock Centre, 120 Belsize Lane, London NW3 5BA, UK. Tel: +447947213589; Fax: + 39-0805593058; E-mail: rcosta@tavi-port.nhs.uk

Conflict of Interest: The author(s) report no conflicts of interest.

Statement of Authorship

Category 1

(a) Conception and Design
Rosalia Costa, Michael Dunsford, Elin Skagerberg, Victoria Holt, Polly Carmichael, Marco Colizzi

(b) Acquisition of Data
Rosalia Costa, Michael Dunsford, Elin Skagerberg, Victoria Holt, Polly Carmichael, Marco Colizzi

(c) Analysis and Interpretation of Data
Rosalia Costa, Polly Carmichael, Marco Colizzi

Category 2

(a) Drafting the Article
Rosalia Costa, Polly Carmichael, Marco Colizzi

(b) Revising It for Intellectual Content
Rosalia Costa, Michael Dunsford, Elin Skagerberg, Victoria Holt, Polly Carmichael, Marco Colizzi

Category 3

(a) Final Approval of the Completed Article
Rosalia Costa, Michael Dunsford, Elin Skagerberg, Victoria Holt, Polly Carmichael, Marco Colizzi

References

1 American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 5th edition. Washington, DC: American Psychiatric Association; 2013.
2 de Vries AL, McGuire JK, Steensma TD, Wagenaar EC, Doreleijers TA, Cohen-Kettenis PT. Young adult psychological outcome after puberty suppression and gender reassignment. Pediatrics 2014;134:696–704.
3 Colizzi M, Costa R, Pace V, Todarello O. Hormonal treatment reduces psychobiological distress in gender identity disorder, independently of attachment style. J Sex Med 2013;10:3049–58.
4 Coleman E, Bockting W, Botzer M, Cohen-Kettenis P, DeCuypere G, Feldman J, Fraser L, Green J, Knudson G, Meyer WJ, Monstrey S, Adler RK, Brown GR, Devor AH, Ehrbar R, Ettner R, Eyler E, Garofalo R, Karasic DH, Lev AI, Mayer G, Meyer-Bahlburg H, Hall BP, Pfaefflin F, Rachlin K, Robinson B, Schechter LS, Tangpricha W, van Trotsenburg M, Vitale A, Winter S, Whittle S, Wylie KR, Zucker K. Standards of care (SOC) for the health of transsexual, transgender, and gender nonconforming people, 7th version. Int J Transgender 2012;13:165–232.
5 Colizzi M, Costa R, Todarello O. Transsexual patients' psychiatric comorbidity and positive effect of cross-sex hormonal treatment on mental health: Results from a longitudinal study. Psychoneuroendocrinology 2014;39:65–73.
6 Colizzi M, Costa R, Todarello O. Dissociative symptoms in individuals with gender dysphoria: Is the elevated prevalence real? Psychiatry Res 2015;226:173–80.
7 Giordano S. Lives in a chiaroscuro. Should we suspend the puberty of children with gender identity disorder? J Med Ethics 2008;34:580–4.
8 Delemarre-van de Waal HA, Cohen-Kettenis PT. Clinical management of gender identity disorder in adolescents: A protocol on psychological and paediatric endocrinology aspects. Eur J Endocrinol 2006;155(suppl 1):S131–7.
9 De Vries AL, Steensma TD, Doreleijers TA, Cohen-Kettenis PT. Puberty suppression in adolescents with gender identity disorder: A prospective study. J Sex Med 2011;8:2276–83.
10 Kreukels BP, Cohen-Kettenis PT. Puberty suppression in gender identity disorder: The Amsterdam experience. Nat Rev Endocrinol 2011;7:466–72.
11 Di Ceglie D. Management and therapeutic aims with children and adolescents with gender identity disorders and their families. In: Di Ceglie D, Freedman D, eds. A stranger in my own body: Atypical gender identity development and mental health. London: Karnac; 1998:185–97.
12 Cohen-Kettenis PT, Delemarre-van de Waal HA, Gooren LJ. The treatment of adolescent transsexuals: Changing insights. J Sex Med 2008;5:1892–7.
13 Cohen-Kettenis PT, van Goozen SHM. Sex reassignment of adolescent transsexuals: A follow-up study. J Am Acad Child Adolesc Psychiatry 1997;36:263–71.
14 Shaffer D, Gould MS, Brasic J, Ambrosini P, Fisher P, Bird H, Aluwahlia S. A Children's Global Assessment Scale (CGAS). Arch Gen Psychiatry 1983;40:1228–31.
15 Schorre BE, Vandvik IH. Global assessment of psychosocial functioning in child and adolescent psychiatry. A review of three unidimensional scales (CGAS, GAF, GAPD). Eur Child Adolesc Psychiatry 2004;13:273–86.
16 Lundh A, Forsman M, Serlachius E, Lichtenstein P, Landén M. Outcomes of child psychiatric treatment. Acta Psychiatr Scand 2013;128:34–44.
17 Prince M, Glozier N, Sousa R, Dewey M. Measuring disability across physical, mental and cognitive disorders. In: Regier DA, Narrow WE, Kuhl EA, Kupfer DJ, eds. The conceptual evolution of DSM-5. Washington, DC: American Psychiatric Publishing; 2011:189–227.
18 Sheldon KM, Kasser T. Getting older, getting better? Personal striving and psychological maturity across the life span. Dev Psychol 2001;37:491–501.

DE GRUYTER J Pediatr Endocrinol Metab 2019; 32(11): 1235–1239

Shanlee M. Davis*, Jill L. Kaar, Brandy M. Ringham, Christine W. Hockett, Deborah H. Glueck and Dana Dabelea

Sex differences in infant body composition emerge in the first 5 months of life

https://doi.org/10.1515/jpem-2019-0243
Received June 4, 2019; accepted July 26, 2019; previously published online September 4, 2019

Abstract

Background: Sex differences in body composition are appreciated throughout the lifespan with probable contributions from sex steroids: testosterone and estrogen. The purpose of this longitudinal observational study was to determine if sex differences in body composition emerge during the first months of life in healthy infants, corresponding to the age at which male infants produce endogenous testosterone.
Methods: Linear growth and body composition parameters using air displacement plethysmography were obtained from 602 healthy infants after birth and again at 5 months of age. Rate of change in body composition parameters were compared between sexes.
Results: Sex differences in length, total mass, fat free mass (FFM), and percent fat mass (%FM) were present both at birth and at 5 months ($p < 0.001$ for all), with males having greater total mass and FFM but lower %FM. Gain in %FM over the first 5 months was significantly lower in males ($p = 0.0004$). This difference was secondary to a gain of 17 g/week more in FFM in males compared to females.
Conclusions: Sex differences in body composition emerge in the first months of life, with lower adiposity accumulation in males. Endogenous testosterone production in males ~1–4 months of age may account for findings and may have lifelong implications for sex differences in body composition.

Keywords: adipose tissue; body composition; body fat percentage; critical window; mini-puberty; PEA POD.

***Corresponding author: Shanlee M. Davis,** MD, MS, University of Colorado, Department of Pediatrics, Denver, CO 80203, USA; and Children's Hospital Colorado, 13123 East 16th Ave B265, Aurora, CO 80045, USA, Phone: +720-777-6128, Fax: +720-777-7301, E-mail: Shanlee.davis@childrenscolorado.org.
https://orcid.org/0000-0002-0304-9550
Jill L. Kaar and Deborah H. Glueck: University of Colorado, Department of Pediatrics, Aurora, CO, USA
Brandy M. Ringham: University of Colorado, Lifecourse Epidemiology of Adiposity and Diabetes Center, Aurora, CO, USA
Christine W. Hockett: University of Colorado, Department of Epidemiology, Aurora, CO, USA
Dana Dabelea: University of Colorado, Lifecourse Epidemiology of Adiposity and Diabetes Center, Aurora, CO, USA; and University of Colorado, Department of Epidemiology, Aurora, CO, USA

Background

Sex differences in body composition are recognized throughout the lifespan, with females having significantly lower lean mass and greater percent fat mass (%FM) than males. Sex steroids (testosterone and estrogen) are assumed to be largely responsible for these differences; however, sex differences are observed in childhood well before puberty occurs [1]. While some studies report that sex differences are present at birth [2–4] others have found no differences between males and females [5, 6]. By 1 year of age, females have greater %FM and this difference seems to persist until puberty at which time it becomes further exaggerated. Although sex differences in body composition are known to exist, the underlying etiology and timing of emergence are surprisingly understudied.

The mini-puberty period of infancy is a biologically plausible time during which sex differences in body composition may emerge. The mini-puberty period refers to temporary activation of the hypothalamic-pituitary-gonadal axis resulting in near-adult concentrations of testosterone in boys from around 1–4 months of age [7]. Testosterone is an anabolic steroid that acts to increase protein synthesis and lean mass while decreasing fat mass (FM). Although the existence of the mini-puberty period is well known by pediatric endocrinologists, the short and long-term consequences of this temporary testosterone exposure are just beginning to be unveiled. It was recently shown that there are sex differences in linear growth velocity during the first 6 months of life associated with testosterone concentrations [8]. This difference in infancy was calculated to be substantial enough to account for 15% of the final height differences between men and women [9]. Animal models have found that blocking the post-natal testosterone surge in male mice results in greater adiposity lifelong [10]. The effects of sex steroid exposure during

the mini-puberty period on body composition in humans have not been studied.

The objective of this secondary data analysis was to determine if sex differences in body composition emerge during the first months of life at a sensitive time period at which sex steroid exposure is known to be substantially different between the sexes. We hypothesize that differences in body composition emerge during the first 5 months of life, with males developing more lean mass and lower %FM in a pattern consistent with testosterone exposure.

Materials and methods

Between 2009 and 2014, the Healthy Start Study enrolled 1410 mother-offspring dyads before 24 weeks' gestation from obstetrics clinics at the University of Colorado Hospital. Detailed assessment of the study participants and methods have been previously published [11–14]. In brief, sociodemographic data, maternal medical history, pregnancy and birth history, and infant feeding methods were collected. Infant growth and body composition parameters were obtained at birth and again at 4–6 months of age. For this analysis, infants had to be born term (≥37 weeks gestation) and have complete data on body composition available for both the neonatal visit (<3 days of age) and the infant visit (~5 months of age, range 3–7 months). The study was approved by the Colorado Multiple Institutional Review Board and all women provided written informed consent for study participation. The Healthy Start Study was registered at clinicaltrials.gov (NCT02273297).

Mothers self-reported their race/ethnicity, tobacco use, and infant feeding status. Maternal race was categorized as Non-Hispanic White, Hispanic, Non-Hispanic Black, and Other, which included Asian or Pacific Islander and American Indian or Alaskan Native. The total number of breastfeeding months was calculated as previously described [12]. Pre-pregnancy body mass index (BMI), maternal weight gained during pregnancy, gestational age at birth, and infant birthweight were abstracted from medical records.

Infant length was measured supine on a length board to the nearest tenth of a centimeter (cm) by two trained research personnel. Body composition was assessed using air displacement plethysmography (PEA POD, COSMED, Rome, Italy). The PEA POD is a validated instrument that measures total body mass in grams, total body volume, and estimates FM, fat-free mass (FFM), and %FM [15]. Each infant was measured twice, with a third measurement obtained when %FM differed by >2.0%; the average of the two closest readings was used for analysis. The rate of change for each of the following parameters: length, %FM, FM, and FFM, was calculated as the measurement at ~5 months minus the measurement at birth divided by the amount of time between measures in weeks.

Statistical analysis

Descriptive statistics were used to summarize baseline characteristics of the sample as well as body composition parameters at birth and at 5 months of age stratified by sex. Change in body composition parameters between time points were assessed with a paired t-test. Sex differences were evaluated using t-tests for continuous variables and the Cochran-Mantel-Haenszel test for categorical variables.

Separate general linear univariate multivariable models were fit for each outcome with sex of the offspring as the main predictor. *A priori* covariates included race and the measure of the outcome near birth (i.e. %FM near birth was included as covariate when the rate of change of %FM was the outcome). Although several other covariates are known to affect infant adiposity, including maternal BMI, weight gain during pregnancy, smoking, and infant feeding source, these variables are not associated with infant sex or on the causal pathway between sex and change in body composition so they were not included in the model. We assessed the significance of the covariates using an F test at an α level of 0.05. We examined the studentized residuals to ensure we met the model assumptions of normality and homoscedasticity. All model assumptions were met and there were no overly influential observations. The final model for each outcome included race and the measure of the outcome at birth as covariates. The estimates of association, p-values, and 95% confidence intervals are presented. Statistical analyses were conducted in SAS 9.4 (SAS Institute, Cary, NC, USA).

Results

Out of the full 1410 infant cohort, we excluded infants who did not have outcome measures at both time points (n = 673), as well as infants born before 37 weeks' gestation (n = 11), infants who had their initial PEA POD measurement after 3 days of age, or second PEA POD more than 209 days after birth (n = 124). Therefore, the analytic cohort included the remaining n = 602 participants. Maternal and infant characteristics were similar between the full cohort and the analytic cohort, similar to that previously reported [12]. The mean age of the cohort at the first visit was 1.5 days and at the second visit was 5 months. Demographic data were similar between males and females (Table 1). Sex differences in length, FM, FFM, and %FM were present at birth for all parameters ($p < 0.001$, Table 2). As expected, length, FM, FFM, and %FM all increased between birth and 5 months in both males and females ($p < 0.001$ for all, Table 1). At 5 months of age, FFM and %FM remained significantly different between males and females, with relative total body adiposity now 8.7% lower in males compared to females [(25.47–23.25)/25.47, $p < 0.0001$]. FM was no longer significantly different between sexes at 5 months of age ($p = 0.12$).

Sex differences in the rate of change per week in body composition parameters are shown in Table 3. On average, after adjusting for race/ethnicity, FFM increased 17 g more per week in males compared to females ($p < 0.0001$, 95% CI: 14, 20), yielding a difference of 410 g of FFM between sexes at 5 months. The rate of FM gain was not significantly

Table 1: Participant demographics (mean ± SD, n [%]).

	Total sample (n = 602)	Females (n = 306)	Males (n = 296)
Maternal characteristics			
Race			
Hispanic	147 (24%)	74 (24%)	73 (25%)
Non-Hispanic White	341 (57%)	173 (57%)	168 (57%)
Non-Hispanic Black	83 (14%)	42 (14%)	41 (14%)
Other	31 (5%)	17 (6%)	14 (5%)
Pre-pregnancy BMI, kg/m^2	25.84 ± 6.36	25.68 ± 6.20	26.00 ± 6.52
Gestational weight gain, kg	13.65 ± 6.38	13.65 ± 6.68	13.65 ± 6.06
Smoking during pregnancy, n	38 (6%)	18 (6%)	20 (7%)
Infant characteristics			
Gestational age, weeks	39.54 ± 1.15	39.57 ± 1.16	39.50 ± 1.13
Birthweight, g	3277 ± 426	3220 ± 410	3330 ± 428
Age at PEA POD #1, days	1.12 ± 0.53	1.17 ± 0.56	1.08 ± 0.49
Age at PEA POD #2, months	4.94 ± 0.93	4.98 ± 0.92	4.89 ± 0.94
Breastfeeding months	3.48 ± 1.81	3.47 ± 1.80	3.49 ± 1.83

Table 2: Body composition measures (mean ± SD) at birth and at 5 months by infant sex.

	Total sample (n = 602)	Females (n = 306)	Males (n = 296)	p-Value
Body composition at birth				
Total body mass, kg	3.12 ± 0.41	3.07 ± 0.40	3.17 ± 0.41	0.001
Length, cm	49.18 ± 2.04	48.75 ± 1.97	49.63 ± 2.02	<0.0001
Fat free mass, kg	2.83 ± 0.33	2.76 ± 0.31	2.91 ± 0.33	<0.0001
Fat mass, kg	0.29 ± 0.14	0.31 ± 0.14	0.27 ± 0.13	0.0002
% Fat mass	9.02 ± 3.72	9.83 ± 3.85	8.18 ± 3.39	<0.0001
Body composition at 5 months				
Total body mass, kg	6.74 ± 0.83	6.57 ± 0.82	6.92 ± 0.81	<0.0001
Length, cm	63.74 ± 2.66	63.09 ± 2.54	64.44 ± 2.62	<0.0001
Fat free mass, kg	5.08 ± 0.57	4.88 ± 0.53	5.29 ± 0.54	<0.0001
Fat mass, kg	1.66 ± 0.50	1.70 ± 0.50	1.63 ± 0.50	0.12
% Fat mass	24.38 ± 5.42	25.47 ± 5.32	23.25 ± 5.30	<0.0001

Table 3: Sex differences in the rate of change per week in body composition parameters.

Outcome	β Coefficient (males relative to females)	Standard error	p-Value
Rate of change in percent fat mass, %	−0.09	0.02	0.0004
Rate of change in fat free mass, g/week	16.59	1.65	<0.0001
Rate of change in fat mass, g/week	−1.07	2.00	0.59
Rate of change in length, cm/week	0.05	0.01	<0.0001

different between sexes. For males, %FM increased 0.09 percentage points less per week than females (p = 0.0004, 95% CI: 0.04, 0.13), which equates to an 8% difference between sexes in adiposity gain during the first 5 months of life. The length increased 0.05 cm more per week (2.6 cm/year) in males compared to females (p < 0.0001, 95% CI: 0.04, 0.07) during the first 5 months of life.

Discussion

In this large, ethnically diverse cohort of term infants, we have shown that while sex differences in body composition are already present at birth, the magnitude of these differences increases during the first 5 months of life. The gain in adiposity was 8% lower in males compared to females, and this was due to a greater gain in absolute FFM in males vs. females. We also confirmed that the linear growth velocity is greater in males during this time period. These early differences in growth patterns between males and females are important to consider as we further define the critical role of programming

in the early post-natal period on future risk for adult diseases [16].

Sex differences in adiposity are well recognized in adults, with males having less total adipose tissue but disproportionately more visceral adipose tissue [17]. This sexual dimorphism in body composition has been demonstrated even in pre-pubertal children, although to a much smaller extent than after puberty [1, 18, 19]. Adipose tissue was once thought to be a storage depot for lipids but it is now recognized as an endocrine organ that regulates metabolism [20]. Adipose tissue dysfunction is strongly associated with systemic inflammation, insulin resistance, and cardiovascular disease. Furthermore, accumulation of adiposity in the early infancy period has been associated with later childhood, adolescent, and adult obesity, supporting the model that programming during critical windows of development and plasticity influence lifelong metabolism [21, 22]. Despite this, there has been minimal investigation comparing differences in early growth between sexes, and our understanding of the origins of later sex differences is largely speculative. This study specifically compares longitudinal changes in body composition during this critical time period between males and females.

The mini-puberty period of infancy refers to temporary activation of the hypothalamic-pituitary-gonadal axis and production of sex steroids (testosterone and estrogen) in infancy [7]. The time course is more well-defined in males, peaking around 6–8 weeks of age and typically ending by 4–5 months of age. Testosterone is an anabolic hormone that increases lean mass as well as linear growth velocity. Kiviranta et al. found sex differences in growth velocity in 18,570 infants from the UK, with males growing 2–4 cm/year more than females in the first 6 months of life and no sex differences in growth velocity after 6 months of age [8]. Our data corroborate these findings with a difference between sexes of 2.6 cm/year in the first 5 months of life. These results support the importance of the first few months of life in establishing early sex differences in height. Beyond height, the body composition changes we observed in males in this study are identical to what is expected from testosterone exposure: greater FFM, lower %FM, and higher linear growth velocity compared to females. To confirm these observed changes in body composition parameters were not secondary to the change in length, the analyses were repeated with change in length as a covariate in the model and the outcomes were unchanged. Although we cannot definitively conclude that these differences are due to testosterone, the body composition changes we observed during the time course of the normal mini-puberty period are consistent with the known effects of testosterone. The implications of the mini-puberty period and its coincidental timing with the hypothesized critical programming in the first months of life are just beginning to be investigated.

The major strengths of this study are the large, ethnically diverse cohort of infants with a rigorous assessment of body composition. Limitations of this study include no measures to directly attribute the observed sex differences to testosterone and data limited at this time to the first 6 months of life (though additional follow-up is underway). In addition, as the PEA POD assesses total body adiposity we are unable to evaluate sex differences in adiposity depots, such as visceral vs. subcutaneous fat, which is recognized to be sex specific and have a strong correlation with cardiometabolic disease states in adults [23]. Despite these limitations, this study is the first to demonstrate that sex differences in body composition increase during these early months of life and provide justification for further investigation into the mechanisms and clinical implications of these findings.

In conclusion, our study confirms there are sex differences emerging in the rate of linear growth during early infancy, and also provides novel data supporting emerging sex differences in body composition. Sex differences in body composition are present at birth but significantly widen over the first 5 months of life. FFM increases significantly more in males than females, while FM gain is similar between sexes resulting in lower adiposity gain in males. Although we cannot confirm this is secondary to the testosterone surge occurring in healthy male infants during this time period with our dataset, this pattern of change in body composition is congruent with what would be expected from testosterone exposure. The finding supports further study of the physical sex differences emerging during the mini-puberty period of infancy that may have lifelong implications in health and disease.

Author contributions: Davis wrote the first draft of the manuscript and each co-author listed contributed to and approved the final manuscript. All authors agree to be accountable for all aspects of the work.
Research funding: This work was funded by NIH/NIDDK Funder Id: http://dx.doi.org/10.13039/100000062, R01DK076648 (Dabelea), NIH/OD Funder Id: http://dx.doi.org/10.13039/100000052, UH3OD023248 (Dabelea), NIH/NIGMS Funder Id: http://dx.doi.org/10.13039/100000057, R01GM121081 (Glueck, Muller, Dabelea), and NIH/NICHD Funder Id: http://dx.doi.org/10.13039/100009633, K23HD092588 (Davis).
Employment or leadership: None declared.
Honorarium: None declared.

Competing interests: The funding organization(s) played no role in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the report for publication.

Statement of ethics: Parents of participants have given their written informed consent. The study was approved by the research institute's committee on human research. The authors have no ethical conflicts to disclose.

Disclosure statement: The authors have no conflicts of interest to declare.

References

1. Hanks LJ, Casazza K, Alvarez JA, Fernandez JR. Does fat fuel the fire: independent and interactive effects of genetic, physiological, and environmental factors on variations in fat deposition and distribution across populations. J Pediatr Endocrinol Metab 2010;23:1233–44.
2. Butte NF, Hopkinson JM, Wong WW, Smith EO, Ellis KJ. Body composition during the first 2 years of life: an updated reference. Pediatr Res 2000;47:578–85.
3. Fields DA, Gilchrist JM, Catalano PM, Gianni ML, Roggero PM, et al. Longitudinal body composition data in exclusively breast-fed infants: a multicenter study. Obesity 2011;19:1887–91.
4. Fomon SJ, Nelson SE. Body composition of the male and female reference infants. Annu Rev Nutr 2002;22:1–17.
5. Andersen GS, Girma T, Wells JC, Kaestel P, Leventi M, et al. Body composition from birth to 6 mo of age in Ethiopian infants: reference data obtained by air-displacement plethysmography. Am J Clin Nutr 2013;98:885–94.
6. Jain V, Kurpad AV, Kumar B, Devi S, Sreenivas V, et al. Body composition of term healthy Indian newborns. Eur J Clin Nutr 2016;70:488–93.
7. Rey RA. Mini-puberty and true puberty: differences in testicular function. Ann Endocrinol (Paris) 2014;75:58–63.
8. Kiviranta P, Kuiri-Hanninen T, Saari A, Lamidi ML, Dunkel L, et al. Transient postnatal gonadal activation and growth velocity in infancy. Pediatrics 2016;138:e2015356.
9. Copeland KC, Chernausek S. Mini-puberty and growth. Pediatrics 2016;138:e20161301.
10. Nohara K, Zhang Y, Waraich RS, Laque A, Tiano JP, et al. Early-life exposure to testosterone programs the hypothalamic melanocortin system. Endocrinology 2011;152:1661–9.
11. Perng W, Ringham BM, Glueck DH, Sauder KA, Starling AP, et al. An observational cohort study of weight- and length-derived anthropometric indicators with body composition at birth and 5 mo: the Healthy Start study. Am J Clin Nutr 2017;106:559–67.
12. Sauder KA, Kaar JL, Starling AP, Ringham BM, Glueck DH, et al. Predictors of infant body composition at 5 months of age: the Healthy Start Study. J Pediatr 2017;183:94–9.e1.
13. Harrod CS, Chasan-Taber L, Reynolds RM, Fingerlin TE, Glueck DH, et al. Physical activity in pregnancy and neonatal body composition: the Healthy Start study. Obstet Gynecol 2014;124(2 Pt 1):257–64.
14. Harrod CS, Reynolds RM, Chasan-Taber L, Fingerlin TE, Glueck DH, et al. Quantity and timing of maternal prenatal smoking on neonatal body composition: the Healthy Start Study. J Pediatr 2014;165:707–12.
15. Ma G, Yao M, Liu Y, Lin A, Zou H, et al. Validation of a new pediatric air-displacement plethysmograph for assessing body composition in infants. Am J Clin Nutr 2004;79:653–60.
16. Koletzko B, Brands B, Chourdakis M, Cramer S, Grote V, et al. The Power of Programming and the EarlyNutrition project: opportunities for health promotion by nutrition during the first thousand days of life and beyond. Ann Nutr Metab 2014;64:187–96.
17. Palmer BF, Clegg DJ. The sexual dimorphism of obesity. Mol Cell Endocrinol 2015;402:113–9.
18. Karlsson AK, Kullberg J, Stokland E, Allvin K, Gronowitz E, et al. Measurements of total and regional body composition in preschool children: a comparison of MRI, DXA, and anthropometric data. Obesity 2013;21:1018–24.
19. Staiano AE, Broyles ST, Gupta AK, Katzmarzyk PT. Ethnic and sex differences in visceral, subcutaneous, and total body fat in children and adolescents. Obesity 2013;21:1251–5.
20. Kershaw EE, Flier JS. Adipose tissue as an endocrine organ. J Clin Endocrinol Metab 2004;89:2548–56.
21. Wells JC, Chomtho S, Fewtrell MS. Programming of body composition by early growth and nutrition. Proc Nutr Soc 2007;66:423–34.
22. Varvarigou AA. Intrauterine growth restriction as a potential risk factor for disease onset in adulthood. J Pediatr Endocrinol Metab 2010;23:215–24.
23. Fried SK, Lee MJ, Karastergiou K. Shaping fat distribution: new insights into the molecular determinants of depot- and sex-dependent adipose biology. Obesity 2015;23:1345–52.

OPEN
International Journal of Obesity (2014) 38, S57–S66

www.nature.com/ijo

ORIGINAL ARTICLE

Physical fitness reference standards in European children: the IDEFICS study

P De Miguel-Etayo[1,2,3,16], L Gracia-Marco[1,4,16], FB Ortega[5,6], T Intemann[7], R Foraita[7], L Lissner[8], L Oja[9], G Barba[10], N Michels[11], M Tornaritis[12], D Molnár[13], Y Pitsiladis[14], W Ahrens[7,15] and LA Moreno[1,2] on behalf of the IDEFICS consortium

BACKGROUND/OBJECTIVES: A low fitness status during childhood and adolescence is associated with important health-related outcomes, such as increased future risk for obesity and cardiovascular diseases, impaired skeletal health, reduced quality of life and poor mental health. Fitness reference values for adolescents from different countries have been published, but there is a scarcity of reference values for pre-pubertal children in Europe, using harmonised measures of fitness in the literature. The IDEFICS study offers a good opportunity to establish normative values of a large set of fitness components from eight European countries using common and well-standardised methods in a large sample of children. Therefore, the aim of this study is to report sex- and age-specific fitness reference standards in European children.

SUBJECTS/METHODS: Children (10 302) aged 6–10.9 years (50.7% girls) were examined. The test battery included: the flamingo balance test, back-saver sit-and-reach test (flexibility), handgrip strength test, standing long jump test (lower-limb explosive strength) and 40-m sprint test (speed). Moreover, cardiorespiratory fitness was assessed by a 20-m shuttle run test. Percentile curves for the 1st, 3rd, 10th, 25th, 50th, 75th, 90th, 97th and 99th percentiles were calculated using the General Additive Model for Location Scale and Shape (GAMLSS).

RESULTS: Our results show that boys performed better than girls in speed, lower- and upper-limb strength and cardiorespiratory fitness, and girls performed better in balance and flexibility. Older children performed better than younger children, except for cardiorespiratory fitness in boys and flexibility in girls.

CONCLUSIONS: Our results provide for the first time sex- and age-specific physical fitness reference standards in European children aged 6–10.9 years.

International Journal of Obesity (2014) **38,** S57–S66; doi:10.1038/ijo.2014.136

INTRODUCTION

Physical fitness (hereinafter called just fitness) has been related to a person's ability to perform physical activities that require aerobic capacity, endurance, strength or flexibility, which seems to be linked to inherited and environmental factors.[1] Fitness has been considered a powerful marker of health, both, in childhood and in adulthood,[2] independent of physical activity.[3,4] A low fitness status during childhood and adolescence is associated with important health-related outcomes, such as increased future risk for obesity[5] and cardiovascular diseases,[6] impaired skeletal health,[7] reduced quality of life[8] and poor mental health.[9] In spite of the healthy benefits of a high fitness, children and adolescents performance in fitness tests has declined over the last three decades.[10]

The most frequently evaluated fitness component is cardiorespiratory fitness (CRF).[11] However, in the last years other fitness components such as flexibility, muscular fitness and speed/agility were evaluated and associated with health outcomes.[9]

Nowadays, fitness reference values in adolescents from different countries in America,[12–14] Asia,[10] Africa[11] and Europe[9] have been published. However, there is a scarcity of reference values for pre-pubertal children in Europe and in other continents, using harmonised measures of fitness in the literature.[14] Reference values are necessary to classify children and to monitor the fitness status of the population. The IDEFICS study (Identification and prevention of dietary and lifestyle-induced health effects in children and infants)[15] offers a good opportunity to establish normative values for a large set of fitness components using common and well-standardised methods in a large sample of children from eight European countries. Therefore, the aim of this study is to provide sex- and age-specific fitness reference standards for European children aged 6–9 years.

MATERIALS AND METHODS

Study design

The IDEFICS study[15] is a prospective cohort study with an embedded controlled intervention aiming to prevent childhood obesity in a community-oriented approach. IDEFICS is a multi-centre study on lifestyle

[1]GENUD (Growth, Exercise, NUtrition and Development) Research Group, Universidad de Zaragoza, Zaragoza, Spain; [2]Department of Psychiatry and Nursing, Faculty of Health Sciences, University of Zaragoza, Zaragoza, Spain; [3]Department of Paediatrics, Faculty of Medicine, Universidad de Zaragoza, Zaragoza, Spain; [4]CHERC (Children's Health and Exercise Research Centre), College of Life and Environmental Sciences, Sport and Health Sciences, University of Exeter, Exeter, UK; [5]PROFITH (PROmoting FITness and Health through physical activity) Research Group, School of Sports Science, University of Granada, Granada, Spain; [6]Department of Biosciences and Nutrition, Karolinska Institutet, Stockholm, Sweden; [7]Leibniz Institute for Prevention Research and Epidemiology—BIPS, Bremen, Germany; [8]Department of Public Health and Community Medicine, University of Gothenburg, Gothenburg, Sweden; [9]National Institute for Health Development, Tallinn, Estonia; [10]Epidemiology and Population Genetics, Institute of Food Science, National Research Council, Avellino, Italy; [11]Department of Public Health, Ghent University, Ghent, Belgium; [12]Research and Education Institute for Child Health, Strovolos, Cyprus; [13]Department of Pediatrics, University of Pécs Medical School, Pécs, Hungary; [14]Centre for Sport and Exercise Science and Medicine (SESAME), University of Brighton, Eastbourne, UK and [15]Institute of Statistics, Faculty of Mathematics and Computer Science, University of Bremen, Bremen, Germany. Correspondence: P De Miguel-Etayo, GENUD (Growth, Exercise, NUtrition and Development) Research Group, Universidad de Zaragoza, C/Pedro Cerbuna, 12, Zaragoza 50009, Spain.
E-mail: pilardm@unizar.es
[16]These authors contributed equally to this work.

and nutrition among children aged 2–9 years from eight European countries (Sweden, Germany, Hungary, Italy, Cyprus, Spain, Belgium and Estonia).

In the first survey, data collection took place from September 2007 to June 2008 (T0, baseline survey); 2 years later the follow-up survey was conducted from September 2009 to May 2010 (T1, follow-up survey), where the present analysis is based on cross-sectional data only. A detailed description of IDEFICS sampling and recruitment approaches, standardisation and harmonisation process, data collection, analysis strategies, quality control measures and inclusion/exclusion criteria have been already published.[15] The study was approved by the Research Ethics Committees of each institution and region involved. Parents signed an informed consent.

Study population

A cohort of 18 745 children aged 2–10.9 years was established including all children recruited at baseline and children newly recruited at thr first follow-up. Fitness was measured in children older than 6 years ($N = 10\,302$, 50.7% girls) fulfilling the inclusion criteria (having complete data on weight, height, age and sex). Because not all children took part in all components of the physical fitness battery, analysis groups and sample sizes vary for the different physical fitness tests (Figure 1). We compared these varying analysis groups with the overall study population of children older than 6 years. Since the groups of children who participated in the back-saver sit-and-reach test, the standing long jump or the handgrip strength test were nearly identical, we considered these children as one analysis group. The second analysis group consisted of children who participated in the 40-m sprint test; the third group of children who participated in the Flamingo balance test; and the children who conducted the shuttle run test formed the fourth group. Mean age varied between 7.6 and 7.7 years, and the percentage of boys was highest in the total study population with 49.4% and smallest (47.3%) in the third analysis group (Flamingo balance test). The prevalence of overweight and obese children

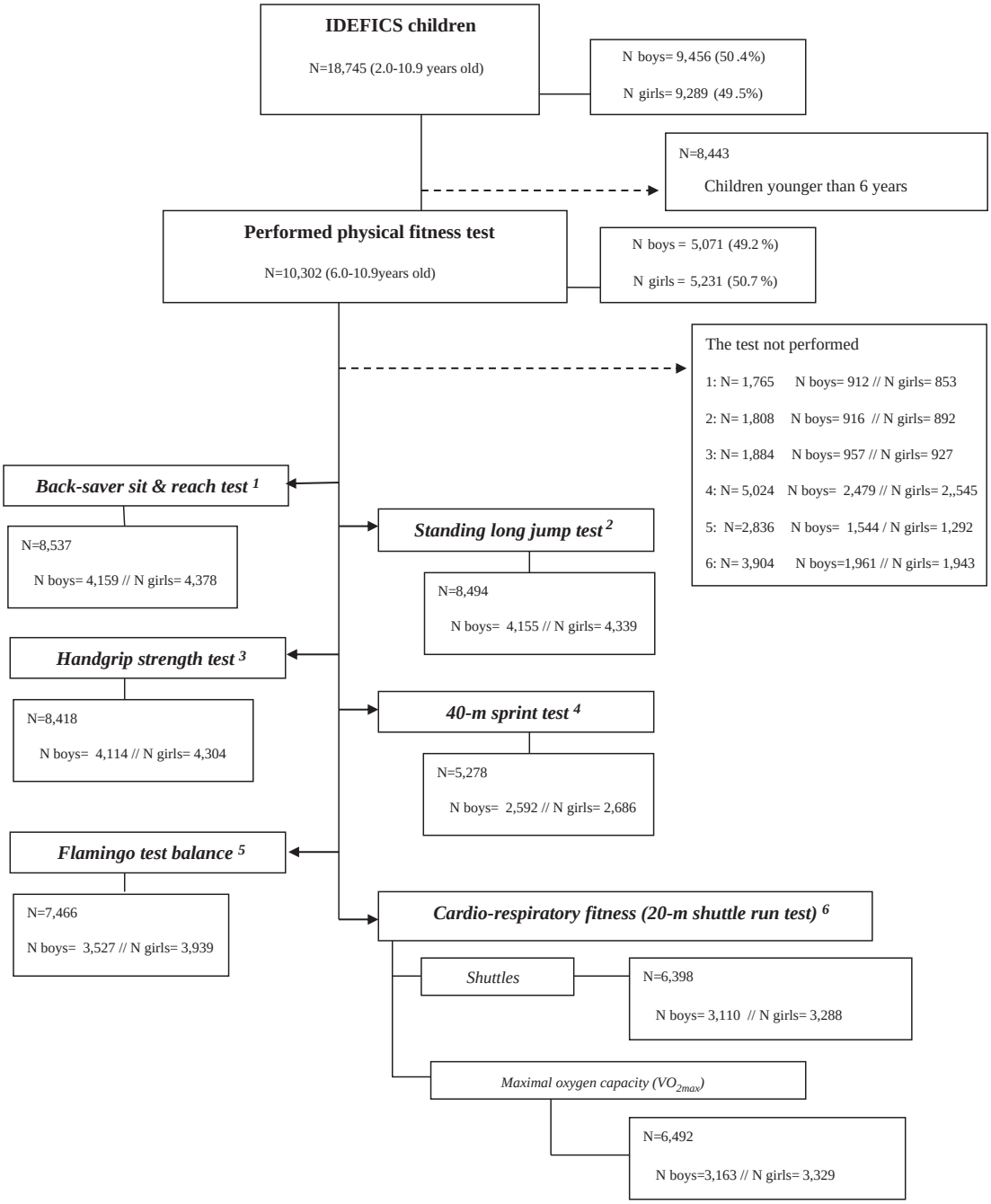


Figure 1. Flow chart of the population involved in this study from the IDEFICS study.

was highest (25.3%) for children of group 1 compared with 24.5% in the total study population and smallest (20.9%) in the children participating in the shuttle run test. The most pronounced differences were seen with respect to the ISCED (International Standard Classification of Education) level where the maximum level of both parents was considered. The distribution of ISCED levels was rather similar in the overall study population, analysis groups 1 and 3, but fewer children from families with ISCED level 0–2 participated in the 40-m sprint (4.5%; analysis group 2) and in the shuttle run test (6.8%; analysis group 4) as compared with 10.5% in the overall study population, 10.3% in group 1 and 10.0% in group 3.

Physical fitness

Components of the fitness tests were mostly adapted from the ALPHA health-related fitness test battery, and based on a published review their reliability has been shown in children and adolescents.[9,16–18] The test battery thus included: the flamingo balance test, back-saver sit-and-reach test, handgrip strength test, standing long jump test and 40-m sprint test. Moreover, CRF was assessed by the 20-m shuttle run test (20mSRT).[19]

The protocols used for fitness testing are described in detail below:

The flamingo balance test measures the ability to balance successfully on a single leg. The child has to bend his/her free leg backwards and grip the back foot with his/her hand on the same side, and stand like this for 1 min. The child is given one try before to become familiar with the test. Then the number of attempts needed to stand on one leg for 1 min is counted for each leg. Children were excluded if they had put down the other foot 15 times or more within the first 30 s. The test score is calculated as the sum of attempts with both legs; lower scores indicate better performance.

The back-saver sit-and-reach test measures the flexibility of the hamstring muscles and it was conducted in the FITNESSGRAM battery.[20] The test is performed with a standard box with a scale on the top. The participant is required to sit with the untested leg bent at the knee; the tested leg is placed straight with the foot placed against the box. In the back-saver sit-and-reach test, only one leg was evaluated at a time. The participant slowly reached forward as far as possible. The back-saver sit-and-reach test is similar to the traditional sit-and-reach test, except that the measurement is performed on one side at a time, so a specific score is obtained from each side. The score is calculated as the average of both sides; higher scores indicate better performance.

The handgrip strength test measures the maximal isometric force that can be generated mainly by the forearm. The child stays in a standard bipedal position with the arms in complete extension holding the dynamometer (TKK 5101; Takei, Tokyo, Japan) without touching any part of the body with it. The dynamometer is adjusted to sex and hand size for each child.[21] The measurement scale starts with 5 kg. Children who did not reach this minimum ($N = 5$; 0.1%) were scored as 2.5 kg (average 0–5 kg). The score is calculated as the average of right and left handgrip strength; higher scores indicate better performance.

The standing long jump test[22] assesses lower-limb explosive strength. The child jumps as far as possible off the stand, trying to land with both feet together and maintaining the equilibrium once landed (it was not allowed to put the hands on the floor). The score was obtained by measuring the distance between the last heel-mark and the take-off line. Two tries were allowed and the best score was retained. Higher scores indicate better performance.

The 40-m sprint test measures the maximum running speed of the child. This test is carried out along 40 m delimited by five marker cones aligned, within a distance of 10 m between each two neighbouring cones. With 3-m distance, five more marker cones are placed in parallel marking the running track. The child is instructed to run as quick as he/she can, after the starting signal. Two tries were allowed, the best score was retained. In this test lower scores indicate better performance.

The 20mSRT estimates the CRF (aerobic capacity). Children run back and forth between two lines 20-m apart following beep signals played from a pre-recorded CD. The test is finished when the child stops owing to fatigue or when he/she does not reach the line in time with the audio signal on two consecutive occasions. During the fieldwork, 20mSRT was performed using four different versions: multistage fitness test, sports coach UK (applied in Germany, Estonia and Cyprus); Leger test, CAEP Quebec Faca (applied in Spain and Hungary); multistage fitness test (applied in Sweden) and Uithouding shuttle run test (applied in Belgium). During the fieldwork, 20mSRT was assessed in shuttles. Results were unified according to the Leger test protocol. The initial speed in the Leger test starts in 8.5 km h^{-1}, with progressive increases of 0.5 km h^{-1}. Taking into account the speed in the other three protocols, we estimated the equivalent shuttles for every protocol. Shuttles were converted to stages in order to calculate maximal oxygen consumption (VO_{2max}) using Leger´s equation.[19] A greater number of shuttles indicate better performance. This test was not performed in Italian children ($N = 2440$; 51.8% girls).

Anthropometric measurements

International guidelines for anthropometry in children were used in the IDEFICS study.[23,24] Body weight (kg) and height (cm) were measured in barefoot children, clothed in underwear, using an adapted version of electronic scale TANITA BC 420 SMA (Tanita Europe BV, Amsterdam, The Netherlands), precision 100 g, range 0–150 kg and a portable stadiometer (seca 225, seca, Birmingham, UK), precision 0.1 cm, range 70–200 cm, respectively. Body mass index was calculated as body weight (kg) divided by height (m) squared.

Table 1. Chosen GAMLSS models to calculate the physical fitness reference values

Variable	*Sex*	*Model distribution*	*Parameters*			
			μ	log(σ)	ν	log(τ)
Back-saver sit-and-reach test (cm)	Girls	BCCG	1	Age	1	—
	Boys	BCCG	Age	Age	1	—
Handgrip strength test (kg)	Girls	BCCG	Age	1	Age	—
	Boys	BCCG	Age	1	1	—
Standing long jump test (cm)	Girls	BCCG	Age	Age	1	—
	Boys					
40-m sprint test (s)	Girls	BCPE	Age	1	1	1
	Boys					
20-m shuttle run test (shuttles)	Girls	BCPE	Age	1	1	1
	Boys					
			log(μ)	log(σ)	logit(ν)	—
Flamingo balance test (attempts)	Girls	DEL	Age	Age	Age	—
	Boys					
			log(μ)	log(σ)	—	—
20-m shuttle run test (VO_{2max}, ml kg^{-1} min^{-1})	Girls	IG	Age	Age	—	—
	Boys					

Abbreviations: BCCG, Box–Cox Cole and Green; BCPE, Box–Cox power exponential; DEL, Delaporte distribution; GAMLSS, General Additive Model for Location Scale and Shape; IG, inverse Gaussian.

Table 2. Percentiles of coordination/equilibrium; flexibility and speed/agility calculated with GAMLSS in normal-weight children

		Percentile for girls										*Percentile for boys*									
	Age	*1*	*3*	*10*	*25*	*50*	*75*	*90*	*97*	*99*		*Age*	*1*	*3*	*10*	*25*	*50*	*75*	*90*	*97*	*99*
Flamingo balance test (attempts)											Flamingo balance test (attempts)										
	6– < 6.5	1	1	3	4	8	13	20	28	36		6– < 6.5	1	2	4	6	11	17	24	33	41
	6.5– < 7	0	1	2	3	6	10	17	26	34		6.5– < 7	1	1	3	5	9	14	21	30	38
	7– < 7.5	0	1	2	3	5	8	14	23	32		7– < 7.5	0	1	2	4	7	12	19	27	36
	7.5– < 8	0	0	1	2	4	7	12	20	29		7.5– < 8	0	1	2	3	5	10	16	25	33
	8– < 8.5	0	0	1	2	3	5	9	17	25		8– < 8.5	0	1	1	3	5	8	13	22	30
	8.5– < 9	0	0	1	2	3	5	7	14	22		8.5– < 9	0	0	1	2	4	6	11	19	27
Back-saver sit-an-reach test (cm)											Back-saver sit-and-reach test (cm)										
	6– < 6.5	10.5	12.9	16.0	18.9	21.9	24.8	27.3	29.7	31.4		6– < 6.5	7.5	10.1	13.3	16.5	19.8	23.1	25.9	28.6	30.6
	6.5– < 7	9.7	12.3	15.6	18.7	21.9	25.0	27.6	30.1	31.9		6.5– < 7	7.1	9.7	13.0	16.2	19.6	22.9	25.8	28.6	30.6
	7– < 7.5	8.9	11.7	15.2	18.5	21.9	25.1	27.9	30.6	32.5		7– < 7.5	6.6	9.2	12.7	15.9	19.4	22.8	25.7	28.5	30.6
	7.5– < 8	8.1	11.1	14.8	18.3	21.9	25.3	28.3	31.1	33.1		7.5– < 8	6.2	8.8	12.3	15.7	19.2	22.6	25.6	28.5	30.6
	8– < 8.5	7.3	10.4	14.4	18.1	21.9	25.5	28.6	31.6	33.7		8– < 8.5	5.7	8.4	12.0	15.4	19.0	22.5	25.5	28.4	30.6
	8.5– < 9	6.5	9.7	13.9	17.8	21.9	25.7	29.0	32.1	34.4		8.5– < 9	5.3	8.1	11.6	15.1	18.8	22.4	25.4	28.4	30.6
40-m sprint (s)											40-m sprint (s)										
	6– < 6.5	8.1	8.5	8.7	9.2	9.9	10.7	11.5	12.4	13.1		6– < 6.5	7.6	7.8	8.3	8.8	9.4	10.2	11.0	11.8	12.6
	6.5– < 7	7.9	8.3	8.5	9.0	9.6	10.4	11.2	12.1	12.8		6.5– < 7	7.4	7.7	8.1	8.6	9.2	10.0	10.8	11.6	12.3
	7– < 7.5	7.7	8.0	8.3	8.7	9.4	10.1	10.9	11.8	12.5		7– < 7.5	7.2	7.5	7.9	8.4	9.0	9.8	10.5	11.3	12.0
	7.5– < 8	7.5	7.8	8.1	8.5	9.1	9.9	10.6	11.4	12.1		7.5– < 8	7.1	7.3	7.7	8.2	8.8	9.5	10.3	11.1	11.7
	8– < 8.5	7.3	7.6	7.8	8.3	8.9	9.6	10.3	11.1	11.8		8– < 8.5	6.9	7.2	7.5	8.0	8.6	9.3	10.0	10.8	11.4
	8.5– < 9	7.1	7.4	7.6	8.0	8.6	9.3	10.0	10.8	11.5		8.5– < 9	6.7	7.0	7.4	7.8	8.4	9.1	9.8	10.5	11.2

Abbreviation: GAMLSS, General Additive Model for Location Scale and Shape.[25] The values corresponded to the 1st, 3rd, 10th, 25th, 50th, 75th, 90th, 97th and 99th age- and sex-specific percentiles.

Table 3. Percentiles of upper-limb maximal strength and lower -limb explosive strength calculated with GAMLSS in normal-weight children

		Percentile for girls										*Percentile for boys*								
	Age	*1*	*3*	*10*	*25*	*50*	*75*	*90*	*97*	*99*	*Age*	*1*	*3*	*10*	*25*	*50*	*75*	*90*	*97*	*99*
Handgrip strength (kg)											Handgrip strength (kg)									
	6– < 6.5	5.0	5.5	6.2	7.0	8.1	9.3	10.6	12.1	13.2	6– < 6.5	5.4	6.0	6.9	7.9	9.1	10.4	11.7	13.0	14.1
	6.5– < 7	5.4	6.0	6.8	7.8	9.0	10.3	11.7	13.1	14.4	6.5– < 7	5.9	6.6	7.6	8.7	10.0	11.4	12.8	14.3	15.4
	7– < 7.5	5.8	6.4	7.4	8.5	9.8	11.2	12.7	14.2	15.4	7– < 7.5	6.4	7.2	8.3	9.4	10.9	12.4	13.9	15.5	16.8
	7.5– < 8	6.1	6.9	8.0	9.2	10.6	12.2	13.7	15.3	16.5	7.5– < 8	7.0	7.8	8.9	10.2	11.7	13.4	15.1	16.8	18.1
	8– < 8.5	6.5	7.3	8.6	9.9	11.5	13.1	14.7	16.3	17.6	8– < 8.5	7.5	8.3	9.6	11.0	12.6	14.4	16.2	18.0	19.5
	8.5– < 9	6.7	7.7	9.1	10.6	12.3	14.1	15.7	17.4	18.6	8.5– < 9	8.0	8.9	10.3	11.7	13.5	15.4	17.3	19.3	20.8
Standing long jump test (cm)											Standing long jump test (cm)									
	6– < 6.5	46.7	56.6	69.1	81.1	93.8	106.1	116.7	127.0	134.4	6– < 6.5	52.0	63.5	77.3	90.0	103.0	115.0	125.3	134.9	141.8
	6.5– < 7	51.1	61.0	73.6	85.7	98.6	111.1	121.9	132.3	139.9	6.5– < 7	56.6	68.1	82.1	94.9	108.0	120.3	130.7	140.6	147.6
	7– < 7.5	55.6	65.5	78.2	90.4	103.5	116.0	127.0	137.6	145.2	7– < 7.5	61.3	72.8	86.8	99.8	113.1	125.6	136.2	146.2	153.4
	7.5– < 8	60.2	70.1	82.8	95.2	108.3	121.0	132.1	142.7	150.5	7.5– < 8	66.1	77.6	91.7	104.7	118.2	130.8	141.6	151.7	159.0
	8– < 8.5	64.9	74.8	87.5	99.9	113.1	125.9	137.1	147.9	155.7	8– < 8.5	71.0	82.4	96.5	109.7	123.3	136.0	146.9	157.2	164.6
	8.5– < 9	69.6	79.5	92.3	104.7	118.0	130.8	142.1	152.9	160.8	8.5– < 9	75.9	87.3	101.5	114.6	128.3	141.2	152.2	162.6	170.1

Abbreviation: GAMLSS, General Additive Model for Location Scale and Shape.[25] The values corresponded to the 1st, 3rd, 10th, 25th, 50th, 75th, 90th, 97th and 99th age- and sex-specific percentiles.

Table 4. Percentiles of cardiorespiratory fitness: 20-m shuttle run test (shuttles and VO_2max (ml kg^{-1} min^{-1})) calculated with GAMLSS in normal-weight children

		Percentile for girls										*Percentile for boys*								
	Age	*1*	*3*	*10*	*25*	*50*	*75*	*90*	*97*	*99*	*Age*	*1*	*3*	*10*	*25*	*50*	*75*	*90*	*97*	*99*
20mSRT (shuttles)											20mSRT (shuttles)									
	6– < 6.5	3.9	4.9	6.5	8.6	11.8	15.9	20.7	26.7	32.1	6– < 6.5	3.5	4.6	6.6	9.4	13.7	19.3	25.3	32.3	38.2
	6.5– < 7	4.4	5.4	7.3	9.7	13.2	17.8	23.2	29.9	35.9	6.5– < 7	4.0	5.3	7.5	10.7	15.6	21.9	28.8	36.8	43.5
	7– < 7.5	4.8	6.0	8.1	10.7	14.6	19.8	25.7	33.1	39.8	7– < 7.5	4.5	5.9	8.5	12.0	17.4	24.6	32.3	41.2	48.7
	7.5– < 8	5.3	6.6	8.8	11.8	16.0	21.7	28.2	36.3	43.7	7.5– < 8	4.9	6.5	9.4	13.3	19.3	27.2	35.8	45.7	54.0
	8– < 8.5	5.8	7.2	9.6	12.8	17.5	23.6	30.7	39.6	47.6	8– < 8.5	5.4	7.2	10.3	14.6	21.2	29.9	39.3	50.1	59.2
	8.5– < 9	6.2	7.8	10.4	13.9	18.9	25.5	33.2	42.8	51.5	8.5– < 9	5.9	7.8	11.2	15.9	23.1	32.5	42.8	54.5	64.5
20mSRT (VO_{2max} (ml kg^{-1} min^{-1}))											20mSRT (VO_{2max} (ml kg^{-1} min^{-1}))									
	6– < 6.5	42.9	43.7	44.9	46.0	47.4	48.7	50.0	51.3	52.3	6– < 6.5	42.8	43.8	45.1	46.5	48.1	50.6	51.2	52.7	53.9
	6.5– < 7	42.1	43.0	44.2	45.5	47.0	48.5	49.9	51.3	52.4	6.5– < 7	42.1	43.1	44.6	46.1	47.8	50.6	51.3	53.0	54.3
	7– < 7.5	41.2	42.2	43.5	45.0	46.6	48.3	49.8	51.4	52.7	7– < 7.5	41.3	42.4	44.0	45.6	47.5	50.6	51.4	53.3	54.7
	7.5– < 8	40.3	41.4	42.8	44.4	46.2	48.1	49.8	51.6	53.0	7.5– < 8	40.4	41.7	43.4	45.2	47.3	50.7	51.5	53.6	55.2
	8– < 8.5	39.3	40.5	42.1	43.8	45.8	47.9	49.8	51.8	53.4	8– < 8.5	39.5	40.9	42.7	44.7	47.0	50.8	51.7	54.0	55.9
	8.5– < 9	38.3	39.6	41.3	43.2	45.4	47.7	49.9	52.1	53.9	8.5– < 9	38.6	40.0	42.0	44.2	46.7	50.9	51.9	54.5	56.6

Abbreviation: GAMLSS, General Additive Model for Location Scale and Shape[25]; 20mSRT, 20-m shuttle run test. The values corresponded to the 1st, 3rd, 10th, 25th, 50th, 75th, 90th, 97th and 99th age- and sex-specific percentiles.

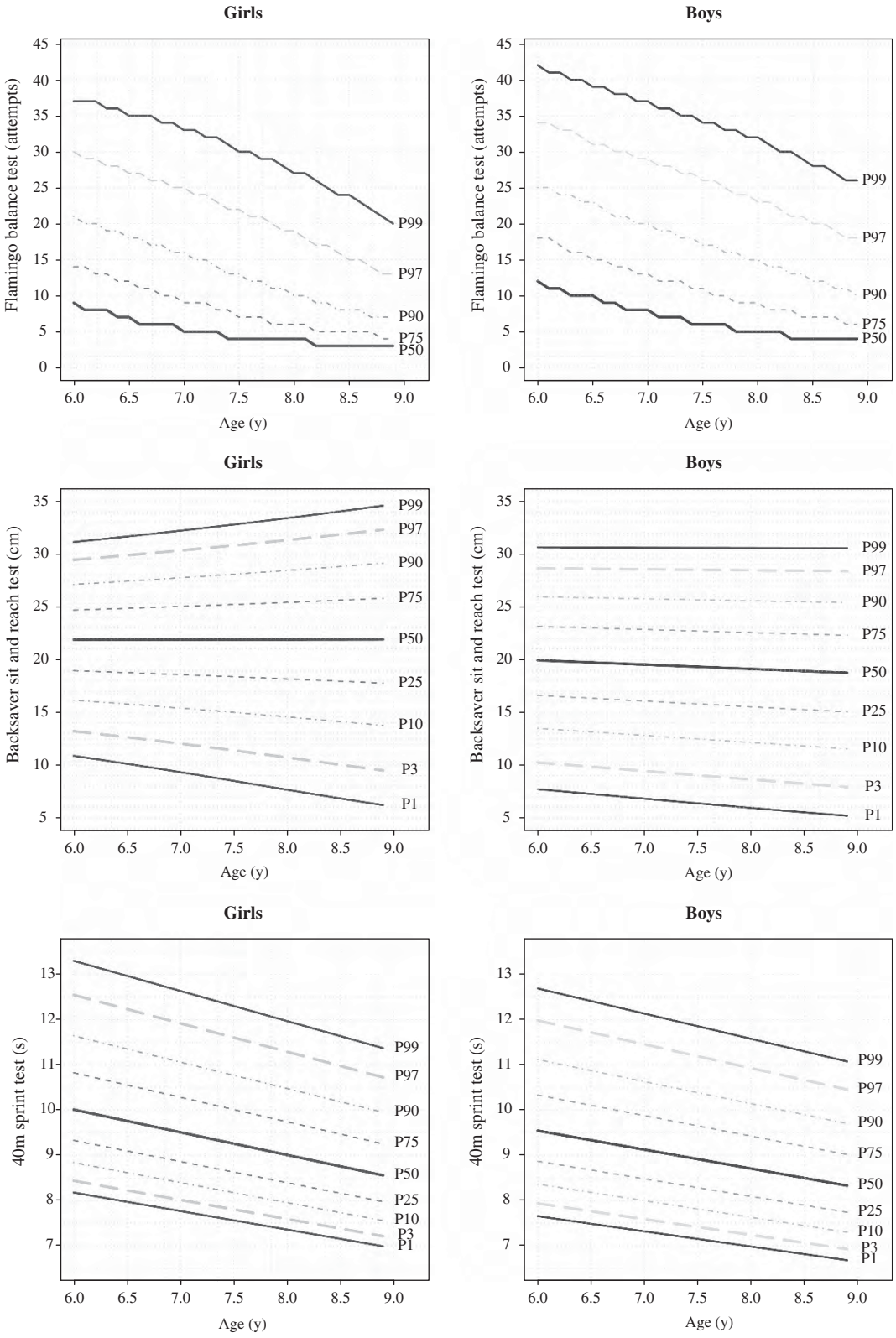


Figure 2. Percentile curves of the flamingo balance test (attempts), back-saver sit-and-reach test (cm) and 40-m sprint test (s) at median age in normal-weight children.

Statistical analysis

We calculated percentile curves of the physical fitness outcome variables as a function of the covariate age stratified by sex using the General Additive Model for Location Scale and Shape (GAMLSS) method. The GAMLSS method is an extension of the LMS method that models three parameters depending on one explanatory variable: M accounts for the median of the outcome variable and the coefficient of variation, (S) accounts for the variation around the mean and adjusts for non-uniform

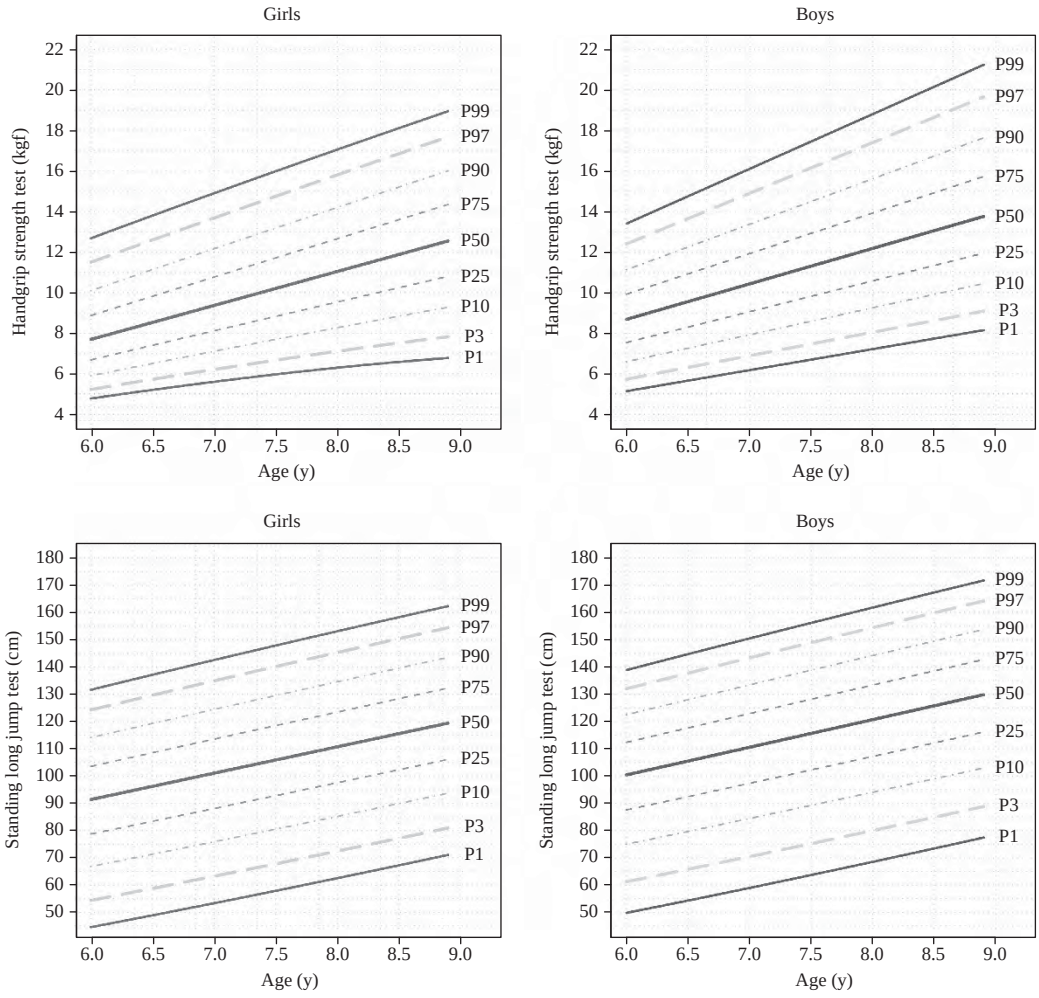


Figure 3. Percentile curves of handgrip strength test (kg) and standing long jump test (cm) at median age in normal-weight children.

dispersion, whereas the skewness (L) accounts for the deviation from a normal distribution using a Box–Cox transformation. The GAMLSS method is able to particularly model the kurtosis using other distributions and to include >1 covariate. We used the gamlss package (version 4.2-6)[25] of the statistical software *R* (version 3.0.1).[26] Different distributions, that is, the Box–Cox power exponential (BCPE), Box–Cox Cole and Green, inverse Gaussian and Delaporte distribution were fitted to the observed distribution of the physical fitness outcome variables. Moreover, the influence of age on the parameters of the considered distributions were modelled either as a constant, as a linear function or as a cubic spline. Goodness of fit was assessed by the Bayesian Information Criterion and Q–Q plots to select the final model including the fitted distribution of the physical fitness outcome variables. Worm plots were used as a diagnostic tool to assess whether adjustment for kurtosis was required.[27] Finally, percentile curves for the 1st, 3rd, 10th, 25th, 50th, 75th, 90th, 97th and 99th percentiles were calculated based on the model that showed the best goodness of fit.[25,28] For the Flamingo balance test only the 50th, 75th, 90th, 97th and 99th percentile curves are presented, because lower percentile curves of the best model did not correspond to the percentage of values below the percentile curves, that is, for example, instead of the targeted 25% of the girls' values there were only 12.6% of the girls' values below the 25th percentile curve. Similar results were observed for the 3rd, 10th and 25th percentile curves of the flamingo balance test in boys and girls.

The chosen GAMLSS models in boys and girls are listed in Table 1; for example, the best model for the 40-m sprint test in boys and girls was achieved with a Box–Cox power exponential distribution where the four parameters were modelled as follows: the location parameter μ linearly, the scale parameter log(σ) and the shape parameters ν and log(τ) as constants. Taking account the discrete distribution of the data from the flamingo balance test in contrast to all other models, the discrete Delaporte distribution was chosen.

RESULTS

Tables 2–4 and Figures 2–4 show the sex- and age-specific reference values (P_1, P_3, P_{10}, P_{25}, P_{50}, P_{75}, P_{90}, P_{97} and P_{99}) for the different fitness tests in European children.

Boys performed better than girls in speed/agility (40-m sprint), muscular strength (handgrip strength and standing long jump) and CRF (20mSRT; shuttles and VO_{2max}). Nevertheless, girls performed better than boys in balance (flamingo) and flexibility (back-saver sit-and-reach).

Older children performed better than younger children in all tests, except $\leqslant P_{50}$ in back-saver sit-and-reach (in both sexes) and $< P_{90}$ in 20mSRT (VO_{2max} in both sexes).

Finally, for all fitness tests, the range of fitness levels between P_3 and P_{99} is wider for a given age than the range of fitness levels across age groups and between sexes.

DISCUSSION

The main findings of this study are (1) boys performed better than girls in speed, lower- and upper-limb strength and CRF; (2) girls performed better in balance and flexibility; (3) overall, older children performed better than younger children; and (4) the

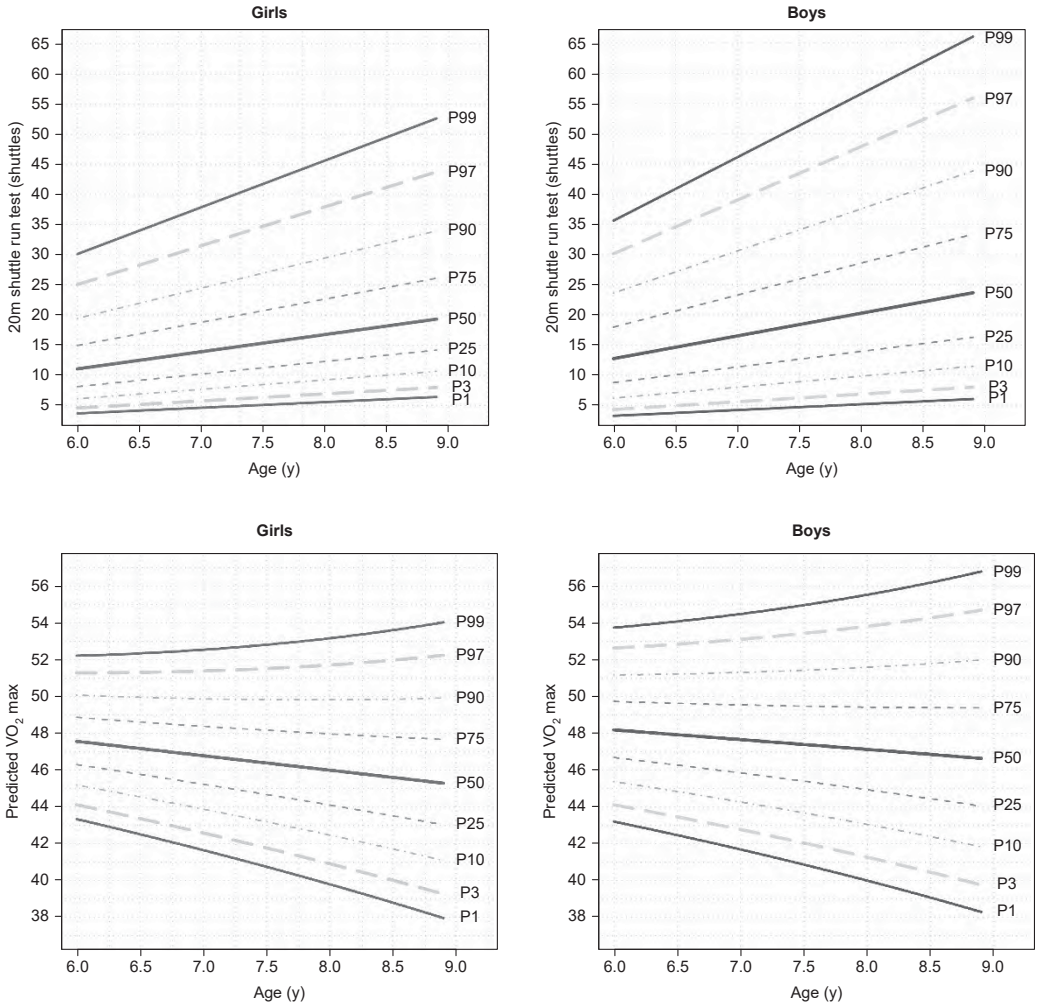


Figure 4. Percentile curves of the 20-m shuttle run test (20mSRT) estimates the CRF (aerobic capacity) at median age in normal-weight children.

range of fitness levels between P_3 and P_{99} is wider for a given age than the range of fitness levels across age groups and between sexes.

The majority of the fitness tests used in the IDEFICS study was selected based on the scientific evidence that showed associations with different health outcomes.[29] These tests have also been shown to be valid, reliable and feasible for health-monitoring purposes at the population level.[16] Fitness has been identified as a predictor of morbidity and mortality for all causes.[30–33] In this regard, various studies have shown that low fitness in children and adolescents is associated with adiposity,[7,34,35] cardiovascular risk factors[36,37] and skeletal[38] and mental health.[39]

In 2009, Ruiz *et al.*[9] systematically reviewed whether fitness in childhood was a predictor of cardiovascular disease risk factors, events and syndromes, quality of life and low back pain later in life. The authors found moderate evidence for the association between changes in CRF and cardiovascular risk factors, and between CRF and the risk of developing metabolic syndrome and arterial stiffness. In addition, CRF has also been widely studied in relation to metabolic risk and adiposity.[36,37,38] Interestingly, handgrip strength in Swedish male adolescents (aged 16–19 years) has been identified as a risk factor for the major causes of death in adulthood (<55 years).[6] Moreover, results from the Amsterdam Growth and Health Study, following adolescents aged 13 years until the age of 27 years, showed that the longitudinal improvements in VO_{2max} were related to a healthy cardiovascular risk profile.[41]

To the best of our knowledge, our study is the first in providing cutoff values of sex- and age-specific fitness for pre-pubertal European children. These values may be useful in identifying children being at higher risk for developing unfavourable health outcomes owing to their low fitness level. Previously, Casajus *et al.*[42] published physical fitness levels in children from Aragon (Spain) aged 7–12 years. Our results are roughly comparable to these with respect to flexibility (they used sit-and-reach), muscular strength (handgrip strength and standing long jump) and CRF (20mSRT, VO_{2max}) only for those children aged 7–9 years. Similar values were observed in both studies for these fitness tests in boys and girls. For example, P_{50} of sit-and-reach was 17.3cm vs 19.4 cm for Spanish and European boys, respectively, and 19.3cm vs 21.9 cm for Spanish and European girls, respectively. P_{50} of handgrip as average of the right and left side was 12.7 kg vs 11.8 kg for Spanish and European boys, respectively, and 11.6 kg vs 10.8 kg for Spanish and European girls, respectively. P_5 of standing long jump was 117.2 cm vs 120.5 cm for Spanish and European boys, respectively, and 104.5 cm vs 110.7 cm for Spanish and European girls, respectively. Finally, P_{50} of 20mSRT was 48.1 ml kg^{-1} min^{-1} vs 47.3 ml kg^{-1} min^{-1} for Spanish and European boys, respectively, and 46.7 ml kg^{-1} min^{-1} vs 46.2 ml kg^{-1} min^{-1} for Spanish and European girls, respectively.

Castro-Pinero *et al.*[44] published percentile values for muscular strength (standing long jump)[43] and CRF (20mSRT, stages) in children and adolescents from Cadiz (Spain) aged 6–17 years. Our results are also comparable to those from Castro-Pinero *et al.*[43,44] for those children aged 6–9 years. For example, P_{50} of standing long jump was 111.6 cm vs 120.5 cm for Spanish and European boys, respectively, and 114.8 cm vs 110.7 cm for Spanish and European girls, respectively. In addition, P_{50} of 20mSRT (stage) was 2.7 vs 2.0 for Spanish and European boys, respectively, and 1.9 vs 2.0 for Spanish and European girls, respectively.

In adolescent populations, Ortega *et al.*[45] first published European fitness reference values for 12.5–17.5-year-old youths from 10 cities (HELENA study); reporting sex- and age-specific physical fitness levels. In this line, Haugen *et al.*[46] recently reported normative fitness data for Norwegian 13–15-year-old adolescents. A systematic review[10] identified Australian studies reporting normative fitness data for children and adolescents aged 9–17 years. In addition, Tremblay *et al.*[14] reported normative data for aerobic fitness, flexibility and muscular strength for Canadian 6–19-year-old youths; however, only the 50th percentile was reported. The present results are comparable to those obtained by Tremblay *et al.*[14] for flexibility (sit-and-reach) and muscular fitness (handgrip strength) in Canadian children (for the age range 6–10 years). For these fitness dimensions, European children performed worse than the Canadian ones. For example, in the study of Tremblay *et al.*,[14] the performance in sit-and-reach at P_{50} of 6–10-year-old children was 25 cm and 30 cm in boys and girls, respectively. In the present study, the performance in back-saver sit-and-reach at P_{50} of 6–9-year-old children was 19.4 cm and 21.9 cm for boys and girls, respectively. In addition, in the study of Tremblay *et al.*,[14] the performance in handgrip strength at P_{50} was 25 kg and 22 kg in boys and girls, respectively, calculated as the sum of the best right- and left-hand attempt. In our study, the performance in handgrip test at P_{50} of 6–9-year-old children was 23.2 kg and 21 kg for boys and girls, respectively. The differences observed in flexibility might be explained by slightly different methodologies. Tremblay *et al.*[14] measured flexibility by using sit-and-reach and using the best performance of two attempts. In our study, the back-saver sit-and-reach values were calculated as the average of the left and right sides.

There is scarcity of data on reference standards of fitness for children. Hence, the presented European data are relevant in various respects. A fitness level below P_5 may be considered as potentially pathologic since, for example, low scores on CRF and handgrip tests are associated with cardiovascular risk. Therefore, it may be recommended to monitor children with a fitness level below this percentile for cardiovascular risk markers.[47,48] For a practical use of these data, schools, sport clubs and so on may consider to take into account a Likert type scale to classify children´s performance X, as follows: very poor ($X < P_{10}$); poor ($P_{10} \leqslant X < P_{25}$); medium ($P_{25} \leqslant X < P_{75}$); good ($P_{75} \leqslant X < P_{90}$); and very good ($X \geqslant P_{95}$).

Some limitations should be considered when interpreting the findings in this study. As already mentioned before, the study was not planned to be representative with respect to the broad range of variables that we investigated, but given the population-based approach our study sample should be considered as an unselected population. As the children were free to opt out of examination modules, the analysis groups and the sample sizes of children participating in the various components of the test battery to assess physical fitness varied from test to test. However, comparing the respective analysis groups with the overall sample of 6–10.9-year-old children, no major differences were detected with respect to mean age, sex distribution and prevalence of overweight and obesity. It would have been interesting to investigate whether the reliability of the fitness tests differs in apparently pre-pubertal children and adolescents. However, this was not feasible within our IDEFICS sample but may be investigated in later follow-ups of the cohort. The main strengths of this study are the large and heterogeneous sample of European children, the standardised use of well-known and validated health-related fitness tests and a strong statistical method to obtain normative values of fitness tests.

In summary, our results provide sex- and age-specific physical fitness reference standards in European children. These data in children aged 6.0–9.9 years complement the study published by Ortega *et al.*[45] in adolescents aged 12.5–17.5 years. Unfortunately, there is a gap between 9.9 and 12.4 years without updated fitness reference data at the European level, which has to be filled. In the meanwhile, country-specific data such as those commented above should be used in children of this age.

CONFLICT OF INTEREST

The authors declare no conflct of interest.

ACKNOWLEDGEMENTS

This work has been done as part of the IDEFICS Study (http://www.idefics.eu). We gratefully acknowledge the financial support of the European Community within the Sixth RTD Framework Programme Contract No. 016181 (FOOD). We are grateful to the Volkswagen Foundation that financially supported the production of this supplement. This analysis was also supported by a grant from the Spanish Ministry of Science and Innovation (RYC-2011-09011). LL received grant support from VR, FORTE and FORMAS (research councils in Sweden). The authors wish to thank the IDEFICS children and their parents who generously, volunteered have participated in the IDEFICS study. We also acknowledge all the staff members involved in the fieldwork for their efforts and great enthusiasm. The information in this document reflects the authors' view and is provided as is.

STATEMENT OF ETHICS

We certify that all applicable institutional and governmental regulations concerning the ethical use of human volunteers were followed during this research. Approval by the appropriate Ethics Committees was obtained by each of the 8 centres doing the fieldwork. Study children did not undergo any procedure before both they and their parents had given consent for examinations, collection of samples, subsequent analysis and storage of personal data and collected samples. Study subjects and their parents could consent to single components of the study while abstaining from others.

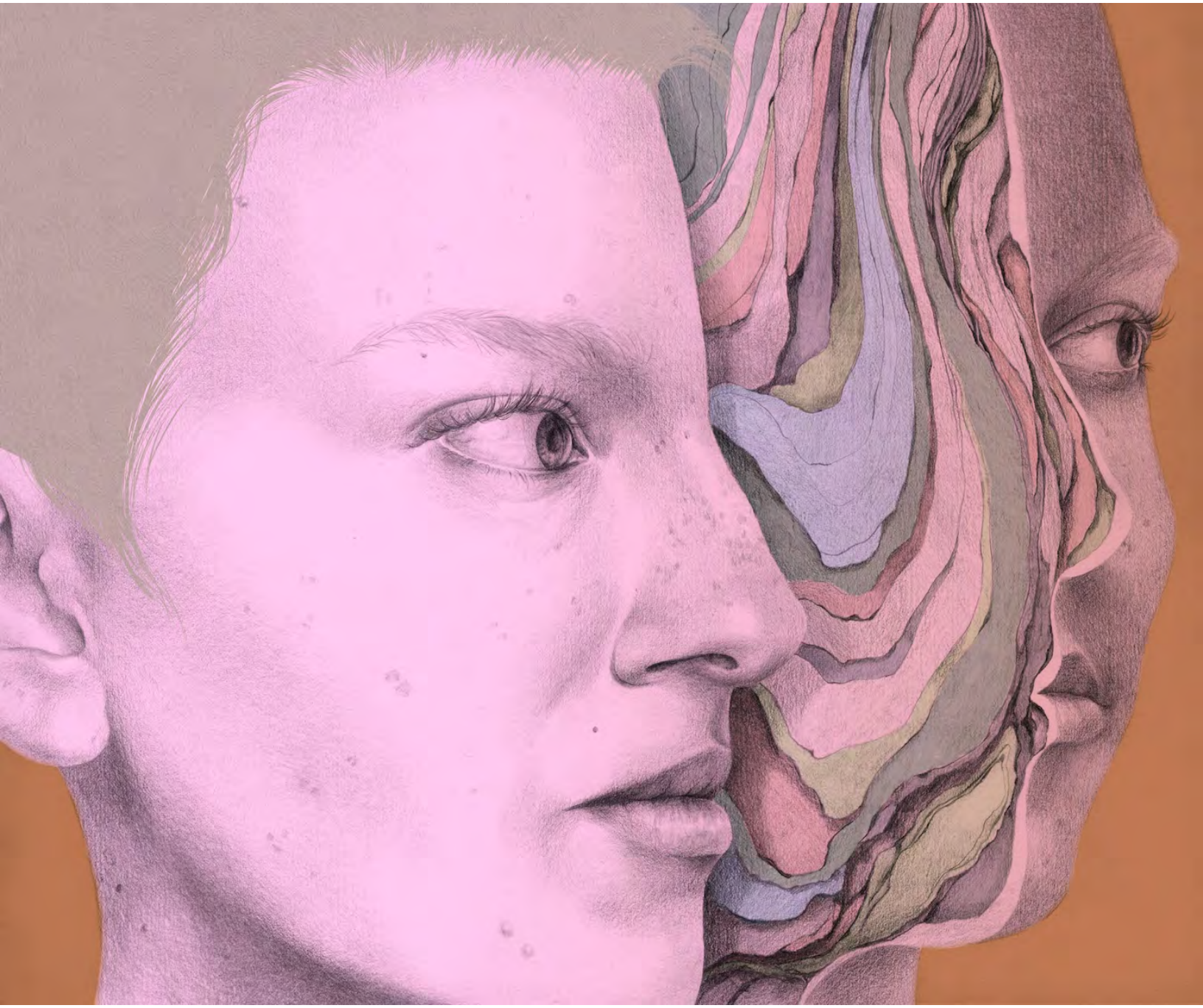
REFERENCES

1 Caspersen CJ, Powell KE, Christenson GM. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. *Public Health Rep* 1985; **100**: 126–131.
2 Ortega FB, Ruiz JR, Castillo MJ, Sjostrom M. Physical fitness in childhood and adolescence: a powerful marker of health. *Int J Obes (Lond)* 2008; **32**: 1–11.
3 Ekelund U, Anderssen SA, Froberg K, Sardinha LB, Andersen LB, Brage S. Independent associations of physical activity and cardiorespiratory fitness with metabolic risk factors in children: the European youth heart study. *Diabetologia* 2007; **50**: 1832–1840.
4 Blaes A, Baquet G, Fabre C, Van Praagh E, Berthoin S. Is there any relationship between physical activity level and patterns, and physical performance in children? *Int J Behav Nutr Phys Act* 2011; **8**: 122.
5 Ortega FB, Labayen I, Ruiz JR, Kurvinen E, Loit HM, Harro J *et al.* Improvements in fitness reduce the risk of becoming overweight across puberty. *Med Sci Sports Exerc* 2011; **43**: 1891–1897.
6 Ortega FB, Silventoinen K, Tynelius P, Rasmussen F. Muscular strength in male adolescents and premature death: cohort study of one million participants. *BMJ* 2012; **345**: e7279.
7 Moliner-Urdiales D, Ortega FB, Vicente-Rodriguez G, Rey-Lopez JP, Gracia-Marco L, Widhalm K *et al.* Association of physical activity with muscular strength and fat-free mass in adolescents: the HELENA study. *Eur J Appl Physiol* 2010; **109**: 1119–1127.
8 Morales PF, Sanchez-Lopez M, Moya-Martinez P, Garcia-Prieto JC, Martinez-Andres M, Garcia NL *et al.* Health-related quality of life, obesity, and fitness in schoolchildren: the Cuenca study. *Qual Life Res* 2013; **22**: 1515–1523.
9 Ruiz JR, Castro-Pinero J, Artero EG, Ortega FB, Sjostrom M, Suni J *et al.* Predictive validity of health-related fitness in youth: a systematic review. *Br J Sports Med* 2009; **43**: 909–923.

10 Catley MJ, Tomkinson GR. Normative health-related fitness values for children: analysis of 85347 test results on 9-17-year-old Australians since 1985. *Br J Sports Med* 2013; **47**: 98–108.
11 Olds T, Tomkinson G, Leger L, Cazorla G. Worldwide variation in the performance of children and adolescents: an analysis of 109 studies of the 20-m shuttle run test in 37 countries. *J Sports Sci* 2006; **24**: 1025–1038.
12 Pate RR, Wang CY, Dowda M, Farrell SW, O'Neill JR. Cardiorespiratory fitness levels among US youth 12 to 19 years of age: findings from the 1999-2002 National Health and Nutrition Examination Survey. *Arch Pediatr Adolesc Med* 2006; **160**: 1005–1012.
13 Carrel AL, Bowser J, White D, Moberg DP, Weaver B, Hisgen J *et al.* Standardized childhood fitness percentiles derived from school-based testing. *J Pediatr* 2012; **161**: 120–124.
14 Tremblay MS, Shields M, Laviolette M, Craig CL, Janssen I, Connor Gorber S. Fitness of Canadian children and youth: results from the 2007-2009 Canadian Health Measures Survey. *Health Rep* 2010; **21**: 7–20.
15 Ahrens W, Bammann K, Siani A, Buchecker K, De Henauw S, Iacoviello L *et al.* IDEFICS consortium. The IDEFICS cohort: design, characteristics and participation in the baseline survey. *Int J Obes (Lond)* 2011; **35**: S3–S15.
16 Ruiz JR, Castro-Pinero J, Espana-Romero V, Artero EG, Ortega FB, Cuenca MM *et al.* Field-based fitness assessment in young people: the ALPHA health-related fitness test battery for children and adolescents. *Br J Sports Med* 2011; **45**: 518–524.
17 Artero EG, Espana-Romero V, Castro-Pinero J, Ortega FB, Suni J, Castillo-Garzon MJ *et al.* Reliability of field-based fitness tests in youth. *Int J Sports Med* 2011; **32**: 159–169.
18 Castro-Pinero J, Artero EG, Espana-Romero V, Ortega FB, Sjostrom M, Suni J *et al.* Criterion-related validity of field-based fitness tests in youth: a systematic review. *Br J Sports Med* 2010; **44**: 934–943.
19 Leger LA, Mercier D, Gadoury C, Lambert J. The multistage 20 metre shuttle run test for aerobic fitness. *J Sports Sci* 1988; **6**: 93–101.
20 Meredith, M Welk, G. *Fitnessgram-Activitygram Test Administration Manual*. Human Kinetics: Champaign, IL, USA, 2007.
21 Ruiz JR, Espana-Romero V, Ortega FB, Sjostrom M, Castillo MJ, Gutierrez A. Hand span influences optimal grip span in male and female teenagers. *J Hand Surg Am* 2006; **31**: 1367–1372.
22 Castro-Pinero J, Ortega FB, Artero EG, Girela-Rejon MJ, Mora J, Sjostrom M *et al.* Assessing muscular strength in youth: usefulness of standing long jump as a general index of muscular fitness. *J Strength Cond Res* 2010; **24**: 1810–1817.
23 Bammann K, Peplies J, Sjöström M, Lissner L, De Henauw S, Galli C *et al.* Assessment of diet, physical activity and biological, social and environmental factors in a multi-centre European project on diet- and lifestyle-related disorders in children (IDEFICS). *J Public Health* 2006; **14**: 279–289.
24 Stomfai S, Ahrens W, Bammann K, Kovacs E, Marild S, Michels N *et al.* IDEFICS consortium. Intra- and inter-observer reliability in anthropometric measurements in children. *Int J Obes (Lond)* 2011; **35**: S45–S51.
25 Stasinopoulos DM, Rigby R. Generalized additive models for location scale and shape (GAMLSS) in R. *J Stat Softw* 2007; **23**: 1–46.
26 R Core Team. R: *A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing: Vienna, Austria, 2013. http://www.R-project.org/.
27 Van Buuren S, Fredriks A. Worm plot: a simple diagnostic device for modelling growth reference curves. *Stat Med* 2001; **20**: 1259–1277.
28 Cole TJ, Stanokevic S, Stocks J, Coates A, Hankinson J, Wade A. Age-and size-reñated reference ranges: A case study of spirometry through childhood and adulthood. *Stat Med* 2009; **28**: 880–898.
29 Ruiz JR, Ortega FB, Gutierrez A, Meusel D, Sjöström M, Castillo MJ. Health-related fitness assessment in childhood and adolescence: a European approach based on the AVENA, EYHS and HELENA studies. *J Public Health* 2006; **14**: 269–277.
30 Metter EJ, Talbot LA, Schrager M, Conwit R. Skeletal muscle strength as a predictor of all-cause mortality in healthy men. *J Gerontol A Biol Sci Med Sci* 2002; **57**: B359–B365.
31 Blair SN, Kohl 3rd HW, Paffenbarger Jr RS, Clark DG, Cooper KH, Gibbons LW. Physical fitness and all-cause mortality. A prospective study of healthy men and women. *JAMA* 1989; **262**: 2395–2401.
32 Myers J, Prakash M, Froelicher V, Do D, Partington S, Atwood JE. Exercise capacity and mortality among men referred for exercise testing. *N Engl J Med* 2002; **346**: 793–801.
33 Mora S, Redberg RF, Cui Y, Whiteman MK, Flaws JA, Sharrett AR *et al.* Ability of exercise testing to predict cardiovascular and all-cause death in asymptomatic women: a 20-year follow-up of the lipid research clinics prevalence study. *JAMA* 2003; **290**: 1600–1607.
34 Ara I, Moreno LA, Leiva MT, Gutin B, Casajus JA. Adiposity, physical activity, and physical fitness among children from Aragon, Spain. *Obesity (Silver Spring)* 2007; **15**: 1918–1924.
35 Nassis GP, Psarra G, Sidossis LS. Central and total adiposity are lower in overweight and obese children with high cardiorespiratory fitness. *Eur J Clin Nutr* 2005; **59**: 137–141.
36 Ruiz JR, Ortega FB, Rizzo NS, Villa I, Hurtig-Wennlof A, Oja L *et al.* High cardiovascular fitness is associated with low metabolic risk score in children: the European Youth Heart Study. *Pediatr Res* 2007; **61**: 350–355.
37 Artero EG, Ruiz JR, Ortega FB, Espana-Romero V, Vicente-Rodriguez G, Molnar D *et al.* Muscular and cardiorespiratory fitness are independently associated with metabolic risk in adolescents: the HELENA study. *Pediatr Diabetes* 2011; **12**: 704–712.
38 Vicente-Rodriguez G, Urzanqui A, Mesana MI, Ortega FB, Ruiz JR, Ezquerra J *et al.* Physical fitness effect on bone mass is mediated by the independent association between lean mass and bone mass through adolescence: a cross-sectional study. *J Bone Miner Metab* 2008; **26**: 288–294.
39 Castelli DM, Hillman CH, Buck SM, Erwin HE. Physical fitness and academic achievement in third- and fifth-grade students. *J Sport Exerc Psychol* 2007; **29**: 239–252.
40 Moliner-Urdiales D, Ruiz JR, Vicente-Rodriguez G, Ortega FB, Rey-Lopez JP, Espana-Romero V *et al.* Associations of muscular and cardiorespiratory fitness with total and central body fat in adolescents: the HELENA study. *Br J Sports Med* 2011; **45**: 101–108.
41 Twisk JW, Kemper HC, van Mechelen W. Tracking of activity and fitness and the relationship with cardiovascular disease risk factors. *Med Sci Sports Exerc* 2000; **32**: 1455–1461.
42 Casajus JA, Ortega FB, Vicente-Rodriguez G, Leiva MT, Moreno LA, Ara I. [Physical fitness, fat distribution and health in school-age children (7 to 12 years).] *Rev Int Med Cienc Act Fís Deporte* 2010; **12**: 523–537.
43 Castro-Pinero J, Gonzalez-Montesinos JL, Mora J, Keating XD, Girela-Rejon MJ, Sjostrom M *et al.* Percentile values for muscular strength field tests in children aged 6 to 17 years: influence of weight status. *J Strength Cond Res* 2009; **23**: 2295–2310.
44 Castro-Pineiro J, Ortega FB, Keating XD, Gonzalez-Montesinos JL, Sjostrom M, Ruiz JR. Percentile values for aerobic performance running/walking field tests in children aged 6 to 17 years: influence of weight status. *Nutr Hosp* 2011; **26**: 572–578.
45 Ortega FB, Artero EG, Ruiz JR, Espana-Romero V, Jimenez-Pavon D, Vicente-Rodriguez G *et al.* Physical fitness levels among European adolescents: the HELENA study. *Br J Sports Med* 2011; **45**: 20–29.
46 Haugen T, Hoigaard R, Seiler S. Normative data of BMI and physical fitness in a Norwegian sample of early adolescents. *Scand J Public Health* 2014; **42**: 67–73.
47 Hasselstrom H, Hansen SE, Froberg K, Andersen LB. Physical fitness and physical activity during adolescence as predictors of cardiovascular disease risk in young adulthood. Danish Youth and Sports Study. An eight-year follow-up study. *Int J Sports Med* 2002; **23**: S27–S31.
48 Twisk JW, Kemper HC, van Mechelen W. Prediction of cardiovascular disease risk factors later in life by physical activity and physical fitness in youth: general comments and conclusions. *Int J Sports Med* 2002; **23**: S44–S49.

The mental health establishment is failing trans kids

Gender-exploratory therapy is a key step. Why aren't therapists providing it?



Daryn Ray for The Washington Post

By Laura Edwards-Leeper and Erica Anderson

November 24, 2021 at 5:54 p.m. EST

CORRECTION

A previous version of this essay said that a quarter of study subjects who reversed their gender transitions did not report this change to their doctors. In fact, three-quarters did not share the information.

At 13, Patricia told her parents she was a transgender boy. She had never experienced any gender dysphoria — distress at a disconnect between gender identity and the sex assigned at birth — she said. But a year earlier, she'd been sexually assaulted by an older girl. Soon after this trauma, she met another older girl who used they/them pronouns and introduced her to drugs, violent pornography and the notion of dissociation from her body. Her lingering psychic wounds, coinciding with a raft of new and unsettling ideas, plunged her into depression and anxiety. Patricia's parents took her to a therapist so she could talk through her shifting identity and acute mood swings.

The job of a mental health provider here should have been clear: Perform an assessment, ask how long she'd experienced dysphoria and investigate how mental health issues and any other changes in her life might be contributing to it. Instead, on first meeting, the therapist simply affirmed her new identity, a step that can lead to hormonal and eventually surgical treatments. Was Patricia ready for these next steps — or, her parents wondered, was this a normal bout of teenage confusion stemming from a recent trauma? The therapist instructed them to "support" their child's trans self-diagnosis and to socially transition her. If they didn't, Patricia might end her own life: 41 percent of unsupported children commit suicide, they were told. Would Patricia's parents rather have a dead child or a trans one?

They sought another therapist, one who was more curious and less certain, one who listened closely. After a year of exploring who she was, Patricia no longer felt she was a boy. She decided to stop binding her breasts and wearing boys' clothes.

We are both psychologists who have dedicated our careers to serving transgender patients with ethical, evidence-based treatment. But we see a surge of gender dysphoria cases like Patricia's — cases that are handled poorly. One of us was the founding psychologist in 2007 of the first pediatric gender clinic in the United States; the other is a transgender woman. We've held recent leadership positions in the World Professional Association for Transgender Health (WPATH), which writes the standards of care for transgender people worldwide. Together, across decades of doing this work, we've helped hundreds of people transition their genders. This is an era of ugly moral panic about bathrooms, woke indoctrination and identity politics in general. In response, we enthusiastically support the appropriate gender-affirming medical care for trans youth, and we are disgusted by the legislation trying to ban it.

But the number of adolescents requesting medical care is skyrocketing: Now 1.8 percent of people under 18 identify as transgender, double the figure from five years earlier, according to the Trevor Project. A flood of referrals to mental health providers and gender medical clinics, combined with a political climate that sees the treatment of each individual patient as a litmus test of social tolerance, is spurring many providers into sloppy, dangerous care. Often from a place of genuine concern, they are hastily dispensing medicine or recommending medical doctors prescribe it — without following the strict guidelines that govern this treatment. Canada, too, is following our lead: A study of 10 pediatric gender clinics there found that half do not require psychological assessment before initiating puberty blockers or hormones.

The standards of care recommend mental health support and comprehensive assessment for all dysphoric youth before starting medical interventions. The process, done conscientiously, can take a few months (when a young person's gender has been persistent and there are no simultaneous mental health issues) or up to several years in complicated cases. But few are trained to do it properly, and some clinicians don't even believe in it, contending without evidence that treating dysphoria medically will resolve other mental health issues. Providers and their behavior haven't been closely studied, but we find evidence every single day, from our peers across the country and concerned parents who reach out, that the field has moved from a more nuanced, individualized and developmentally appropriate assessment process to one where every problem looks like a medical one that can be solved quickly with medication or, ultimately, surgery. As a result, we may be harming some of the young people we strive to support — people who may not be prepared for the gender transitions they are being rushed into.

American opinions about transgender youth have shifted dramatically in the past 15 years. The pendulum has swung from a vile fear and skepticism around ever treating adolescents medically to what must be described, in some quarters, as an overcorrection. Now the treatment pushed by activists, recommended by some providers and taught in many training workshops is to affirm without question. "We don't actually have data on whether psychological assessments lower regret rates," Johanna Olson-Kennedy, a pediatrician at Children's Hospital in Los Angeles who is skeptical of therapy requirements and gives hormones to children as young as 12 (despite a lack of science supporting this practice, as well), told the Atlantic. "I don't send someone to a therapist when I'm going to start them on insulin." This perspective writes off questions about behavioral and mental health, seeing them as a delaying tactic or a dodge, a way of depriving desperate people of the urgent care they clearly need.

But comprehensive assessment and gender-exploratory therapy is the most critical part of the transition process. It helps a young person peel back the layers of their developing adolescent identity and examine the factors that contribute to their dysphoria. In this stage, patients reflect on the duration of the dysphoria they feel; the continuum of gender; the intersection with sexual orientation; what medical interventions might realistically entail; social media, Internet and peer influences; how other factors (e.g., autism, trauma, eating disorders/body image concerns, self-esteem, depression, anxiety) may help drive dysphoria, rather than assuming that they are always a result of dysphoria; family dynamics and social/peer relationships; and school/academic challenges. The messages that teens get from TikTok and other sources may not be very productive for understanding this constellation of issues.

There are several reasons the process can move too quickly and hurtle toward medical treatment. For one, the stigma around mental health in general, along with the trauma caused to transgender adults by the health-care field in the past (yes, including conversion therapy), has made our peers extremely skeptical of becoming "gatekeepers" — experts who deny the needed help because they supposedly know best. Slowing down the process and encouraging deeper, thoughtful exploration is considered, many tell us, unnecessary and unaffirming. Providers may also be afraid of being cast as transphobic bigots by their local colleagues and referral sources if they engage in gender exploring therapy with patients, as some have equated this with conversion therapy. We've personally experienced this backlash at professional conferences.

All this means only that the purpose of assessment is improperly understood. The approach WPATH recommends is collaborative and aims to provide a developmentally appropriate process that involves the parents and takes the complexities of adolescence into consideration. (The constituency of agitated parents who feel excluded is also growing rapidly. These are not conservative evangelicals who don't believe trans people exist or deserve treatment. They're usually progressive, educated, loving people who all say, *If our kid is really trans, we'll fully support them. We just want to be as sure as possible, and we can't find a provider who will actually engage in gender exploring therapy. Instead, doctors and psychologists and social workers are ready to start hormones after one short visit.*)

Another reason that teens can receive substandard mental health care is that gender clinics are disastrously overwhelmed. Most have a single social worker who completes a brief "intake," relying instead on other mental health clinicians in the community to assess patients and offer their conclusions. Frequently, those community clinicians, just like the parents, assume that a more comprehensive assessment will occur in the gender specialty clinic. But in our experience, and based on what our colleagues share, this is rarely the case. Most clinics appear to assume that a referral means a mental health provider in the community has diagnosed gender dysphoria and thereby given the green light for medical intervention.

When working in gender clinics, we've also both received letters from therapists who had "assessed" patients they were referring to us. An astonishing number of these were nothing but a paragraph that stated the youth identified as trans, had dysphoria and wanted hormones, so that course was recommended. There are nearly 200,000 members of the American Psychological Association and the American Psychiatric Association. Add to that the clinical social workers, marriage counselors and family therapists. The overwhelming majority of those well-intentioned professionals receive limited or no training in the assessment of gender-diverse youth. (We receive requests frequently from people eager for more comprehensive, nuanced trainings, which we both deliver.) In simple terms, the demand for competent care has outstripped the supply of competent providers.

In professional circles, we hear from pediatric endocrinologists and others who prescribe hormones for trans youth. Many openly discuss how they use the adult informed-consent model of care with their teen patients, which almost always means no mental health involvement and sometimes no parent input, either. "If you are trans, I believe you," says A.J. Eckert, the medical director of Anchor Health Initiative in Connecticut. Eckert is wary of psychologists who follow the guidelines by completing a comprehensive assessment before recommending medical intervention for youths. "Gender-affirming medicine," Eckert holds, means that "you are best equipped to make decisions about your own body," full stop. These providers do not always realize they've confessed to ignoring the standards of care. (Contacted by The Post for comment on this essay, Eckert said that "no medical or surgical interventions are provided to anyone who has not started puberty" but added that, as Anchor Health sees it, "Therapy is not a requirement in this approach because being trans is not a pathology.")

Some providers may move quickly because they believe that an adolescent's clarity around their gender identity is no different than that of transgender adults, whose care is now typically based on simple informed consent. Some assume that a person with gender dysphoria who declares they are transgender is transgender and needs medical interventions immediately. Yet we know this is not always true. In a recent study of 100 detransitioners, for instance, 38 percent reported that they believed their original dysphoria had been caused by "something specific, such as trauma, abuse, or a mental health condition." Fifty-five percent said they "did not receive an adequate evaluation from a doctor or mental health professional before starting transition."

A handful of studies supposedly showing the suicide risk of gender minority youth who are not supported are also not entirely conclusive. The term "support," for instance, is defined differently across studies, and it is never defined as "starting medical interventions." Supporting trans youth may include using the correct name/pronouns or allowing the young person to present in a way that aligns with their affirmed gender (e.g., clothing, hairstyle). These studies also show correlations between teen-transition hurdles and suicidality, but not causal relationships. Suicide is a horrifying outcome for too many gender-diverse youth, but its specter should not be used to push forward unrelated medical treatment without professional care or attention for each patient.

Longer-term longitudinal studies are needed to better understand the role of medical interventions on lifetime psychological health, particularly with the newer subset of adolescents presenting with no childhood dysphoria and significant mental health concerns. Research is needed to help determine whether quick medical treatment or a more cautious approach is best in these cases. Based on our experience with patients, we suspect that there will be variability based on age, when gender identity questions first emerged and other factors — which is why an individualized approach with careful assessment is so critical.

Trans youth, more than most patients in the health-care system, require an interdisciplinary approach: Their doctors rely on mental health colleagues for direction, and it is crucial that those therapists take the reins. Without proper assessment, many youths are being rushed toward the medical model, and we don't know if they will be liberated or restrained by it. National figures do not yet exist, but the rising number of detransitioners that clinicians report seeing (they are forming support groups online) indicates that this approach can backfire. This is not the most common outcome of a transition process, but it is hardly unheard of, either. These are typically youth who experienced gender dysphoria and other complex mental health issues, rushed to medicalize their bodies and regretted it later. Only a quarter of them told their doctors they had reversed their transitions, making this population especially hard to track.

Many trans activists want to silence detransitioners or deny their existence, because those cases do add fuel to the conservative agenda that is pushing to deny medical treatment to all transgender young people. (Those conservative views are unacceptable, and medically unsound.) Instead, we should be learning from them and returning to the empirically supported careful assessment model recommended by WPATH. And none of this means that we shouldn't be listening to the views of gender-diverse teens; it only means that we should listen in the fullest and most probing way possible.

The pressure by activist medical and mental health providers, along with some national LGBT organizations to silence the voices of detransitioners and sabotage the discussion around what is occurring in the field is unconscionable. Not only is it harmful to detransitioned young people — to be made to feel as if their lived experiences are not valid, the very idea that the gender-transition treatment is meant to remedy — but it will undoubtedly raise questions regarding the objectivity of our field and our commitment to help trans people. The fact that some people detransition does not mean that transgender people should not receive the services they need.

The energy currently spent fighting this political battle would be much better directed toward improving care for all gender-diverse young people. They deserve nothing less.

ORIGINAL ARTICLE

Maximum oxygen uptake and objectively measured physical activity in Danish children 6–7 years of age: the Copenhagen school child intervention study

S Eiberg, H Hasselstrom, V Grønfeldt, K Froberg, J Svensson, L B Andersen

Br J Sports Med 2005;**39**:725–730. doi: 10.1136/bjsm.2004.015230

See end of article for authors' affiliations

Correspondence to: Professor Andersen, Department of Health, Norwegian University of Sport and Physical Education, Oslo 0806, Norway; lars.bo.andersen@nih.no

Accepted 29 March 2005

Objectives: To provide normative data on maximum oxygen uptake ($\dot{V}O_2MAX$) and physical activity in children 6–7 years of age and analyse the association between these variables.
Methods: $\dot{V}O_2MAX$ was measured in 366 boys (mean (SD) 6.8 (0.4) years of age) and 332 girls (6.7 (0.4) years of age) from preschool classes in two suburban communities in Copenhagen, during a progressive treadmill exercise. Habitual physical activity was measured with accelerometers.
Results: Boys had higher $\dot{V}O_2MAX$ both in absolute values (1.19 (0.18) *v* 1.06 (0.16) litres/min (+11%), $p<0.001$) and relative to body weight (48.5 (6.0) *v* 44.8 (5.6) ml/kg/min (+8%); $p<0.001$) than girls. The difference in $\dot{V}O_2MAX$ between boys and girls decreased to +2% when expressed relative to lean body mass (LBM). Absolute $\dot{V}O_2MAX$ was related to LBM, body mass, and stature (all $p<0.001$). Boys were more physically active than girls (mean counts +9.4%, $p<0.001$), and even when boys and girls with the same $\dot{V}O_2MAX$ were compared, boys were more active. The difference in physical activity between the sexes was higher when sustained activity of higher intensity was compared.
Conclusions: $\dot{V}O_2MAX$ is higher in boys than girls (+11%), even when related to body mass (+8%) and LBM (+2%). Most of the difference in $\dot{V}O_2MAX$ relative to body mass was explained by the larger percentage body fat in girls. When boys and girls with the same $\dot{V}O_2MAX$ were compared, boys engaged in more minutes of exercise of at least moderate intensity.

The fitness level in children in the western world has declined, but few data are available.[1] Even though children do not suffer from lifestyle diseases, it is recommended that they stay physically active and fit, because a high physical activity (PA) level may prevent future illness.[2] A high level of physical fitness has been associated with a decreased risk of cardiovascular disease (CVD) in the adult population.[3] Even in children, risk factors for CVD have been identified,[4] and physical fitness also seems to have an effect on the level of risk factors in children.[5] Even though children do not suffer from CVD, it is recommended that children are physically active and fit, as this may help to prevent the disease in the future. Therefore it is important to assess physical fitness at an early age. Maximum oxygen uptake ($\dot{V}O_2MAX$) is probably the best index of physical fitness and has been studied intensively in adults.[6] However, measurements of $\dot{V}O_2MAX$ in 6–7 year old children are sparse, and representative data do not exist.[6–11] Even in older children and adolescents, very few population based studies exist.[10 12 13] The reasons for this include difficulties in testing young children. Ethical considerations, safety factors, and equipment made for adults render testing with young children more challenging. Another health related measure of children is their level of PA, as PA has a positive effect on metabolism.[14] Most studies have assessed PA by self report, but accelerometers provide a robust measure of habitual PA.[15] No population study has analysed the association between $\dot{V}O_2MAX$ and PA assessed by accelerometry.

$\dot{V}O_2MAX$ is consistently higher in boys than in girls even before puberty.[16] This has been attributed to a different body composition and a larger stroke volume of the heart in boys. However, even though boys have been shown to be more active than girls, the difference is too small to explain the difference in $\dot{V}O_2MAX$.[8 17 18]

The aim of this study was to provide population data on $\dot{V}O_2MAX$ and PA measured using accelerometry in 6–7 year old children from Copenhagen, and also explore potential sex differences and analyse the association between PA and $\dot{V}O_2MAX$. We hypothesised that boys have higher $\dot{V}O_2MAX$ than girls, and that this difference may be the result of differences in their levels of fat and PA.

METHODS

Children from 46 preschool classes (6–7 years of age) in 18 schools in two suburban communities in the Copenhagen area were invited to participate in the Copenhagen school child intervention study. In 2000 the community of Ballerup (10 schools, 27 classes) increased the number of PE lessons from two to four a week for the first three years of school. The community of Taarnby (eight schools, 19 classes) was chosen as a control as it resembles Ballerup in sociodemographics. A total of 706 children (69% of those eligible) volunteered for the study, and written informed consent was obtained from the parents/guardians. Of these 706 children, 415 from Ballerup and 291 from Taarnby participated. The ethics committee of Copenhagen county approved the study. The tests were performed from December 2001 until May 2002 at the 18 different schools involved. The exercise test was performed using permanently installed equipment in a camper trailer. All other physiological tests were performed in a gym or a classroom. All tests were performed before noon (0800–1200).

PA was generally measured one week after the other tests. At schools with more than 50 participating pupils, two weeks of PA measurements were necessary. The 31% not

Abbreviations: BMI, body mass index; CVD, cardiovascular disease; LBM, lean body mass; PA, physical activity; SFS, sum of four skinfolds; $\dot{V}O_2MAX$, maximum oxygen uptake

participating were analysed about one year later. The first systematic medical examination of children in Denmark takes place one and a half years after they start school, so we were unable to gather data from non-participants in their first year. There were no significant differences between groups with respect to age, height, weight, and body mass index (BMI) for either sex. The analysis included 612 participants and 277 non-participants.

Height was measured by a Harpenden stadiometer to the nearest 1 mm. Body weight was measured to the nearest 0.1 kg using a SECA electronic scale. Bicipital, tricipital, subscapular, and suprailiac skinfolds were measured with a Harpenden skinfold caliper according to criteria presented by De Lorenzo *et al*.[19] The dominant side of the body was determined by asking the child to take a pen and write his/her name. The data shown in this study represent the mean of three measurements taken on the non-dominant side of the body. The sum of four skinfolds (SFS) was used as an indicator of body fatness. Fat mass, fat percentage, and lean body mass (LBM) were derived as:

Fat mass (kg) $= 0.38 \times$ body weight + $(0.30 \times$ triceps) + $(0.87 \times G) - 9.42$

where for boys G = 1 and for girls G = 2.[20]

Habitual PA was measured by the MTI 7164 activity monitor (Manufactory Technology Inc, Fort Walton Beach, Florida, USA). The monitor has been validated in several studies and has shown both high mechanical reproducibility[15] and good validity with respect to free living conditions in children.[21] The monitor samples acceleration at 10 Hz and integrates counts over a time period (epoch) defined by the user. In children, PA is characterised by short bursts of activity. Therefore we chose an epoch of 10 seconds. The choice of epoch limited the registration to three full days and 19 hours. To allow familiarisation, the children had the MTI monitor put on one day before recording. It was secured directly to the skin at the lower back using an elastic belt. The children were instructed to wear the monitor continuously except during water based activity or when sleeping. To distinguish true zeros that arise as a result of below threshold activity from zeros recorded when the MTI monitor was not worn, the data were cleaned as follows: all MTI files were screened for sustained periods of zero activity. Periods of 10 minutes or more with zero counts were interpreted as ''MTI not worn'' and removed from the file. In spite of instructions, some children slept with the monitor on and therefore had activity recorded even late at night. Therefore we chose to control for nocturnal activity (2230 to 0600). Given these criteria, the data were included if the child had accumulated more than eight hours of activity a day for at least three days. In the end, 466 children had four valid days and 96 children had three valid days, of whom 82 had two week days and one weekend day and 14 had one week day and two weekend days. Fifty eight children had less than three valid days. A mean count was calculated for each child. Furthermore, the number of minutes in periods longer than five or 10 minutes of sustained activity above 2000 counts/min, which is about three METS, was calculated.

Measurement of $\dot{V}O_2MAX$

$\dot{V}O_2$ was determined with an AMIS 2001 Cardiopulmonary Function Test System (Innovision, DK 5260 Odense, Denmark). This system has been validated against the Douglas bag system.[22] We used a Hans Rudolph mouth piece with a volume of 15 ml especially designed for children. Heart rate was measured continuously every fifth second (Polar Sport Tester, Kempele, Finland).

We conducted a pilot study to establish a protocol that enabled most of the children to reach exhaustion. We chose a continuous walking and running protocol on a treadmill. The velocity on the treadmill was initially set to 4 km/h without inclination and kept there for the first three minutes to allow familiarisation. At three minutes, the velocity was increased to 8 km/h, and at five minutes the inclination was raised to 3%. At seven, nine, and 11 minutes, the inclination was increased to 6%, 9%, and 11% respectively. If the child could endure more, the velocity was increased to 9 km/h after 13 minutes and then 10 km/h at 15 minutes. No child completed the last work load.

The children were instructed to run until exhaustion. One subjective and three objective criteria were used to determine if the test was maximal. Every child had to meet the subjective criterion and at least one of the three objective criteria. Criteria were chosen according to Rowland's recommendations.[23] The physiological criteria were: heart rate >200 beats/min; respiratory exchange ratio $\geq$0.99; a defined plateau of $\dot{V}O_2$ (an increase less than 2.1 ml/kg/min). Subjective criteria were signs of intense effort such as unsteady running pattern, sweating, facial flushing, and clear unwillingness to continue running in spite of repeated strong verbal encouragement.

Twenty 6–7 year old children outside the sample were invited to do a test-retest on the peak $\dot{V}O_2$ test. Written informed consent was obtained from the parents/guardians. The children were retested either five or seven days after the first test. Four children did not meet the criteria on one or both occasions, and their results were not used in the later analysis, leaving 16 children with two valid tests. A Bland-Altman plot showed no systematic difference between the first and second test, suggesting no learning effect from the first to the second test. The plot also showed that the size of the difference was not dependent on the absolute level of $\dot{V}O_2MAX$. The typical error of measurement on the difference between $\dot{V}O_2MAX$ between tests was 0.02 litre/kg.

Statistical analysis

The data were stored and analysed using SPSS 11.5.0. Data that were not normally distributed were log transformed, and the mean of transformed values was back transformed to obtain the geometric mean as suggested by Altman.[24] $\dot{V}O_2MAX$ relative to body mass was split into deciles by sex to compare differences over the whole distribution, and $\dot{V}O_2MAX$ (ml/kg/min) was calculated for each of the deciles. For a comparison of differences between sexes, we adjusted for body size by two methods of allometric scaling. Data from this study were used to find the scaling factors. In the analyses, the scaling exponent b was identified in the allometric equation $Y = a_1 + a_2X^b$, where Y is the physiological variable ($\dot{V}O_2MAX$) and X is the anthropometric scaling variable (weight or LBM). To obtain b, both the Y and X were log transformed, and least squares regression identified the b in the equation $\ln(Y) = \ln(a_2) + b\ln(X)$.

Differences between sexes were tested using Student's *t* test. The relation between $\dot{V}O_2MAX$ and PA was assessed by linear regression. A significance level of $p<0.05$ was chosen.

RESULTS

Table 1 shows the total number of boys and girls that entered the study, the number of valid tests, the number of children that stopped before exhaustion, the number of children not fulfilling the criteria, and children who refused to do the treadmill test. A valid $\dot{V}O_2MAX$ measurement was not assessed in 114 children, because 37 failed to comply with the approval criteria and 57 did not want to wear the $\dot{V}O_2$ equipment or did not wear the equipment satisfactorily. Ten children were unwilling to perform the test, and 10 children were absent on all the test days. The final number of accepted tests was 592. Analysis of children with non-valid tests was performed for each sex separately. Time to exhaustion, heart

Table 1 Age, height, body weight, and maximal exercise data for groups with different test status for the maximal exercise test

	Boys				Girls			
	Valid	Non-valid	No $\dot{V}O_2$ data†	Not willing to be tested	Valid	Non-valid	No $\dot{V}O_2$ data†	Not willing to be tested
Number	309	20	28	7	283	17	29	3
Age (years)	6.8 (0.4)	6.7 (0.3)	6.7 (0.4)	6.9 (0.4)	6.7 (0.4)	6.6 (0.3)	6.6 (0.4)	6.4 (0.4)
Stature (m)	1.24 (0.05)	1.23 (0.05)	1.23 (0.06)	1.23 (0.05)	1.22 (0.05)	1.20 (0.04)*	1.23 (0.04)	1.21 (0.01)
Mass (kg)‡	24.6 (3.7)	24.4 (3.3)	24.7 (3.2)	24.3 (4.6)	23.9 (3.2)	22.6 (3.2)	25.0 (4.5)	22.8 (1.3)
Heart rate (beats/min)	196 (9)	174 (8) *	193 (9)		199 (9)	179 (6) *	196 (11)	
Test time (min:s)	8:06 (1:55)	5:53 (1:16)*	7:39 (1:39)		7:28 (1:36)	4:50 (0:55)*	6:46 (1:33)*	
$\dot{V}O_2MAX$ (litres/min)	1.19 (0.18)	0.99 (0.24)*			1.06 (0.16)	0.76 (0.16)*		

Values are mean (SD).
*Significant difference between valid test groups and other groups: $p<0.05$.
†Children performing a maximal test with no recording of $\dot{V}O_2$.
‡Mass was not a Gaussian distribution. Therefore data were log transformed and a mean was calculated and then back transformed to a geometric mean.

rate, and $\dot{V}O_2MAX$ were lower for the 20 boys and 17 girls that did not meet the criteria ($p<0.001$). The girls that did not comply with the criteria were 2.6 cm shorter than girls who did comply ($p<0.05$). Twenty eight boys and 29 girls had a recorded time to exhaustion, but had no valid $\dot{V}O_2$ measurements. The girls without a valid $\dot{V}O_2$ recording ran 42 seconds less ($p<0.05$) than girls with a valid test. Differences between all other age, anthropometric, and $\dot{V}O_2$ test data were non-significant. Only data for the 592 children with a valid test were included in the following analysis. Of the 592 children that met the subjective criteria, 210 met the levelling off criterion (35%), 283 (47%) met the pulse criterion, and 507 (86%) met the respiratory exchange ratio criterion.

Table 2 shows descriptive data for age, anthropometry, and PA. Except for BMI, differences were found between boys and girls for all variables. The girls had larger SFS (15%, $p<0.001$) and fat percentage (23%, $p<0.001$). The boys were older (2%, $p<0.001$), heavier (3%, $p<0.05$), taller (1%, $p<0.001$), had a greater LBM (8%, $p<0.001$), a higher level of PA (mean counts/min 8%, $p<0.001$), and more minutes in periods of five or 10 minutes of sustained activity at a level of 2000 counts/min or greater (60% and 103% respectively, $p<0.001$). It should be noted that all boys had five minute periods of sustained activity, whereas three girls did not have any. Further, 38 boys and 83 girls did not have any 10 minute periods of sustained activity.

The absolute differences in the treadmill data between boys and girls were moderate (table 3). The girls had higher maximum heart rate (2%, $p<0.001$) and respiratory exchange ratio (5%, $p<0.001$), and the boys had higher absolute $\dot{V}O_2MAX$ (11%, $p<0.001$), $\dot{V}O_2MAX$ relative to body mass (8%, $p<0.001$), and $\dot{V}O_2MAX$ relative to LBM (2%, $p<0.05$). $\dot{V}O_2MAX$ remained different (9%, $p<0.001$) after allometric scaling using $(\text{body mass})^{0.712}$ as scaling, but did not differ when $LBM^{1.105}$ was used as the scaling factor.

Deciles in $\dot{V}O_2MAX$ (ml/kg/min) by sex were constructed and then $\dot{V}O_2MAX$ was compared between boys and girls for each decile (fig 1). This was done to see if the difference between the sexes appeared over the whole distribution or if only the lower or upper part of the distribution differed. The whole distribution was shifted to the right in boys compared with girls. Differences were found between all groups within sexes using Bonferroni post hoc analysis of variance ($p<0.001$).

Relation between PA and fitness

A scatterplot of $\dot{V}O_2MAX$ relative to body weight and the PA variables showed a linear relation. A linear regression was performed with mean count of PA, the number of minutes above 2000 counts/min in five minute periods, and sex as independent variables, and $\dot{V}O_2MAX$ relative to body weight as a dependent variable.

The following equations where G = 1 for boys and G = 2 for girls were used:

Model (1) $\dot{V}O_2MAX$ (l/kg/min) $= 49.9 - 3.8 \times G + 0.02 \times PA_{\text{mean count}}$ ($p<0.001$)

Model (2) $\dot{V}O_2MAX$ (l/kg/min) $= 49.6 - 3.1 \times G + 0.02 \times PA_{\text{in 5 min periods}}$ ($p<0.001$).

Partial correlations for model (1) were −0.38 and 0.12 for sex and $PA_{\text{mean count}}$ respectively. Partial correlations for model (2) were −0.25 and 0.24 for sex and $PA_{\text{in 5 min periods}}$ respectively.

To explore the relation $\dot{V}O_2MAX$ versus PA and fatness further, all subjects were ranked into six fitness groups by $\dot{V}O_2MAX$ (independent of sex) and the mean sums of PA level ($PA_{\text{in 5 min periods}}$) and SFS were calculated for each sex

Table 2 Anthropometric and physical activity data in 6–7 year old children by sex

	Boys		Girls		
	No	Mean	No	Mean	p Value
Age (years)	309	6.8 (0.4)	283	6.7 (0.4)	<0.05
Mass (kg)	309	24.4 (19.8 to 34.9)	283	23.7 (18.4 to 32.0)	<0.001
Stature (m)	309	1.24 (0.05)	283	1.22 (0.05)	<0.001
BMI (kg/m^2)	309	15.9 (13.5 to 20.2)	283	15.9 (13.2 to 20.5)	NS
SFS (mm)	307	23.3 (15.3 to 51.2)	281	27.5 (17.6 to 54.4)	<0.001
Fat percentage	307	16.5 (5.2)	281	21.4 (5.0)	<0.001
LBM (kg)	307	20.4 (1.8)	281	18.7 (1.9)	<0.001
PA (mean counts/min)	291	743 (452 to 1308)	269	679 (397 to 1062)	<0.001
PA (minutes in 5 min bouts)	291	122 (17 to 334)	269	76 (0 to 194)	<0.001
PA (minutes in 10 min bouts)	291	53 (0 to 194)	269	26 (0 to 94)	<0.001

Values in parentheses are SD or 95% confidence interval. Mass, BMI, SFS, and PA were not normally distributed. Therefore data were log transformed and a mean was calculated and then back transformed to a geometric mean as described in the text. For these variables a 95% confidence interval is given.
BMI, Body mass index; SFS, sum of four skinfolds; LBM, lean body mass; PA, physical activity.

Table 3 Maximal exercise data on 6–7 year old children by sex

	Boys (n = 309)	Girls (n = 283)	p Value
Heart rate (beats/min)	196 (9)	199 (9)	<0.001
RER	1.05 (0.11)	1.10 (0.11)	<0.001
$\dot{V}O_2MAX$ (litres/min)	1.19 (0.18)	1.06 (0.16)	<0.001
$\dot{V}O_2MAX$ relative to body mass (ml/kg/min)	48.5 (6.0)	44.8 (5.6)	<0.001
$\dot{V}O_2MAX$ relative to LBM (ml/kg/min)	58.2 (6.8)	57.0 (6.5)	<0.05
$\dot{V}O_2MAX$ relative to body mass (ml/$kg^{0.71}$/min)	121.7 (14.5)	111.2 (12.9)	<0.001
$\dot{V}O_2MAX$ relative to LBM (ml/$kg^{1.105}$/min)	42.4 (4.9)	41.9 (4.8)	NS

Values are mean (SD).
RER, Respiratory exchange ratio.

within each fitness group. When PA was compared between boys and girls with the same level of fitness, boys were more active than girls except for the least fit decile (fig 2). In a linear regression, the association between sex and PA level was highly significant ($r = 0.20$, $p<0.001$) even after adjustment for fitness ($r = 0.26$, $p<0.001$).

When fatness was compared between boys and girls with the same level of fitness, girls had a larger SFS than boys except for the least fit decile (fig 3). In a linear regression, the association between sex and SFS was borderline significant ($r = 0.09$, $p = 0.05$) after adjustment for fitness ($r = 0.27$, $p<0.001$). After adjustment for both fitness and PA, there was no difference between sexes with respect to SFS.

There were only two girls in the most fit group, and no standard error was plotted for this group in fig 3. The opposite analysis was also performed—that is, mean $\dot{V}O_2MAX$ was calculated for boys and girls with the same activity level and the same SFS respectively. The boys had a 3–5 ml/kg/min higher fitness level in each category of activity ($p<0.001$) and a 1–6 ml/kg/min higher fitness in each category of SFS ($p<0.01$).

DISCUSSION

The main findings in this study are that, in a large sample of young children, boys had a higher $\dot{V}O_2MAX$ and PA level than girls. Most of the difference in fitness between the sexes could be accounted for partly by body composition and partly by PA. As no differences in haemoglobin or sex hormones have been reported in this age group,[17] it is likely that the difference in $\dot{V}O_2MAX$ relative to body weight is due to body composition and PA. However, when children with the same $\dot{V}O_2MAX$ were compared, boys were still more active, and in boys and girls with the same PA level, boys were fitter.

Only a few studies in the literature report measured values of $\dot{V}O_2MAX$ for children less than 8 years of age and only for a small number of children.[8 25–27] Mean values of 39–53 ml/kg/min have been reported. Differences in $\dot{V}O_2MAX$ between studies seem to be due, at least partly, to differences in protocol. The lowest mean value reported (39 ml/kg/min) came from a field test.[25] The highest mean values found were 53 and 52 ml/kg/min for 5–7 year old boys and girls respectively.[26] The highest values were calculated from the original data of Åstrand.[26] Testing was carried out on a treadmill with a continuous running protocol in a laboratory setting, and subjects were highly selected. Compared with earlier studies,[8 26] in which a difference in $\dot{V}O_2MAX$ relative to body weight of less than 2% between the sexes in prepubertal children was reported, our data show a considerable difference (8%). Rowland[23] suggested that at least some of the sex difference in children comes from the difference in fat percentage. In our study, the difference was 23%. As the boys were 3% heavier than the girls, this indicates that at least part of the sex difference in $\dot{V}O_2MAX$ is related to difference in body composition. The much smaller 2% difference we found in $\dot{V}O_2MAX$ relative to LBM between boys and girls supports this. Further, when allometric scaling was performed with LBM as scaling factor, the difference was only 1% and statistically insignificant. This implies that the sex difference in $\dot{V}O_2MAX$ is due mainly to a difference in body composition.

Vinet *et al*[28] and Rowland *et al*[29] came to different conclusions. Vinet *et al* found no difference in stroke volume and $\dot{V}O_2MAX$ after allometric scaling and stated that body composition alone (and not cardiac functional capacity) could account for the sex difference. Rowland *et al* found a difference between sexes even after allometric scaling and

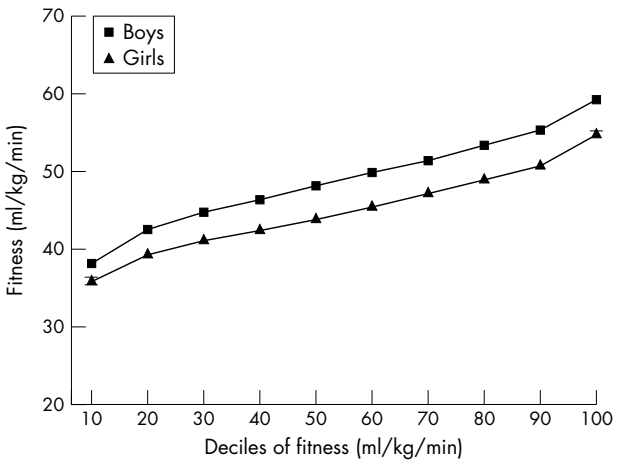


Figure 1 Subjects were ranked into 10 centiles of $\dot{V}O_2MAX$ by sex. The mean $\dot{V}O_2MAX$ was calculated for boys and girls separately in each centile. Standard errors are so small that they are not visible.

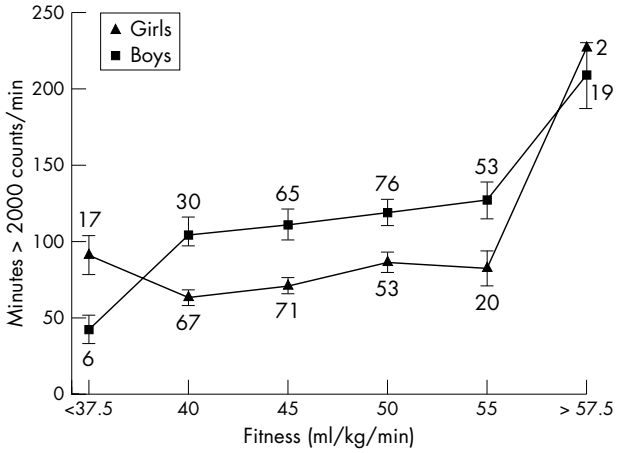


Figure 2 All subjects together were ranked into six groups of $\dot{V}O_2MAX$ relative to body weight. Physical activity level was plotted with SE for boys and girls with the same fitness level. Numbers of girls and boys are shown for each group.

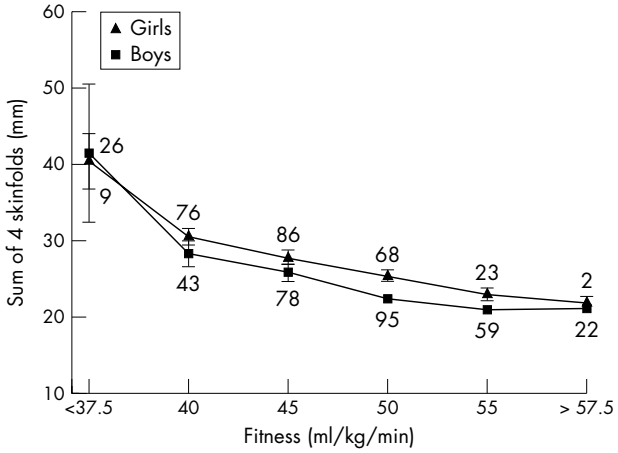


Figure 3 All subjects together were ranked into six groups of $\dot{V}O_2MAX$ relative to body weight. Sum of four skinfolds was plotted with SE for boys and girls with the same fitness level. Numbers of girls and boys are shown for each group.

stated that both body composition and cardiac functional capacity accounted for the difference between boys and girls. Further, Rowland *et al* found that anthropometric and aerobic physiological factors cannot entirely explain the magnitude of the sex differences.

Physical activity

We found a large sex difference (60%) in PA with regard to sustained activity, and sustained periods of PA explained 9% of the variance in $\dot{V}O_2MAX$ relative to body weight. It was tempting to claim that habitual PA may be part of the explanation for the difference in $\dot{V}O_2MAX$ between boys and girls. This is also the case, but even within the same stratum of $\dot{V}O_2MAX$ the boys had a higher PA level, so differences in PA seem only partly to explain the differences in $\dot{V}O_2MAX$. Sundberg[30] compared $\dot{V}O_2MAX$ (ml/kg/min) in sighted children and blind children, who are not able to participate in vigorous play, and found 22% and 26% higher values in sighted boys and girls respectively compared with blind children. From Sundberg's data, it could be expected that much of the difference between fitness levels in prepubertal children can be explained by differences in PA. It is therefore surprising in the present study that boys were more active than girls even within the same stratum of fitness. It is perhaps less surprising that boys have smaller SFS compared with girls in the same stratum of fitness. These two observations support the view of Rowland *et al*[29] that the higher level of fitness in boys is due to both the higher level of PA and the lower level of fatness.

Representativity

The number of children in this study is relatively high and represents 69% of all children in preschool in two suburban communities of Copenhagen. There were no significant differences between participants and non-participants in age, height, weight, and BMI in the study, so it seems safe to conclude that the participant cohort is representative of the two communities. Further, we have compared our sample with a sample from nationwide Danish studies with measurements of height and weight in children.[31 32] The differences between these two studies and our study with regard to mean height and BMI are less than 1%. It seems reasonable to conclude that with regard to height and BMI our cohort is representative of Danish children aged 6–7 years. The strengths of this study are that a large number of subjects were examined using objective and direct measurements and that the sample was representative.

The number of rejected tests in this study is relatively high. Even though there were no differences in age and size between subjects with valid tests and rejected tests, the tendency was that younger and smaller children chose not to run or stopped prematurely. We also chose to reject the data for 57 children who actually ran a subjectively approved test, but did not have a $\dot{V}O_2MAX$ recording. In the end, 84% of the participating children completed a valid test, and we consider this satisfactory, especially as some authors have claimed that the reliability or validity of testing $\dot{V}O_2MAX$ in children aged less than 8 years is questionable.[33 34] Our own validation of the protocol showed good reliability, and it seems reasonable to conduct reliable physiological testing on 6–7 year old children with satisfying results. It should be noted that care must be taken to carefully examine every child's measurements to ensure that the appropriate criteria have been met, and to discard non-valid tests. In conclusion, the absolute levels of $\dot{V}O_2MAX$ relative to body weight were 48.5 ml/kg/min and 44.8 ml/kg/min in girls, $\dot{V}O_2MAX$ is larger in boys than girls (+11%), also when related to body mass (+8%) and to LBM (+2%). Most of the difference in $\dot{V}O_2MAX$ relative to body mass was due to the greater percentage body fat in girls and lower level of PA. When boys and girls with the same $\dot{V}O_2MAX$ were compared, boys engaged in more minutes of at least moderately intense activity and had a smaller SFS.

What is already known on this topic

No study has assessed population data in 6–7 year old children on aerobic fitness and physical activity using direct measurement of $\dot{V}O_2MAX$ and objective measurement of physical activity.

What this study adds

- Even in 6–7 year old children, boys had 8% higher fitness levels than girls, and the difference in fitness may mainly be explained by differences in physical activity level and body composition
- Moderate intensity physical activity may not influence physical fitness but still change body composition

Authors' affiliations
S Eiberg, H Hasselstrom, V Grønfeldt, K Froberg, J Svensson, Institute for Exercise and Sport sciences, University of Copenhagen, Denmark
L B Andersen, Norwegian University of Sport and Physical Education, Oslo, Norway

Competing interests: none declared

REFERENCES

1 **Wedderkopp N**, Froberg K, Hansen HS, *et al*. Secular trends in physical fitness and fatness in Danish 9-year old girls and boys. Odense School child Study and Danish substudy of The European Youth Heart Study. *Scand J Med Sci Sports Exerc* 2004;**14**:1–6.
2 **Berenson GS**, Wattigney WA, Tracy RE, *et al*. Atherosclerosis of the aorta and coronary arteries and cardiovascular risk factors in persons 6 to 30 years and studied at necropsy (the Bogalusa Heart Study). *Am J Cardiol* 1992;**70**:851–8.
3 **Hasselstrøm H**, Hansen SE, Froberg K, *et al*. Physical fitness and physical activity during adolescence as predictors of cardiovascular disease risk in young adulthood. *Int J Sports Med* 2002;**23**:S27–31.

4 **Webber LS**, Voors AW, Srinivasan SR, *et al*. Occurrence in children of multiple risk factors for coronary artery disease: the Bogalusa Heart Study. *Prev Med* 1979;**8**:407–18.
5 **Wedderkopp N**, Froberg K, Hansen HS, *et al*. Cardiovascular risk factors cluster in children and adolescents with low physical fitness. *Pediatr Exerc Sci* 2003;**15**:419–22.
6 **Armstrong N**, Williams J, Balding J, *et al*. The peak oxygen uptake of British children with reference to age, sex and sexual maturity. *Eur J Appl Physiol* 1991;**62**:369–75.
7 **Cooper DM**, Weiler-Ravell D, Whipp BJ, *et al*. Growth-related changes in oxygen uptake and heart rate during progressive exercise in children. *Pediatr Res* 1984;**18**:845–51.
8 **Lemura LM**, von Duvillard SP, Cohen SL, *et al*. Treadmill and cycle ergometry testing in 5- to 6-year-old children. *Eur J Appl Physiol* 2001;**85**:472–8.
9 **McMiken DF**. Maximum aerobic power and physical dimensions of children. *Ann Hum Biol* 1976;**3**:141–7.
10 **Wong TW**, Yu TS, Wang XR, *et al*. Predicted maximal oxygen uptake in normal Hong Kong Chinese schoolchildren and those with respiratory diseases. *Pediatr Pulmonol* 2001;**31**:126–32.
11 **Shuleva KM**, Hunter GR, Hester DJ, *et al*. Exercise oxygen uptake in 3- through 6-year-old children. *Pediatr Exerc Sci* 1990;**2**:130–9.
12 **Andersen LB**, Henckel P, Saltin B. Maximal oxygen uptake in Danish adolescents 16–19 years of age. *Eur J Appl Physiol* 1987;**56**:74–82.
13 **Knuttgen HG**. Aerobic capacity of adolescents. *J Appl Physiol* 1968;**22**:655–8.
14 **Zhu S**, St Onge MP, Heshka S, *et al*. Lifestyle behaviors associated with lower risk of having the metabolic syndrome. *Metabolism* 2004;**53**:1503–11.
15 **Brage S**, Wedderkopp N, Franks PW, *et al*. Reexamination of validity and reliability of the CSA monitor in walking and running. *Med Sci Sports Exerc* 2003;**35**:1447–54.
16 **Krahenbuhl GS**, Skinner JS, Kohrt WM. Developmental aspects of maximal aerobic power in children. *Exerc Sport Sci Rev* 1985;**13**:503–38.
17 **Armstrong N**, Kirby BJ, McManus AM, *et al*. Aerobic fitness of prepubescent children. *Ann Hum Biol* 1995;**22**:427–41.
18 **Armstrong N**, Welsman J. *Young people and physical activity*. London: Health Education Authority, 1997.
19 **De Lorenzo A**, Bertini I, Candeloro N, *et al*. Comparison of different techniques to measure body composition in moderately active adolescents. *Br J Sports Med* 1998;**32**:215–19.
20 **Dezenberg CV**, Nagy TR, Gower BA, *et al*. Predicting body composition from anthropometry in pre-adolescent children. *International Journal of Obesity* 1999;**23**:253–9.
21 **Ekelund U**, Sjöström M, Yngve A, *et al*. Physical activity assessed by activity monitor and doubly labeled water in children. *Med Sci Sports Exerc* 2001;**33**:275–81.
22 **Jensen K**, Jorgensen S, Johansen L. A metabolic cart for measurement of oxygen uptake during human exercise using inspiratory flow rate. *Eur J Appl Physiol* 2002;**87**:202–6.
23 **Rowland TW**. *Developmental exercise physiology*. Champaign, IL: Human Kinetics, 1996:1–269.
24 **Altman DG**. *Practical statistics for medical research*. London: Chapman & Hall/CRC, 1991:36–7.
25 **Yoshida T**, Ishiko I, Muraoka I. Effects of endurance training on cardiorespiratory functions of 5-year old children. *Int J Sports Med* 1980;**1**:91–4.
26 **Åstrand P-O**. Experimental studies of physical working capacity in relation to age and sex. 1952:1–171.
27 **Robinson S**. Experimental studies of physical fitness in relation to age. *Arbeitsphysiologie* 1938;**10**:251–323.
28 **Vinet A**, Mandigout S, Nottin S, *et al*. Influence of body composition, hemoglobin concentration, and cardiac size and function of gender differences in maximal oxygen uptake in prepubertal children. *Chest* 2003;**124**:1494–9.
29 **Rowland T**, Goff D, Martel L, *et al*. Influence of cardiac functional capacity on gender differences in maximal oxygen uptake in children. *Chest* 2000;**117**:629–35.
30 **Sundberg S**. Maximal oxygen uptake in relation to age in blind and normal boys and girls. *Acta Pædiatr Scand* 1982;**71**:603–8.
31 **Rasmussen S**, Petersen TA, Madsen M. [Body height of 6–15-year-old school children measured in the period 1986/1987 to 1996/1997. Compared with Danish measurements in 1971/1972]. *Ugeskr Laeger* 2002;**164**:5011–15.
32 **Petersen A-G**, Rasmussen S, Madsen M. Danske skolebørns BMI målt i Perioden 1986/87–1996/97 sammenlignet med danske målinger fra 1971/72. *Ugeskr Læger* 2002;**164**:5006–10.
33 **Malina RM**, Bouchard C. *Growth, maturation and physical activity*. Champaign, IL: Human Kinetics, 1991.
34 **Armstrong N**, Welsman JR. Assessment and interpretation of aerobic fitness in children and adolescents. *Exerc Sport Sci Rev* 1994;**22**:435–76.

Persistent Postconcussion Symptoms After Injury

Linda Ewing-Cobbs, PhD,[a,b] Charles S. Cox Jr, MD,[c] Amy E. Clark, MS,[d] Richard Holubkov, PhD,[d] Heather T. Keenan, MDCM, PhD[d]

abstract

OBJECTIVES: We examined whether preinjury, demographic, and family factors influenced vulnerability to postconcussion symptoms (PCSs) persisting the year after mild traumatic brain injury (mTBI).

METHODS: Children with mTBI (n = 119), complicated mild traumatic brain injury (cmTBI) (n = 110), or orthopedic injury (OI) (n = 118), recruited from emergency departments, were enrolled in a prospective, longitudinal cohort study. Caregivers completed retrospective surveys to characterize preinjury demographic, child, and family characteristics. PCSs were assessed using a validated rating scale. With multivariable general linear models adjusted for preinjury symptoms, we examined predictors of PCSs 3, 6, and 12 months after injury in children ages 4 to 8, 9 to 12, and 13 to 15 years at injury. With logistic regression, we examined predictors of chronic PCSs 1 year after traumatic brain injury.

RESULTS: Postinjury somatic, emotional, cognitive, and fatigue PCSs were similar in the mTBI and cmTBI groups and significantly elevated compared with the OI group. PCS trajectories varied with age and sex. Adolescents had elevated PCSs that improved; young children had lower initial symptoms and less change. Despite similar preinjury PCSs, girls had elevated symptoms across all time points compared with boys. PCS vulnerability factors included female sex, adolescence, preinjury mood problems, lower income, and family discord. Social capital was a protective factor. PCSs persisted in 25% to 31% of the traumatic brain injury group and 18% of the OI group at 1 year postinjury. The odds of chronic PCSs were almost twice as high in girls as in boys and were >4 times higher in young children with cmTBI than in those with mTBI.

CONCLUSIONS: A significant minority of children with mTBI and OI have PCSs that persisted 1 year after injury.

WHAT'S KNOWN ON THIS SUBJECT: After mild traumatic brain injury (mTBI), ~15% to 30% of children have postconcussion symptoms (PCSs) for several months. There is little consensus regarding which injury-related, child demographic, preinjury, and family factors confer vulnerability to or protect against PCSs persisting during the first year.

WHAT THIS STUDY ADDS: Vulnerability factors used to predict PCSs persisting during the year after uncomplicated mTBI (25%), complicated mTBI (31%), or orthopedic injury (18%) included preinjury affective problems, female sex, adolescence, and family stresses. Girls had twice the odds of having chronic PCS compared with boys.

To cite: Ewing-Cobbs L, Cox CS, Clark AE, et al. Persistent Postconcussion Symptoms After Injury. *Pediatrics.* 2018; 142(5):e20180939

[a]*Children's Learning Institute and Departments of* [b]*Pediatrics and* [c]*Pediatric Surgery, John P. and Katherine G. McGovern Medical School, The University of Texas Health Science Center at Houston, Houston, Texas; and* [d]*Department of Pediatrics, The University of Utah, Salt Lake City, Utah*

Drs Ewing-Cobbs and Keenan conceptualized and designed the study, drafted the initial manuscript, and reviewed and revised the manuscript; Dr Cox made substantial contributions to the acquisition of data and critically reviewed and revised the manuscript for important intellectual content; Ms Clark and Dr Holubkov made substantial contributions to the analysis and interpretation of data and revised the manuscript for important intellectual content; and all authors approved the final manuscript as submitted and agree to be accountable for all aspects of the work.

DOI: https://doi.org/10.1542/peds.2018-0939

Accepted for publication Aug 29, 2018

Address correspondence to Linda Ewing-Cobbs, PhD, Department of Pediatrics, Children's Learning Institute, The University of Texas Health Science Center at Houston, 7000 Fannin St, Suite 2401, Houston, TX 77030. E-mail: linda.ewing-cobbs@uth.tmc.edu

Postconcussion symptoms (PCSs) are cognitive, physical, and affective symptoms, such as difficulty concentrating, headache, and irritability, that occur in ~30% children with mild traumatic brain injury (mTBI) seen in the emergency department (ED). Although PCSs resolve in many children with mTBI within 1 week to 1 month,[1,2] symptoms persist for ≥1 month in 11% to 30% and negatively impact functioning at home and school.[2–5] PCSs or "concussionlike symptoms" are relatively nonspecific and are endorsed after traumatic brain injury (TBI), to a lesser extent by children experiencing bodily injuries,[3,6–8] and by some children without injuries.[9] Even in children whose PCSs resolve, persistent reductions in health-related quality of life, particularly in physical or academic areas, are documented in children who are managed up to 1 year after injury.[10,11]

The literature is inconsistent regarding injury and noninjury factors that may place children at high risk for prolonged PCSs. Greater injury severity and positive computed tomography (CT) scan findings are often used to predict acute PCSs; however, preinjury characteristics of the child and family, including increasing age, female sex, poorer preinjury child adjustment, and family dysfunction, may be used to predict more chronic PCSs.[2,5,12–15] Persistent PCSs have a major impact on both health care and school systems. With 1 million to 2 million US children sustaining a concussion from just sport and recreation participation annually,[16] it is essential to identify injury and noninjury factors that either promote or hinder recovery from mTBI.

To address gaps in the literature, we examined parent ratings of PCSs in a prospective, longitudinal cohort study of recovery from pediatric mTBI relative to an orthopedic injury (OI) comparison group. We hypothesized that elevated parent ratings of PCSs during the first year after a mTBI would be associated with vulnerability factors of greater injury severity (including loss of consciousness and the presence of neuroimaging abnormalities), demographic variables (including older age and female sex), preinjury child factors (including learning and psychological health difficulties), and family circumstances (including poverty and poorer family functioning). Social capital, or a family's connectedness to the community, was expected to buffer the effects of injury on PCSs.

METHODS

Participants were children ages 4 to 15 years who sustained a mTBI or OI and were recruited from the ED at 2 level 1 pediatric trauma centers (Children's Memorial Hermann Hospital and the University of Texas Health Science Center at Houston [Houston, TX] and Primary Children's Hospital [Salt Lake City, UT]) as part of a larger cohort study from January 2013 through September 2015. Recruitment was sequential and stratified by age at injury (4–5, 6–11, and 12–15 years), type of injury, and TBI severity. Children with severe psychiatric disorders or developmental delay were excluded because of difficulty assessing the impact of injury on outcomes. Institutional review board approval was received from each institution. Parents provided consent. Children ≥8 years of age provided assent.

Definitions

TBI severity was based on the lowest ED Glasgow Coma Scale (GCS) score.[17] mTBI was defined on the basis of the World Health Organization[18] and Centers for Disease Control and Prevention[19] criteria of a GCS score of 13 to 15 on presentation for health care with 1 or more of the following: confusion or disorientation, loss of consciousness for ≤30 minutes, posttraumatic amnesia for <24 hours, the presence or absence of a skull fracture, and/or other transient neurologic abnormalities. Complicated mild traumatic brain injury (cmTBI) met the above criteria but included an intracranial contusion or hemorrhage diagnosed by using a CT scan.[20] CT imaging in the ED was performed for clinical indication only. Those in an OI comparison group sustained an extremity fracture but no head injury. The Abbreviated Injury Scale[21] score and Injury Severity Score were assigned by trauma registrars.

Procedure

Parents or legal guardians completed a survey as soon as possible after injury to characterize preinjury family structure, sociodemographic, and child characteristics. Follow-up surveys were scheduled for 3, 6, and 12 months after injury. English-speaking families completed surveys either online or by telephone interview; Spanish-speaking families completed telephone interviews with bilingual study coordinators.

Child and Family Measures

Child Behavior Checklist

We used the attention-deficit/hyperactivity disorder (ADHD), affective disorder, and anxiety disorder scales yielding t scores normed for age and sex; higher scores indicate more problems.[22] The Child Behavior Checklist (CBCL) has excellent test-retest reliability (r = 0.7–0.8) and internal consistency (α = 0.90–0.94) at 1 year.

Postconcussion Symptom Inventory–Parent

The Postconcussion Symptom Inventory–Parent (PCSI-P) scale is a validated parent-report measure used to provide a total score and physical, cognitive, emotional, and fatigue subscores.[23] It has 20 developmentally appropriate items

that are used to discriminate children who are concussed from those who are uninjured ages 5 to 15 years; we extended the age range to include 4-year-olds for consistency with other survey measures. A global question asked to what degree the child acted differently than before the injury. The PCSI-P total score has favorable internal consistency (α = 0.94). Higher scores indicate more symptoms.

The presence or absence of PCSs was dichotomized on the basis of the *International Classification of Diseases, 10th Revision* (ICD-10) criterion of at least 1 symptom being present (or, for follow-up, increasing relative to preinjury) in at least 3 of the following areas: cognitive, emotional, somatic, and sleep and/or fatigue.

Family Environment Covariates

Preinjury family function was assessed by using the McMaster Family Assessment Device (FAD)–General Functioning Scale.[24] The FAD has 12 items scored 1 to 4; higher scores represent worse functioning. The Social Capital Index is used to measure perceptions of personal, family, neighborhood, and spiritual community support; higher total scores indicate greater support.[25] Families self-reported race, ethnicity, and income by family size; we calculated income relative to the poverty level using federal norms.

Statistical Approach

All children with outcomes available at preinjury and at least 1 follow-up time point were included in the analysis. Generalized linear mixed models, in which maximum likelihood estimation is used to incorporate all available outcome data, were fit with an unstructured covariance matrix and empirical estimates of the SE for model parameters for PCSI-P total scores and subscores by using SAS PROC MIXED (SAS Institute, Inc, Cary, NC). Potentially clinically important covariates were selected a priori, including the 3-way and 2-way interactions between injury group, time, and age (4–8, 9–12, and 13–15 years). Additional candidate covariates, including enrollment site, injury factors (previous concussion and loss of consciousness), child characteristics (sex; race and/or ethnicity; preinjury learning, behavioral, or developmental delay; and preinjury CBCL scores), and parent and/or family factors (respondent education, poverty level, preferred language, FAD score, social capital) were initially screened in a model controlled for preinjury PCSI-P score, injury group, time, and injury group by time interaction. To develop the final reported longitudinal models, a full model that included all candidate covariates with $P < .20$ in initial screening, and these 4 factors, was iteratively reduced by removing variables (excluding preinjury PCSI-P score and main effects of injury and time) with $P > .1$. To identify predictors of chronic PCSs (yes or no) at the 1-year follow-up, multivariable logistic regression models were constructed by using an analogous approach. In all analyses, we used a significance level of .05.

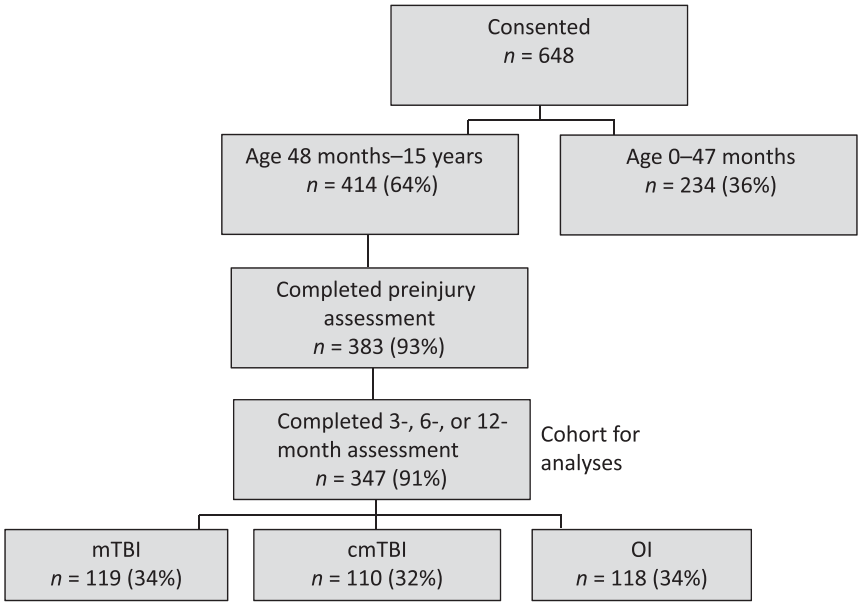


FIGURE 1
Flow diagram of cohort recruitment.

RESULTS

Study Population

Of the 414 children who consented to participate, 383 (93%) completed the initial survey in which we assessed retrospective ratings of preinjury PCSs and child and family functioning. The final cohort contained 347 (91%) children completing at least 1 postinjury assessment: 119 children with mTBI, 110 with cmTBI, and 118 with OI (Fig 1). In Supplemental Table 5, we compare key variables for children who did and did not have complete data; retention was lower in Hispanic children from the Houston site. Most families (76%) completed surveys online and had an English language preference (90%). The injury groups did not differ significantly on age, sex, race, or parental employment; however, parent income and education were lower in the mTBI group. Preinjury child psychological health and PCS estimates did not differ significantly across groups (Tables 1 and 2) or by sex.

TABLE 1 Child, Family, and Injury Characteristics by Injury Group

	mTBI (*N* = 119)	cmTBI (*N* = 110)	OI (*N* = 118)	*P*[a]
Child and family characteristics				
Enrollment site Texas (versus Utah), *n* (%)	50 (42)	44 (40)	52 (44)	.82
Prefer to complete surveys online, *n* (%)	82 (69)	93 (85)	90 (76)	.02
Preferred language Spanish (versus English), *n* (%)	14 (12)	4 (4)	15 (13)	.04
Age at injury, y, mean (SD)	10.3 (3.7)	10.5 (3.5)	9.7 (3.7)	.16
Child sex female, *n* (%)	44 (37)	33 (30)	44 (37)	.43
Child race, *n* (%)				.14
American Indian or Alaskan native	2 (2)	0 (0)	0 (0)	
Asian American	1 (1)	5 (5)	2 (2)	
African American	11 (9)	4 (4)	6 (5)	
Native Hawaiian or other Pacific Islander	1 (1)	0 (0)	0 (0)	
White	90 (76)	93 (85)	95 (82)	
Multiracial	13 (11)	7 (6)	13 (11)	
Child ethnicity Hispanic or Latino, *n* (%)	31 (26)	17 (16)	33 (28)	.06
Married parents, *n* (%)	85 (73)	90 (83)	83 (72)	.11
Either caregiver currently employed, *n* (%)	112 (94)	105 (95)	106 (90)	.21
Respondent education, *n* (%)				.001
Less than high school	18 (15)	7 (6)	16 (14)	
High school	29 (24)	16 (15)	14 (12)	
Vocational and/or some college	36 (30)	59 (54)	42 (36)	
Bachelor's degree or more	36 (30)	28 (25)	46 (39)	
Income at or below poverty level, *n* (%)	37 (34)	11 (11)	21 (19)	<.001
Insurance type, *n* (%)				.29
None	9 (8)	10 (9)	5 (4)	
Medicaid and/or CHIP	43 (36)	31 (28)	33 (28)	
Commercial, private, and/or military	67 (56)	68 (62)	80 (68)	
Previous concussion with ED or doctor visit, *n* (%)	10 (8)	8 (7)	5 (4)	.41
Developmental, learning, or behavioral problem, *n* (%)	17 (14)	13 (12)	9 (8)	.26
CBCL affective t score, mean (SD)	55.3 (7.5)	54.4 (6.6)	54.2 (6.0)	.38
CBCL anxiety t score, mean (SD)	53.5 (6.2)	53.8 (6.6)	53.7 (5.4)	.95
CBCL ADHD t score, mean (SD)	55.0 (7.2)	53.5 (5.7)	53.8 (6.0)	.16
FAD general functioning scale, mean (SD)	1.5 (0.5)	1.5 (0.4)	1.5 (0.5)	.66
Social Capital Index, mean (SD)	3.5 (1.1)	3.6 (1.0)	3.7 (1.0)	.44
Injury characteristics				
Injury mechanism, *n* (%)				<.001
Pedestrian or bicycle	21 (18)	15 (14)	4 (3)	
Motorized vehicle	30 (25)	24 (22)	11 (9)	
Fall	44 (37)	54 (49)	83 (70)	
Struck by or against	12 (10)	10 (9)	6 (5)	
Organized sport	10 (8)	4 (4)	12 (10)	
Other	2 (2)	3 (3)	2 (2)	
Loss of consciousness (yes), *n* (%)	45 (38)	49 (45)	0 (0)	<.001
ED GCS (lowest postresuscitation), median (Q1, Q3)	15 (15, 15)	15 (14, 15)	—	.09
Maximum AIS excluding head, median (Q1, Q3)	1 (0, 2)	1 (0, 1)	2 (2, 2)	<.001
ISS score, median (Q1, Q3)	5 (1, 10)	10 (9, 16)	4 (4, 5)	<.001
Head imaging in ED (CT), *n* (%)	104 (87)	110 (100)	—	<.001
Injuries seen on brain imaging, *n* (%)				
Skull fracture	27 (23)	74 (67)	—	<.001
Cortical contusion	0 (0)	32 (29)	—	—
Hemorrhage	0 (0)	97 (88)	—	—
Admission type, *n* (%)				<.001
ED and/or observation only	53 (45)	5 (5)	78 (66)	
Hospital but not PICU	50 (42)	51 (46)	39 (33)	
PICU	16 (13)	54 (49)	1 (1)	
Hospital LOS, d, median (Q1, Q3)	2 (1, 4)	2 (2, 3)	2 (1, 3)	.80

AIS, Abbreviated Injury Scale; CHIP, Children's Health Insurance Program; ISS, Injury Severity Score; Q1, first quartile; Q3, third quartile; —, not applicable.

[a] *P* values reflect tests of association with injury group, specifically the χ^2 test for categorical variables, analysis of variance for continuous variables summarized by using the mean, and the Kruskal-Wallis test for continuous variables summarized by using the median.

TABLE 2 Unadjusted PCSI-P Outcomes and *ICD-10* Concussion Criteria

Injury Group	PCSI Cluster and Total Scores					PCSI Global Outcome	*ICD-10*
	Somatic	Emotional	Cognitive	Fatigue	Total	Act Differently	Concussion Criteria[a]
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	*n* (%)	*n* (%)
mTBI							
Preinjury	2.1 (4.6)	1.6 (3.4)	2.0 (4.6)	0.5 (1.4)	6.1 (12.1)	—	26 (22)
Month 3	3.7 (6.1)	3.1 (4.6)	2.9 (4.8)	1.6 (3.1)	11.3 (15.7)	54 (48)	35 (31)
Month 6	3.4 (5.6)	3.6 (4.6)	3.1 (5.2)	1.5 (2.9)	11.5 (16.0)	44 (40)	41 (37)
Month 12	3.2 (5.7)	2.5 (4.4)	2.5 (4.6)	1.0 (2.3)	9.3 (13.8)	40 (39)	25 (25)
cmTBI							
Preinjury	1.4 (3.5)	1.4 (3.0)	1.0 (2.9)	0.7 (2.2)	4.4 (9.8)	—	17 (15)
Month 3	4.0 (6.7)	3.4 (5.4)	2.7 (5.2)	1.8 (3.2)	12.0 (18.2)	67 (64)	38 (37)
Month 6	2.9 (4.7)	4.0 (5.0)	2.8 (4.5)	1.6 (2.8)	11.3 (15.2)	48 (47)	38 (37)
Month 12	2.5 (5.0)	2.9 (4.5)	2.3 (4.1)	1.3 (3.1)	9.0 (14.7)	45 (46)	30 (31)
OI							
Preinjury	1.8 (3.8)	1.3 (2.8)	1.1 (3.4)	0.5 (1.7)	4.7 (9.8)	—	20 (17)
Month 3	1.4 (3.5)	1.9 (3.1)	0.9 (2.1)	0.7 (1.8)	4.9 (8.2)	38 (36)	17 (16)
Month 6	1.6 (3.5)	2.0 (3.1)	1.4 (2.9)	0.7 (1.8)	5.7 (9.2)	21 (19)	23 (21)
Month 12	1.3 (3.0)	1.3 (2.9)	0.8 (2.2)	0.6 (1.4)	4.0 (7.4)	23 (21)	19 (18)

—, not applicable.

[a] Preinjury: 1 symptom in at least 3 clusters; postinjury: ≥1 new or elevated postinjury symptom in at least 3 clusters.

Injury mechanism differed, with the youngest age group sustaining the mildest injuries, primarily from falls. Loss of consciousness was reported in 11% of younger and 33% to 41% of older children.

Outcomes

Injury Group and Time Since Injury

Before injury, 18% of children had symptoms that were consistent with *ICD-10* concussion diagnostic criteria (Table 2). Figure 2 includes the unadjusted PCSI-P group means from preinjury through the 12-month follow-up. Children were similar at baseline except for slightly higher cognitive scores seen in children with mTBI. Children with TBI had elevated scores at all follow-up time points that did not return to baseline. Table 3 includes multivariable model results across the 3 time points that were adjusted for preinjury PCSI-P ratings. Children with cmTBI (6 points) and mTBI (3.5 points) had higher adjusted scores compared with those with OI on the total postinjury PCSI-P and all subscales; however, those with mTBI and cmTBI did not differ significantly from each other.

Time from injury was important. Emotional and cognitive symptoms increased from 3 to 6 months but then fell at 12 months. Total, somatic, and fatigue symptoms resolved differently over time depending on age; 4- to 8-year-olds had lower scores at 3 months than older children, which either did not change or increased across follow-up (Fig 3). Older children had higher total and somatic scores than the 4- to 8-year-old group at 3 months, but their symptoms decreased over time. Adolescents had the highest fatigue symptoms, which had decreased by 1 year after injury.

Demographic, Child, and Family Predictors of PCSs

Girls had higher unadjusted postinjury symptoms than boys in all areas of the PCSI-P despite similar preinjury PCSs (Fig 4). In the adjusted analysis, total scores remained 3.4 points higher for girls than for boys across the follow-up.

Preexisting affective problems, as measured by using the CBCL, were associated with elevated PCSs. Preinjury CBCL affective, anxiety, and ADHD scores were significantly associated with postinjury PCSs in univariable analyses. Only the affective score remained significant in multivariable analyses.

Family characteristics, including lower income, were associated with higher symptom burden; poorer family functioning predicted greater emotional and cognitive symptoms. Hispanic ethnicity was protective for emotional symptoms; Spanish language preference was protective for both total and somatic symptoms. Higher social capital was associated with lower symptom burden.

PCSs 12 Months After Injury

Despite significant improvement in PCSI-P scores over time, the PCSI-P global outcome question revealed that 21%, 39%, and 46% of the OI, mTBI, and cmTBI groups, respectively, continued to act differently 1 year after injury (Table 2). Most changes were mild to moderate; however, 3%, 8%, and 10% of the groups, respectively, showed significant to major differences.

Chronic postinjury concussion symptoms, defined as ≥1 symptom increasing in at least 3 areas relative to preinjury at 12 months postinjury, were identified in 18%, 25%, and

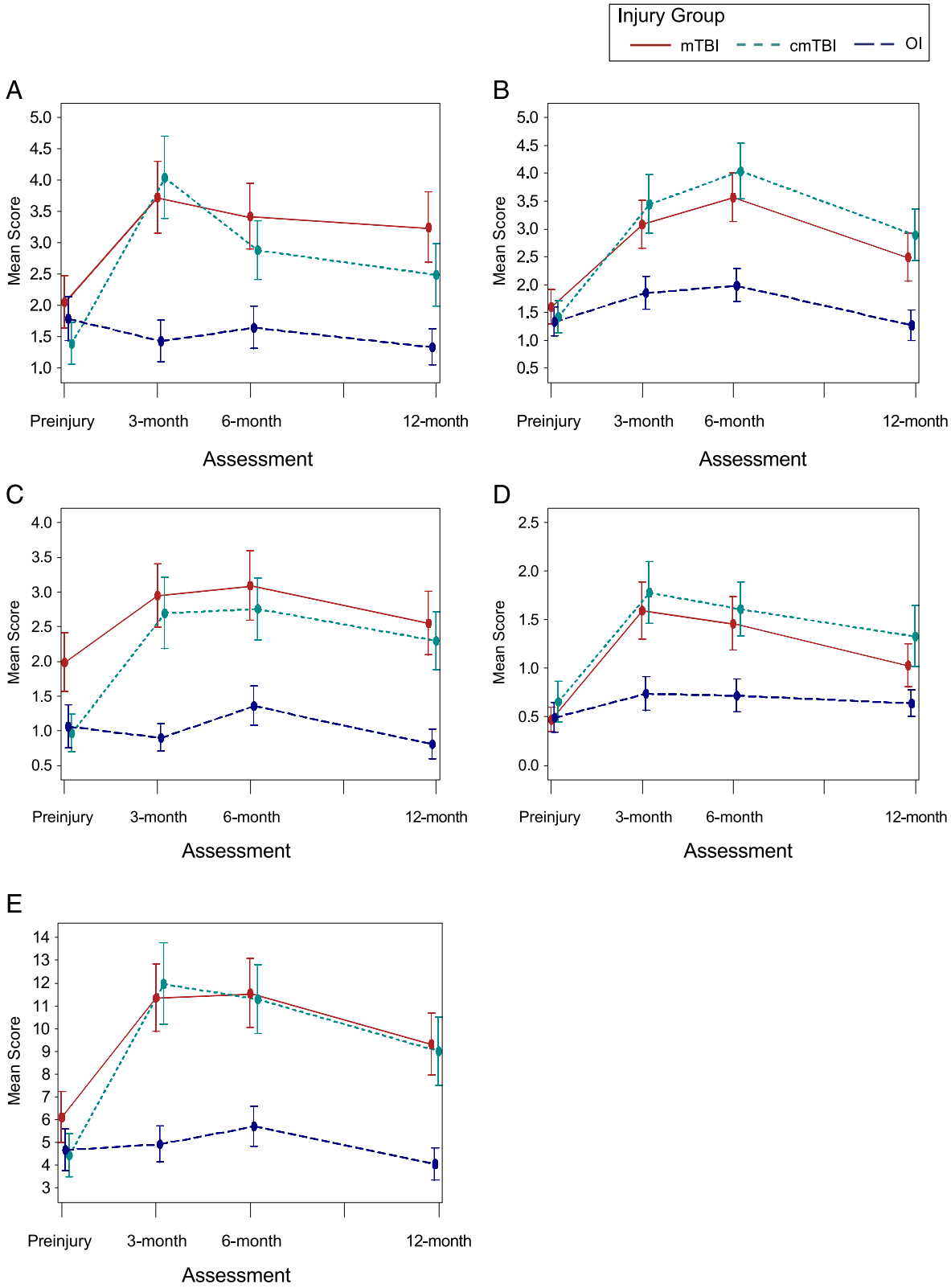


FIGURE 2

Unadjusted longitudinal PCSI-P total and subscore means (±1 SE) by injury group. Despite similar preinjury ratings, both the mTBI and cmTBI groups showed increases in all PCSI-P scores at the 3-month time point relative to the OI group that persisted across the follow-up. A, Somatic. B, Emotional. C, Cognitive. D, Fatigue. E, PCSI total.

TABLE 3 Multivariable Model Results for PCSI-P Outcomes

	Somatic ($N = 315$)		Emotional[a] ($N = 313$)		Cognitive[b] ($N = 314$)		Fatigue ($N = 323$)		Total[c] ($N = 315$)	
	Coefficient (SE)	*P*	Coefficient (SE)	*P*	Coefficient (SE)	*P*	Coefficient (SE)	*P*	Coefficient (SE)	*P*
Preinjury rating of outcome	0.46 (0.13)	<.001	0.40 (0.07)	<.001	0.31 (0.08)	<.001	0.34 (0.14)	.02	0.45 (0.08)	<.001
Injury severity, time										
Injury severity (versus OI)		.001		<.001		<.001		.003		<.001
mTBI	1.23 (0.41)		0.88 (0.34)		0.96 (0.31)		0.45 (0.20)		3.52 (1.03)	
cmTBI	1.67 (0.51)		1.67 (0.42)		1.67 (0.39)		0.85 (0.27)		5.97 (1.42)	
Time from injury (versus 3 mo)	Figure 2	.02		<.001		.01	Figure 2	.008	Figure 2	<.001
6 mo			0.54 (0.22)		0.22 (0.21)					
12 mo			−0.55 (0.22)		−0.38 (0.21)					
Demographics										
Age at injury	Figure 2	.34	—	—	—	—	Figure 2	.02	Figure 2	.62
Age by time interaction	Figure 2	.003	—	—	—	—	Figure 2	.01	Figure 2	.02
Female sex	1.17 (0.42)	.006	0.94 (0.36)	.01	0.84 (0.33)	.01	0.61 (0.23)	.008	3.43 (1.19)	.004
Spanish preferred language	−1.44 (0.58)	.01	—	—	—	—	−0.51 (0.27)	.06	−4.03 (1.36)	.003
Preexisting problems										
CBCL affective problems score	0.11 (0.04)	.002	0.10 (0.03)	.002	0.08 (0.04)	.06	0.09 (0.02)	<.001	0.38 (0.10)	<.001
Family environment										
Income relative to poverty level	−0.29 (0.09)	.002	−0.21 (0.09)	.02	−0.18 (0.07)	.01	−0.17 (0.05)	<.001	−0.80 (0.25)	.002
FAD family function	—	—	1.55 (0.44)	<.001	0.86 (0.43)	.046	—	—	2.89 (1.30)	.03
Social capital	−0.68 (0.23)	.004	−0.54 (0.20)	.006	−0.56 (0.19)	.003	—	—	−1.78 (0.66)	.008

—, variable not included in the final model for the outcome.

[a] The emotional outcome model also included race ($P = .001$); Hispanic participants had a lower score on average than non-Hispanic white participants (effect estimate -1.48 ± 0.38).

[b] The cognitive outcome model also included the CBCL ADHD score ($P = .09$) and previous diagnosis of learning, behavioral, or developmental delay ($P = .08$).

[c] Example interpretation of the total outcome model effect estimates: the total PCSI outcome was higher, on average, for those with mTBI (3.5 points) and cmTBI (6.0 points) compared with those with an OI in the year after an injury. The effect of age at injury differed by time (Fig 2). Total scores were 3.4 points higher for girls than boys, on average, and 4.0 points lower for those with Spanish as a preferred language. Higher preinjury affective problems, lower income, poorer family function, and lower social capital were all associated with increased total PCSI scores.

31% of children with OI, mTBI, and cmTBI, respectively (Table 2). In a multivariable model adjusted for preinjury PCSs, the odds of chronic PCSs were higher for girls and children with poorer family function and lower social capital. The odds of chronic PCSs were increased in 4- to 8-year-olds with cmTBI relative to both those with mTBI and OI but not for older children with mTBI or cmTBI relative to those with OI (Table 4).

DISCUSSION

In the current study, we examined injury characteristics as well as demographic, preinjury, and family predictors of persistent PCSs during the first year after TBI in a broadly generalizable cohort of children. Key findings include the striking persistence of PCSs, particularly in girls; the differences in PCS trajectories by age; and the strong association of preinjury PCS and psychological heath symptoms with persistent PCSs. Different characteristics of the family environment influenced PCSs and served as either protective or vulnerability factors. One year after injury, parents rated >40% of children with TBI as acting differently than before the injury, and 25% to 31% had postinjury symptoms meeting concussion diagnostic criteria. Our results converge with findings in other longitudinal studies in which researchers recruited children from EDs and reported ~20% to 30% of children with new or reemerging symptoms persisting at 3 or 12 months after injury.[2,3,6]

The high rate of chronic concussion symptoms is of concern because of the strong relation between persisting PCSs and reduced health-related quality of life.[10–12,26] PCSs vary over time; physical effects occur immediately after injury, cognitive symptoms occur throughout, and emotional symptoms develop later.[27]

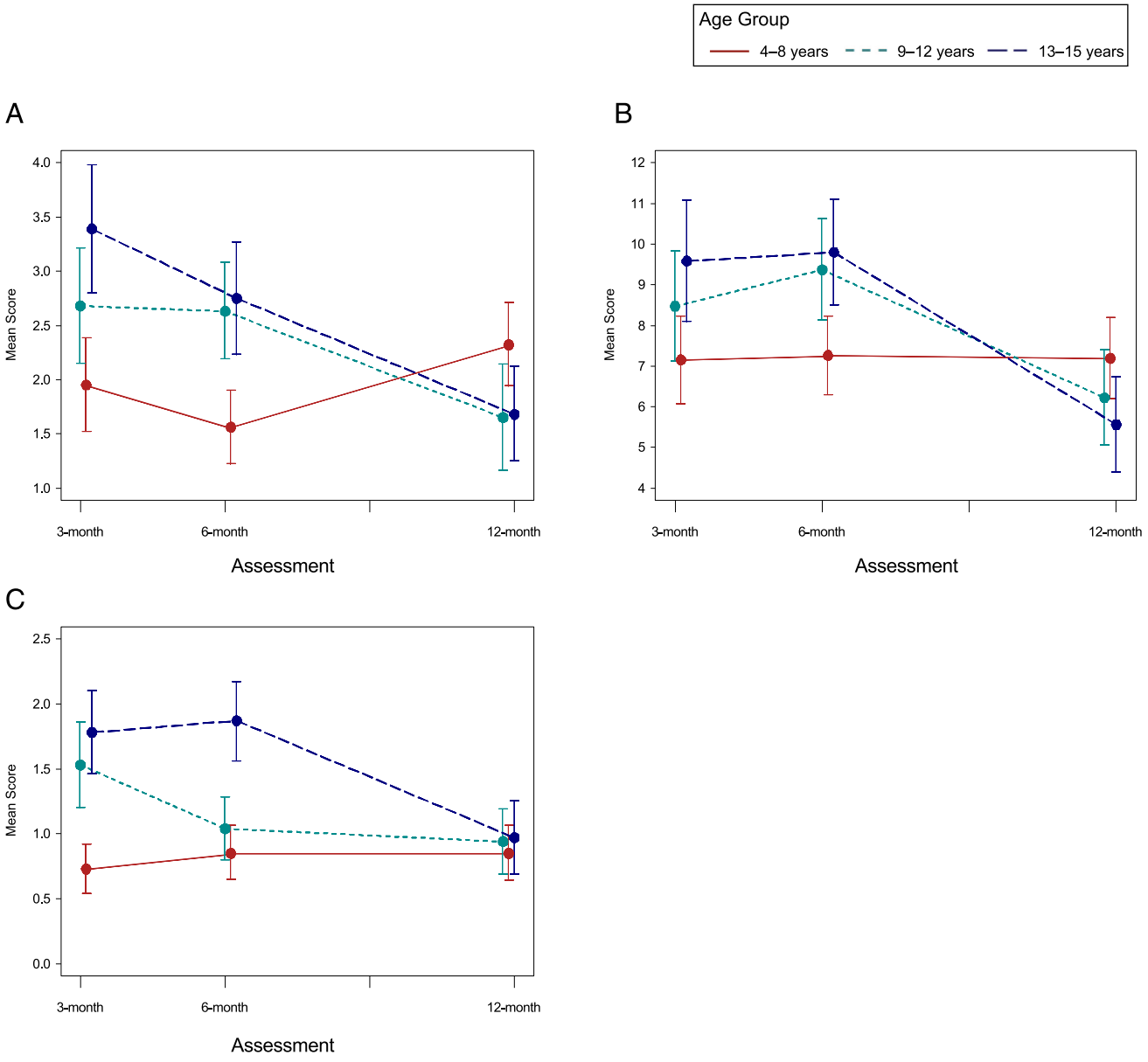


FIGURE 3

Interactions between age and time for PCSI-P scores. Least squares means (±1 SE) for somatic, fatigue, and total scores, assuming average values of preinjury scores and other covariates, are shown. Children 4 to 8 years old had the lowest scores at 3 months and less change over time; 9- to 15-year-olds had higher initial total and somatic scores, and adolescents had the highest fatigue symptoms that decreased over time. A, Somatic. B, PCSI total. C, Fatigue.

Persisting cognitive and emotional PCSs likely contribute to reduced school functioning[10] and changes in psychological health.[28] Little is known about how psychological characteristics, such as negative attributions, or physiologic changes in stress response and neural systems contribute to PCSs.[29–32] Recently, structural imaging revealed associations of brain network abnormalities in children with persistent PCSs after mTBI that improved with aerobic training.[29] This reveals both neural changes after injury and their potential response to interventions.

The nonspecific nature of PCSs is underscored by the substantial rates of PCSs in children with no brain injury. Before injury, 18% of our sample met *ICD-10* criteria for a concussion diagnosis; 1 year after injury, 18% of those in the OI group had postinjury-onset PCSs. Yeates et al[6] also found that children with mTBI and OI had a comparable rate of PCSs of moderate severity across the first year after injury. It is becoming increasingly clear that there is a general effect of injury on PCSs as well as on neurocognitive outcomes in children with mTBI

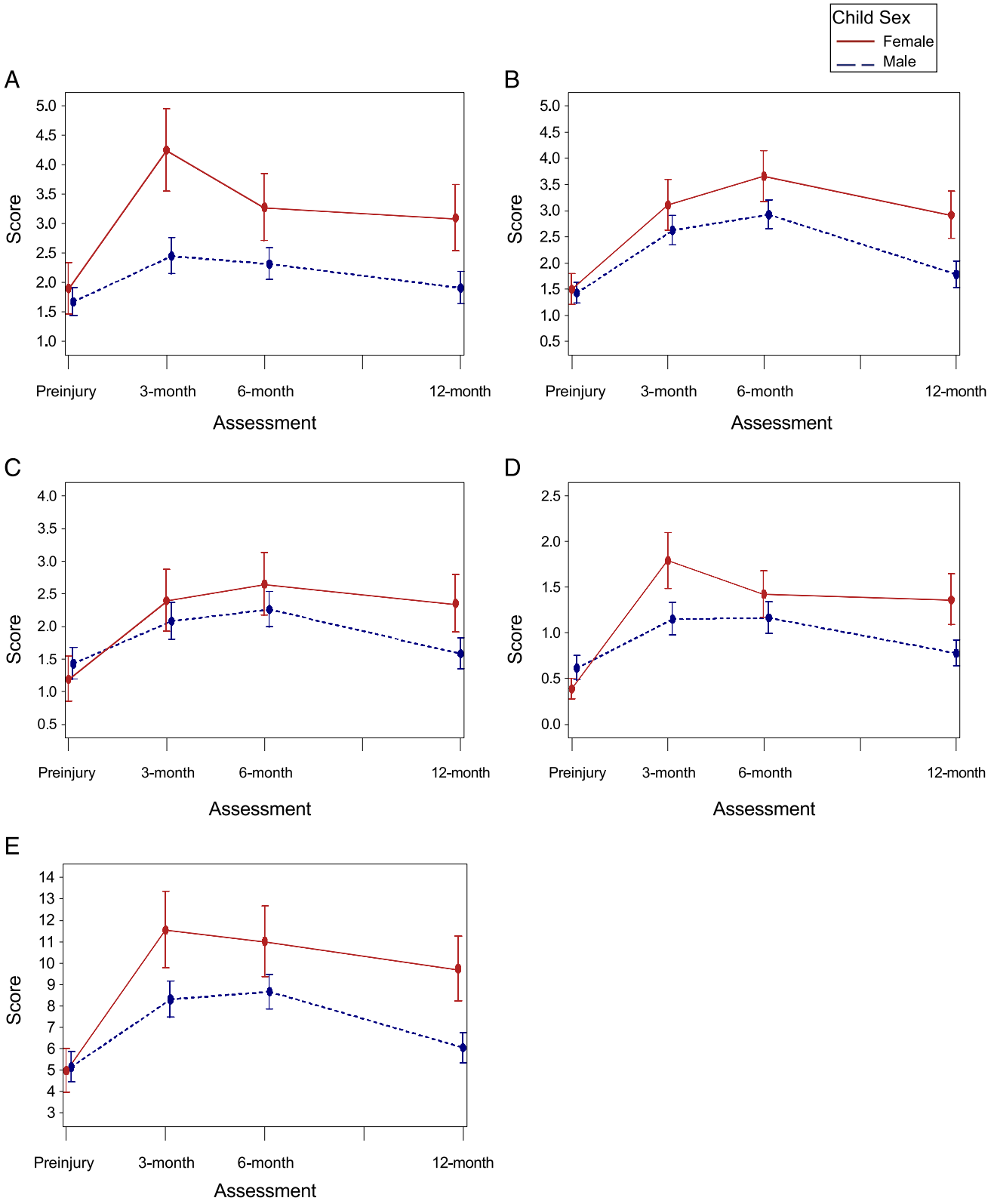


FIGURE 4
Influence of child sex on longitudinal PCSI-P scores (means ± 1 SE). Girls had higher unadjusted postinjury total and subscale PCSI-P symptoms than boys despite similar preinjury PCSs. A, Somatic. B, Emotional. C, Cognitive. D, Fatigue. E, PCSI total.

TABLE 4 Logistic Regression Results for Chronic Postinjury Concussion Symptoms at 12 Months After Injury ($N = 300$)

	Adjusted Odds Ratio (95% CI)	*P*
Preinjury concussion[a]	1.92 (0.93–3.96)	.08
Injury severity by age at injury, y		.04
4–8[b]		
mTBI versus OI	1.57 (0.41–5.99)	
cmTBI versus OI	7.01 (2.10–23.4)	
9–12		
mTBI versus OI	0.60 (0.17–2.04)	
cmTBI versus OI	1.03 (0.31–3.40)	
13–15		
mTBI versus OI	2.60 (0.69–9.83)	
cmTBI versus OI	1.37 (0.33–5.66)	
Female (versus male) sex	1.94 (1.05–3.59)	.04
Family function (1-point increase)	2.42 (1.23–4.74)	.01
Social capital (1-point increase)	0.66 (0.50–0.88)	.01

One or more symptoms increased relative to preinjury at the 12-mo follow-up in at least 3 of the following areas: cognitive, emotional, somatic, and sleep and/or fatigue. CI, confidence interval.

[a] One or more symptoms at preinjury in at least 3 of the following areas: cognitive, emotional, somatic, and sleep and/or fatigue.

[b] In children 4 to 8 y old, having a cmTBI was associated with increased odds of chronic concussion relative to mTBI (odds ratio 4.47; 95% confidence interval 1.38–14.5).

and OI.[33] The development of next-generation PCS measures will help better discriminate symptom profiles in children with brain versus bodily injuries.[29–32,34]

Girls were almost twice as likely as boys to have persistent PCSs. Despite similar preinjury PCSs and psychological health symptoms, girls developed more elevated postinjury PCSs than boys in all areas that did not resolve. Elevated PCSs have been reported in girls recruited from both ED and sport samples.[3,9,12,13,35,36] However, the basis for sex differences and their relation with age at injury is unknown. In sport-related concussion samples, girls report more symptoms before and after a concussive event and have a slower recovery trajectory than do boys.[35] These samples contain predominantly postpubertal girls, which has raised the possibility that altered hypothalamic-pituitary-gonadal axis or other physiologic sex-linked differences may contribute to female vulnerability to PCSs.[35,37] Although this mechanism may help in explaining symptom burden in pubertal and postpubertal girls, girls in our sample had elevated PCSs irrespective of age. PCSs are also related to psychological health problems that may occur more frequently in girls than in boys, including anxiety and posttraumatic stress symptoms.[38,39]

Adolescents and young children had different PCS patterns across the first year after injury. Similar to previous studies, adolescence was a susceptibility factor for total PCSs[3,12,14,40,41] and somatic and fatigue problems. This symptom elevation may be due to higher injury severity because more adolescents had a loss of consciousness and higher-velocity injuries. Although adolescents had more PCSs, their symptoms tended to improve over time. Children in the 4- to 8-year-old group tended to have a lower symptom burden at 3 months, but their symptoms did not improve over time. The odds of chronic PCSs were >4 times higher in children 4 to 8 years old with cmTBI than in those with mTBI and were 7 times higher than in the OI group.

Assessment of functioning before injury is necessary to dissociate postinjury from preexisting symptoms.[2,9,12] Preinjury PCS predicted persistent postinjury PCSs across time points. Similarly, preinjury CBCL affective problem scores predicted elevated postinjury somatic, emotional, and fatigue PCSI scores. The preinjury PCSI-P total score was strongly related to each CBCL score, indicating that these measures share variance related to preinjury adjustment. Although the affective problems score had the most consistent relation with PCSs, a variety of preinjury psychological health issues may influence persistent PCSs.

Family factors exerted independent effects on PCSs, with low income and less adaptive family functioning being associated with greater PCS burden. Hispanic ethnicity and/or preference for Spanish language usage were protective factors for emotional, somatic, and total symptoms. Hispanic ethnicity has been associated with health disparities[42] and lower receipt of outpatient psychological health services after pediatric TBI.[43,44] However, Hispanic families may have cultural features, such as extended family support, that promote resilience.[45] Greater social capital was associated with lower rates of somatic, emotional, and cognitive PCSs. Families with greater social networks and connection to community resources may better access to support services, buffer health inequalities, and reduce the risk of adverse outcomes after injury.

We identified several vulnerability factors for prolonged PCSs that may put children at risk for decreased school participation. Consistent with American Academy of Pediatrics guidelines, children with PCSs should be served under return to learning initiatives in which the collaboration of medical, family, and school teams is emphasized.[46] The goal is to target symptoms and institute accommodations to return the children to full participation in school and community activities without significant symptom exacerbation.[47] Academic accommodations range from informal academic adjustments

to services mandated under federal statutes, such as Section 504. Although evidence-based information regarding interventions is lacking, physical and psychological health interventions ranging from graduated exercise to medication management of headache and mood, cognitive behavioral therapy, and family services that are effective in other populations are likely candidates.[48,49]

LIMITATIONS AND STRENGTHS

Limitations of this study include that data were collected via parent report, which may be subject to bias[50,51] and possible under- or overestimation of PCSs relative to self-report. We did not assess PCSs in the initial weeks after injury and may have lost information regarding characteristics of children who recovered quickly. We did not measure pubertal development or litigation status. Our sample was recruited from the ED, and it may not be generalizable to the larger group of children with mTBI who seek community treatment.[52] Although 18% did not complete all time points, multivariable model results for PCSI-P total scores were similar when they were rerun, including only the cases with complete data.

Our multicenter study had several notable strengths, including a prospective, longitudinal cohort design with broad racial and ethnic representation and an injury comparison group. Careful evaluation of preinjury psychological and physical heath by using validated measures allowed for the dissociation of new postinjury symptoms from preexisting symptoms as well as the identification of the subgroups that were at elevated risk for chronic PCSs.

CONCLUSIONS

Clinical management of children with mTBI, as well as children with bodily injuries, may be enhanced by understanding which children are at risk for persistent PCSs. Because emotional and cognitive symptoms may emerge over time, children with symptoms persisting at 1 month after injury should be managed clinically to monitor symptom course and refer for any needed physical, cognitive, or psychological health interventions. The consistent importance of family functioning and social capital on PCS resolution reveals that family support services should be considered as an adjunctive intervention.

ABBREVIATIONS

ADHD: attention-deficit/hyperactivity disorder
CBCL: Child Behavior Checklist
cmTBI: complicated mild traumatic brain injury
CT: computed tomography
ED: emergency department
FAD: McMaster Family Assessment Device
GCS: Glasgow Coma Scale
ICD-10: *International Classification of Diseases, 10th Revision*
mTBI: mild traumatic brain injury
OI: orthopedic injury
PCS: postconcussion symptom
PCSI-P: Postconcussion Symptom Inventory–Parent
TBI: traumatic brain injury

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275).

FINANCIAL DISCLOSURE: The authors have indicated they have no financial relationships relevant to this article to disclose.

FUNDING: Supported by the Centers for Disease Control and Prevention (Cooperative Agreement U01/CE002188). Contents are solely the responsibility of the authors and do not necessarily represent the official views of the Centers for Disease Control and Prevention. Dr Keenan was also supported by the National Institutes of Health *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (K24HD072984). Neither the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development nor the Centers for Disease Control and Prevention played a role in the design or conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the article; or decision to submit the article for publication. Funded by the National Institutes of Health (NIH).

POTENTIAL CONFLICT OF INTEREST: The authors have indicated they have no potential conflicts of interest to disclose.

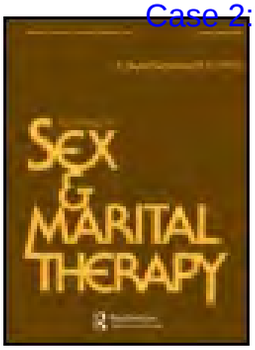
REFERENCES

1. Belanger HG, Vanderploeg RD. The neuropsychological impact of sports-related concussion: a meta-analysis. *J Int Neuropsychol Soc*. 2005;11(4):345–357
2. Barlow KM, Crawford S, Brooks BL, Turley B, Mikrogianakis A. The incidence of postconcussion syndrome remains stable following mild traumatic brain injury in children. *Pediatr Neurol*. 2015;53(6):491–497
3. Babcock L, Byczkowski T, Wade SL, Ho M, Mookerjee S, Bazarian JJ. Predicting postconcussion syndrome after mild traumatic brain injury in children and adolescents who present to the emergency department. *JAMA Pediatr*. 2013;167(2):156–161
4. Grubenhoff JA, Currie D, Comstock RD, Juarez-Colunga E, Bajaj L, Kirkwood MW. Psychological factors associated with delayed symptom resolution in children with concussion. *J Pediatr*. 2016;174:27–32.e1
5. Zemek RL, Farion KJ, Sampson M, McGahern C. Prognosticators of persistent symptoms following pediatric concussion: a systematic review. *JAMA Pediatr*. 2013;167(3):259–265

6. Yeates KO, Taylor HG, Rusin J, et al. Longitudinal trajectories of postconcussive symptoms in children with mild traumatic brain injuries and their relationship to acute clinical status. *Pediatrics*. 2009;123(3):735–743
7. Ponsford J, Willmott C, Rothwell A, et al. Cognitive and behavioral outcome following mild traumatic head injury in children. *J Head Trauma Rehabil*. 1999;14(4):360–372
8. Mittenberg W, Wittner MS, Miller LJ. Postconcussion syndrome occurs in children. *Neuropsychology*. 1997;11(3):447–452
9. Iverson GL, Silverberg ND, Mannix R, et al. Factors associated with concussion-like symptom reporting in high school athletes. *JAMA Pediatr*. 2015;169(12):1132–1140
10. Novak Z, Aglipay M, Barrowman N, et al; Pediatric Emergency Research Canada Predicting Persistent Postconcussive Problems in Pediatrics (PERC 5P) Concussion Team. Association of persistent postconcussion symptoms with pediatric quality of life. *JAMA Pediatr*. 2016;170(12):e162900
11. Moran LM, Taylor HG, Rusin J, et al. Quality of life in pediatric mild traumatic brain injury and its relationship to postconcussive symptoms. *J Pediatr Psychol*. 2012;37(7):736–744
12. McNally KA, Bangert B, Dietrich A, et al. Injury versus noninjury factors as predictors of postconcussive symptoms following mild traumatic brain injury in children. *Neuropsychology*. 2013;27(1):1–12
13. Iverson GL, Gardner AJ, Terry DP, et al. Predictors of clinical recovery from concussion: a systematic review. *Br J Sports Med*. 2017;51(12):941–948
14. Bernard CO, Ponsford JA, McKinlay A, McKenzie D, Krieser D. Predictors of post-concussive symptoms in young children: injury versus non-injury related factors. *J Int Neuropsychol Soc*. 2016;22(8):793–803
15. Hung R, Carroll LJ, Cancelliere C, et al. Systematic review of the clinical course, natural history, and prognosis for pediatric mild traumatic brain injury: results of the International Collaboration on Mild Traumatic Brain Injury Prognosis. *Arch Phys Med Rehabil*. 2014;95(suppl 3):S174–S191
16. Bryan MA, Rowhani-Rahbar A, Comstock RD, Rivara F; Seattle Sports Concussion Research Collaborative. Sports- and recreation-related concussions in US youth. *Pediatrics*. 2016;138(1):e20154635
17. Teasdale G, Jennett B. Assessment of coma and impaired consciousness. A practical scale. *Lancet*. 1974;2(7872):81–84
18. Carroll LJ, Cassidy JD, Peloso PM; WHO Collaborating Centre Task Force on Mild Traumatic Brain Injury, et al. Prognosis for mild traumatic brain injury: results of the WHO collaborating centre task force on mild traumatic brain injury. *J Rehabil Med*. 2004;(suppl 43):84–105
19. National Center for Injury Prevention and Control. *Report to Congress on Mild Traumatic Brain Injury in the US: Steps to Prevent a Serious Public Health Problem*. Atlanta, GA: Centers for Disease Control and Prevention; 2003
20. Levin HS, Hanten G, Roberson G, et al. Prediction of cognitive sequelae based on abnormal computed tomography findings in children following mild traumatic brain injury. *J Neurosurg Pediatr*. 2008;1(6):461–470
21. Gennarelli T A, Wodzin E. *Abbreviated Injury Scale 2005*. Barrington, IL: Association for the Advancement of Automotive Medicine; 2005
22. Achenbach T. *Manual for the Child Behavior Checklist/4-18 and 1991 Profile*. Burlington, VT: University of Vermont College of Medicine; 1991
23. Sady MD, Vaughan CG, Gioia GA. Psychometric characteristics of the postconcussion symptom inventory in children and adolescents. *Arch Clin Neuropsychol*. 2014;29(4):348–363
24. Miller IW, Bishop DS, Epstein NB, Keitner GI. The McMaster Family Assessment Device: reliability and validity. *J Marital Fam Ther*. 1985;11(4):345–356
25. Runyan DK, Hunter WM, Socolar RR, et al. Children who prosper in unfavorable environments: the relationship to social capital. *Pediatrics*. 1998;101(1, pt 1):12–18
26. Zonfrillo MR, Durbin DR, Koepsell TD, et al. Prevalence of and risk factors for poor functioning after isolated mild traumatic brain injury in children. *J Neurotrauma*. 2014;31(8):722–727
27. Eisenberg MA, Meehan WP III, Mannix R. Duration and course of post-concussive symptoms. *Pediatrics*. 2014;133(6):999–1006
28. Max JE, Friedman K, Wilde EA, et al. Psychiatric disorders in children and adolescents 24 months after mild traumatic brain injury. *J Neuropsychiatry Clin Neurosci*. 2015;27(2):112–120
29. Yuan W, Wade SL, Quatman-Yates C, Hugentobler JA, Gubanich PJ, Kurowski BG. Structural connectivity related to persistent symptoms after mild TBI in adolescents and response to aerobic training: preliminary investigation. *J Head Trauma Rehabil*. 2017;32(6):378–384
30. McCarthy MT, Kosofsky BE. Clinical features and biomarkers of concussion and mild traumatic brain injury in pediatric patients. *Ann N Y Acad Sci*. 2015;1345:89–98
31. Kamins J, Bigler E, Covassin T, et al. What is the physiological time to recovery after concussion? A systematic review. *Br J Sports Med*. 2017;51(12):935–940
32. Ewing-Cobbs L, Prasad MR, Cox CS Jr, Granger DA, Duque G, Swank PR. Altered stress system reactivity after pediatric injury: relation with post-traumatic stress symptoms. *Psychoneuroendocrinology*. 2017;84:66–75
33. Babikian T, Satz P, Zaucha K, Light R, Lewis RS, Asarnow RF. The UCLA longitudinal study of neurocognitive outcomes following mild pediatric traumatic brain injury. *J Int Neuropsychol Soc*. 2011;17(5):886–895
34. Laborey M, Masson F, Ribéreau-Gayon R, Zongo D, Salmi LR, Lagarde E. Specificity of postconcussion symptoms at 3 months after mild traumatic brain injury: results from a comparative cohort study. *J Head Trauma Rehabil*. 2014;29(1):E28–E36
35. Covassin T, Savage JL, Bretzin AC, Fox ME. Sex differences in sport-related concussion long-term

outcomes [published online ahead of print September 18, 2017]. *Int J Psychophysiol*.10.1016/j.ijpsycho.2017.09.010

36. Rabinowitz AR, Li X, McCauley SR, et al. Prevalence and predictors of poor recovery from mild traumatic brain injury. *J Neurotrauma*. 2015;32(19):1488–1496
37. Snook ML, Henry LC, Sanfilippo JS, Zeleznik AJ, Kontos AP. Association of concussion with abnormal menstrual patterns in adolescent and young women. *JAMA Pediatr*. 2017;171(9):879–886
38. Albanese BJ, Boffa JW, Macatee RJ, Schmidt NB. Anxiety sensitivity mediates gender differences in post-concussive symptoms in a clinical sample. *Psychiatry Res*. 2017;252:242–246
39. Truss K, Godfrey C, Takagi M, et al. Trajectories and risk factors for post-traumatic stress symptoms following pediatric concussion. *J Neurotrauma*. 2017;34(14):2272–2279
40. Barlow KM, Crawford S, Stevenson A, Sandhu SS, Belanger F, Dewey D. Epidemiology of postconcussion syndrome in pediatric mild traumatic brain injury. *Pediatrics*. 2010;126(2). Available at: www.pediatrics.org/cgi/content/full/126/2/e374
41. Zemek R, Barrowman N, Freedman SB; Pediatric Emergency Research Canada (PERC) Concussion Team, et al. Clinical risk score for persistent postconcussion symptoms among children with acute concussion in the ED. *JAMA*. 2016;315(10):1014–1025
42. Jimenez N, Ebel BE, Wang J, et al. Disparities in disability after traumatic brain injury among Hispanic children and adolescents. *Pediatrics*. 2013;131(6). Available at: www.pediatrics.org/cgi/content/full/131/6/e1850
43. Moore M, Jimenez N, Graves JM, et al. Racial disparities in outpatient mental health service use among children hospitalized for traumatic brain injury. *J Head Trauma Rehabil*. 2018;33(3):177–184
44. Jimenez N, Quistberg A, Vavilala MS, Jaffe KM, Rivara FP. Utilization of mental health services after mild pediatric traumatic brain injury. *Pediatrics*. 2017;139(3):e20162462
45. Gallo LC, Penedo FJ, Espinosa de los Monteros K, Arguelles W. Resiliency in the face of disadvantage: do Hispanic cultural characteristics protect health outcomes? *J Pers*. 2009;77(6):1707–1746
46. Halstead ME, McAvoy K, Devore CD, Carl R, Lee M, Logan K; Council on Sports Medicine and Fitness; Council on School Health. Returning to learning following a concussion. *Pediatrics*. 2013;132(5):948–957
47. Lumba-Brown A, Yeates KO, Sarmiento K, et al. Centers for Disease Control and Prevention guideline on the diagnosis and management of mild traumatic brain injury in children [published online ahead of print September 4, 2018]. *JAMA Pediatr*.10.1001/jamapediatrics.2018.2853
48. Graham R, Rivara FP, Ford MA, Spicer CM, eds. *Sports-Related Concussions in Youth: Improving the Science, Changing the Culture*. Washington, DC: Institute of Medicine and National Research Council of the National Academies; 2013
49. McCarty CA, Zatzick D, Stein E, Wang J, Hilt R, Rivara FP; Seattle Sports Concussion Research Collaborative. Collaborative care for adolescents with persistent postconcussive symptoms: a randomized trial. *Pediatrics*. 2016;138(4):e20160459
50. Gunstad J, Suhr JA. "Expectation as etiology" versus "the good old days": postconcussion syndrome symptom reporting in athletes, headache sufferers, and depressed individuals. *J Int Neuropsychol Soc*. 2001;7(3):323–333
51. Brooks BL, Kadoura B, Turley B, Crawford S, Mikrogianakis A, Barlow KM. Perception of recovery after pediatric mild traumatic brain injury is influenced by the "good old days" bias: tangible implications for clinical practice and outcomes research. *Arch Clin Neuropsychol*. 2014;29(2):186–193
52. Zogg CK, Haring RS, Xu L, et al. The epidemiology of pediatric head injury treated outside of hospital emergency departments. *Epidemiology*. 2018;29(2):269–279



Routledge
Taylor & Francis Group

Journal of Sex & Marital Therapy

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/usmt20

A Typology of Gender Detransition and Its Implications for Healthcare Providers

Pablo Expósito-Campos

To cite this article: Pablo Expósito-Campos (2021): A Typology of Gender Detransition and Its Implications for Healthcare Providers, Journal of Sex & Marital Therapy, DOI: 10.1080/0092623X.2020.1869126

To link to this article: https://doi.org/10.1080/0092623X.2020.1869126

Published online: 10 Jan 2021.

JOURNAL OF SEX & MARITAL THERAPY
https://doi.org/10.1080/0092623X.2020.1869126

A Typology of Gender Detransition and Its Implications for Healthcare Providers

Pablo Expósito-Campos

Clinical and Health Psychology and Research Methods, University of the Basque Country, Donostia, Spain

ABSTRACT

Gender detransition is an emerging yet poorly understood phenomenon in our society. In the absence of research, clinicians and researchers have applied the concept of detransition differently, leading to inconsistencies in its use. The article suggests a typology of gender detransition based on the cessation or the continuation of a transgender identity to address this issue. Implications of this typology for healthcare providers are discussed, emphasizing the increasing necessity of developing clinical guidelines for detransitioners. Finally, the article reflects on the possibilities of preventing detransition, which underlines the challenges that clinicians face when treating individuals with gender dysphoria.

Introduction

Gender detransition, i.e., the process of reidentifying with one's birth sex after having undergone a gender transition, has captured the attention of the scientific community, the media, and the public in the last few years. Despite not being a genuinely novel phenomenon from a historical perspective—psychiatrist Harry Benjamin described one such case in his 1966 book *The Transsexual Phenomenon*—, research on detransition has been absent from the academic literature until recently. As a consequence, our understanding of this issue is still limited and primarily based on anecdotal evidence, which comes from a variety of sources such as personal testimonies shared on the internet (e.g., González, 2019; Palmer, 2020), parent reports (e.g., Barnes, 2020), informal surveys carried out by detransitioners (Stella, 2016), media outlets (e.g., Dodsworth, 2020; Herzog, 2017), support groups (e.g., Post-Trans, n.d.; The Detransition Advocacy Network, n.d.), documentaries (e.g., BBC Newsnight, 2019), case studies (e.g., Cain & Velasco, 2020; Expósito-Campos, 2020; Levine, 2018b; Pazos-Guerra et al., 2020; Turban & Keuroghlian, 2018), and the experiences of clinicians who work with this cohort (e.g., Graham, 2017; Marchiano, 2020).

Gender detransition is as scientifically fascinating as socially controversial, for it poses significant professional and bioethical challenges for those clinicians working in the field of gender dysphoria (henceforth "GD"). However, the scarcity of information, along with the lack of formal recognition of detransitioners and their experiences—although this trend seems to be changing (e.g., Butler & Hutchinson, 2020; Entwistle, 2020)—, has contributed to a state of things in which we fall short of a shared and scientifically consolidated language to approach detransition. This gap has favored the proliferation of inconsistent usages of the concept, thus adding to the confusion and unclarity.

CONTACT Pablo Expósito-Campos pablo.exposito.campos@gmail.com Clinical and Health Psychology and Research Methods, University of the Basque Country, Tolosa Hiribidea 70, Donostia 20018, Spain.

This article has been republished with minor changes. These changes do not impact the academic content of the article.

The term "detransition" has been used to describe two types of situations. In the first, a person stops identifying as transgender[1] after having socially, legally, or medically transitioned. This decision usually involves halting and reversing the transition process, for instance, by stopping taking hormones and going back to the pre-transition name and pronouns. In the second, a person stops transitioning due to health concerns, lack of societal/familial support, or dissatisfaction with the results—among many other reasons—but does not cease to identify as transgender. That person would not have decided to stop transitioning had the circumstances been different.

There is a fundamental distinction to make between the two scenarios just delineated. In the first scenario, the person stops transitioning because he/she no longer identifies as transgender. He/she may still experience some symptoms of GD (Lev, 2019), but concludes that being transgender is not the reason underlying his/her distress and body discomfort. In this case, detransition is fundamentally driven by the cessation of a transgender identity, which renders the process of transitioning not desirable or necessary anymore. In the second scenario, the person stops transitioning for reasons beyond their control, but not because they do not identify as transgender. In this case, detransition is motivated by external forces that make transitioning difficult, risky, or a too-heavy load to bear. The critical factor that differentiates these two situations is the *cessation of a transgender identity through the reidentification with one's birth sex*.

If we abided by the definition of detransition given initially—i.e., reidentifying with one's birth sex after having transitioned—, the second scenario described above would not constitute a "genuine" instance of detransition. Nevertheless, many clinicians have used—and use—the concept as including examples of that sort. For instance, Turban and Keuroghlian (2018) describe the case of "Lupita," a transgender woman who stops and reverses her social and medical transition due to continuous harassment and institutional disregard but resumes the process after finding herself in a more favorable and accepting environment. Pazos-Guerra et al. (2020), on their part, report the case of a 16-year-old transgender man who stops his hormonal treatment with testosterone after considering that it brings no more benefit to his identity. He expresses feeling less gender dysphoric and comfortable with the experienced physical changes. At present, he keeps identifying, living, and presenting to others as a man. In these cases, the decision to stop medically transitioning is not driven by the cessation of a transgender identity, but by social discrimination and satisfaction with the already achieved physical changes, respectively.

If not by the term "detransition," how do we refer to these particular situations in which the reason to stop transitioning does not relate to a reidentification with one's birth sex? More importantly, is it feasible to maintain such a concrete definition when the term is already being used—and will most probably keep being used—with a different connotation? Instead of adopting a prescriptive stance, these difficulties could be resolved by (1) Accepting a rather general definition of detransition, e.g., as the *interruption or reversal of a gender transition process*; and (2) Delineating a typology of gender detransition based on the cessation or the continuation of a transgender identity, which is the *core or primary* reason behind people's desire to transition in the first place.

A typology of gender detransition

Typologies are useful because they allow clinicians to discriminate between situations that may appear to be similar but, in reality, have entirely distinct causes, trajectories and, more importantly, demand different therapeutic approaches. For example, in the past, researchers have widely used typologies of GD based on the age of onset or the individual's sexual orientation to describe, classify, and understand people who receive the diagnosis (see Lawrence, 2010). By having a typology of detransition, it would be possible to maintain a unified definition of detransition—thus overcoming the disparities in its use—while also signifying the primary rationale underlying each particular case. The typology proposed in this article distinguishes between two main types of detransitions: *core*—or primary—and *non-core*—or secondary—detransitions.

Core gender detransitions

In core or primary detransitions, the decision to detransition is primarily motivated by the cessation of a transgender identity. This category potentially includes anyone who identified as transgender, socially or medically transitioned, and later returned to identifying with his/her birth sex. The reasons behind core or primary detransitions are multifarious, and may comprise: realizing that transitioning does not alleviate GD (Dodsworth, 2020; Herzog, 2017; Lev, 2019; Marchiano, 2020), finding alternative ways to cope with GD (Herzog, 2017; Stella, 2016), mental health concerns (Post-Trans, n.d.; Stella, 2016), solving previous psychological/emotional problems that contributed to GD (Butler & Hutchinson, 2020; Stella, 2016), the remission of GD itself over time (Stella, 2016), understanding how past trauma, internalized sexism, and other psychological difficulties influenced the experience of GD (Dodsworth, 2020; González, 2019; Herzog, 2017; McFadden, 2017; Post-Trans, n.d.; Stella, 2016; Yoo, 2018); the reconciliation with one's sexuality (Marchiano, 2020; GNC Centric, 2019; Pazos-Guerra et al., 2020; Post-Trans, n.d.); and a change in individual, political, social, or religious views that leads the person to question his/her transgender status (Dodsworth, 2020; Expósito-Campos, 2020; Herzog, 2017; Kermode, 2019; Stella, 2016; Turban & Keuroghlian, 2018).

One particular subcase within core detransitions concerns people with autism spectrum disorders (ASD). Anecdotal reports (e.g., Barnes & Cohen, 2019; Post-Trans, n.d.; Prestidge, n.d.) indicate that the rate of detransitioned individuals who fall within the autistic spectrum is higher than one would expect in the general population. In this regard, emerging evidence suggests a co-occurrence of GD and ASD (de Vries, Noens, Cohen-Kettenis, van Berckelaer-Onnes, & Doreleijers, 2010; Glidden, Bouman, Jones, & Arcelus, 2016; van der Miesen, Hurley, & de Vries, 2016), which may be related to an elevation in intense/obsessional interests around gender-related themes (VanderLaan et al., 2014; Zucker et al., 2017). The high number of individuals with GD who appear to fall in the autistic spectrum may explain why a significant number of core detransitioners also present autistic traits.

One would expect the likelihood of future retransitioning—i.e., resuming or reinitiating one's gender transition—after a core detransition to be low. Nevertheless, given how fluid and changeable some individuals' identities are, this possibility should not be completely ruled out. In either case, any hypothesis relative to the developmental trajectories of core detransitioners would need to be verified by in-depth, preferably longitudinal research into their life experiences.

Non-core gender detransitions

In non-core or secondary detransitions, the decision to detransition is influenced by reasons other than the cessation of a transgender identity. This category potentially includes anyone who stops or reverses their gender transition but continues to identify as transgender. The reasons behind non-core or secondary detransitions are also diverse and extend to: health concerns, including medical complications and the appearance of undesired side-effects (Danker, Narayan, Bluebond-Langner, Schechter, & Berli, 2018); disappointment or dissatisfaction with the results of medical—hormonal or surgical—treatments (Cain & Velasco, 2020; Graham, 2017; Pinkston, 2017); lack of societal support and lack of financial resources (e.g., Rei, 2018); pressure from family members or spiritual counselors (James et al., 2016); social discrimination/harassment (e.g., Kanner, 2018; Rose, 2018; see James et al., 2016); having trouble getting a job (James et al., 2016); feeling already comfortable with the acquired physical changes and thus not wanting to go any further (Graham, 2017); and the desire to become a parent (e.g., Americo, 2018) or undertake fertility preservation procedures (e.g., White, 2018). Non-core or secondary detransitions also include those who stop medically transitioning due to a change in gender identity yet maintain a transgender identity. Cain and Velasco (2020), for instance, report the case of "Gray," a natal

female with ASD who initially identifies as a transgender man and begins hormone replacement therapy (HRT) with testosterone, but later decides to stop HRT and detransitions to a non-binary identity.

In many of the cases above, detransition has a temporary character (see, e.g., James et al., 2016), and the likelihood of future retransitioning may be higher, given that the underlying identitarian motivation to transition—be it socially or medically—remains.

Further clarifications

Some might have noticed that this typology does not include *desisters*, i.e., those who desist in their gender dysphoric feelings—and, in some cases, also from a transgender identity—*before undergoing any kind of gender transition*, be it social or medical. The reason behind this exclusion is purely conceptual: it is essential to separate between desistance and detransition, which are two closely related but qualitatively distinct phenomena. The difference between both concepts is twofold. First, desistance, as it has been described in the literature, involves the remission of GD (e.g., Steensma, Biemond, de Boer, & Cohen-Kettenis, 2011), while detransition does not. Many detransitioners experience symptoms of GD long after having detransitioned (Lev, 2019). Second, desistance occurs without there being a gender transition process, while detransition occurs after having socially, legally, or medically transitioned.

It is also important to note that this typology does not suggest two clear-cut categories, for a secondary detransition can lead to a primary detransition—but not vice versa. In *r/detrans* (https://www.reddit.com/r/detrans/), a subreddit for detransitioners to share their experiences with more than 16.000 members, one can find several stories of people who call their transgender status into question after stopping transitioning due to medical complications or feeling dissatisfied with their treatment results. Furthermore, some individuals initially detransition to a non-binary identity to later end up reidentifying with their birth sex (e.g., tuffsofty, 2020). In some of these cases, a non-binary identity may be functioning as a "transitional period" before taking the definitive decision to stop identifying as transgender.

Admittedly, the relevance of the typology will depend on how well it encapsulates the experiences of all those who detransition, something that demands much more investigation—both quantitative and qualitative—into the phenomenon. However, based on our present knowledge, understanding the differences between core and non-core detransitions, as well as the reasons that might lead an individual to make such a decision, has particularly crucial implications for clinicians working in the field of GD.

Implications for healthcare providers

Core or primary detransitions underline Zucker's (2018) important—and often missed—remark that "a transgender identity is not isomorphic with a diagnosis of gender dysphoria" (p. 232). The former is the result of a subjective process of self-labeling and self-determination—that may be shared by others, including friends, family members, and clinicians—; the latter must be based on rigorous and comprehensive psychological assessments, which include attempts at differential diagnosis (Byne et al., 2018) and screening for any other potentially associated psychological issue (see Coleman et al., 2012, pp. 180–181).

On the one hand, this highlights the importance of not drawing *solely* on people's identities as the basis for decision making in clinical settings. Identities can be fluid and variable over time and thus do not constitute a reliable ground for clinicians to judge the best therapeutic approach for each patient. In this regard, some of the detransitioners interviewed by Yoo (2018) regret not having received a sufficient exploration of their previous psychological and emotional problems before transitioning, which may have played a significant role in their experience of GD. Others

express having been too enthusiastically "affirmed" in their identities by their clinicians, which led to a poor understanding of the medical procedures and the consequences of those changes. Clinicians have the "epistemological responsibility" (Van Baalen & Boon, 2015) of constructing a comprehensive and coherent picture of their patients by gathering all sorts of information to ensure that their treatment decision is indeed the best possible.

On the other hand, Zucker's (2018) observation evinces the importance of offering individuals various alternatives to address their concerns instead of promoting a one-and-only therapeutic approach. Some patients might prefer to deal with their GD in a non-affirmative manner but might be unaware of how or where to get that kind of help. They deserve to be supported in that decision and have their needs served to the best of the clinicians' abilities. For instance, many detransitioners in Yoo's (2018) study "wished their providers […] had initiated a discussion about other ways to address, treat, or live with gender dysphoria" (p. 184). Thus, developmentally informed, ethical, exploratory psychotherapy should be equally available for individuals—along with biomedical interventions—as a first-line treatment to ameliorate their feelings of GD (D'Angelo et al., 2020). This is especially significant insofar as considerable gaps in knowledge still exist regarding the impact and safety of gender-affirmative medical interventions for youth with GD (Olson-Kennedy et al., 2016), a circumstance that requires clinicians to be open about different therapeutic approaches instead of fostering a single view, for doing so "is worse than admitting uncertainty" (Lenzer, Hoffman, Furberg, & Ioannidis, 2013, p. 2).

Non-core or secondary detransitions emphasize the importance of communication between healthcare providers and their gender dysphoric patients. Clinicians should avoid creating unrealistic or unattainable expectations around the impacts and benefits of gender-affirmative treatments (GAT). Several studies have found GAT to improve gender dysphoric individuals' mental health and psychosocial functioning (e.g., Cohen-Kettenis, Schagen, Steensma, de Vries, & Delemarre-van de Waal, 2011; Costa & Colizzi, 2016; de Vries et al., 2014), but others have reported no significant differences (see Costa & Colizzi, 2016). In this regard, it is crucial to bear in mind that one size does not fit all (D'Angelo et al., 2020), meaning that GAT may not be a panacea for every individual with GD. Moreover, GAT do not necessarily bring benefits to other domains of the individual's life, so clinicians must address any potential issue in these areas in addition to GD. Finally, the physical results of biomedical GAT are not equal for everyone and, in some cases, they may involve side effects and medical complications (Chew, Anderson, Williams, May, & Pang, 2018; Scahrdein, Zhao, & Nikolavsky, 2019).

However, the insistence and pressure to initiate GAT that some adolescents with GD put upon clinicians can hinder the conducting of adequate psychological assessments (Becerra Fernández, 2020) and the fluid communication during the whole process. Therefore, it is more important than ever that clinicians be honest and transparent with their patients about the *known* benefits, risks—biological, social, and psychological—, and long-term consequences associated with each treatment option, which is the only way to ensure the obtention of meaningful informed consent (Levine, 2018a).

Additionally, non-core detransitions highlight the crucial role of clinicians in providing comprehensive psychosocial support through the process of gender transition, which may function as a protective factor for those individuals facing societal, institutional, or workplace discrimination, as well as pressure to detransition from their families or spiritual advisors. Regular and frequent follow-ups ensure transgender individuals' adjustment and well-being and, as such, they should always be an essential feature of a high-quality service for patients with GD.

The increasing necessity of developing clinical guidelines for detransitioners

The rising numbers of detransitioners (Lane, 2019; Marchiano, 2020) who are publicly sharing their experiences speaks to the necessity of developing and implementing new clinical guidelines for clinicians working in the field of GD (Butler & Hutchinson, 2020). These would preferably

need to address the differential—though sometimes overlapping—necessities of both core and non-core detransitioners.

For core detransitioners, these may include (1) Obtaining information on how to safely stop HRT; (2) Finding alternative, non-medical ways to cope with GD; (3) Securing ongoing psychological support to deal with the possible distress, anxiety, shame or regret associated with the experience of detransition; (4) Securing ongoing psychological support to address any other existing mental health issue; (5) Understanding the origins of GD and the role that identifying as transgender and transitioning played in that person's life; (6) Receiving counseling on how to announce detransition to family and friends; (7) Obtaining information about the possibilities of reversing some the physical changes derived from HRT and/or sex reassignment surgeries (SRS); (8) Obtaining information on the possibility to change back one's legal name and sex on the civil registry; and (9) Accessing legal support in cases of possible medical malpractice—organizations such as the Gender Care Consumer Advocacy Network (GCCAN; https://www.gccan.org/) have been created with this purpose.

For non-core detransitioners, these may include (1) Obtaining information on how to safely stop—and resume, in case of retransitioning—HRT; (2) In case of side effects or medical complications, receiving counseling on how to cope with these fallouts, as well as obtaining information on the possibilities of undergoing a less invasive GAT; (3) Securing ongoing psychological support to deal with discrimination, anxiety, uncertainty, or any other negative experience associated with being transgender; (4) In case of dissatisfaction with the results of the GAT, obtaining information about the possibilities of reversal; (5) Receiving counseling on how to announce detransition—and retransition, given the case—to family and friends; (6) Exploring how detransitioning might affect their experienced gender identity; and (7) Accessing social and legal support in cases of possible medical malpractice.

When facing a person who decides to detransition, clinicians must always adopt a non-judgmental, compassionate stance. Detransitioning can be as difficult as transitioning due to societal lack of understanding, social isolation, fear, shame, trauma, and the paucity of answers and resources for those who take that path. Even when the person has only undergone a social transition, going back to living according to one's birth sex can be troublesome (Steensma et al., 2011). Many core detransitioners lose the social support they had during their transition process (Kermode, 2019; Marchiano, 2020), leading to feelings of loneliness and helplessness. For some of them, their clinicians could be one of the primary sources of support in their lives. For this reason, guidelines should stress the importance of regular and long-term follow-ups to ensure that every detransitioner gets adequate care through the process of detransitioning. This point acquires critical relevance as many anecdotal accounts online point to detransitioners not going back to their gender therapists to inform them of their decision to detransition (see, e.g., GCCAN, 2020), be it out of resentment, mistrust, or the conviction that it will not make things better.

Furthermore, clinicians should not approach detransition exclusively through the monolithic lens of regret since regret and detransition are not always synonymous. For example, some core detransitioners express that transitioning was part of their own gender exploration process and that they could not know whether it was the right decision until they did it (Graham, 2017; Kermode, 2019; Turban & Keuroghlian, 2018). Detransition processes are as multiple and diverse as transition processes, so clinicians must avoid applying a homogeneous prism of interpretation.

Is it possible to prevent detransition?

One of the most crucial questions that the study of detransition poses to healthcare providers is whether detransition can be prevented and, if so, how this could be achieved. There is no easy answer to this inquiry. However, some would argue that a focus on *preventing* detransition is laden with negative values judgments about detransition and that researchers and clinicians

should instead concentrate on *supporting* detransitioners by looking at their unique life experiences (Hildebrand-Chupp, 2020). This distinction between preventing and supporting detransition might be useful from a theoretical point of view, but it does not fit so well in a real-life clinical context, where healthcare providers have the responsibility to ensure that their patients' decisions are thoughtful, well-informed, and beneficial in the long term. Imagine that a clinician identifies other issues, concerns, or factors that could be influencing one person's GD and that may jeopardize the benefits of transitioning, ultimately leading to a detransition. In such a case, it would seem highly unethical to leave those matters unaddressed and not to be cautious before making a decision.

This is not an argument for restricting access to gender-related healthcare. Instead, it intends to highlight how important it is for healthcare providers to develop an integrated view of each patient by carrying out comprehensive exploratory assessments. There is a variety of reasons to support this point. First, because there are different pathways to GD (see Zucker, 2019), which demands from clinicians an individualized approach that allows discerning its possible causes, developmental trajectories, and potential outcomes. Second, because individuals with GD may present with a range of additional concerns relating to sexuality, gender, family, and friendships (Bewley, Clifford, McCartney, & Byng, 2019) that may play an important role in the experience of GD and during the whole gender transition process. Third, because GD may come in associated with other complex psychological issues, such as mood, anxiety, and eating disorders, ASD, substance abuse, deliberate self-harm, suicidal ideation, and suicide attempts (e.g., Bechard, VanderLaan, Wood, Wasserman, & Zucker, 2017; de Graaf et al., 2020; de Vries, Doreleijers, Steensma, & Cohen- Kettenis, 2011; Donaldson et al., 2018; Holt, Skagerberg, & Dunsford, 2016; Kaltiala-Heino, Sumia, Työläjärvi, & Lindberg, 2015; Khatchadourian, Amed, & Metzger, 2014; Olson, Schrager, Belzer, Simons, & Clark, 2015; Peterson, Matthews, Copps-Smith, & Conard, 2017; Reisner et al., 2015; Sevlever & Meyer-Bahlburg, 2019; Spack et al., 2012). It is vital to explore whether these problems precede or follow the onset of GD and, more importantly, their possible relationship with GD. Finally, because other cultural, societal, and psychological factors may be influencing young people's identities and decisions to seek gender-related healthcare (see, e.g., Pang et al., 2020). In this regard, clinicians must be aware of how their young patients navigate a world of continuous changes and challenges and how these affect and shape their gender-related experiences.

Nevertheless, even when comprehensive exploratory assessments have been carried out, some individuals might decide to detransition in the future (see, e.g., Pazos-Guerra et al., 2020). *Prevention does not equate to prescience*: it is very complicated—if not impossible—to know what will happen in each particular case. Some people may detransition after a few months on their gender transition; for others, it may take several years (e.g., Dhejne, Öberg, Arver, & Landén, 2014). All of them need to be listened to and fully supported in their processes by all means, since *prevention and support are not exclusionary terms*. The logic of prevention primarily responds to an attempt to avoid any potential harm that detransition may come with, such as the irreversibility of some physical changes—derived from HRT and SRS—, trauma, shame, or social isolation, for this is inherent to the task of ensuring individuals' well-being in the long term. However, this does not mean that detransition is a clinical "failure" or that clinicians should stop their patients from detransitioning. *Life after detransition can be livable, meaningful, and fulfilling*. The role of healthcare providers is, precisely, to work toward that end.

Conclusion

Gender detransition is an emerging yet poorly understood phenomenon in our society, which poses significant professional and bioethical challenges for clinicians working in the field of GD. The absence of systematic research around detransition has given rise to inconsistencies in its

conceptual use and application, adding to the unclarity and confusion. A typology of gender detransition based on the cessation or the continuation of a transgender identity could address these issues, while offering clinicians a framework to reflect on their therapeutic endeavor when treating patients with GD. Furthermore, recognizing the disparities between core and non-core detransitioners could also help develop clinical guidelines, thus assisting healthcare providers to accommodate their different needs and demands. The conducting of comprehensive exploratory assessments can prove to be a useful tool to ensure thoughtful decision-making and prevent any potential harm associated with the experience of detransition. In conclusion, detransitioners are an underserved population whose experiences we need to listen to and understand if we truly aim to improve healthcare for people with GD. This will require extensive research to learn more about their unique experiences, motivations, needs, and demands.

Note

1. In this article, "transgender" is used as an umbrella term to include a wide range of gender identities that depart from the sociocultural expectations associated with one's birth sex (e.g., man, woman, transgender man, transgender woman, transsexual, non-binary, genderqueer, agender, etcetera) (see, e.g., Davidson, 2007).

Declaration of interest statement

This article's open access fee was paid by the Society for Evidence-Based Gender Medicine (SEGM).

ORCID

Pablo Expósito-Campos http://orcid.org/0000-0001-5825-0211

References

Americo, L. (2018, May 13). *I'm a trans woman who detransitioned to become a mom*. Them. https://www.them.us/story/im-a-trans-woman-who-detransitioned-to-become-a-mom

Barnes, R. [RivalMaverick]. (2020, August 13). *DETRANSITION: A parents perspective* [Video]. YouTube. https://www.youtube.com/watch?v=pIRK0at261s&ab_channel=RivalMaverick

Barnes, H., & Cohen, D. (2019, November 26). *How do I go back to the Debbie I was?* BBC Newsnight. https://www.bbc.com/news/health-50548473

BBC Newsnight. (2019, November 26). *Detransitioning: Reversing a gender transition – BBC Newsnight* [Video]. YouTube. https://www.youtube.com/watch?v=fDi-jFVBLA8&ab_channel=BBCNewsnight

Becerra Fernández, A. (2020). Disforia de género/incongruencia de género: Transición y detransición, persistencia y desistencia [Gender dysphoria/gender incongruence: Transition and detransition, persistence and desistence]. *Endocrinología, Diabetes y Nutrición*, *67*(9), 559−561. doi:10.1016/j.endinu.2020.03.011

Bechard, M., VanderLaan, D. P., Wood, H., Wasserman, L., & Zucker, K. J. (2017). Psychosocial and psychological vulnerability in adolescents with gender dysphoria: A "Proof of Principle" study. *Journal of Sex & Marital Therapy*, *43*(7), 678−688. doi:10.1080/0092623X.2016.1232325

Benjamin, H. (1966). *The transsexual phenomenon*. New York, NY: Julian Press.

Bewley, S., Clifford, D., McCartney, M., & Byng, R. (2019). Gender incongruence in children, adolescents, and adults. *British Journal of General Practice*, *69*(681), 170−171. doi:10.3399/bjgp19X701909

Butler, C., & Hutchinson, A. (2020). Debate: The pressing need for research and services for gender desisters/detransitioners. *Child and Adolescent Mental Health*, *25*(1), 45−47. doi:10.1111/camh.12361

Byne, W., Karasic, D. H., Coleman, E., Eyler, A. E., Kidd, J. D., Meyer-Bahlburg, H. F. L., Pleak, R. R., & Pula, J. (2018). Gender dysphoria in adults: An overview and primer for psychiatrists. *Transgender Health*, *3*(1), 57. doi:10.1089/trgh.2017.0053

Cain, L. K., & Velasco, J. C. (2020). Stranded at the intersection of gender, sexuality, and autism: Gray's story. *Disability & Society*. Advance online publication. doi:10.1080/09687599.2020.1755233.

Chew, D., Anderson, J., Williams, K., May, T., & Pang, K. (2018). Hormonal treatment in young people with gender dysphoria: A systematic review. *Pediatrics*, *141*(1), e20173742. doi:10.1542/peds.2017-3742

Cohen-Kettenis, P. T., Schagen, S. E. E., Steensma, T. D., de Vries, A. L. C., & Delemarre-van de Waal, H. A. (2011). Puberty suppression in a gender-dysphoric adolescent: A 22-year follow-up. *Archives of Sexual Behavior*, *40*(4), 843–847. doi:10.1007/s10508-011-9758-9

Coleman, E., Bockting, W., Botzer, M., Cohen-Kettenis, P. T., DeCuypere, G., Feldman, J., … Zucker, K. J. (2012). Standards of care for the health of transsexual, transgender, and gender-nonconforming people, Version 7. *International Journal of Transgenderism*, *13*(4), 165–232. doi:10.1080/15532739.2011.700873

Costa, R., & Colizzi, M. (2016). The effect of cross-sex hormonal treatment on gender dysphoria individuals' mental health: A systematic review. *Neuropsychiatric Disease and Treatment*, *12*, 1953–1966. doi:10.2147/NDT.S95310

D'Angelo, R., Syrulnik, E., Ayad, S., Marchiano, L., Kenny, D. T., & Clarke, P. (2020). One size does not fit all: In support of psychotherapy for gender dysphoria. *Archives of Sexual Behavior*. Advance online publication. doi:10.1007/s10508-020-01844-2.

Danker, S., Narayan, S. K., Bluebond-Langner, R., Schechter, L. S., & Berli, J. U. (2018). Abstract: A survey study of surgeons' experience with regret and/or reversal of gender-confirmation surgeries [Abstract]. *Plastic and Reconstructive Surgery - Global Open*, *6*(9), 189. doi:10.1097/01.GOX.0000547077.23299.00

Davidson, M. (2007). Seeking refuge under the umbrella: Inclusion, exclusion, and organizing within the category transgender. *Sexuality Research & Social Policy*, *4*(4), 60–80. doi:10.1525/srsp.2007.4.4.60

de Graaf, N. M., Steensma, T. D., Carmichael, P., Vanderlaan, D. P., Aitken, M., Cohen-Kettenis, P. T., … Zucker, K. J. (2020). Suicidality in clinic-referred transgender adolescents. *European Child & Adolescent Psychiatry*. Advance online publication. 10.1007/s00787-020-01663-9.

de Vries, A. L. C., Doreleijers, T. A., Steensma, T. D., & Cohen- Kettenis, P. T. (2011). Psychiatric comorbidity in gender dysphoric adolescents. *Journal of Child Psychology and Psychiatry*, *52*(11), 1195–1202. doi:10.1111/j.1469-7610.2011.02426.x

de Vries, A. L. C., McGuire, J. K., Steensma, T. D., Wagenaar, E. C. F., Doreleijers, T. A. H., & Cohen-Kettenis, P. T. (2014). Young adult psychological outcome after puberty suppression and gender reassignment. *Pediatrics*, *134*(4), 696–704. doi:10.1542/peds.2013-2958

de Vries, A. L. C., Noens, I. L. J., Cohen-Kettenis, P. T., van Berckelaer-Onnes, I. A., & Doreleijers, T. A. H. (2010). Autism spectrum disorders in gender dysphoric children and adolescents. *Journal of Autism and Developmental Disorders*, *40*(8), 930–936. doi:10.1007/s10803-010-0935-9.

Dhejne, C., Öberg, K., Arver, S., & Landén, M. (2014). An analysis of all applications for sex reassignment surgery in Sweden, 1960-2010: Prevalence, incidence, and regrets. *Archives of Sexual Behavior*, *43*(8), 1535–1545. doi:10.1007/s10508-014-0300-8

Dodsworth, L. (2020, August 18). *The Detransitioners*. Medium. https://medium.com/@barereality/the-detransitioners-72a4e01a10f9

Donaldson, A. A., Hall, A., Neukirch, J., Kasper, V., Simones, S., Gagnon, S., Reich, S., & Forcier, M. (2018). Multidisciplinary care considerations for gender nonconforming adolescents with eating disorders: A case series. *International Journal of Eating Disorders*, *51*(5), 475–479. doi:10.1002/eat.22868

Entwistle, K. (2020). Debate: Reality check – Detransitioner's testimonies require us to rethink gender dysphoria. *Child and Adolescent Mental Health*. Advance online publication. 10.1111/camh.12380

Expósito-Campos, P. (2020). *Destransición de género: Una historia de vida* [Gender detransition: A life-story] [Unpublished master thesis]. University of Granada.

Gender Care Consumer Advocacy Network (GCCAN). (2020, February 10). Interview with A: On detransition, obsessive thoughts, and "really trans." https://www.gccan.org/blog/interview-with-a-detransition-obsessive-thoughts-and-the-concept-of-really-trans

Glidden, D., Bouman, W. P., Jones, B. A., & Arcelus, J. (2016). Gender dysphoria and autism spectrum disorder: A systematic review of the literature. *Sexual Medicine Reviews*, *4*(1), 3–14. doi:10.1016/j.sxmr.2015.10.003

GNC Centric. (2019, July 19). *Internalized homophobia is more powerful than you know* [Video]. YouTube. https://youtu.be/k6Xe2P9c5x0

González, S. (2019, June 24). *DETRANSITION: A short documentary by Silas González*. [Video]. YouTube. https://youtu.be/j7rtj6xtThU

Graham, J. (2017, October 13 – 15). *Detransition, retransition: What providers need to know* [Conference session]. Advancing Excellence for Transgender Health, Boston, MA.

Herzog, K. (2017, June 28). *The detransitioners: They were transgender, until they weren't*. The Stranger. https://www.thestranger.com/features/2017/06/28/25252342/the-detransitioners-they-were-transgender-until-they-werent

Hildebrand-Chupp, R. (2020). More than "canaries in the gender coal mine": A transfeminist approach to research on detransition. *The Sociological Review*, *68*(4), 800–816. doi:10.1177/0038026120934694

Holt, V., Skagerberg, E., & Dunsford, M. (2016). Young people with features of gender dysphoria: Demographics and associated difficulties. *Clinical Child Psychology and Psychiatry*, *21*(1), 108–118. doi:10.1177/1359104514558431

James, S. E., Herman, J. L., Rankin, S., Keisling, M., Mottet, L., & Anafi, M. (2016). *The report of the 2015 US transgender survey*. National Center for Transgender Equality.

Kaltiala-Heino, R., Sumia, M., Työläjärvi, M., & Lindberg, N. (2015). Two years of gender identity service for minors: Overrepresentation of natal girls with severe problems in adolescent development. *Child and Adolescent Psychiatry and Mental Health*, *9*(1), 1–9. doi:10.1186/s13034-015-0042-y

Kanner, R. (2018, June 22). *I detransitioned. But not because i wasn't trans*. The Atlantic. https://www.theatlantic.com/family/archive/2018/06/i-detransitioned-but-not-because-i-wasnt-trans/563396/

Kermode, J. (2019). *Supporting transgender and non-binary people with disabilities or illnesses*. London, UK: Jessica Kingsley Publishers.

Khatchadourian, K., Amed, S., & Metzger, D. L. (2014). Clinical management of youth with gender dysphoria in Vancouver. *The Journal of Pediatrics*, *164*(4), 906–911. doi:10.1016/j.jpeds.2013.10.068

Lane, B. (2019, October 14). *Regretful 'detransitioners' on rise*. The Australian. https://www.theaustralian.com.au/nation/regretful-detransitioners-on-rise/news-story/627a9cc0f42d700be7dfab435c0522a9

Lawrence, A. A. (2010). Sexual orientation versus age of onset as bases for typologies (subtypes) for gender identity disorder in adolescents and adults. *Archives of Sexual Behavior*, *39*(2), 514–545. doi:10.1007/s10508-009-9594-3

Lenzer, J., Hoffman, J. R., Furberg, C. D. & Ioannidis, J. P. A. (2013). Ensuring the integrity of clinical practice guidelines: A tool for protecting patients. *The BMJ*, *347*, f5535. 10.1136/bmj.f5535

Lev, A. I. (2019). Introduction. In A. I. Lev & A. R. Gottlieb (Eds.), *Families in transition: Parenting gender diverse children, adolescents, and young adults* (pp. 11–36). New York, NY: Harrington Park Press.

Levine, S. B. (2018a). Informed consent for transgendered patients. *Journal of Sex and Marital Therapy*, *45*(3), 218–229. doi:10.1080/0092623X.2018.1518885

Levine, S. B. (2018b). Transitioning back to maleness. *Archives of Sexual Behavior*, *47*(4), 1295–1300. doi:10.1007/s10508-017-1136-9

Marchiano, L. (2020, January 2). *The ranks of gender detransitioners are growing: We need to understand why*. Quillette. https://quillette.com/2020/01/02/the-ranks-of-gender-detransitioners-are-growing-we-need-to-understand-why/

McFadden, J. (2017, September 16). *Transition caused more problems than it solved*. The Guardian. https://www.theguardian.com/lifeandstyle/2017/sep/16/transition-caused-more-problems-than-it-solved

Olson, J., Schrager, S. M., Belzer, M., Simons, L. K., & Clark, L. F. (2015). Baseline physiologic and psychosocial characteristics of transgender youth seeking care for gender dysphoria. *Journal of Adolescent Health*, *57*(4), 374–380. doi:10.1016/j.jadohealth.2015.04.027

Olson-Kennedy, J., Cohen-Kettenis, P. T., Kreukels, B. P. C., Meyer-Bahlburg, H. F. L., Garofalo, R., Meyer, W., & Rosenthal, S. M. (2016). Research priorities for gender nonconforming/transgender youth: Gender identity development and biopsychosocial outcomes. *Current Opinion in Endocrinology, Diabetes & Obesity*, *23*(2), 172–179. doi:10.1097/MED.0000000000000236

Palmer, E. (2020, January 9). *Why I transitioned and detransitioned* [Video]. YouTube. https://youtu.be/n0pVuZ0CT7Q

Pang, K. C., de Graaf, N. M., Chew, D., Hoq, M., Keith, D. R., Carmichael, P., & Steensma, T. D. (2020). Association of media coverage of transgender and gender diverse issues with rates of referral of transgender children and adolescents to specialist gender clinics in the UK and Australia. *JAMA Network Open*, *3*(7), e2011161. doi:10.1001/jamanetworkopen.2020.11161

Pazos-Guerra, M., Gómez Balaguer, M., Gomes Porras, M., Hurtado Murillo, F., Solá Izquierdo, E., & Morillas Ariño, C. (2020). Transexualidad: Transiciones, detransiciones y arrepentimientos en España [Transsexuality: Transitions, detransitions and regrets in Spain]. *Endocrinología, Diabetes y Nutrición*, *67*(9), 562–567. doi:10.1016/j.endinu.2020.03.008

Peterson, C. M., Matthews, A., Copps-Smith, E., & Conard, L. A. (2017). Suicidality, self-harm, and body dissatisfaction in transgender adolescents and emerging adults with gender dysphoria. *Suicide and Life-Threatening Behavior*, *47*(4), 475–482. doi:10.1111/sltb.12289

Pinkston, D. (2017, December 5). *Why do people consider de-transitioning? 3 Reasons Why* [Video]. YouTube. https://www.youtube.com/watch?v=SZ0-6TC05NE&feature=youtu.be

Post-Trans. (n.d). *Detransition stories*. https://post-trans.com/Detransition-Stories-English

Prestidge, G. (n.d.). *Gender issues for Autistic girls*. Bayswater Support Group. https://www.bayswatersupport.org.uk/gender-issues-for-autistic-girls#

Rei, M. (2018, February 7). *My detransition: How it happened* [Video]. YouTube. https://youtu.be/JUoY15ceUp4

Reisner, S. L., Vetters, R., Leclerc, M., Zaslow, S., Wolfrum, S., Shumer, D., & Mimiaga, M. J. (2015). Mental health of transgender youth in care at an adolescent urban community health center: A matched retrospective cohort study. *Journal of Adolescent Health*, *56*(3), 274–279. doi:10.1016/j.jadohealth.2014.10.264

Rose, I. (2018, February 27). *Why I detransitioned (Not Clickbait)* [Video]. YouTube. https://youtu.be/lLVUdGpWQ0w

Scahrdein, J. N., Zhao, L. C., & Nikolavsky, D. (2019). Management of vaginoplasty and phalloplasty complications. *Urologic Clinics of North America*, *46*, 605–618. doi:10.1016/j.ucl.2019.07.012

Sevlever, M., & Meyer-Bahlburg, H. F. L. (2019). Late-onset transgender identity development of adolescents in psychotherapy for mood and anxiety problems: Approach to assessment and treatment. *Archives of Sexual Behavior*, *48*(7), 1993–2001. doi:10.1007/s10508-018-1362-9

Spack, N. P., Edwards-Leeper, M., Feldman, H. A., Leibowitz, S., Mandel, F., Diamond, D. A., & Vance, S. R. (2012). Children and adolescents with gender identity disorder referred to a pediatric medical center. *Pediatrics*, *129*(3), 418–425. doi:10.1542/peds.2011-0907

Steensma, T. D., Biemond, R., de Boer, F., & Cohen-Kettenis, P. T. (2011). Desisting and persisting gender dysphoria after childhood: A qualitative follow-up study. *Clinical Child Psychology and Psychiatry*, *16*(4), 499–516. doi:10.1177/1359104510378303

Stella, C. (2016, September 3). *Female detransition and reidentification: Survey results and interpretation* [Post]. Tumblr. https://guideonragingstars.tumblr.com/post/149877706175/female-detransition-and-reidentification-survey

The Detransition Advocacy Network. (n.d.). *About us*. https://www.detransadv.com/about

tuffsofty. (2020, October 10). *i'm going back on Testosterone! MTFTM Detransition* [Video]. YouTube. https://youtu.be/-9EtxIjhbR4

Turban, J. L., & Keuroghlian, A. S. (2018). Dynamic gender presentations: Understanding transition and "de-transition" among transgender youth. *Journal of the American Academy of Child & Adolescent Psychiatry*, *57*(7), 451–453. doi:10.1016/j.jaac.2018.03.016

Van Baalen, S., & Boon, M. (2015). An epistemological shift: From evidence-based medicine to epistemological responsibility. *Journal of Evaluation in Clinical Practice*, *21*(3), 433–439. doi:10.1111/jep.12282

van der Miesen, A. I., Hurley, H., & de Vries, A. L. C. (2016). Gender dysphoria and autism spectrum disorder: A narrative review. Int Rev Psychiatry, *28*(1), 70–80. doi:10.3109/09540261.2015.1111199

VanderLaan, D. P., Postema, L., Wood, H., Singh, D., Fantus, S., Hyun, J., … & Zucker, K. J. (2014). Do children with gender dysphoria have intense/obsessional interests? *The Journal of Sex Research*, *52*(2), 213–219. doi:10.1080/00224499.2013.860073

White, B. (2018, November 21). *I'm detransitioning* [Video]. YouTube. https://youtu.be/FZnAsk5vWzE

Yoo, A. (2018). Transition regret and detransition. In C. Stewart (Ed.), *Lesbian, gay, bisexual, and transgender Americans at risk: Problems and solutions. Volume 2: Adults, Generation X, and Generation Y* (pp. 181–192). Santa Barbara, CA: Praeger.

Zucker, K. J. (2018). The myth of persistence: Response to "A critical commentary on follow-up studies and 'desistance' theories about transgender and gender non-conforming children" by Temple Newhook et al. (2018). *International Journal of Transgenderism*, *19*(2), 231–245. doi:10.1080/15532739.2018.1468293

Zucker, K. J. (2019). Adolescents with gender dysphoria: Reflections on some contemporary clinical and research issues. *Archives of Sexual Behavior*, 48(7), 1983–1992. doi:10.1007/s10508-019-01518-8

Zucker, K. J., Nabbijohn, A. N., Santarossa, A., Wood, H., Bradley, S. J., Matthews, J., & VanderLaan, D. P. (2017). Intense/obsessional interests in children with gender dysphoria: A cross-validation study using the Teacher's Report Form. *Child and Adolescent Psychiatry and Mental Health*, *11*, 51. 10.1186/s13034-017-0189-9

Gershoni and Pietrokovski *BMC Biology* (2017) 15:7
DOI 10.1186/s12915-017-0352-z

BMC Biology

RESEARCH ARTICLE Open Access

The landscape of sex-differential transcriptome and its consequent selection in human adults

Moran Gershoni* and Shmuel Pietrokovski

Abstract

Background: The prevalence of several human morbid phenotypes is sometimes much higher than intuitively expected. This can directly arise from the presence of two sexes, male and female, in one species. Men and women have almost identical genomes but are distinctly dimorphic, with dissimilar disease susceptibilities. Sexually dimorphic traits mainly result from differential expression of genes present in both sexes. Such genes can be subject to different, and even opposing, selection constraints in the two sexes. This can impact human evolution by differential selection on mutations with dissimilar effects on the two sexes.

Results: We comprehensively mapped human sex-differential genetic architecture across 53 tissues. Analyzing available RNA-sequencing data from 544 adults revealed thousands of genes differentially expressed in the reproductive tracts and tissues common to both sexes. Sex-differential genes are related to various biological systems, and suggest new insights into the pathophysiology of diverse human diseases. We also identified a significant association between sex-specific gene transcription and reduced selection efficiency and accumulation of deleterious mutations, which might affect the prevalence of different traits and diseases. Interestingly, many of the sex-specific genes that also undergo reduced selection efficiency are essential for successful reproduction in men or women. This seeming paradox might partially explain the high incidence of human infertility.

Conclusions: This work provides a comprehensive overview of the sex-differential transcriptome and its importance to human evolution and human physiology in health and in disease.

Keywords: Sex-differential expression, Sex-differential selection, Sexual dimorphism

Background

Sexual reproduction is present in nearly all multicellular eukaryotes [1]. In all cases, males and females have identical genetic information across most of their genomes, but harbor many distinct sex-specific characteristics. For example, mammalian offspring depend on maternal lactation in their early life. Lactation is thus a key factor in mammalian reproduction, and its associated genetic system is expected to be under tight selection. However, genes involved in lactation are also carried by males, who do not express this trait [2]. Different selection constraints are thus expected on these genes in males and females. Such cases can lead to reduced purifying selection on genes that otherwise are expected to be highly conserved [3]. In the same manner, many genes that are associated with sexually dimorphic traits might undergo differential selection, which will likely impact reproduction, evolution, and even speciation events [4]. Human sexual dimorphism has been demonstrated for diverse traits, such as brain anatomy and development [5–7], behavior [8], mortality, longevity and morbidity [9, 10], and distribution and metabolism of fat biogenesis [11, 12]. Physical performance capabilities and pain response have also been shown to differ between men and women [13–15]. Previous work found that about 15% of the expression quantitative trait loci (eQTLs) identified in B-lymphocytes have a sex-biased impact on gene expression [16]. That work also reported an overlap of eQTLs and genome-wide association study single

* Correspondence: moran.gershoni@weizmann.ac.il
Department of Molecular Genetics, Weizmann Institute of Science, Rehovot, Israel

nucleotide polymorphisms that are associated with sex-biased diseases. Moreover, a recent work reported sex-specific genetic architecture in complex traits [17]. It is therefore not surprising that men and women differ in their predisposition to many diseases, in disease courses, and in drug response [18, 19]. Manifestations of all these differences are likely associated with the biology of sexual reproduction.

Sexual dimorphism was suggested to evolve due to differential selection on equally expressed traits that become sexually dimorphic and even sex-limited traits [20]. This can lead to the accumulation of genes with different effects on males and females. It is thus expected that the vast majority of sexually dimorphic traits are due to differential expression of genes that are present in both sexes [21]. While carried by both males and females, such genes are expected to undergo sex-biased selection. This can lead to diverse selection patterns, including sexual antagonism where alleles increasing the fitness in one sex reduce it in the other [21]. In population genetics terms, the cost of sexual dimorphisms might be reflected in the elevated frequency of an allele with deleterious effects only on one sex. Hence, a mutation causing congenital disease in only one sex can propagate to a high population frequency due to reduced selective constraints or neutrality in half of the population (i.e., in the other sex). This might contribute to sex specificity in the susceptibility to common diseases, and provide a partial explanation to the phenomenon of "missing heritability" [18]. Indeed, differential selection due to sexual dimorphism was suggested and modeled as a mechanism that contributes to the propagation of deleterious mutations in the population [22, 23]. We recently showed first evidence that this occurs in humans. We found that deleterious mutations in testis-exclusive genes tended to accumulate more than expected, likely due to reduced selective constraints in women [24]. However, a more general demonstration of the association between sex-differential gene expression and sex-differential selection is limited to model organisms [25], mainly due to poor mapping of the sex genetic architecture and the unavailability of large-scale transcriptome sequencing in humans [24, 26].

Mapping sex-differential selection and gene expression are fundamental for understanding human evolution and biology, in health and disease. Recent advances in DNA sequencing technologies with steadily dropping costs have made such aims feasible. The release of the Genotype-Tissue Expression (GTEx) project data, which currently includes 53 tissue samples from 544 donors [27, 28], has paved the way for such progress, and preliminary results for sex-differential gene expression are already available [28].

Here, by rigorous analysis of RNA-sequencing (RNA-seq) data from the GTEx project [27, 28], we have comprehensively mapped, for the first time, human adults sex-differential gene expression over 45 tissues common to both sexes. We then identified highly and moderately sex-specific genes while considering the complete panel of 53 tissues. Such genes are expected to have general sex-differential roles, thus suggesting differential selection. We thus hypothesized that deleterious mutations in these genes will propagate in the population more than expected by chance, due to the reduced impact of purifying selection [22, 24, 29]. By analyzing the signature of selection in these genes, we have found, for the first time, reduced selective constraints and differential rates of accumulation of deleterious mutations in both men and women sex-specific genes. The expression and function of these genes are associated with several tissues and biological pathways, including ones common to both sexes, suggesting a general phenomenon that directly arises from sex-differential selection. Moreover, many of these sex-differentially expressed genes were enriched in sexually dimorphic systems. Finally, some of these genes suggest new insights into the pathophysiology of several human diseases.

Results

We examined human gene expression from RNA-seq data of the GTEx project version 6 (October 2015 release), including 8555 samples comprising 53 tissues from 357 men and 187 women post-mortem donors aged 20–79 years old [30]. Gene expression data for each tissue was grouped by sex. This created 98 sets with 45 tissues common to men and women and eight tissues specific to one of the sexes.

Sex-differentially expressed genes

Sex-differential expression (SDE) was tested in each of the 45 common tissues by comparing the individual expression values of 18,670 out of 19,644 informative protein-coding genes in men versus women. To identify SDE we used the NOISeqBIO method [31, 32] to compare gene expression in the common tissues between men and women. The results were further analyzed to produce a relative SDE score for each gene in each common tissue using a metric we devised (Additional file 1: Figure S1).

On the background of similar expression in most tissues of most genes (Additional file 2: Figure S2; Additional file 3: Table S1), there are over 6500 protein-coding genes with significant SDE in at least one tissue. Most of these genes have SDE in just one tissue, but about 650 have SDE in two or more tissues, 31 have SDE in more than five tissues, and 22 have SDE in nine

or more tissues (Additional file 4: Figure S3 and Additional file 5: Table S2). As expected, Y-linked genes that are normally carried only by men show SDE in many tissues. Nevertheless, 16 out of the 244 X-linked SDE genes also have widespread SDE (across six or more tissues, Additional file 5: Table S2) in either men or women. We found that three of these X-linked genes are located at pseudo-autosomal region 1 (PAR1), which undergoes relatively frequent recombination between the X and Y chromosomes and is known to escape X-inactivation [33] (Additional file 5: Table S2; Additional file 6: Figure S4). It is noteworthy that these PAR1 genes have men-biased expression.

The most sex-differentiated tissue, with 6123 SDE protein-coding genes, is the breast mammary glands (Fig. 1; Additional file 2: Figure S2), as previously noted [28]. This suggests major differences in the physiology and sex genetic architecture of this tissue. We found 1145 genes to be SDE in non-mammary gland tissues. The most differentiated of these tissues, with over 100 SDE genes, are the skeletal muscle, two skin tissues, subcutaneous adipose, anterior cingulate cortex, and heart left ventricle (Figs. 1 and 2). Most GTEx tissues (46 out of 53) have more than seventy samples (70–361). This sample-size variation can affect the number of identified SDE genes per tissue. The Pearson correlation coefficient between the sample size and the number of identified SDE genes is 0.10 for the 45 analyzed tissues common to men and women, and 0.57 when the mammary glands tissue is excluded. This suggests that sample size contributes to the differences in the number of identified SDE genes per tissue, although several tissues noticeably deviate from this trend (e.g., the breast and whole blood tissues, Fig. 1). Besides the number of SDE genes, the tissues can also be clustered by the patterns of gene SDE scores. This analysis found the two skin, adipose-subcutaneous, and stomach tissues to deviate the most from all other tissues, and that seven of the thirteen

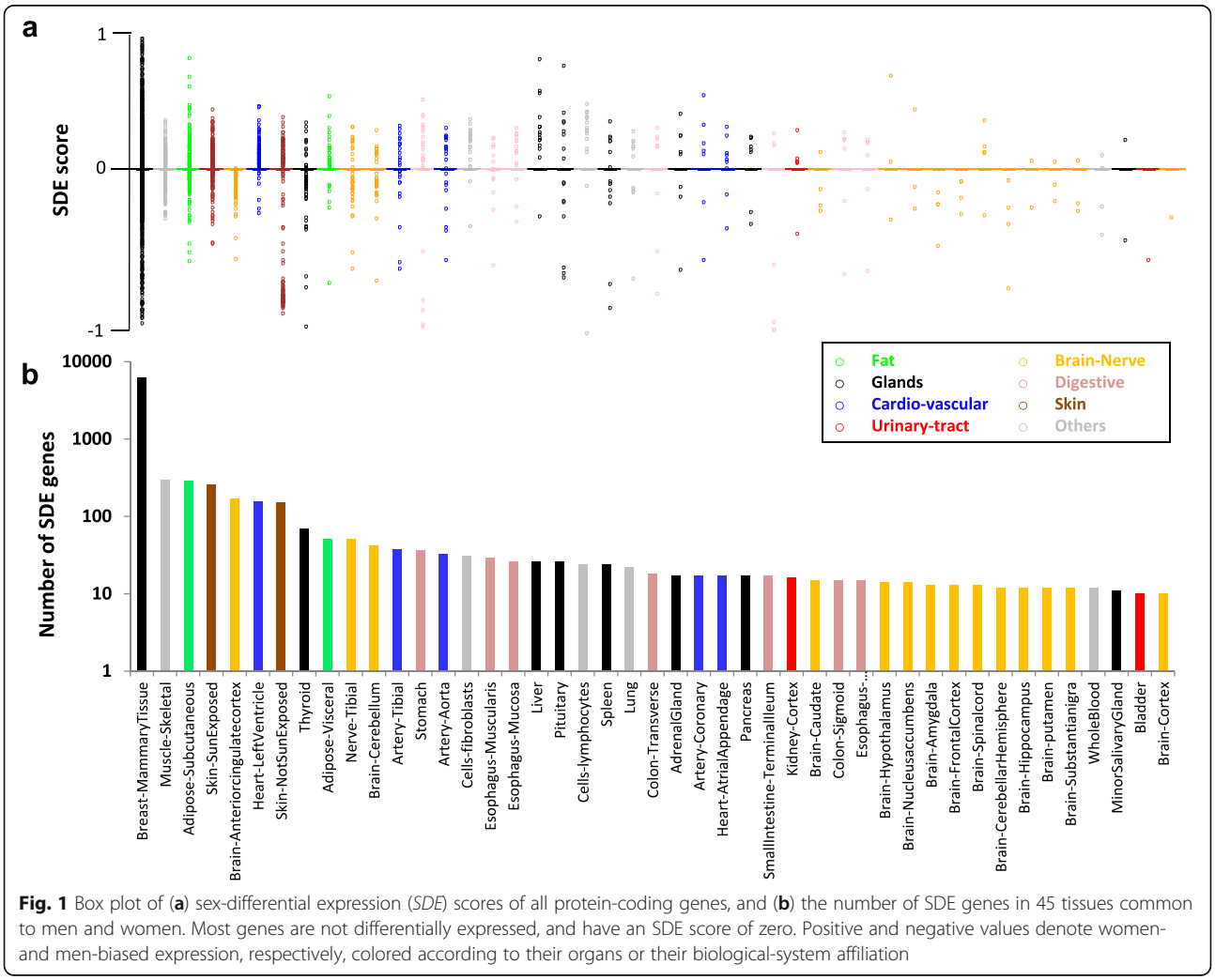


Fig. 1 Box plot of (**a**) sex-differential expression (*SDE*) scores of all protein-coding genes, and (**b**) the number of SDE genes in 45 tissues common to men and women. Most genes are not differentially expressed, and have an SDE score of zero. Positive and negative values denote women- and men-biased expression, respectively, colored according to their organs or their biological-system affiliation

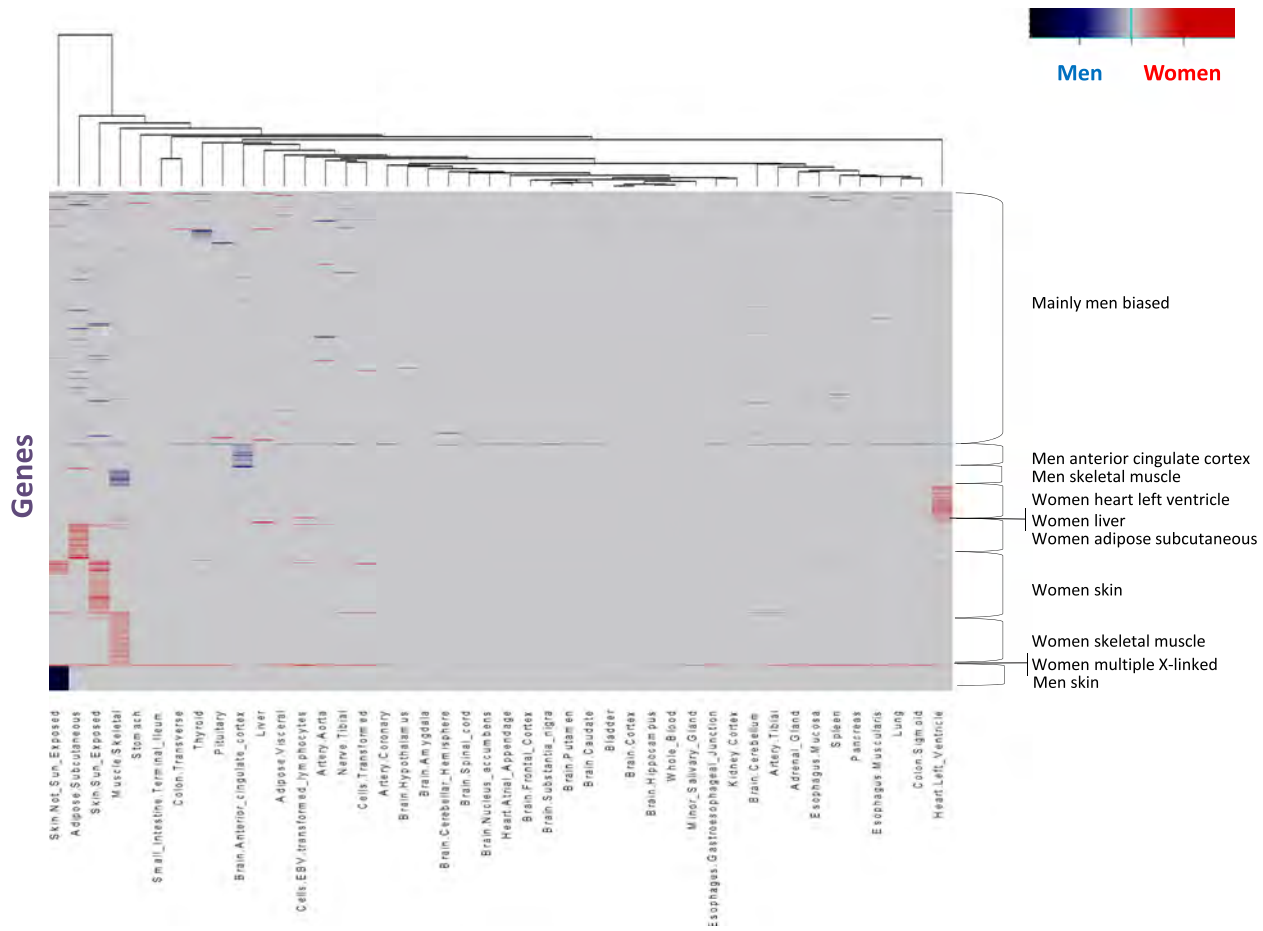


Fig. 2 Heatmap of sex-differential expression (SDE) scores of all genes with at least one SDE in non-mammary gland tissue. *Red* denotes women specificity and *blue* denotes men specificity. The genes are grouped according to principal component analysis clusters (Additional file 8: Figure S6). Tissues are grouped using hierarchical clustering

brain tissues clustered together (Fig. 2, Additional file 7: Figure S5) [34].

Clustering genes by their SDE patterns across tissues revealed 10 groups (Fig. 2, Additional file 8: Figure S6), nine of which can be described as follows:

1. Three groups of men-biased expression in the skin, skeletal muscle, or cingulate cortex tissues (e.g., *MYH1*; Fig. 3).
2. Five groups of women-biased expression in the liver, heart left ventricle, skin, skeletal muscle, or adipose subcutaneous tissues (e.g., *NPPB*; Fig. 3).
3. A group of mostly X-linked genes with SDE in various tissues, mainly with women-biased expression (e.g., *ZFX*; Fig. 3).

Other genes, such as *TSHB*, show tissue-specific expression bias (Additional file 9: Figure S7), and a few genes present an alternating pattern of expression biases, such as *MUCL1* that is overexpressed in men skin tissue and in women mammary glands (Additional file 9: Figure S7). To detect differential expression in genes with complex modes of expression we used an additional analysis approach, which is more sensitive to such cases. This analysis uncovered 241 additional genes in non-mammary gland tissues that were clearly not detected in the first approach (see "Methods" and Additional file 10: Table S3, supplementary results). For instance, we found a likely age-related gene overexpression in women brain tissue (Additional files 11 and 12: Figures S8 and S9).

Genes found to have SDE were analyzed for gene enrichment in different types of terms (e.g., diseases, Gene Ontology (GO) terms, pathways [35]). Genes with women-biased expression were associated with obesity, muscular diseases, and cardiomyopathy. In addition, overexpressed women-biased genes were enriched in glucose metabolism and adipogenesis pathways (Additional file 13: Table S4). Interestingly, 15 out of 20 genes found to be associated with cardiomyopathy also showed a women overexpression bias in heart tissue, as in the

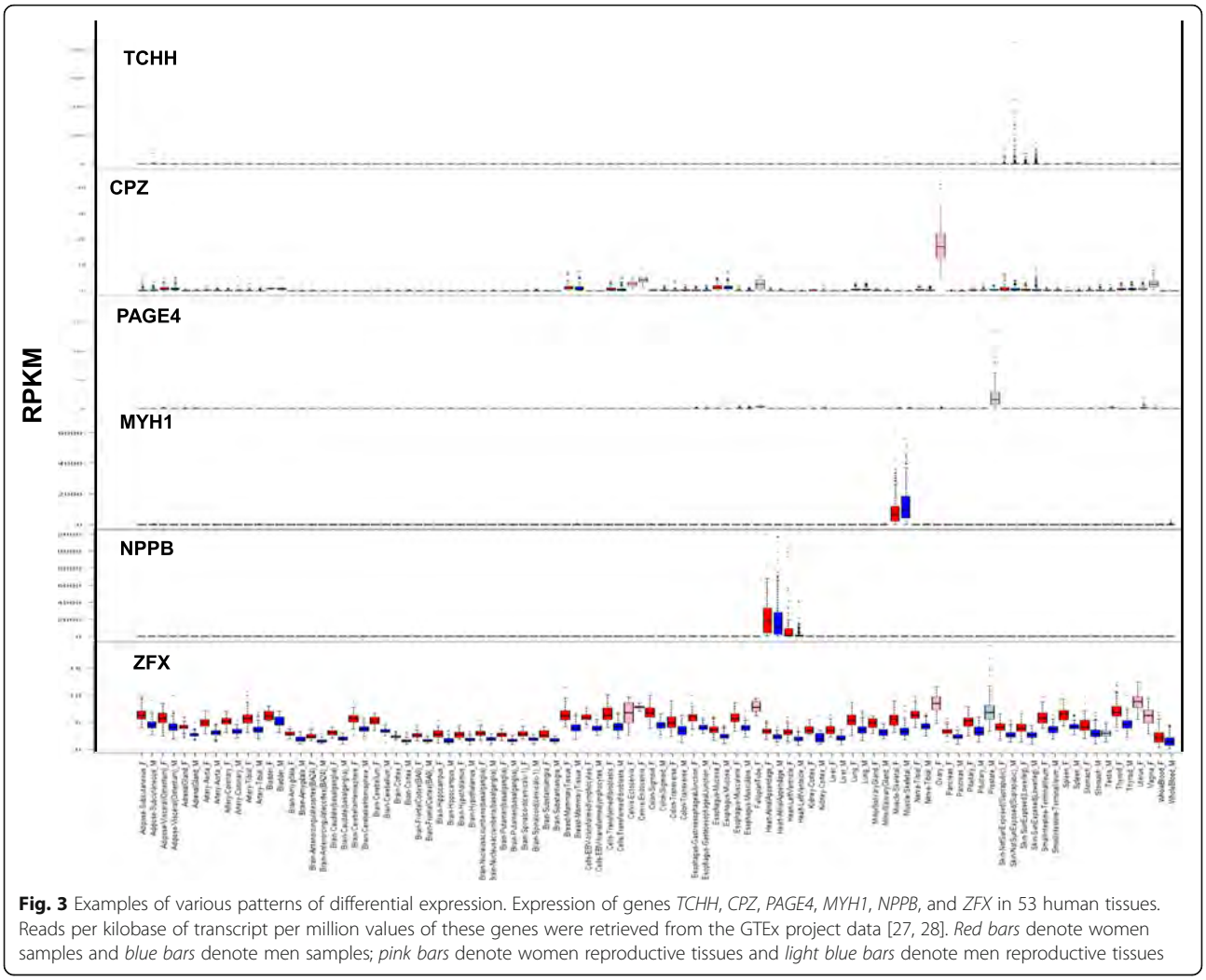


Fig. 3 Examples of various patterns of differential expression. Expression of genes *TCHH*, *CPZ*, *PAGE4*, *MYH1*, *NPPB*, and *ZFX* in 53 human tissues. Reads per kilobase of transcript per million values of these genes were retrieved from the GTEx project data [27, 28]. *Red bars* denote women samples and *blue bars* denote men samples; *pink bars* denote women reproductive tissues and *light blue bars* denote men reproductive tissues

natriuretic peptide B-secreted cardiac hormone gene *NPPB* (Fig. 3), supporting previous evidence on its involvement in sex-differential cardiovascular phenotypes [36, 37]. Genes with men-biased expression also showed enrichment in glucose metabolism pathways, but the gene sets differed, suggesting alternative pathways in glucose metabolism between men and women (Additional file 14: Table S5). A muscle-contraction pathway was also associated with genes overexpressed in men (Additional file 14: Table S5). This might be related to the physiological differences in muscle tissues and in physical features between men and women [38, 39].

Identification of sex-specific genes

Beyond genes that have SDE in one or several tissues are more extreme cases of genes with overall exclusive or high expression-specificity in one sex [40]. Such sex-specific genes are more likely to have global sex-differential functional roles, and are thus expected to present measurable sex-differential selection that can be reflected by a reduction in purifying selection [24]. A gene was considered sex specific if its maximal expression value in one sex was significantly higher from its expression values in all tissues of the other sex. In addition, genes were considered as non-SDE if their maximal expression values in men and women differed by no more than 10% (≤1.1 fold). We identified 1559 sex-specific and moderately sex-specific genes. Of these genes, 1288 (82.6%) were men-specific and overexpressed in the testis (Additional file 15: Table S6; Additional file 16: Figure S10). Aside from these 1559 genes, we found 26 women-specific and 114 moderately women-specific genes, and 82 non-testis men-specific and 49 moderately men-specific genes (Fig. 4; Additional file 17: Table S7). Over 8000 genes were identified as non-SDE (see "Methods" and Additional file 3: Table S1).

The sex-specific and moderately sex-specific genes could be grouped by their expression patterns into six major categories (Fig. 4; Additional file 16: Figure S10):

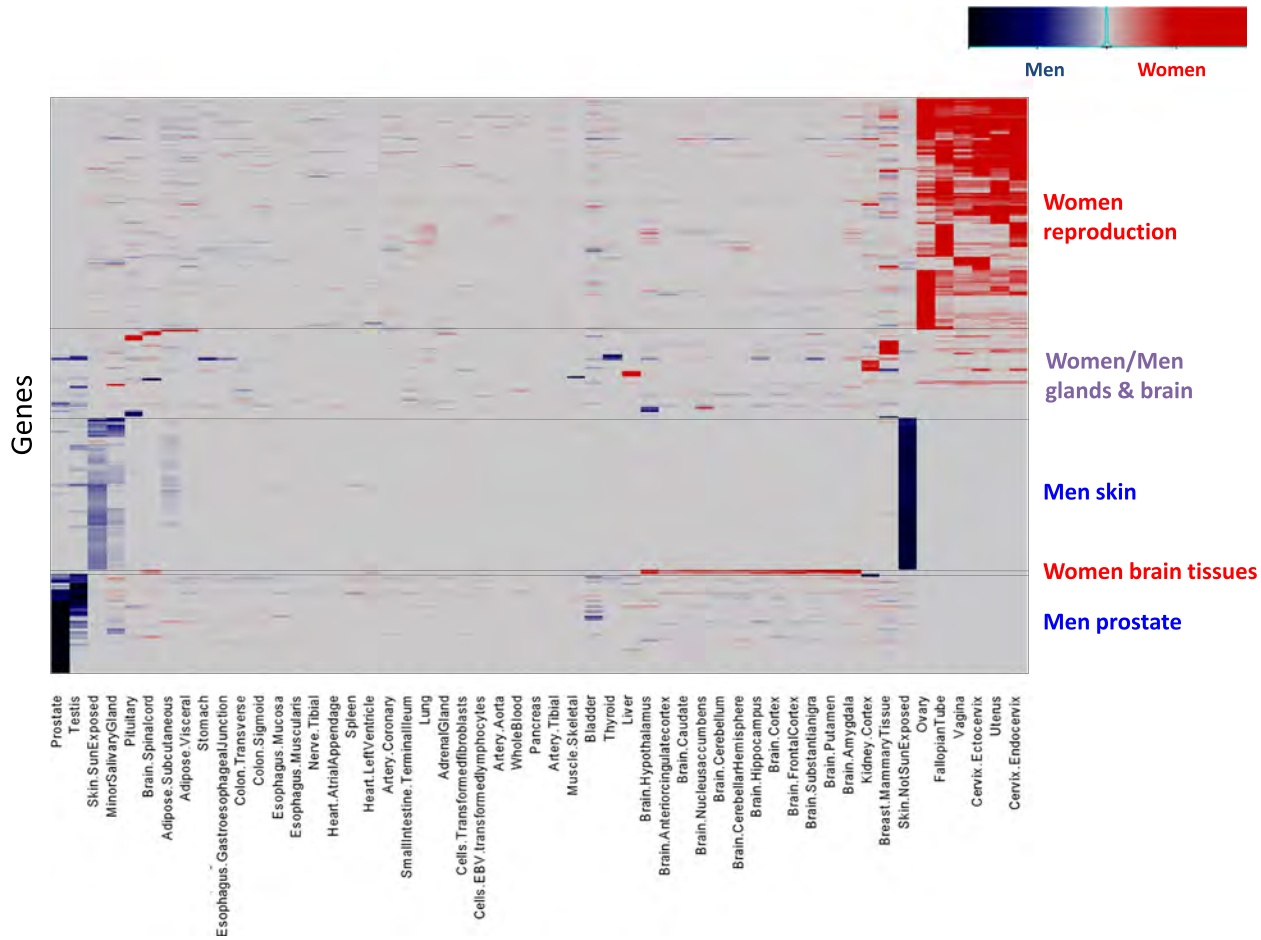


Fig. 4 Heatmap of sex-differential expression (*SDE*) scores of the sex-specific and moderately sex-specific genes, colored as in Fig. 2. *Red*, *blue*, and *purple boxes* denote major women, men, and combined gene clusters, respectively

1) Testis overexpressed genes in men (Additional file 16: Figure S10)
2) Prostate overexpressed genes in men (e.g., *PAGE4*, Fig. 3)
3) Reproductive system overexpressed genes in women (e.g., *CPZ*, Fig. 3)
4) Skin-specific overexpressed genes in men (e.g., *TCHH*, Fig. 3)
5) Brain tissue overexpressed genes in women
6) Mainly gland and brain tissue overexpressed genes, in men or women (e.g., *TSHB*, Additional file 17: Table S7).

Overall, sex-specific genes are mainly expressed in the reproductive system, emphasizing the notable physiological distinction between men and women. However, scores of genes that are not known to directly associate with reproduction were also found to have sex-specific expression (e.g., the men-specific skin genes).

Selection analysis

We calculated the numbers of observed (1000 Genomes Project [41]) and possible deleterious non-synonymous (DNS), stop-gain, and synonymous (S) single-nucleotide variants (SNVs) for each gene. This allowed us to quantify the selection pressure and its direction by dDNS/dS and dStop/dS ratios. Similar to dN/dS, these ratios are selection indicators [42, 43]. Ratios close to 1 indicate neutral selection, lower ratios indicate purifying (negative) selection, and significantly higher ratios suggest adaptive (positive) selection (see "Methods").

Natural gene variants have different frequencies, with most of the variation due to alleles with rare to low minor allele frequencies (MAFs) [24, 44]. However, selection is expected to have only a slight effect on the propagation of very rare variations because they are predominantly new while selection is mainly a long-term process [44, 45]. In addition, most phenotypes result from allele and gene interplay, and are thus highly unlikely (except in inbreeding) for rare variations, as in recessive and epistatic models of inheritance [45]. We hence studied the population genetics of the dDNS/dS and dStop/dS to find the proper MAF threshold in which the selection efficiency is maximal. Higher dDNS/

dS ratios are more abundant for SNVs with rare MAFs (<0.005, Fig. 5), indicating that negative selection predominantly affects the propagation of deleterious SNVs for MAFs >0.005, as previously shown [24, 44]. However, dStop/dS ratios are sharply decreased for very rare SNVs (i.e., MAFs <0.001, Fig. 5). We thus further analyzed the effect of selection on DNS and stop-gain using MAF thresholds of >0.005 and >0.001, respectively.

Selection analyses of sex-specific and moderately sex-specific genes

We have previously shown that human testis-exclusive genes are under reduced selection [24]. All 1100 of 1295 men testis-overexpressed genes identified here that are covered in the 1000 Genomes Project were also found to have significantly higher dDNS/dS and dStop/dS ratios (Table 1). This gene set includes 77 out of 95 of the genes we previously identified as testis exclusive [24]. The other 18 out of 95 genes that we previously found to be specifically expressed in testis tissues might not be identified here because these tissues are not present in the GTEx samples. The non-testis men-specific and moderately women-specific genes also had significantly higher dDNS/dS ratios (Table 1, Fig. 6). The significantly higher dDNS/dS ratio of these men-specific genes did not depend on the presence of the 55 keratin genes (Table 1). women-specific genes too had a significantly higher dDNS/dS ratio (Table 1, Fig. 6). Moderately women-specific genes had a higher, yet not significant, dDNS/dS ratio (Table 1). However, when comparing the moderately women-specific genes to non sex-specific genes, we found the dDNS ratios to be significantly higher for the moderately women-specific genes (1.66 fold change, Fisher's exact test p-value $<1 \times 10^{-4}$) but the dS ratios showed no significant change (1.08 fold change, Fisher's exact test p-value $= 1.5 \times 10^{-1}$). Thus, moderately women-specific genes have significantly reduced selection relative to non sex-specific genes. The same analysis for dStop/dS of men- and women-specific genes also found significantly reduced selection (Table 1).

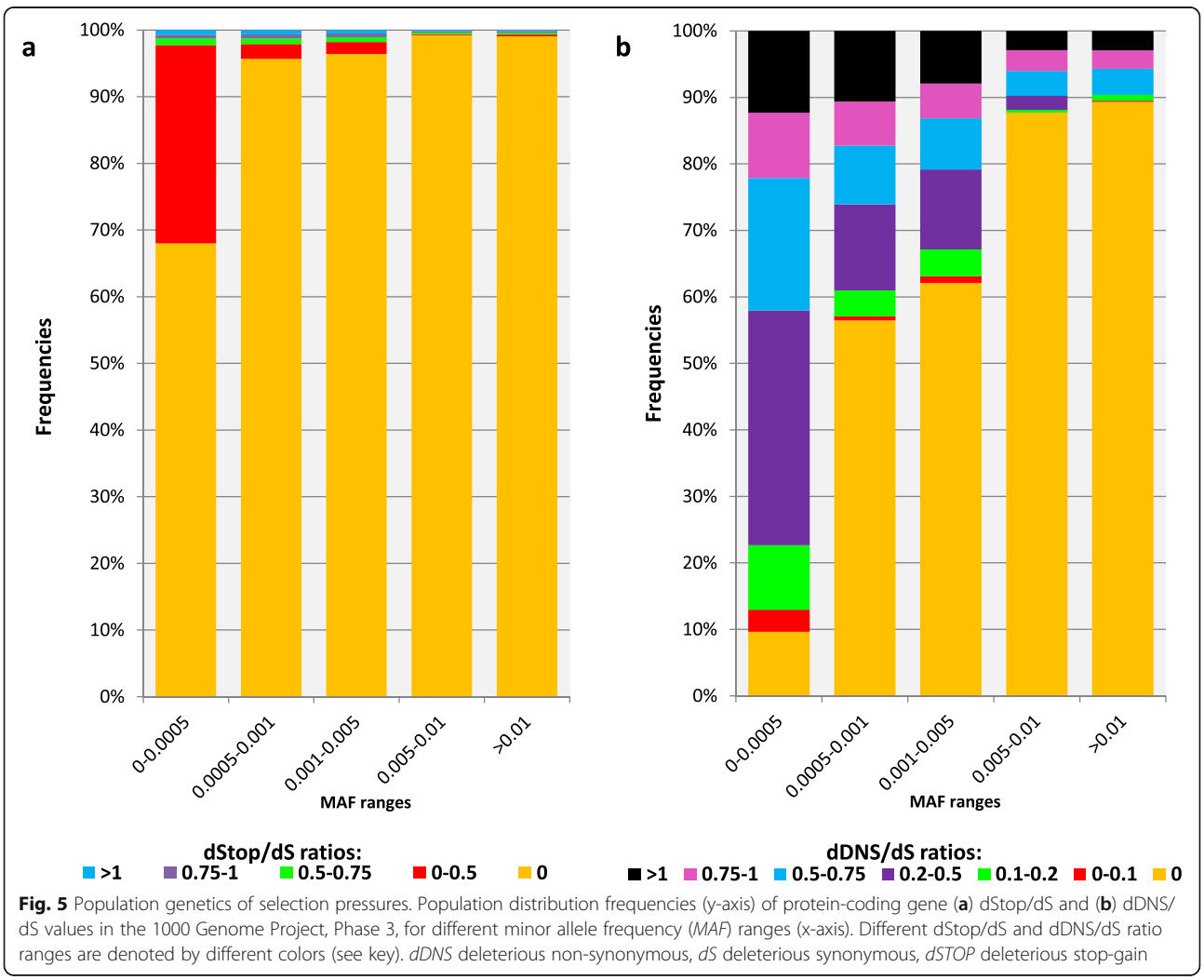


Fig. 5 Population genetics of selection pressures. Population distribution frequencies (y-axis) of protein-coding gene (**a**) dStop/dS and (**b**) dDNS/dS values in the 1000 Genome Project, Phase 3, for different minor allele frequency (*MAF*) ranges (x-axis). Different dStop/dS and dDNS/dS ratio ranges are denoted by different colors (see key). *dDNS* deleterious non-synonymous, *dS* deleterious synonymous, *dSTOP* deleterious stop-gain

Table 1 Selection analysis summary

Gene group	*n*	dDNS/dS (MAF > 0.005)	*p*-value	dStop/dS (MAF > 0.001)	*p*-value
Women-specific	26	0.23	0.02	0.27	0.0117
Men-specific	82	0.30	0.0005	0.29	0.0001
Men-specific; no keratin and keratin-associated genes	27	0.28	0.009	0.22	0.026
Moderately women-specific	114	0.16	0.09	0.13	0.0076
Moderately men-specific	49	0.25	0.005	0.07	0.27
Men testis overexpressed	1100	0.238	<0.0001	0.29	<0.0001

dDNS deleterious non-synonymous, *dS* deleterious synonymous, *dSTOP* deleterious stop-gain, *MAF* minor allele frequency

A significant reduction in purifying selection on sex-specific genes was hence found by independent analyses of selection on DNS and stop-gain mutations on diverse sets of sex-specific genes from both women and men, including sets from non-reproduction-related tissues. It is also notable that although reduced selection was observed for both men- and women-specific genes, it was higher in men-specific genes compared to women-specific genes (Fig. 6, Table 1).

Discussion

Mapping sex-differential gene expression we found more than 6500 protein-coding genes with significant SDE in one tissue or more. The most differentiated tissue was the breast mammary gland, with more than 6000 genes having significant SDE (Fig. 1). This remarkable sex-biased gene expression is likely due to the distinct physiologic properties of this tissue between men and women [2]. In evolutionary terms, differential selection between the sexes of so many genes that are likely involved in lactation, an essential reproductive trait, might inhibit optimal adaptation of this trait due to its distinct importance in men and women.

Almost all SDE genes are sex differentiated in one or just a few tissues. Thirty-one genes have SDE in six or more tissues. Besides Y-linked genes that have men-specific expression, 16 of the other genes are X-linked, with multiple-tissue SDE in either men or women. Three of these X-linked genes are located in the PAR1 region (Additional file 6: Figure S4; Additional file 5: Table S2), which includes genes that undergo recombination with the Y chromosome and also escape X-inactivation [33]. These PAR1 genes have identical sequences in their X and Y copies (Additional file 5: Table S2), but are only classified as X-linked in the GTEx data. While this should have led to similar expression in men and women (as in most autosomal genes), these genes have men-biased expression in multiple tissues. It is possible that although the copies are identical, the regulation of their expression is distinct between the X and Y-

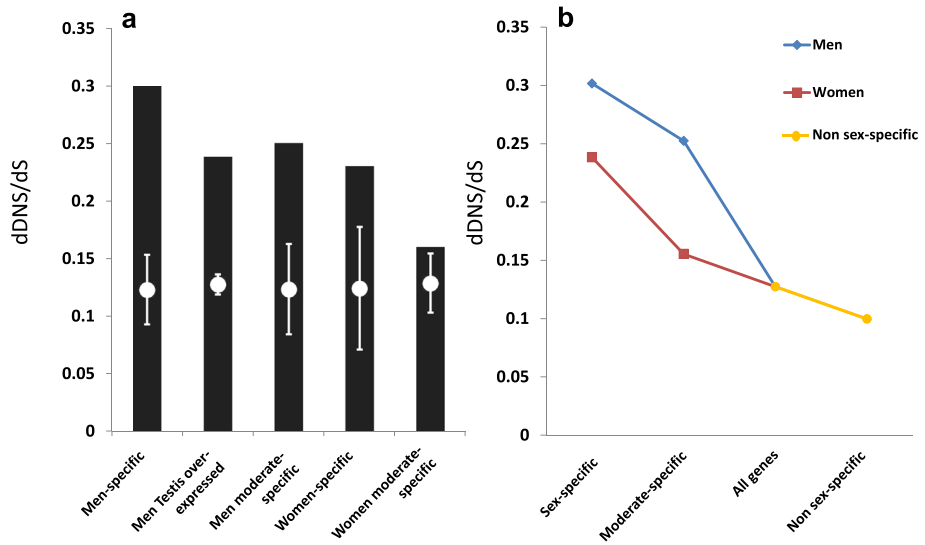


Fig. 6 Sex-specific expression and purifying selection. **a** dDNS/dS ratios of different groups of genes (Table 1, *black bars*) and the mean (*white circles*) and standard deviations (*lines*) of 10,000 random control sets with the corresponding number of genes. **b** Inverse correlation between sex specificity and selection efficiency. *dDNS* deleterious non-synonymous, *dS* deleterious synonymous

chromosomes. Besides the PAR1 genes, X-linked SDE genes in multiple tissues were found to only have women-biased expression (Additional file 6: Figure S4). In several cases we found that such genes have an active paralog on the Y chromosome and it is therefore likely that these genes escape X-inactivation and both X alleles are expressed in women, while men have only one X-linked allele.

Aside from the mammary glands, the adipose, skeletal muscle, skin, and heart tissues have over a one hundred SDE genes. This indicates substantial differences in the physiology, or alternate biological pathways, in these tissues between adult men and women. However, the differences in the number of SDE genes per tissue should be carefully assessed because the variability in tissue sample sizes could contribute to the number of SDE genes per tissue that we can identify. Functional terms analysis of SDE genes suggests sexual dimorphism in fat biogenesis, muscle contraction, and cardiomyopathy (Additional files 13 and 14: Tables S4 and S5). Tissues with few identified SDE genes might have overall similar function between men and women, yet even very few SDE genes can have extensive physiological impacts on the organism. For instance, the pituitary gland has only 26 identified SDE genes (Figs. 1 and 2), but two of them are the FSHB (women-biased) and TSHB (men-biased) gonadotropin hormones that have wide-ranging roles in human reproduction and metabolism [46, 47]. Another example is the CYP3A4 and CYP2B6 cytochrome P450 enzymes, which have women-biased expression in liver. Cytochrome P450 (P450, CYP) enzymes are associated with drug metabolism and other essential catabolic processes [48], and might be involved in sex-differential drug responses, as previously reported [49]. Other identified specific genes might shed new light on the pathophysiology of human diseases. For instance, the *NPPB* gene, which is mainly overexpressed in young women's hearts (Additional file 18: Figure S13), is related to cardiovascular homeostasis [36, 37]. Variations in this gene are associated with postmenopausal osteoporosis, a health condition mainly affecting women [50]. Thus, a sexually dimorphic effect of this gene on both phenotypes would be interesting to assess.

To evaluate the association between SDE and selection we identified sex-specific genes. Such genes are likely to possess different roles between the sexes and therefore are likely to undergo different selection pressures in each sex. The vast majority of sex-specific genes we found are overexpressed in the testis. We previously showed reduced selection and accumulation of damaging mutation in such genes. Here we confirmed our previous findings, extended them to many more testis-overexpressed genes, and to sex-specific genes of other men and women tissues. Many of the non-testis sex-specific genes are also related to the reproductive system, including genes expressed in tissues common to both sexes, such as gonadotropin hormones expressed in the pituitary (e.g., *FSHB* and *CGB7*). Dozens of genes with no direct association to reproduction were also identified as sex specific. Many of these genes are expressed in skin tissues, are linked to hairiness (Additional files 13 and 14: Tables S4 and S5), and are likely involved in hair dimorphism in women and men. Other non-reproductive genes do not seem to share common features with each other, but are each interesting on their own, for example, the moderately men-specific growth hormone GHRH and the men-specific calcitonin-related polypeptide alpha (CALCA) (Additional file 17: Table S7). The latter is involved in calcium regulation and functions as a vasodilator [51, 52]. The genes fro both seem specific to adult men, although they are related to apparently general biological processes.

Analyzing selection on highly and moderately men- and women-specific genes, we found a significant association with reduced selection efficiency, as reflected in their dDNS/dS and dStop/dS ratios (Table 1, Fig. 6). The reduced purifying selection efficiency was also correlated with the level of sex specificity. This suggests that higher sex specificity indicates greater distinction in the functional importance for each sex, and reduced selection efficiency. This in turn enables the propagation of damaging alleles through the non-expressing sex lineages. The resulting relatively high population frequencies of these alleles can enhance the prevalence of different human diseases.

Although we found reduced selection on both men- and women-specific genes, it is notable that reduced selection was more prevalent in men-specific genes (Fig. 6). This supports our previous expectations to find men-specific genes to be under less selection than women-specific genes [24]. We suggest that the basis for this could be the practically unlimited numbers of available male gametes compared to the restricted number of available women gametes, as suggested in the Bateman principle [53]. Thus, the ability of women to pass on alleles that cause men-specific lethality will less affect the number of fertile men required to sustain the population, but not vice versa.

In this work we focused on protein-coding genes, because currently there is a broad functional knowledge on these genes and extensive experience in analyzing and quantifying the selection trends these genes have undergone. However, the importance of non-coding RNA genes for the regulation and execution of sexual dimorphism was not ignored. For instance, the function of the *XIST* long non-coding RNA gene in the sex-specific X-inactivation process is well documented (Additional file 19: Figure S11) [54]. Our preliminary observations of

the RNA gene differential transcriptome support a global role of these genes in the sex genetic architecture (Additional file 20: Figure S12). Hence, this work and the data it provides might trigger further in-depth studies on the contribution of RNA genes to sexual dimorphism.

Finally, the vast majority of sex-specific genes we found are associated with the reproductive system. Damaging mutations in many reproductive genes can hence propagate to high population frequencies. We suggest that sex-specific genes are major contributors to the high incidence of infertility in men and women.

Our results are delimited by the scope of the data in the GTEx study. This study includes 53 tissues from adult humans. All tissues are composed of several cell types and a few are represented in fewer than 15 men or women donors. We believe our statistical and analysis measures excluded most false-positive results. However, the distinct age limits of the samples are acutely pertinent to sexual dimorphism and we do not know how much of our findings can be extended beyond adults. Examining comparable data from puberty and during embryonic stages of sex determination will likely augment the genes and phenomena described here.

After submitting this work for review, two studies on sexual dimorphism in human gene expression were made public. Kassam *et al.* examined the sex-specific genetic architecture of autosomal gene expression in whole blood samples from about one thousand men and one thousand women using DNA arrays [55]. No differences between men and women were found in autosomal genetic control of gene expression. We too did not identify autosomal genes with different expression between men and women in the GTEx whole blood tissue (Fig. 1; Additional file 3: Table S1). Chen *et al.* posted to bioRxiv a non-peer-reviewed preprint analyzing the GTEx data for gene expression sexual dimorphism and regulatory networks [56]. They report sexually dimorphic patterns of gene expression involving as many as 60% of autosomal genes. Similar to our findings, they reported breast, skin, adipose, heart, and skeletal muscle as the most sexually dimorphic tissues. The studies vary in their analyses procedures and emphasize different contexts of SDE. These studies are complementary works with different insights.

The mode of gene expression is very complex, depending on the gene's genomic and chromatin contexts, activity of other genes, expressing tissue, the individual's developmental stage, and external factors such as exposure to pathogens, diet, and temperature. The expression level of genes thus varies temporally (in scales of minutes to decades) and across tissues, and is a multidimensional system. This is the key challenge in evaluating differential gene expression between populations.

SDE between men and women stems from any deviations of gene activity in place (i.e., organs, tissues, and cells) and time (e.g., developmental stage, age, cell cycle point, or periodic processes). The overall distribution of gene expression values in two populations could be highly similar, and distinct in only a minor subset of samples that represents a genuine biological difference in time and/or place. For instance, a gene can have similar basal expression in men and women, but upon sex-specific induction its expression will be altered only in one sex. Thus, only a small fraction of one population in any one time might differentially express this gene. Identifying differential expression is thus a challenging problem. In addition, sex-specific expression is a particular case of SDE, in which genes present a global bias in their mode of expression in one sex compared to the other.

We applied several approaches to identify SDE and sex-specific expression. Besides analyzing differences according to the population variance (NOISeqBIO), we also used an approach that gave weight to a subset of samples that notably deviated from all other samples (using count trimmed means and NOISeq-sim). The DESeq2 method was also used to validate the results in selected datasets. In addition we used a new normalized measure for gene differential expression between pairs of sample populations. This differential expression measure takes into account the expression difference between the sexes and the maximal expression of the gene in all tissues, placing the difference in specific tissues in the context of the gene overall mode of expression. This measure is general and can be used in other population-based differential gene expression studies (Additional file 1: Figure S1). Combining these approaches increased our ability to identify differential expression from various modes of gene expression. Accumulation of many more samples from different donors and conditions will uncover the full spectra of gene modes of expression and improve the resolution of differential expression analyses.

Conclusions

This work comprehensively mapped for the first time the sex-specific genetic architecture of human adults. We identified hundreds of genes with women and men SDE, and showed the relation of these genes to several sexually dimorphic features, to human diseases, and to human evolution. Our results can facilitate the understanding of diverse biological characteristics in the context of sex. We also demonstrated the increased propagation of deleterious mutations in many men- and women-specific genes and thus the likely contribution of SDE genes to the occurrence of common human diseases.

Methods

Data sources

RNA-seq data were retrieved from the GTEx project version 6 [27, 28]. Population variation data were retrieved from the 1000 Genomes Project, Phase 3 (n = 2504 individuals, [41]). The human GRCh37 release coding exome coordinates and sequences were retrieved from Ensembl [57].

Variation analysis

The AnnoVar software package [58] was used to annotate the reported variations from the 1000 Genomes Project, and all possible variations (relative to the GRCh37; h19 reference genome) in every human protein-coding position documented in GRCh37. For each variation we determined its specific protein-transcript consequences, its population frequency, and its predicted functional likelihood (using both SIFT and PolyPhen algorithms [59, 60]). A non-synonymous (NS) variation was considered functional only when both SIFT and PolyPhen algorithms predicted it as deleterious [24]. Because SIFT mainly uses sequence conservation and PolyPhen mainly uses structural and functional impacts, we found the combination of the two methods to be highly accurate (number of true positive from total positive prediction [24]). This analysis calculated the distribution of all mutation types for each gene as observed in the 1000 Genomes Project population, and the computed distribution of all possible nucleotide substitutions consequences (i.e., NS, DNS, S, and stop-gain) for each protein-coding gene. The obtained data allowed us to calculate the deleterious (dDNS), loss-of-function (dStop-gain), and neutral (dS) mutation rates for each gene or group of genes according to the 1000 Genomes Project data. We examined the use of other available sources of human genetic variations, such as ExAC [61]. However, the number of additional SNVs with population frequencies >0.005, which are predominantly affected by selection, from these sources was negligible relative to the 1000 Genomes Project data (not shown).

Selection analysis

Previously, others and we have shown that the effect of selection on a mutation largely depends on its population frequency. Selection predominantly affects mutations that are have a population frequency >0.005, while very rare mutations (population frequency <0.001) tend to undergo negligible selection [24, 44]. Selection was thus analyzed according to the MAF range of the variations [24]. Selection pressures were assessed by calculating for each gene, or group of genes, the ratios of its functional (DNS and stop-gain) mutation rates to its neutral (S) mutation rate. The rate of a mutation type is the number of observed mutations from a certain type (e.g., S) in the 1000 Genomes Project, Phase 3, divided by all computed possible nucleotide substitutions leading to that type of mutation in the gene. The selection signature is the ratio of the functional rates (dDNS or dStop-gain) divided by the neutral rate (dS), that is, dDNS/dS and dStop-gain/dS. These measures extend the dN/dS type measures, similar to a previous work [43]. As in dN/dS, higher ratios indicate lower purifying selection [42]. To calculate if the dDNS/dS and dStop/dS ratios in a group of sex-specific genes deviated from these ratios in other protein-coding genes, we performed a randomization test: all non-Y-linked, non-testis-specific unique protein-coding human genes for which we have variation data in the 1000 Genomes Project and expression data in the GTEx project were used to create 10,000 random sets for each gene group. The number of genes in each set was the number of genes in the examined gene group. To compare the dDNS and the dS ratios between the two independent groups of moderately women-specific genes and non-sex-specific genes, we performed a Fisher's exact test.

Differential expression

Genes with SDE were detected by two approaches from the NOISeq R package [31, 32] The first approach used the NOISeqBIO algorithm, which treats the sample population as biological replicates in which the computed variability within the population is considered as noise [31, 32]. We used this to compare gene reads per kilobase of transcript per million mapped reads (RPKM) expression values between women and men population samples from corresponding tissues after excluding uninformative genes, that is, genes that did not have at least an expression of 1 RPKM in any sample. A probability cutoff of 0.95 was used to identify genes with significant differential expression, as this cutoff value is considered correct for multiple testing [31, 62]. The NOISeqBIO method provides effective statistics for determining differential expression between two populations. However, this approach regards the population variability as noise and could exclude some genuinely sex-differentiated genes that have complex modes of expression. For instance, genes activated during ovulation are expected to be expressed only in a few women (mainly in women <50 years old and on a few days each month [63]), while not being expressed in most women and in all men samples. The differential expression of such genes will be difficult to identify using a straightforward population analysis. To detect at least some of these cases, we used an additional analysis approach that could identify the difference in such cases.

To overcome this issue, at least partially, a single trimmed mean of all RPKM or read count expression values was calculated for every gene from each tissue

sample and sex (men or women) by removing the two most extreme sample values. This removed samples that could have skewed the mean. Assuming the trimmed means of read counts reflect the population expression of a gene in men or women samples, we then computed their differential expression using the NOISeq-sim algorithm [32]. NOISeq-sim relies on the assumption that read counts follow a multinomial distribution, in which the probability for each feature in the multinomial distribution is the probability of a read to map to that feature. This identified an additional list of genes with differential expression that were not identified by NOISeqBIO but had NOIseq-sim probability scores of at least 0.8 and a NOISeqBIO probability score at least 0.2 smaller.

Finally, to assess the reproducibility of SDE analysis by NOISeqBIO we implemented and used another differential expression method, DESeq2 [64], and analyzed the adipose-subcutaneous and liver datasets. We found that after *p*-value adjustment for multiple-testing correction, >92% of the adipose-subcutaneous and liver genes identified as SDE in NOISeqBIO were also found to be SDE by DESeq2.

The possible impact of the sample size on the number of identified SDE genes per tissue was tested by the Pearson correlation co-efficient (r). To assess a possible bias in the age distribution between men and women samples we used the two-sample Kolmogorov–Smirnov test. We found no significant differences in age distribution between men and women.

Gene and tissue clustering

Patterns of differential expression were analyzed using the following gene differential expression score, calculated for tissues with data for both men and women:

$$\mathrm{SDE} = LOG_2\left\{\left(1 + EXPR_{\boldsymbol{g,t}}{}^{\mathbf{w}}/MAX_g\right)/\left(1 + EXPR_{\boldsymbol{g,t}}{}^{\mathbf{m}}/MAX_g\right)\right\}$$

Where g is a specific gene, t is a specific tissue, and m and w represent men and women, respectively. $\mathrm{EXPR}^{\boldsymbol{w}}_{\boldsymbol{g,t}}$ is the NOIseqBIO-calculated mean RPKM expression value of gene g in tissue t for women (or for men with the m superscript), and MAX_g is the maximal NOIseqBIO calculated mean RPKM expression value of gene g in all tissues (including tissues specific for men or women). This score returns the differential expression value of a gene in a specific tissue, relative to the maximal expression of the gene. The value ranges from 1 (exclusive expression in women) to -1 (exclusive expression in men). This formula gives lower scores when expression in the examined gene and tissue are lower than those of the gene in some other tissue and can be generalized to compare the difference between two populations normalizing by a maximal value (Additional file 1: Figure S1).

Hierarchical cluster analyses and principle component analysis (PCA) were performed on a matrix of sex-differential expression (SDE) scores, with values of 0 given to genes that were not found significant in the NOISeq statistical analyses described above. Heatmap and hierarchical cluster analyses used the hclust method of the heatmap.2 R package and the pvclust method [34]. The PCA and the partitioning around medoids analyses used the CLARA and PAM methods of the R cluster package [65], with Euclidean distance measurements. This analysis allowed us to group genes according to their SDE patterns similarity.

Sex-specific expression

To find genes that are specific or highly specific to one sex, for each non Y-chromosome gene we calculated the ratio of its maximal trimmed mean expression values in one sex to its maximal trimmed mean expression in the other sex. Genes were considered as specific or highly sex-specific for ratios of at least 4-fold, when the lower maximal expression value was at least 1 RPKM. A ratio cutoff of 2-fold was used when the higher maximal expression was at least 1 RPKM but the other lower maximal expression value was very low (<1 RPKM). Other genes with sex ratios of 2–4-fold were considered as having moderately sex-specific expression, and genes with ratios of 1.1–0.9-fold were considered as having sex-similar expression. The statistical significance of the highly sex-specific gene expression was tested using the NOISeqBIO method, comparing samples from the tissue with the highest expression in one sex to samples from the tissue with the highest expression in the other sex.

Gene enrichment analysis

Gene enrichment analysis was performed using the GeneAnalytics server, which can identify gene enrichment for several terms and data sources, including diseases, pathways, GO terms, and tissue expression [35].

Additional files

Additional file 1: Figure S1. Differential expression score. Landscape of scores for all possible ratios of men and women expression values using a base 2 logarithm. The formula we derived for SDE score can be generalized to compare the difference between two populations (*x* and *y*) in a certain tissue or condition (*t*), normalizing by a gene (*g*) maximal expression value (MAX*g*). This differential expression score (DES) is a logarithm of the normalized ratios, giving scores between −1 and 1. We use a logarithm base of 2, but other bases (*n*) are possible. The general expression is thus $\mathrm{DES} = \mathrm{LOG}n\{(1 + (\mathrm{EXPR}^{x}_{g,t}/\mathrm{MAX}g) * (n-1))/(1 + (\mathrm{EXPR}^{y}_{g,t}/\mathrm{MAX}g) * (n-1))\}$ where $\mathrm{EXPR}^{x}_{g,t}$ is the expression value of gene *g* in tissue/condition *t* for population *x*. This score returns the differential expression value of a gene in specific tissue/condition, relative to the maximal expression of the gene. The value ranges from 1 (exclusive expression in *x*) to −1 (exclusive expression in *y*). Larger logarithm bases (*n*) exponentially increase the

transitions between exclusive expression (1 and −1) and non-differential (0) scores. (PDF 276 kb)

Additional file 2: Figure S2. SDE score heatmap of all protein-coding genes in 45 tissues common to both sexes. Scores are color-coded from blue (strictly men) to red (strictly women), with non-differential expression in white. Most genes are similarly expressed in most tissues with the exception of the breast mammary gland (more than 6000 SDE genes). (PDF 187 kb)

Additional file 3: Table S1. SDE scores for all protein-coding genes in 45 tissues common to men and women. Genes were analyzed by NOISeqBIO with scores of zero given for genes with insignificant differential expression. Other genes have SDE scores below zero for men-biased expression and above zero for women-biased expression. (CSV 2205 kb)

Additional file 4: Figure S3. Occurrence of genes according to number and exact (**a**) or cumulative (**b**) number of tissues they have SDE in. Most SDE genes are differentially expressed in one or few tissues. SDE genes in multiple tissues are mostly linked to the sex chromosomes (Additional file 5: Table S2). (PDF 177 kb)

Additional file 5: Table S2. Genes with SDE in more than five tissues. (DOCX 17 kb)

Additional file 6: Figure S4. SDE score heatmap of 244 protein-coding X-linked genes, ordered by their chromosomal position. Three genes have men-biased expression in multiple tissues, are in the PAR1, and none in PAR2 regions (green boxes). Scores are color-coded from blue (strictly men) to red (strictly women), with non-differential expression in white. (PDF 152 kb)

Additional file 7: Figure S5. Hierarchical clustering of 44 tissues common to men and women (excluding mammary glands) by their gene SDE patterns. Percent *p*-values are Bootstrap-Probability in green, and Approximately-Unbiased in red [34]. The mammary gland tissue was excluded from the analysis because it had an order of magnitude more SDE genes than the other 44 common tissues. (PDF 45 kb)

Additional file 8 Figure S6. The first two components of principle component analysis of all protein-coding genes with SDE in at least one non-mammary gland tissue. Cluster colors denote groups of genes with the similar SDE patterns. See also Fig. 2. (PDF 191 kb)

Additional file 9: Figure S7. Expression of *TSHB* and *MUCL1* genes in 53 human tissues. Box-plots of women samples are in red and men samples in blue. The pituitary-specific gene *TSHB* is significantly overexpressed in men. *MUCL1* is significantly overexpressed in men skin and in women mammary glands. (PDF 199 kb)

Additional file 10 Table S3. Integrated SDE NOISeqBIO and NOISeq-sim based analyses (see "Methods") for all protein-coding genes in 44 tissues common to men and women (excluding mammary glands). Values below or above zero denote men- or women-biased expression, respectively, and zero denotes genes with insignificant SDE. (CSV 2111 kb)

Additional file 11 Figure S8. Non-Y-linked genes partitioning around medoids clustering by the gene SDE patterns in 44 tissues common to men and women (excluding mammary glands). To identify SDE in genes with complex modes of expression we applied the NOISeq-sim approach that weighs groups of outliers (see "Methods"). The mammary gland tissue was excluded from the analysis because it had an order of magnitude more SDE genes than the other common tissues. (PDF 138 kb)

Additional file 12: Figure S9. Sex-biased expression of the *MTRNR2L2* gene. *MTRNR2L2* is notably expressed in women substantia-nigra (**a**, red box) due to overexpression in women older than 60 years. **b** Women under and above 60 years, $n = 12$ and $n = 12$ respectively. Men under and above 60 years, $n = 16$ and $n = 23$, respectively. (PDF 297 kb)

Additional file 13: Table S4. Women-biased protein-coding gene terms enrichment analyses. Multiple Excel worksheets summarizing GeneAnalytics [35] gene enrichments including diseases, GO terms, and pathways. (XLSX 625 kb)

Additional file 14: Table S5. Men-biased protein-coding gene terms enrichment analyses. Multiple Excel worksheets summarizing GeneAnalytics [35] gene enrichments including diseases, GO-terms, and pathways. (XLSX 557 kb)

Additional file 15: Table S6. SDE scores for testis overexpressed protein-coding genes across 53 tissues. Values below or above zero denote male or women-biased expression, respectively, and zero denotes genes with insignificant SDE. (CSV 791 kb)

Additional file 16: Figure S10. SDE score heatmap of testis-specific and moderately specific genes. Red and blue denote women or men specificity, respectively. (PDF 200 kb)

Additional file 17: Table S7. SDE scores for non-testis sex-specific and moderately sex-specific protein-coding genes across 53 tissues. Values below or above zero denote men- or women-biased expression, respectively. (CSV 177 kb)

Additional file 18: Figure S13. Age-related expression of the *NPPB* gene in heart left ventricle show overexpression in young women. Women under and above 60, $n = 54$ and $n = 22$, respectively. Men under and above 60, $n = 103$ and $n = 39$, respectively. (PDF 92 kb)

Additional file 19: Figure S11. Expression of *XIST* gene in 53 human tissues shown as box-plots with women samples in red and men samples in blue. Pink and light blue are women and men reproductive tissues, respectively. (PDF 41 kb)

Additional file 20: Figure S12. SDE score heatmap of non-protein-coding genes. Red and blue denote women or men specificity, respectively. (PDF 221 kb)

Acknowledgements
We thank Dan Mishmar and Tsviya Olender for helpful discussion and advice.

Funding
This research was supported in part by the Ministry of Agriculture of the State of Israel, and Weizmann Institute Forchheimer Center for Molecular Genetics and Crown Human Genome Center. The results shown here are in part based upon data generated by the GTEx project and the 1000 Genomes Project.

Availability of data and materials
All the data is publicly available. The GTEx Analysis V6 RNA-seq (05 Oct 2015) data are available on the GTEx portal (http://www.gtexportal.org/home/datasets). Variation data from the 1000 Genomes Project, Phase 3, can be obtained from http://www.internationalgenome.org/data/.

Authors' contributions
MG and SP designed the experiments, performed the analysis, and wrote the manuscript. Both authors read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests.

Consent for publication
Not applicable.

Ethics approval and consent to participate
This study only analyses existing publicly available data.

Received: 24 October 2016 Accepted: 19 January 2017
Published online: 07 February 2017

References
1. Bachtrog D, Mank JE, Peichel CL, Kirkpatrick M, Otto SP, Ashman T-L, Hahn MW, Kitano J, Mayrose I, Ming R. Sex determination: why so many ways of doing it? PLoS Biol. 2014;12:e1001899.
2. McClellan HL, Miller SJ, Hartmann PE. Evolution of lactation: nutrition v. protection with special reference to five mammalian species. Nutr Res Rev. 2008;21:97–116.
3. McClellan J, King M-C. Genetic heterogeneity in human disease. Cell. 2010; 141:210–7.
4. Connallon T. The geography of sex-specific selection, local adaptation, and sexual dimorphism. Evolution. 2015;69:2333–44.
5. Deaner RO, Shepherd SV, Platt ML. Familiarity accentuates gaze cuing in women but not men. Biol Lett. 2007;3:65–8.
6. Goldstein JM, Holsen L, Handa R, Tobet S. Fetal hormonal programming of sex differences in depression: linking women's mental health with sex differences in the brain across the lifespan. Front Neurosci. 2014;8:247.

7. Giedd JN, Castellanos FX, Rajapakse JC, Vaituzis AC, Rapoport JL. Sexual dimorphism of the developing human brain. Prog Neuro-Psychopharmacol Biol Psychiatry. 1997;21:1185–201.
8. Collaer ML, Hines M. Human behavioral sex differences: a role for gonadal hormones during early development? Psychol Bull. 1995;118:55.
9. Waldron I. Sex differences in human mortality: the role of genetic factors. Soc Sci Med. 1983;17:321–33.
10. Subbaraman M, Goldman-Mellor S, Anderson E, LeWinn K, Saxton K, Shumway M, Catalano R. An exploration of secondary sex ratios among women diagnosed with anxiety disorders. Human Reprod. 2010;25:2084–91.
11. Pulido MR, Rabanal-Ruiz Y, Almabouada F, Díaz-Ruiz A, Burrell MA, Vázquez MJ, Castaño JP, Kineman RD, Luque RM, Diéguez C. Nutritional, hormonal, and depot-dependent regulation of the expression of the small GTPase Rab18 in rodent adipose tissue. J Mol Endocrinol. 2013;50:19–29.
12. Link JC, Chen X, Arnold AP, Reue K. Metabolic impact of sex chromosomes. Adipocyte. 2013;2:74–9.
13. Bartley EJ, Fillingim RB. Sex differences in pain: a brief review of clinical and experimental findings. Br J Anaesth. 2013;111:52–8.
14. Courtright SH, McCormick BW, Postlethwaite BE, Reeves CJ, Mount MK. A meta-analysis of sex differences in physical ability: revised estimates and strategies for reducing differences in selection contexts. J Appl Psychol. 2013;98:623.
15. Tseng LA, Delmonico MJ, Visser M, Boudreau RM, Goodpaster BH, Schwartz AV, Simonsick EM, Satterfield S, Harris T, Newman AB. Body composition explains sex differential in physical performance among older adults. J Gerontol Ser A Biol Med Sci. 2014;69:93–100.
16. Dimas AS, Nica AC, Montgomery SB, Stranger BE, Raj T, Buil A, Giger T, Lappalainen T, Gutierrez-Arcelus M, McCarthy MI. Sex-biased genetic effects on gene regulation in humans. Genome Res. 2012;22:2368–75.
17. Rawlik K, Canela-Xandri O, Tenesa A. Evidence for sex-specific genetic architectures across a spectrum of human complex traits. Genome Biol. 2016;17:166.
18. Gilks WP, Abbott JK, Morrow EH. Sex differences in disease genetics: evidence, evolution, and detection. Trends Genet. 2014;30:453–63.
19. Sandberg K, Verbalis JG. Sex and the basic scientist: is it time to embrace Title IX? Biol Sex Differ. 2013;4:13.
20. Fisher RA. The genetical theory of natural selection: a complete variorum edition. Oxford: Oxford University Press; 1930.
21. Connallon T, Clark AG. The resolution of sexual antagonism by gene duplication. Genetics. 2011;187:919–37.
22. Frank SA, Hurst LD. Mitochondria and male disease. Nature. 1996;383:224.
23. Morrow EH, Connallon T. Implications of sex-specific selection for the genetic basis of disease. Evol Appl. 2013;6:1208–17.
24. Gershoni M, Pietrokovski S. Reduced selection and accumulation of deleterious mutations in genes exclusively expressed in men. Nat Commun. 2014;5:4438.
25. Innocenti P, Morrow EH. The sexually antagonistic genes of Drosophila melanogaster. PLoS Biol. 2010;8:e1000335.
26. Su AI, Wiltshire T, Batalov S, Lapp H, Ching KA, Block D, Zhang J, Soden R, Hayakawa M, Kreiman G. A gene atlas of the mouse and human protein-encoding transcriptomes. Proc Natl Acad Sci U S A. 2004;101:6062–7.
27. Ardlie KG, Deluca DS, Segrè AV, Sullivan TJ, Young TR, Gelfand ET, Trowbridge CA, Maller JB, Tukiainen T, Lek M. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. Science. 2015;348:648–60.
28. Melé M, Ferreira PG, Reverter F, DeLuca DS, Monlong J, Sammeth M, Young TR, Goldmann JM, Pervouchine DD, Sullivan TJ. The human transcriptome across tissues and individuals. Science. 2015;348:660–5.
29. Mank JE. The transcriptional architecture of phenotypic dimorphism. Nature Ecology & Evolution. 2017;1:0006.
30. Carithers LJ, Ardlie K, Barcus M, Branton PA, Britton A, Buia SA, Compton CC, DeLuca DS, Peter-Demchok J, Gelfand ET. A novel approach to high-quality postmortem tissue procurement: The GTEx Project. Biopreservation Biobanking. 2015;13:311–9.
31. Tarazona S, Furió-Tarí P, Turrà D, Di Pietro A, Nueda MJ, Ferrer A, Conesa A. Data quality aware analysis of differential expression in RNA-seq with NOISeq R/Bioc package. Nucleic Acids Res. 2015;43:e140.
32. Tarazona S, García-Alcalde F, Dopazo J, Ferrer A, Conesa A. Differential expression in RNA-seq: a matter of depth. Genome Res. 2011;21:2213–23.
33. Mangs HA, Morris BJ. The human pseudoautosomal region (PAR): origin, function and future. Curr Genomics. 2007;8:129–36.
34. Suzuki R, Shimodaira H. Pvclust: an R package for assessing the uncertainty in hierarchical clustering. Bioinformatics. 2006;22:1540–2.
35. Fuchs SB-A, Lieder I, Stelzer G, Mazor Y, Buzhor E, Kaplan S, Bogoch Y, Plaschkes I, Shitrit A, Rappaport N. GeneAnalytics: an integrative gene set analysis tool for next generation sequencing, RNAseq and microarray data. OMICS. 2016;20:139–51.
36. Holditch SJ, Schreiber CA, Burnett JC, Ikeda Y. Arterial remodeling in B-type natriuretic peptide knock-out females. Sci Rep. 2016;6:25623.
37. Wang TJ, Larson MG, Levy D, Leip EP, Benjamin EJ, Wilson PW, Sutherland P, Omland T, Vasan RS. Impact of age and sex on plasma natriuretic peptide levels in healthy adults. Am J Cardiol. 2002;90:254–8.
38. Clark BC, Collier SR, Manini TM, Ploutz-Snyder LL. Sex differences in muscle fatigability and activation patterns of the human quadriceps femoris. Eur J Appl Physiol. 2005;94:196–206.
39. Russ DW, Kent-Braun JA. Sex differences in human skeletal muscle fatigue are eliminated under ischemic conditions. J Appl Physiol. 2003;94:2414–22.
40. Ellegren H, Parsch J. The evolution of sex-biased genes and sex-biased gene expression. Nat Rev Genet. 2007;8:689–98.
41. 1000 Genomes Project Consortium, Abecasis GR, Altshuler D, Auton A, Brooks LD, Durbin RM, Gibbs RA, Hurles ME, McVean GA. A map of human genome variation from population-scale sequencing. Nature. 2010;467:1061–73.
42. Kryazhimskiy S, Plotkin JB. The population genetics of dN/dS. PLoS Genet. 2008;4:e1000304.
43. Ostrow SL, Barshir R, DeGregori J, Yeger-Lotem E, Hershberg R. Cancer evolution is associated with pervasive positive selection on globally expressed genes. PLoS Genet. 2014;10:e1004239.
44. Tennessen JA, Bigham AW, O'Connor TD, Fu W, Kenny EE, Gravel S, McGee S, Do R, Liu X, Jun G, et al. Evolution and functional impact of rare coding variation from deep sequencing of human exomes. Science. 2012;337:64–9.
45. Wu R, Lin M. Functional mapping - how to map and study the genetic architecture of dynamic complex traits. Nat Rev Genet. 2006;7:229–37.
46. de Moura Souza A, Sichieri R. Association between serum TSH concentration within the normal range and adiposity: a review. Eur J Endocrinol. 2011;165:11–5.
47. Skorupskaite K, George JT, Anderson RA. The kisspeptin-GnRH pathway in human reproductive health and disease. Hum Reprod Update. 2014;20:485–500.
48. Guengerich FP, Waterman MR, Egli M. Recent structural insights into cytochrome P450 function. Trends Pharmacol Sci. 2016;37:625–40.
49. Lamba V, Lamba J, Yasuda K, Strom S, Davila J, Hancock ML, Fackenthal JD, Rogan PK, Ring B, Wrighton SA. Hepatic CYP2B6 expression: gender and ethnic differences and relationship to CYP2B6 genotype and CAR (constitutive androstane receptor) expression. J Pharmacol Exp Ther. 2003;307:906–22.
50. Xiong Q, Jiao Y, Hasty KA, Canale ST, Stuart JM, Beamer WG, Deng H-W, Baylink D, Gu W. Quantitative trait loci, genes, and polymorphisms that regulate bone mineral density in mouse. Genomics. 2009;93:401–14.
51. Brain S, Williams T, Tippins J, Morris H, MacIntyre I. Calcitonin gene-related peptide is a potent vasodilator. Nature. 1985;313:54–6.
52. Gangula PR, Zhao H, Supowit SC, Wimalawansa SJ, Dipette DJ, Westlund KN, Gagel RF, Yallampalli C. Increased blood pressure in α-calcitonin gene–related peptide/calcitonin gene knockout mice. Hypertension. 2000;35:470–5.
53. Bateman AJ. Intra-sexual selection in Drosophila. Heredity. 1948;2:349–68.
54. Cerase A, Pintacuda G, Tattermusch A, Avner P. Xist localization and function: new insights from multiple levels. Genome Biol. 2015;16:1.
55. Kassam I, Lloyd-Jones L, Holloway A, Small KS, Zeng B, Bakshi A, Metspalu A, Gibson G, Spector TD, Esko T. Autosomal genetic control of human gene expression does not differ across the sexes. Genome Biol. 2016;17:248.
56. Chen C-Y, Lopes-Ramos CM, Kuijjer ML, Paulson JN, Sonawane AR, Fagny M, Platig J, Glass K, Quackenbush J, DeMeo DL. Sexual dimorphism in gene expression and regulatory networks across human tissues. bioRxiv 2016. Epub ahead of print. doi:10.1101/082289.
57. Herrero J, Muffato M, Beal K, Fitzgerald S, Gordon L, Pignatelli M, Vilella AJ, Searle SM, Amode R, Brent S. Ensembl comparative genomics resources. Database. 2016;2016:bav096.
58. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 2010;38:e164.

59. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc. 2009;4:1073–81.
60. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR. A method and server for predicting damaging missense mutations. Nat Methods. 2010;7:248–9.
61. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, O'Donnell-Luria AH, Ware JS, Hill AJ, Cummings BB, et al. Analysis of protein-coding genetic variation in 60,706 humans. Nature. 2016;536:285–91.
62. Efron B, Tibshirani R, Storey JD, Tusher V. Empirical Bayes analysis of a microarray experiment. J Am Stat Assoc. 2001;96:1151–60.
63. Ferin M, Jewelewicz R. The menstrual cycle: physiology, reproductive disorders, and infertility. New York: Oxford University Press; 1993.
64. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15:1.
65. Kaufman L, Rousseeuw PJ. Finding groups in data: an introduction to cluster analysis, vol. 344. Hoboken, New Jersey: John Wiley & Sons; 2009.

The New York Times https://www.nytimes.com/2022/01/13/health/transgender-teens-hormones.html

Doctors Debate Whether Trans Teens Need Therapy Before Hormones

Clinicians are divided over new guidelines that say teens should undergo mental health screenings before receiving hormones or gender surgeries.

By Azeen Ghorayshi
Jan. 13, 2022

An upsurge in teenagers requesting hormones or surgeries to better align their bodies with their gender identities has ignited a debate among doctors over when to provide these treatments.

An international group of experts focused on transgender health last month released a draft of new guidelines, the gold standard of the field that informs what insurers will reimburse for care.

Many doctors and activists praised the 350-page document, which was updated for the first time in nearly a decade, for including transgender people in its drafting and for removing language requiring adults to have psychological assessments before getting access to hormone therapy.

But the guidelines take a more cautious stance on teens. A new chapter dedicated to adolescents says that they must undergo mental health assessments and must have questioned their gender identity for "several years" before receiving drugs or surgeries.

Experts in transgender health are divided on these adolescent recommendations, reflecting a fraught debate over how to weigh conflicting risks for young people, who typically can't give full legal consent until they are 18 and who may be in emotional distress or more vulnerable to peer influence than adults are.

Some of the drug regimens bring long-term risks, such as irreversible fertility loss. And in some cases, thought to be quite rare, transgender people later "detransition" to the gender they were assigned at birth. Given these risks, as well as the increasing number of adolescents seeking these treatments, some clinicians say that teens need more psychological assessment than adults do.

"They absolutely have to be treated differently," said Laura Edwards-Leeper, a child clinical psychologist in Beaverton, Ore., who works with transgender adolescents.

Dr. Edwards-Leeper was one of seven authors of the new adolescent chapter, but the organization that publishes the guidelines, the World Professional Association for Transgender Health, did not authorize her to comment publicly on the draft's proposed wording.

On the other side of the debate are clinicians who say the guidelines are calling for unnecessary barriers to urgently needed care. Transgender teens have a high risk of attempting suicide, according to the Centers for Disease Control and Prevention. And preliminary studies have suggested that adolescents who receive drug treatments to affirm their gender identity have improved mental health and well-being. Considering those data, some clinicians are opposed to any mental health requirements.

"I'm really not a believer in requiring that for people," said Dr. Alex Keuroghlian, a clinical psychiatrist at Fenway Health in Boston and the director of the Massachusetts General Hospital Psychiatry Gender Identity Program. "Being trans isn't a mental health problem," he later added.

The public is invited to comment on the guidelines until Sunday, with a final version expected by spring.

As clinicians debate the intricacies of the new health standards, state legislatures across the country are trying to ban gender-affirming medical care for adolescents. According to the Williams Institute at the UCLA School of Law, 21 states introduced such bills last year. Gov. Greg Abbott of Texas has described gender surgeries as "genital mutilation" and "child abuse."

Professional medical groups and transgender health specialists have overwhelmingly condemned these legal attempts as dangerous. So far, two have passed into law, in Tennessee and Arkansas, though the latter has temporarily been blocked because of legal appeals.

Some clinicians worry that public disagreement over the best way to care for transgender adolescents will add fuel to this simmering political movement.

"It's a stressful environment to be in," said Dr. Cassie Brady, a pediatric endocrinologist at Vanderbilt University Medical Center who provided legislative testimony to make the bill in Tennessee less restrictive. "It not only puts fear in us as providers, but way more fear I think for families who might be trying to balance this."

Dr. Cassie Brady, a pediatric endocrinologist at Vanderbilt University Medical Center. William DeShazer for The New York Times

“A real shift”

The first version of the guidelines, called the Standards of Care, was released by a small group of doctors at a San Diego meeting in 1979. At the time, there was little public acknowledgment of transgender people, and they had scant options for medical care.

The document “was a real shift,” said Beans Velocci, a historian at the University of Pennsylvania.

But those first guidelines characterized gender nonconformity as a psychological disorder. They stated that transgender people could be delusional or unreliable, and required two letters from psychiatrists before adults could access surgeries. That focus on psychology set a lasting precedent, experts said.

“The establishment medical world didn’t even understand it — they were still treating it as a mental health concern — just 20 years ago,” said Dr. Joshua Safer, an endocrinologist and executive director of the Center for Transgender Medicine and Surgery at Mt. Sinai, who contributed to the guidelines’ chapter on hormone therapy.

Children and teenagers struggling with their gender identities did not get much attention from the medical community until the 1990s, when two contrasting models emerged.

In one approach, clinicians in the Netherlands suggested that parents wait for puberty to make decisions about their children transitioning to another gender, pioneering the use of drugs that suppress the production of hormones like testosterone and estrogen. The Dutch model argued that these puberty blockers, which are reversible, would buy adolescents time to further explore their gender before starting hormones with more lasting consequences.

In another model, which was developed in Canada and is now considered a form of “conversion therapy,” children were pushed to live in the gender they were assigned at birth, in order to avoid drugs or surgeries down the line for those who might later change their minds.

Around the late 2000s, clinicians in the United States introduced the “gender affirming” approach, which has since been endorsed by several major medical groups. Its basic philosophy: Minors should be able to live out their gender identities freely, without clinicians or parents imposing unnecessary delays. Their path might involve medications and surgeries, or no medical treatments at all.

“Children are not short adults — but they have autonomy as well, and they can know their gender,” said Dr. Diane Ehrensaft, director of mental health at the University of California, San Francisco Child and Adolescent Gender Center. Dr. Ehrensaft is one of the key early proponents of the gender-affirming model and helped write a new chapter on prepubescent children in the draft guidelines.

Data on the number of transgender or gender nonconforming adolescents and adults in the United States are limited. About 1.8 percent of high school students surveyed in 19 state or urban school districts in 2017 described themselves as transgender, according to the C.D.C.

Adolescent gender clinics like Dr. Ehrensaft’s have seen a rapid growth in referral rates, and more sites have sprung up to meet demand. Today there are more than 50 such specialty clinics in the United States, she said, up from just four in 2012.

Few studies have followed adolescents receiving puberty blockers or hormones into adulthood. Dr. Ehrensaft and others are now working on large, long-term studies of patients in the United States.

An emerging divide

The new standards state that clinicians should facilitate an “open exploration” of gender with adolescents and their families, without pushing them in one direction or another. But the guidelines recommend restricting the use of medications and surgeries, partly because of their medical risks.

Puberty blockers, for example, can impede bone development, though evidence so far suggests it resumes once puberty is initiated. And if taken in the early phase of puberty, blockers and hormones lead to fertility loss. Patients and their families should be counseled about these risks, the standards say, and if preserving fertility is a priority, drugs should be delayed until a more advanced stage of puberty.

The guidelines suggest minimum ages, lower than those in the previous version, for each treatment: 14 for starting hormone therapy, 15 for chest masculinization and at least 17 for more invasive genital operations.

But the most contentious parts of the new standards among clinicians are the mental health requirements. Before discussing any medical treatment, they say, adolescents must get a “comprehensive assessment” led by mental health providers, and must have consistently questioned their gender identity for “several years.”

Dr. Alex Keuroghlian, a clinical psychiatrist at Fenway Health in Boston, said, “I’m really not a believer in requiring” therapy before transitioning. “Being trans isn’t a mental health problem,” he said. M. Scott Brauer for The New York Times

Although mental health counseling should be offered as needed, it should not be a requirement for medical care, said Dr. Keuroghlian of Fenway Health. He pointed out that therapy is not required for cisgender patients who get breast augmentation, hysterectomies or rhinoplasties.

“To make that a requirement for everybody is inherently unnecessary gatekeeping and also stigmatizing and pathologizing and a waste of resources,” he said.

What’s more, some of the mental health problems commonly seen in trans adolescents, such as depression and anxiety, may resolve after gender-affirming medical care, Dr. Keuroghlian said.

And some doctors have also argued that waiting several years to initiate medical treatments could itself be harmful.

“Forcing trans and gender diverse youth to go through an incongruent puberty can cause long-term trauma and physical harm,” said Dr. AJ Eckert, medical director of Anchor Health Initiative’s Gender and Life-Affirming Medicine Program in Stamford, Conn.

But other trans health specialists are concerned by the sharp increase in adolescents who are referred to gender clinics, and worry that the desire for hormones and surgeries may be driven partly by peer influence on social media platforms like TikTok and YouTube.

“The kids presenting these days are very different than what I was seeing in the early days,” said Dr. Edwards-Leeper, who in 2007 helped set up one of the first youth gender clinics in the United States, in Boston.

Dr. Edwards-Leeper said that now she was more likely to see adolescents who had recently begun to question their gender, whereas a decade ago her patients were more likely to have longstanding distress about their bodies.

These seemingly abrupt changes — as well as other mental health issues or a history of trauma — should be flags for providers to slow down, she said. Instead, some gender clinics with long wait lists are “blindly affirming” adolescent patients, she said, offering them hormones without taking these potential issues seriously.

And although it’s unclear how often it happens, some people who transitioned as teenagers have reported detransitioning later on. Although some people who detransition continue living with a more fluid gender identity, others are upset about living with the irreversible changes caused by hormones or surgeries.

“These issues of inadequate assessment and what I sometimes called hasty or sloppy care have resulted in potential harm,” said Erica Anderson, a clinical psychologist who works with transgender adolescents in Berkeley, Calif.

Dr. Anderson, 70, said she understood the trauma of being denied care. She first realized she was transgender in her 30s, but didn’t approach an endocrinologist about hormone treatments until age 45. “The doctor’s response was, ‘I can’t help you,’” she said. Despondent, she waited several more years before pursuing a medical transition again.

“I don’t want any young person to go without the care that they need,” Dr. Anderson said. “But the question is, are there new things going on that weren’t going on 10 or 15 years ago?”

Other doctors say they haven’t seen evidence to suggest that clinics are hastily providing medical treatments, or that many patients are experiencing regret about taking hormones. But they agree that teenagers require more mental health precautions than adults do.

“With kids, you’re more conservative,” said Dr. Safer of Mt. Sinai. But, he added, “I guess time and data will tell.”

A version of this article appears in print on , Section D, Page 1 of the New York edition with the headline: A Teen Trans Divide

Concussion Rates in U.S. Middle School Athletes From the 2015-2016 to 2019-2020 School Years

Hacherl SL, Kelshaw PM, Erdman NK, Martin JR, Cortes N, Dunn RE, Lincoln AE, Caswell SV: Sports Medicine Assessment, Research & Testing Laboratory, ACHIEVES Project, George Mason University, Manassas, VA; Athletic Training Program, Department of Kinesiology, University of New Hampshire, Durham, NH; Orthopedics & Sports Medicine Research, MedStar Health Research Institute, Baltimore, MD

Context: Research describing concussion incidence and mechanisms is necessary to inform primary prevention efforts and improve clinical care. There is limited information regarding the incidence of concussions in middle school athletic settings. Our purpose was to investigate the epidemiology of concussions among middle school age athletes from the 2015-16 to 2019-20 school years. **Methods:** As part of the Advancing Healthcare Initiatives for Underserved Students (ACHIEVES) project, athletic trainers recorded injury and athlete exposure (AE) data from public middle schools in Virginia. Concussion rates per 1000 AEs with 95% confidence intervals (CIs) were calculated for 12 school-sponsored sports (baseball, football, wrestling, boys' and girls' basketball, cheerleading, boys and girls' soccer, softball, boys' and girls' track, and volleyball). Due to COVID-19 data was not collected for boys' track, baseball, softball, and girls' soccer for 2019-2020. Injury rate ratios (IRR) were calculated to compare concussion rates between practice and competition. Sex-comparisons were conducted for sports played by boys and girls (e.g., soccer, basketball, track and field, and softball/baseball). IRRs with 95% CI excluding 1.00 were deemed statistically significant, consistent with prior protocols. **Results:** Overall, 339 concussions were reported across the five school years with 98 concussions (31.3%) occurring outside of school-sponsored sport participation and 8 occurring during unknown activities. A total of 233 concussions (68.7%) were reported during 390,562 AEs attributed to school-sponsored sport participation for an overall concussion rate of 0.60/1,000 AE (95% CI=0.56-0.64). Sports with the highest concussion rates were football (1.36/1000 AE, 95% CI= 1.05-1.67), girls' soccer (1.26/1000 AE, 95% CI= 0.77-1.75), and wrestling (1.12/1000 AE, 95% CI= 0.78-1.46). The overall concussion rate was higher in competition than practice (IRRcompetition/practice=2.45, 95% CI= 1.16-3.75). Among sex-comparable sports, concussion rates for girls were more than twice that for boys (overall: 0.49 vs. 0.23/1000 AE; IRR = 2.13, 95% CI = 0.40-3.86; competition: girls: 1.36 vs. boys: 0.66/1000 AE, IRRgirls/boys=2.08, 95% CI= 0.29-3.86; practice: girls: 0.33 vs. boys: 0.16/1000 AE, IRRgirls/boys =2.09, 95% CI= 0.00-4.17). **Conclusions:** Using a large sample of middle school age athletes, we observed football had the highest overall concussion rate among all middle school sports. This finding is in consistent with prior literature that have reported that football has the greatest incidence of concussion in a similar population of athletes. Consistent with previous research, girls suffered concussions at a higher rate than boys when participating in sex-comparable sports. Collectively, our findings suggest that middle school sports have an overall higher rate of concussion than reported in high school and collegiate settings. Our findings reinforce the value and importance of onsite athletic training services within middle school sport settings.

PHYSIOLOGY 30: 30–39, 2015; doi:10.1152/physiol.00024.2014

Sex-Based Differences in Skeletal Muscle Kinetics and Fiber-Type Composition

K. M. Haizlip, B. C. Harrison, and L. A. Leinwand

Department of Molecular, Cellular, and Developmental Biology, BioFrontiers Institute, University of Colorado at Boulder, Boulder, Colorado
leslie.leinwand@colorado.edu

Previous studies have identified over 3,000 genes that are differentially expressed in male and female skeletal muscle. Here, we review the sex-based differences in skeletal muscle fiber composition, myosin heavy chain expression, contractile function, and the regulation of these physiological differences by thyroid hormone, estrogen, and testosterone. The findings presented lay the basis for the continued work needed to fully understand the skeletal muscle differences between males and females.

Cardiac, smooth, and skeletal are the three muscle types in mammals, with skeletal muscle being the most abundant tissue in the human body. Skeletal muscles are composed of different types of fibers which diverge morphologically, biochemically, and functionally. Early studies describing muscle fiber-type composition and development did not address the potential for differences between species and sex. This is often the case in experimental models where sex is thought to confound results and therefore usually only males are studied. Although there is a significant appreciation of sex-based differences in cardiovascular health and disease, much less is known about the effects of sex on the physiology and pathophysiology of skeletal muscle. More than 3,000 genes have been identified as being differentially expressed between male and female skeletal muscle (96). Thus many of the key differences that exist between the sexes that will be presented here are likely to be the result, at least in part, of these differences in gene expression. In this review, we discuss differences in skeletal muscle fiber-type composition and contractility between sexes and how these differences may be hormonally regulated.

The Mammalian Myosin Heavy Chain Gene Family and Muscle Fiber Types

Because the myosin motor protein is so closely linked to muscle function, it is important to consider the effects of sex on myosin gene expression. There are 11 myosin heavy chain (MyHC) genes that encode different myosin isoforms in mammals (77). Table 1 lists these genes and the muscle fibers in which they are expressed. Four of these genes constitute the majority of MyHCs expressed in adult mammalian skeletal muscle: MYH1 (MyHC-IIx), MYH2 (MyHC-IIa), MYH4 (MyHC-IIb), and MYH7 (MyHC-β). Although there are varying degrees of MyHC isoform co-expression in single skeletal muscle fibers, individual fibers often contain a predominant MyHC isoform. Thus individual fibers may be "typed" or classified by MyHC isoform expression, with four main classifications in mammalian skeletal muscle: type-I (MyHC-I/β), type-IIA (MyHC-IIa), type-IIX (MyHC-IIx), and type-IIB (MyHC-IIb).

Functionally, MyHC isoform expression directly affects muscle fiber-type contractile velocity via myosin ATPase activity (74), with relative velocities of I < IIA < IIX < IIB (73, 74). Importantly, MyHC isoform expression closely correlates with fiber-type morphology and enzymatic make-up. In general terms, muscle fibers expressing MyHC-IIb tend to be larger fibers, rich in glycolytic enzymes, whereas fibers expressing MyHC-I/β tend to be smaller with a higher oxidative capacity. In addition to these functional differences across fiber types, there are species differences in MyHC isoform expression and, therefore, differences in skeletal muscle fiber-type composition. For example, rodent skeletal muscle (especially mouse) is predominantly comprised of muscle fibers expressing MyHC-IIb, with an overall MyHC abundance across murine muscles of IIb > IIx > IIa > I/β. In contrast, human skeletal muscle does not express MyHC-IIb protein, a phenomenon that appears to be, at least in part, due to a reduction in the activity of the human MyHC-IIb promoter region (37). In contrast to murine skeletal muscle, the overall MyHC isoform abundance across human skeletal muscles is I/β > IIa > IIx, although there are certainly regional and muscle-specific differences in MyHC isoform expression, as described below.

Sex-Based Differences in Fiber-Type Composition

Overall, evidence to date suggests that skeletal muscle fiber-type composition is dependent on species, anatomical location/function, and sex. In general terms, MyHC isoform expression tends to

1548-9213/15 ©2015 Int. Union Physiol. Sci./Am. Physiol. Soc.

Table 1. *Mammalian myosin genes, the proteins they encode, and where they are expressed*

Gene	Protein	Expression	Fiber Type
MYH1	**MyHC-IIx**	**Fast IIX fibers**	**Type-IIX**
MYH2	**MyHC-IIa**	**Fast IIA fibers**	**Type-IIA**
MYH3	MyHC-embryonic	Developing muscle	
MYH4	**MyHC-IIb**	**Fast IIB fibers**	**Type-IIB**
MYH6	MyHC-α	Masseter, extraocular muscle and heart	
MYH7	**MyHC-β/slow**	**Slow muscle (Type-I/β) and heart**	**Type-I/β**
MYH7B	MyHC-slow/tonic	Extraocular muscle	
MYH8	MyHC-neonatal	Developing muscle	
MYH13	MyHC-EO	Extra-ocular muscle, pharyngeal muscle	
MYH15	MyHC-15	Extra-ocular muscle	
MYH16	MyHC-M	Jaw muscle	

Genes, protein, expression locations, and fiber types in **bold** represent those prevalent in mammalian skeletal muscle.

be "slower" (i.e., higher relative expression of MyHC-I/β and MyHC-IIa) in larger mammals compared with smaller mammals. As mentioned above, these species differences in MyHC isoform expression are most pronounced with respect to the MyHC-IIb gene. Across individual muscle groups (at least in rodents), deeper postural muscles such as the soleus tend to have higher relative expression of MyHC isoforms associated with a more oxidative phenotype (MyHC-I/β and MyHC-IIa), whereas superficial muscle groups, such as the gastrocnemius, tend to be comprised almost entirely of muscle fibers expressing MyHC-IIx and MyHC-IIb. Although studies investigating sex-based differences in muscle fiber type are somewhat limited, there is also evidence of sexual dimorphisms with respect to muscle fiber-type composition (Table 2).

The mouse masseter is composed of high numbers of IIX fibers in both sexes; however, there are threefold fewer IIB fibers and more IIA fibers in females compared with males (27). Another study of male and female mouse masseter also finds threefold fewer fibers expressing MyHC-IIb in females compared with males (22). Male rabbit masseter is comprised of nearly 80% type-IIA fibers, whereas the female masseter contains only ~50% (28). Separating the rabbit masseter into 10 different muscle compartments reveals that two of the compartments consist of more type-IIA fibers in young males vs. more type-I fibers in young females (27). For illustrative purposes, FIGURE 1 is an immunohistochemical analysis that highlights potential sexually dimorphic regional differences in MyHC isoform expression in the mouse hindlimb. For example, MyHC-IIa expression is higher in males vs. females in the soleus (58% in males vs. 36% in females) and tibialis anterior (TA) (39% in males vs. 25% in females), whereas in the plantaris, IIa expression is higher in females (37%) than in males (16%).

Just as in experimental animal models, myosin isoform mRNA and protein composition can be very different between men and women. Many human studies are conducted on the vastus lateralis muscle due to the relative ease of sample collection. Analysis of mRNA from both male and female vastus lateralis biopsies by microarray shows (when normalized to α-actinin) that the female muscle has 35% more MyHC-I/β, 30% less MyHC-IIa, and 15% less MyHC-IIx mRNA than the corresponding male muscle (96). Type-I fibers account for 36% of the total biopsy area in men and 44% in women, whereas type-IIA fibers account for 41% in men and only 34% in women (89). Overall, in this analysis, the average MyHC isoform percentages in the male vastus lateralis are IIx (20%), IIa (46%), and I (34%), whereas the females samples are IIx (23%), IIa (36%), and I (41%) (89). Additionally, all of the fibers measured in men have significantly larger cross-sectional areas (CSA) compared with women as normalized to biopsy section area (89). The difference in CSA seems likely to be due to the overall greater mass in males compared with females since the increase in CSA is nearly proportional to the differences in mass.

Contractility

Studies on chemically skinned human fibers reveal that MyHC isoform composition is the key determinant of muscle fiber contractile velocity and rate of force development (11, 72). Although evidence suggests that maximum unloaded shortening velocity is not different between young male and female fibers (25), sex-based differences are seen during fatigue recovery and endurance testing (Table 3). Analysis of tetanic force in the mouse masseter shows no significant differences between the sexes on force production (22). However, the same study reveals that, during a tetanic force protocol, the maximal rate of force generation is significantly higher in males than in females (516 vs. 382 g/s). Additionally, the rate of relaxation is significantly faster in the male masseter during a tetanus compared with the female masseter (41 vs. 48 ms) (22).

Table 2. *Fiber-type expression and area*

Muscle	Male	Female	Reference
Mouse Masseter (average)	IIX > IIB > IIA	IIX > IIA > IIB	28
Rabbit Masseter (average)	IIA > I	I > IIA	27
Human vastus lateralis (CSA)	IIA > IIX > I	I > IIA > IIX	90
Human VL MHC isoform %	IIa > I > IIx	I > IIa > IIx	90

Differences in fiber-type size and composition of listed muscle bodies. Data are presented as fiber-type contribution in decreasing order. CSA, cross-sectional area; VL, vastus lateralis.

The capacity for fatigue of a muscle is an indicator of recovery capabilities and can differ between species and sex. Generally, male muscles are more fatigable than female muscles. In vivo studies on fatigue focus on the ability to maintain the contractile strength of a muscle. This maintained contractile force is measured as exerted force or maximal voluntary contraction from single or multiple muscle fibers. At the start of a fatigue experiment, the maximal voluntary contraction is determined for each subject and used to quantify the relative force decline during the protocol. Although there are reports of sex-based differences in fatigability, the reasoning for these differences in fatigability could be associated with differences in muscle substrate utilization, neuromuscular activation, or muscle morphology (38). Additionally, fatigue can be induced through impairment of central drive or motivation, neuromuscular propagation, or peripherally through the impairment of excitation-contraction coupling (30, 32).

Studies on elbow flexor and knee extensor muscles show a significant loss of motor unit activation in males vs. females following a standard fatigue protocol (1). Although forces exerted by both young men and women are not significantly different, endurance is lower for men than for women (11 vs. 18 min) (39). In the elbow flexor muscle, endurance time to fatigue is 1,806 s in females and 829 s in males (40). After intermittent bouts of isometric contraction (5 s on and 5 s off) in the adductor pollicis muscle, women fatigue more slowly and to a smaller degree than men (31). Specifically, force falls to ~93% of the maximum in women vs. 80% in men after <1 min of exercise (31). Endurance is longer in women (15 min) compared with men (8 min), and after reaching exhaustion, recovery occurs faster in women than in men (31). Fatigue also appears to be muscle-specific, as evidenced by analysis of both the biceps femoris and lumbar extensors for fatigability. More specifically, women fatigue similarly in both muscles, whereas men fatigue more in the lumbar musculature compared with the biceps femoris (19). These studies suggest that force generation and relaxation are faster during fatigue in men compared with women, whereas endurance is higher and recovery is quicker in women than in men. These differences in muscle performance between men and women may be due to the larger numbers of type-I fibers in women that are characterized by slow oxidative metabolism and thus higher endurance. Overall, these studies lay the groundwork for more extensive basic research on sex differences in skeletal muscle composition and function.

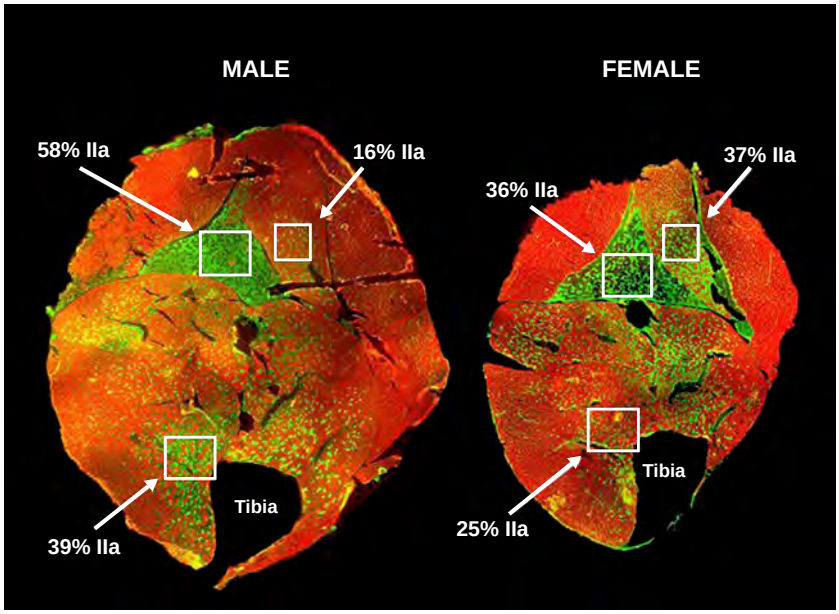


FIGURE 1. **Type-IIa fibers are differentially expressed in male vs. female hindlimb muscle sections and in different muscle bodies**
In the soleus, IIa expression is 58% in the male 36% in the female. In the plantaris, IIa expression is16% in the male and 37% in the female. In the tibialis anterior muscle, IIa expression is 39% in the male and 25% in the female. Green staining identifies IIa fibers, whereas red staining identifies IIb fibers. Figure was generously provided by Brooke Harrison, PhD.

Hormonal Regulation of Myosin Isoforms

It seems logical that hormones contribute to sexual dimorphisms in fiber-type composition and contractility. In fact, hormonal regulation of skeletal muscle development and contractility are well documented (29, 70). Skeletal muscle is a major target of thyroid hormone (42); estrogen exerts its effects on the cardiovascular and musculoskeletal systems (45); and testosterone is heavily studied for its pro-hypertrophic/anabolic effects on skeletal muscle. Fiber-type composition and contractile function can be altered by the presence or removal

Table 3. *Baseline contractility measurements in males and females*

Measurement	Muscle	Result	Reference
Rate of force development	Mouse masseter	Higher in males	22
Rate of relaxation	Mouse masseter	Lower in males	22
Force	Human elbow flexor	Equal in males and females	1
Endurance	Human elbow flexor	Lower in males	41
Endurance	Human adductor pollicis	Lower in males	32

of specific hormones. In the following section, we present findings on the effects of thyroid hormone, estrogen, and testosterone on contractility, fiber type, and the differences that occur between the sexes.

Thyroid Hormone

Thyroid hormone (T3) can regulate MyHC gene expression (42). T3 affects muscle protein expression at the posttranscriptional, translational, and posttranslational levels (for review, see Ref. 15). Individuals with hypothyroidism, or an underactive thyroid, typically display symptoms such as low heart rate, fatigue, weight gain, muscle weakness, a conversion from fast to slow fiber types (41), and a more efficient energy metabolism (5). Hypothyroidism produces a lower type-II fiber percentage in male and female patients compared with healthy patients (63). Hypothyroid females appear to have a higher proportion of type-II fibers compared with hypothyroid males; however, type-II fiber atrophy occurs in hypothyroid female patients but not in males (63). Graves' disease, a common immune disorder, is characterized by an overproduction of T3 or hyperthyroidism. According to the Mayo Clinic, individuals with Graves' disease typically experience symptoms such as difficulty sleeping, fatigue, sensitivity to heat, and a rapid or irregular heartbeat (www.mayoclinic.org). Treatment of euthyroid animals with T3 induces hyperthyroidism, producing a reversible slow-to-fast MyHC isoform transition from I → IIa → IIx → IIb (for review, see Ref. 78). Before treatment with T3, almost all soleus muscle myofibers in male and female rats express type-I fibers (101). Four weeks of T3 treatment induces an increase in type-IIA fibers and a downregulation of type-I fibers in male and female rat soleus muscle (57, 101). Specifically, after treatment with T3, male rat soleus muscle expresses 73% type-I/IIA and 26% type-I/IIAX of the total fibers, whereas females express 37% type-I, 49% type-I/IIA, and only 6% type-I/IIAX fibers (101). Overall, the increase in IIX content from the type-I/IIAX fibers is greater in young T3-treated males (21%) than females (10%), and the upregulation of IIA is greater in young females (43%) than in young males (27%) (101). In another study of euthyroid animals, treatment with T3 induces expression of IIX in the soleus of aged male rats (4–15%), whereas there is no detectable IIX in females at any age (52).

The rat extensor digitorum longus (EDL) contains predominately fast MyHC isoforms (IIa, IIx, and IIb). Hypothyroidism leads to elevated levels of MyHC-IIa mRNA compared with MyHC-IIb mRNA in rat EDL (47). Long-term treatment with T3 in euthyroid rats (for 24 wk) results in a decrease in mRNA levels and the percentage of protein isoforms for MyHC-IIb and MyHC-IIa in the EDL muscle (93). Other studies in rat EDL and TA show that chronic hyperthyroidism enhances MyHC-IIb mRNA, downregulates MyHC-IIa mRNA, and leads to a decrease in IIa protein levels with no change in IIb protein levels compared with euthyroid control animals (47). T3 treatment in euthyroid animals leads to a conversion from IIA to IIB fibers in the EDL of young and old female rats vs. no transition in males (52). Table 4 summarizes the fiber-type changes in distribution as it relates to sex and thyroid hormone levels. These studies highlight the differences in fiber-type conversion as it relates to sex and pinpoints the differences in IIx expression in males and females. These fine differences in fiber-type contribution could effectively change contractile function, endurance, and the response to fatigue.

Contractile regulation by thyroid hormone. T3 is typically implicated in cardiac contractility, but there are also studies showing T3 effects on skeletal muscle contractility (16, 43). In hyperthyroid individuals, ATP turnover rate is faster than in normal individuals, whereas in hypothyroid individuals, ATP turnover is slower (99). Analysis of hyperthyroid canine quadriceps femoris muscle reveals a T3-induced increase in maximal contractile rate and in rate of relaxation over euthyroid levels (65). Given the T3-induced sex-based differences in fiber-type conversion, it is logical to hypothesize that T3 may produce differences in contractility between the sexes. Contractility of rat soleus and EDL is not different in maximal unloaded shortening velocity in isolated fibers (type I, I/IIA, and I/IIAX) between hyperthyroid males and females. However, a faster shortening velocity is observed in soleus fibers from hyperthyroid females (old and young) compared with euthyroid controls (25). On the other hand, other reports have not seen T3-dependent differences in shortening velocity of

Table 4. *Hormonal regulation of fiber-type expression by thyroid hormone*

Insult	Male	Female	Reference
Hypothyroidism in humans	↑ Type-II	↑ ↑ Type-II	64
T3 treatment of rat soleus	I/IIA > I/IIAX	I/IIA > I > I/IIAX	102
T3 treatment on rat soleus	↑ IIX	↑ IIA	53, 102
T3 treatment in rat EDL	No change	IIA → IIB	53

T3 treatment, hyperthyroidism.

EDL muscle and only slight variances in soleus (57). These studies suggest a potential role for T3 in skeletal muscle contractility and in sex-based differences that may exist in the response to changing T3 levels but point to the need for further investigations.

Thyroid hormone receptors and estrogen receptors are suggested to interact in a way that modulates estrogen-sensitive gene expression (26, 102). Thus the effects of estrogen on skeletal muscle fiber-type composition may correlate with the previously described findings on the effects of T3 on skeletal muscle.

Estrogen

The relationship between estrogen and skeletal muscle function and recovery has been analyzed for decades. In the following section, we present the current understanding of the effects of estrogen and its changing levels on skeletal muscle physiology. In humans, a decline in estrogen can be due to conditions such as hypogonadism, hypopituitarism, anorexia nervosa, perimenopause, or menopause. During menopause, the decline in estrogen levels is paralleled with a slight increase in injury risk and a decline in lean body mass (9). Animal models of low estrogen levels, which occur after ovariectomy (OVX), reveal an increase in overall body mass and in the mass of individual muscles (68). Skeletal muscle is an estrogen-responsive tissue such that estrogen receptor (ER) mRNA and protein levels change with circulating estrogen levels (7). Skeletal muscle expresses two forms of ERs (α and β), and the effects of estrogen are mediated, in part, through these receptors (14, 21, 56, 98). In male and female vastus lateralis muscle, expression of ERα mRNA is 180-fold higher than ERβ, with no significant difference in expression levels between males and females (98). Additionally, immunohistochemical analysis highlights ERα and ERβ localization to the nuclei of skeletal muscle equally in males and females (97, 98). ERα-null mice have a decreased muscle regenerative capacity following injury, as evidenced by small myofibers and centrally located nuclei (50). This decline in regenerative capacity is thought to be directly related to findings that estrogen is associated with maintenance of muscle strength in pre- and postmenopausal women (69) and suggests a role for ERs in skeletal muscle maintenance (64, 84, 85, 88).

Mechanism of action. Estrogen is thought to induce cellular effects, such as transcription, through the activation of ERK1/2 and p38 followed by phosphorylation of both cAMP response element-binding protein and elk-1, which play a role in regulating the early c-Fos gene (75). Estrogen has anti-apoptotic effects through the PI3K/Akt pathways (10). Additionally, the IGF-1 pathway, which plays a positive role in skeletal muscle growth and regeneration, is implicated in estrogen signal transduction (36, 53, 54, 94). Studies on muscle repair following exercise-induced injury highlight a sex-based difference in which higher levels of muscle creatine kinase (indicative of injury) are observed in male than in female rats following injury. Following OVX, levels of muscle creatine kinase in females are similar to levels in injured males (3, 4, 8).

Recent studies suggest a beneficial role for hormone replacement therapy (HRT), supplementation with estrogen and progesterone, and the induction of pathways by which estrogen can aide in skeletal muscle maintenance and repair (71). In a test of physical performance, HRT leads to a significant increase in vertical jumping height in women compared with exercise alone and no treatment (85). Ronkainen et al. (76) found that HRT is associated with better mobility, greater muscle power, and favorable body and muscle composition in women ages 54–62 (76). A role for estrogen in the prevention of the inflammatory response, which can exacerbate muscle injury and inhibit recovery, is also suggested (3, 92).

Influence of estrogen on muscle fibers. Estrogen has been shown to influence fiber size, overall muscle weight, muscle regeneration, and contractility, and to induce minimal changes in fiber-type distribution. OVX leads to an increase in overall body weight and muscle weight (68), and a 70% increase in ERα mRNA with no change in ERβ mRNA (7). Soleus muscle weight increases after OVX in 10-wk-old female rats (68, 95). Supplementation with estrogen following OVX leads to a reduction in body weight (62) and attenuates the increase in ERα mRNA (7). In rat soleus and EDL muscles, OVX leads to a slight increase in the mean fiber diameter of type-I, -IIA, and -IIB fi-

Table 5. *Effects of ERβ-null mice*

Measurement	Male WT	Female WT	Male ERβ$^{-/-}$	Female ERβ$^{-/-}$
Force production (soleus)	19.4 mN	35.0 mN	16.9 mN	29.7 mN
Contraction time (soleus)	57.4 ms	90.8 ms	62.6 ms	99.6 ms
Relaxation time (soleus)	100.4 ms	158.2 ms	104.0 ms	191.4 ms
Relaxation time during tetanus (EDL)	39.8 ms	49.6 ms	46.0 ms	60.6 ms

Data are republished from Ref. 34.

bers. Estrogen supplementation decreases fiber diameter in all fiber types in both the soleus and EDL significantly from OVX levels to below baseline levels (90). There is no difference in fiber-type composition of the soleus of OVX mice (67). However, in the mouse plantaris muscle, OVX does lead to a reduction in the relative amount of type-IIX fibers from 38% in sham animals to 33% (70). Supplementation with estrogen increases the type-IIX percentage composition in the plantaris back to 42% (70).

Contractile regulation by estrogen. OVX rats supplemented with estrogen have also been analyzed for changes in contractile function (among other parameters). OVX induces a reduction in time to peak tension and increases time to 50% relaxation in rats, whereas estrogen replacement reverses the OVX-induced reduction in peak tension (62). Analysis of contractile function in OVX mice reveals a significant decrease in maximal isometric tetanic force from sham levels (180 mN) to 168 mN (67, 68). This decline is attenuated to 187 mN with estrogen supplementation (67). In the gonadally intact female rat EDL, twitch tension is 41 g, which increases to 51 g following OVX but then decreases to 31 g in OVX animals treated with estrogen (90). Since a strong actin-myosin bond will produce a greater force, this has been measured in muscles with and without estrogen manipulation. The fraction of strong binding myosin in OVX muscle is 0.263 or ~15% lower than baseline (67, 68). Following estrogen treatment, the fraction of strongly bound myosin increases to 0.311, which is indistinguishable from sham levels (0.300) (67). The relative strength of the bond between actin and myosin is proportional to the energy available for the production of force (44). The differences in actin-myosin bond strength suggest that estrogen may play a role in increasing force production at the myofilament level.

In the studies referenced in the previous sections, there is no change in MyHC isoform expression or overall muscle fiber-type composition, but there are estrogen-related changes in fiber-type diameter, muscle size, contractility, and the actin-myosin interaction. Previous studies highlight the calcium-induced increase in force production in skeletal muscle and the subsequent decline in force following the removal of calcium (58, 91). Because of the OVX-induced decrease in tetanic force, which requires rapid replenishment and utilization of calcium stores, and the increase in fiber diameter, which could compensate for a decline in calcium sensitivity, calcium handling and sensitivity may be playing a role in the regulation of contractile differences in samples with and without estrogen. OVX induces a decline in twitch tension and slows twitch kinetics in the soleus (95). Additionally, there is a decline in maximal tension produced in the soleus of 10-wk-old OVX rats (~19%) and 14-wk-old OVX rats (~20%) compared with sham (95). These changes are not paralleled with changes in calcium sensitivity in OVX animals; however, the decline in the strength of the actin-myosin bond seen by Moran et al. (67) or a decline in calcium storage capacity following OVX remains uninvestigated.

Sex-based differences in fatigue. As stated previously, there is no sexually dimorphic expression in either ERα or ERβ in skeletal muscle (98). Estrogen is present in both males and females (albeit at different levels), and it is suggested to play a role in male sexual behavior and cognition (80). ERα-null mice have decreased tetanic tension of gastrocnemius and TA muscles in females compared with wild-type controls (14). In an estrogen-elimination animal model (aromatase null), tetanic tension produced by the gastrocnemius muscle and TA is lower than in controls (14). The role of ERβ has been investigated through the use of ERβ-null mice (ERβ $^{-/-}$) (49). These data are summarized in Table 5. Force production in soleus muscle declines slightly in ERβ$^{-/-}$ but is not significantly different (33). Contraction duration is significantly longer in the female soleus in both wild-type (91 ms) and null (100 ms) compared with male wild-type (57 ms) and null (63 ms) mice. Relaxation time during a basic contraction in the soleus is significantly shorter in wild-type males (104 ms) than in females (158 ms) and in null males (104 ms) compared with null females (191 ms). In the EDL, half contraction time during a tetanic twitch protocol lasts 40 ms in male wild-type mice and 46 ms in male ERβ$^{-/-}$, whereas in female wild-type, half contraction time lasts 50 ms compared with 61 ms in ERβ$^{-/-}$ females (33). Following a fatiguing protocol on soleus and EDL muscle, force decreases less in ERβ$^{-/-}$ males than in wild-type males, whereas

the reverse is true for females, suggesting a role for ERβ knockout in enhancing endurance in males and reducing endurance in females (33). Overall, ERβ deletion increases the contractile time and decreases the speed of contractile kinetics in both males and females. However, because baseline contractile duration is already longer in females than in males, ERβ deletion simply exacerbates the wild-type phenotypes. During fatigue, where females typically have greater endurance, ERβ elimination decreases the endurance in the female and enhances endurance in the male. This could be due to minute changes in fiber-type contribution, CSA, or changes in calcium-handling proteins. Taken together, these findings on estrogen and the ER-mediated regulation of the contractile response and fiber-type distribution suggest a role for estrogen and its receptors in contractile maintenance and function, with a differential effect of estrogen in males vs. females.

Testosterone

Testosterone is a highly studied androgen that is associated with an increase in muscle mass (66, 83). With advancing age, testosterone levels drop, and testosterone deficiency is associated with early death, primarily from cardiovascular disease, a decrease in muscle mass and/or strength, and sexual dysfunction (46). Testosterone has been shown to induce its effects through binding to intracellular androgen receptors (ARs) (34), which then regulate gene transcription (18, 81). Myoblasts treated with testosterone have enhanced hypertrophic responses via the PI3K pathway (23). Testosterone supplementation leads to an increase in muscle mass and a decrease in fat content in hypogonadal men (13). This response may be through a testosterone-induced increase in the number of satellite cells and thus hypertrophy (82, 83). The testosterone-mediated proliferation and differentiation of satellite cells is thought to be due to an upregulation of follistatin and an inhibition of transforming growth factor-β (12). ARs are widely expressed in myoblasts, myofibers, and motoneurons in both males and females (17, 100), and testosterone injected into the masseter of female guinea pigs leads to an increase in fiber size (35). Additionally, testosterone supplementation in postmenopausal women leads to a 50% increase in protein synthesis rate compared with no effect with estrogen treatment (87). These anabolic effects of testosterone are well characterized and provide the catalysts for further investigation of the response of muscle fiber-type composition to testosterone deficiency and supplementation in males and females.

Influence of testosterone on muscle fiber type and morphology. As mentioned earlier, studies of vastus lateralis biopsies show fiber-type CSA to be larger in men than in women. Specifically, type-I fibers are 19% larger, type-IIA fibers are 59% larger, and type-IIX fibers are 66% larger in men than in women (normalized to total biopsy section analyzed) (89). In the adult guinea pig, castration produces atrophy of several muscles, including the latissimus dorsi, sternomastoid, and spinotrapezius (48). In mice, castration produces a decrease in body weight, which is attenuated with testosterone supplementation, and although overall muscle weight increases, the fiber CSA of the soleus and TA do not significantly change (6). However, there is a correlation between increases in fiber CSA and overall muscle mass. The role of testosterone on muscle fiber-type distribution has been, in part, determined through the analysis of the hypogonadal male and female mouse models. Analysis of fiber-type distribution in hypogonadal males vs. females reveals no significant changes in type-I, -IIA, -IIX, or -IIB fibers in gastrocnemius muscle or in type-I or -IIA fibers in the soleus (79). Additionally fiber-type distribution of the gastrocnemius is relatively unchanged from wild-type mice, with I $<$ IIX $<$ IIA $<$ IIB being the relative contribution of each fiber-type in males and females (79). However, in other instances, IIB fiber-type diameter in male hypogonadal mice significantly declines compared with wild-type males. Surprisingly, in females, the hypogonadal phenotype leads to an increase in IIB fiber diameter compared with wild-type females (79).

Sex-based differences in contractility. Testosterone is not typically associated with enhanced contractile function in that testosterone replacement is not associated with increased endurance. For instance, hypogondal and eugonadal men exhibit similar limb muscle strength and endurance during exercise (51). The increase in strength associated with testosterone supplementation is thought to be due to its anabolic effects. Male AR-null mice have a decrease in muscle mass that is not seen in AR-null females (59). Additionally, in AR-null males, force production decreases in fast-twitch fibers, whereas fatigue resistance increases in slow-twitch fibers to levels similar to wild-type females (59). Unfortunately, the contractile function in AR-null females is not analyzed in this study. Analysis of mRNA reveals an upregulation of genes encoding slow-twitch muscle contractile proteins in AR-null mice (2, 59). Because aromatase converts testosterone to estrogen (60), testosterone can induce its effects via either ERs or ARs, thus this contractile effect is thought to be mediated by the activation of ERs. Overall, the effects of testosterone deficiency are exacerbated in males compared with females. Testosterone deficiency leads to a decline in body mass, a decrease in fast-twitch fiber diameter, a conversion to slow-

twitch fibers, and enhanced fatigue resistance in males but not females.

Conclusions

Sex-based differences in skeletal muscle fiber-type composition and function are apparent in numerous species and are present in specific anatomical locations. Here, we present findings on sexual dimorphisms present in the mammalian musculoskeletal system. There are four main MyHC isoforms present in adult mammalian muscle (MyHC-I, -IIa, -IIx, and -IIb), which increase in contractile speed in the presented order. There is a prevalence of slower type-I and -IIA fibers in females compared with males that parallels the lower contractile velocity in females compared with males. The prevalence of the slower-twitch fibers is also a benefit to female performance in that the slower oxidative fibers and higher oxidative capacity allow for increased endurance and recovery, highlighting the sex-based differences in response to fatigue or muscle tetanus. To explain the potential cause of differences in skeletal muscle performance and fiber-type composition, we also present the differential effects of increases and decreases in levels of thyroid hormone, estrogen, and testosterone. Although thyroid hormone induces a conversion from slow to fast fibers and increases contractile velocity, sex-specific hormones estrogen and testosterone are implicated in skeletal muscle growth, fiber size, and minimally in contractile function. Some reports highlight enhanced contractile function and increased β-oxidative gene expression in men supplemented with estrogen and enhanced muscle growth in females treated with testosterone (61).

The identification of over 3,000 genes differentially regulated in male and female muscle highlights the complex differences that occur in skeletal muscle from both sexes (96). In this study, the authors focus on two particular genes that are upregulated in women compared with men and that are known to code for proteins that are in signaling pathways of growth factors known to regulate muscle mass: growth factor receptor-bound protein 10 (GRB10) and activin receptor type-2A (ACVR2A). GRB10 codes for a protein that suppresses IGF-1, which has an anabolic effect, whereas ACVR2A codes for a myostatin receptor, which has a role in muscle size determination. GRB10 knockout in male and female mice produces an increase in muscle weight and a decrease in body fat percentage and thus enhances muscularity (86). Another study reveals that knockout of ACVR2A causes muscle hypertrophy in female mice (55). Further analysis must be done to validate the role of these newly identified targets in skeletal muscle sex differences. In this review, we present changes in fiber-type composition, size, and contractile function in males and females; however, one or two genes cannot be responsible for altering all of these factors. To fully validate a gene as regulating the sex differences, for instance, in fiber-type composition, a thorough in vitro and in vivo analysis must be completed to first understand how the gene is regulated, the proteins' function and interacting partners, and then how these interactions might lead to sexually dimorphic differences in muscle fiber-type composition and function. The complexity of skeletal muscle and the role of sex adding to that complexity cannot be overlooked.

Future Directions

The lack of studying both males and females in the laboratory has recently attracted the attention of the public and the NIH (20). One recommendation made is that investigators report the sex of the animals or cell lines being studied. For instance, previous studies have identified sex as a determinant for the ability of muscle-derived stem cells to regenerate. Specifically, female muscle-derived stem cells regenerate more efficiently when transplanted into dystrophic mice (24). Sex differences should be accounted for in studies of skeletal muscle composition, function, and adaptive responses to different forms of exercise training and regression. For example, FIGURE 1 depicts differences in fiber-type composition in male and female mouse hindlimb muscles. Further studies of these fiber differences in the context of skeletal muscle adaption should provide insight into the regulatory differences between the sexes. Another neglected area is epigenetic differences in males and females in skeletal muscle, and studies should be aimed at determining the role of hormonal interventions in males and females given their clinical relevance. Numerous skeletal muscle therapies are built based on results from studies in men alone or with only a small subset of women. Having an appreciation for the differences that exist between the sexes is the first step to understanding the mechanisms underlying these sex differences. This review summarizes key findings in skeletal muscle physiology in the hopes of bringing to the forefront areas of future research and sexual disparities in current investigations. ■

This research was supported in part by National Institutes of Health grants to K.H. (GM-029090-30A1S1), B.H. (5K01 AR-55676), and L.L. (GM-29090-31).

No conflicts of interest, financial or otherwise, are declared by the author(s).

Author contributions: K.M.H. and B.C.H. prepared figures; K.M.H. drafted manuscript; K.M.H., B.C.H., and L.A.L. edited and revised manuscript; K.M.H. and L.A.L. approved final version of manuscript.

References

1. Albert WJ, Wrigley AT, McLean RB, Sleivert GG. Sex differences in the rate of fatigue development and recovery. *Dyn Med DM* 5: 2, 2006.
2. Altuwaijri S, Lee DK, Chuang KH, Ting HJ, Yang Z, Xu Q, Tsai MY, Yeh S, Hanchett LA, Chang HC, Chang C. Androgen receptor regulates expression of skeletal muscle-specific proteins and muscle cell types. *Endocrine* 25: 27–32, 2004.
3. Amelink GJ, Bar PR. Exercise-induced muscle protein leakage in the rat. Effects of hormonal manipulation. *J Neurol Sci* 76: 61–68, 1986.
4. Amelink GJ, Koot RW, Erich WB, Van Gijn J, Bar PR. Sex-linked variation in creatine kinase release, and its dependence on oestradiol, can be demonstrated in an in-vitro rat skeletal muscle preparation. *Acta Physiol Scand* 138: 115–124, 1990.
5. Argov Z, Renshaw PF, Boden B, Winokur A, Bank WJ. Effects of thyroid hormones on skeletal muscle bioenergetics. In vivo phosphorus-31 magnetic resonance spectroscopy study of humans and rats. *J Clin Invest* 81: 1695–1701, 1988.
6. Axell AM, MacLean HE, Plant DR, Harcourt LJ, Davis JA, Jimenez M, Handelsman DJ, Lynch GS, Zajac JD. Continuous testosterone administration prevents skeletal muscle atrophy and enhances resistance to fatigue in orchidectomized male mice. *Am J Physiol Endocrinol Metab* 291: E506–E516, 2006.
7. Baltgalvis KA, Greising SM, Warren GL, Lowe DA. Estrogen regulates estrogen receptors and antioxidant gene expression in mouse skeletal muscle. *PLos One* 5: e10164, 2010.
8. Bar PR, Amelink GJ, Oldenburg B, Blankenstein MA. Prevention of exercise-induced muscle membrane damage by oestradiol. *Life Sci* 42: 2677–2681, 1988.
9. Bea JW, Zhao Q, Cauley JA, LaCroix AZ, Bassford T, Lewis CE, Jackson RD, Tylavsky FA, Chen Z. Effect of hormone therapy on lean body mass, falls, and fractures: 6-year results from the Women's Health Initiative hormone trials. *Menopause* 18: 44–52, 2011.
10. Boland R, Vasconsuelo A, Milanesi L, Ronda AC, de Boland AR. 17beta-Estradiol signaling in skeletal muscle cells and its relationship to apoptosis. *Steroids* 73: 859–863, 2008.
11. Bottinelli R, Canepari M, Pellegrino MA, Reggiani C. Force-velocity properties of human skeletal muscle fibres: myosin heavy chain isoform and temperature dependence. *J Physiol* 495: 573–586, 1996.
12. Braga M, Bhasin S, Jasuja R, Pervin S, Singh R. Testosterone inhibits transforming growth factor-beta signaling during myogenic differentiation and proliferation of mouse satellite cells: potential role of follistatin in mediating testosterone action. *Mol Cell Endocrinol* 350: 39–52, 2012.
13. Brodsky IG, Balagopal P, Nair KS. Effects of testosterone replacement on muscle mass and muscle protein synthesis in hypogonadal men: a clinical research center study. *J Clin Endocrinol Metab* 81: 3469–3475, 1996.
14. Brown M, Ning J, Ferreira JA, Bogener JL, Lubahn DB. Estrogen receptor-alpha and -beta and aromatase knockout effects on lower limb muscle mass and contractile function in female mice. *Am J Physiol Endocrinol Metab* 296: E854–E861, 2009.
15. Caiozzo VJ, Haddad F. Thyroid hormone: modulation of muscle structure, function, and adaptive responses to mechanical loading. *Exerc Sport Sci Rev* 24: 321–361, 1996.
16. Carr AN, Kranias EG. Thyroid hormone regulation of calcium cycling proteins. *Thyroid* 12: 453–457, 2002.
17. Chen Y, Zajac JD, MacLean HE. Androgen regulation of satellite cell function. *J Endocrinol* 186: 21–31, 2005.
18. Cinar B, Mukhopadhyay NK, Meng G, Freeman MR. Phosphoinositide 3-kinase-independent nongenomic signals transit from the androgen receptor to Akt1 in membrane raft microdomains. *J Biol Chem* 282: 29584–29593, 2007.
19. Clark BC, Manini TM, The DJ, Doldo NA, Ploutz-Snyder LL. Gender differences in skeletal muscle fatigability are related to contraction type and EMG spectral compression. *J Appl Physiol* 94: 2263–2272, 2003.
20. Clayton JA, Collins FS. Policy: NIH to balance sex in cell and animal studies. *Nature* 509: 282–283, 2014.
21. Couse JF, Lindzey J, Grandien K, Gustafsson JA, Korach KS. Tissue distribution and quantitative analysis of estrogen receptor-alpha (ERalpha) and estrogen receptor-beta (ERbeta) messenger ribonucleic acid in the wild-type and ERalpha-knockout mouse. *Endocrinology* 138: 4613–4621, 1997.
22. Daniels DW, Tian Z, Barton ER. Sexual dimorphism of murine masticatory muscle function. *Arch Oral Biol* 53: 187–192, 2008.
23. Deane CS, Hughes DC, Sculthorpe N, Lewis MP, Stewart CE, Sharples AP. Impaired hypertrophy in myoblasts is improved with testosterone administration. *J Steroid Biochem Mol Biol* 138: 152–161, 2013.
24. Deasy BM, Lu A, Tebbets JC, Feduska JM, Schugar RC, Pollett JB, Sun B, Urish KL, Gharaibeh BM, Cao B, Rubin RT, Huard J. A role for cell sex in stem cell-mediated skeletal muscle regeneration: female cells have higher muscle regeneration efficiency. *J Cell Biol* 177: 73–86, 2007.
25. Degens H, Yu F, Li X, Larsson L. Effects of age and gender on shortening velocity and myosin isoforms in single rat muscle fibres. *Acta Physiol Scand* 163: 33–40, 1998.
26. Dellovade TL, Zhu YS, Krey L, Pfaff DW. Thyroid hormone and estrogen interact to regulate behavior. *Proc Natl Acad Sci USA* 93: 12581–12586, 1996.
27. Eason JM, Schwartz GA, Pavlath GK, English AW. Sexually dimorphic expression of myosin heavy chains in the adult mouse masseter. *J Appl Physiol* 89: 251–258, 2000.
28. English AW, Eason J, Schwartz G, Shirley A, Carrasco DI. Sexual dimorphism in the rabbit masseter muscle: myosin heavy chain composition of neuromuscular compartments. *Cells Tissues Organs* 164: 179–191, 1999.
29. Enns DL, Tiidus PM. The influence of estrogen on skeletal muscle: sex matters. *Sports Med* 40: 41–58, 2010.
30. Enoka RM. Activation order of motor axons in electrically evoked contractions. *Muscle Nerve* 25: 763–764, 2002.
31. Fulco CS, Rock PB, Muza SR, Lammi E, Cymerman A, Butterfield G, Moore LG, Braun B, Lewis SF. Slower fatigue and faster recovery of the adductor pollicis muscle in women matched for strength with men. *Acta Physiol Scand* 167: 233–239, 1999.
32. Gandevia SC. Spinal and supraspinal factors in human muscle fatigue. *Physiol Rev* 81: 1725–1789, 2001.
33. Glenmark B, Nilsson M, Gao H, Gustafsson JA, Dahlman-Wright K, Westerblad H. Difference in skeletal muscle function in males vs. females: role of estrogen receptor-beta. *Am J Physiol Endocrinol Metab* 287: E1125–E1131, 2004.
34. Gustafasson J, Pousette K. Demonstration and partial characterization of cytosol receptors for testosterone. *Biochemistry* 14: 3094–3101, 1975.
35. Gutmann E, Hanzlikova V, Lojda Z. Effect of androgens on histochemical fibre type. Differentiation in the temporal muscle of the guinea pig. *Histochemistry* 24: 287–291, 1970.
36. Hamelers IH, Steenbergh PH. Interactions between estrogen and insulin-like growth factor signaling pathways in human breast tumor cells. *Endocr Related Cancer* 10: 331–345, 2003.
37. Harrison BC, Allen DL, Leinwand LA. IIb or not IIb? Regulation of myosin heavy chain gene expression in mice and men. *Skelet Muscle* 1: 5, 2011.
38. Hicks AL, Kent-Braun J, Ditor DS. Sex differences in human skeletal muscle fatigue. *Exerc Sport Sci Rev* 29: 109–112, 2001.
39. Hunter SK, Critchlow A, Enoka RM. Influence of aging on sex differences in muscle fatigability. *J Appl Physiol* 97: 1723–1732, 2004.
40. Hunter SK, Enoka RM. Sex differences in the fatigability of arm muscles depends on absolute force during isometric contractions. *J Appl Physiol* 91: 2686–2694, 2001.
41. Ianuzzo D, Patel P, Chen V, O'Brien P, Williams C. Thyroidal trophic influence on skeletal muscle myosin. *Nature* 270: 74–76, 1977.
42. Izumo S, Nadal-Ginard B, Mahdavi V. All members of the MHC multigene family respond to thyroid hormone in a highly tissue-specific manner. *Science* 231: 597–600, 1986.
43. Kaasik A, Minajeva A, Paju K, Eimre M, Seppet EK. Thyroid hormones differentially affect sarcoplasmic reticulum function in rat atria and ventricles. *Mol Cell Biochem* 176: 119–126, 1997.
44. Karatzaferi C, Chinn MK, Cooke R. The force exerted by a muscle cross-bridge depends directly on the strength of the actomyosin bond. *Biophys J* 87: 2532–2544, 2004.
45. Katzenellenbogen BS, Montano MM, Le Goff P, Schodin DJ, Kraus WL, Bhardwaj B, Fujimoto N. Antiestrogens: mechanisms and actions in target cells. *J Steroid Biochem Mol Biol* 53: 387–393, 1995.
46. Khaw KT, Dowsett M, Folkerd E, Bingham S, Wareham N, Luben R, Welch A, Day N. Endogenous testosterone and mortality due to all causes, cardiovascular disease, and cancer in men: European prospective investigation into cancer in Norfolk (EPIC-Norfolk) Prospective Population Study. *Circulation* 116: 2694–2701, 2007.
47. Kirschbaum BJ, Kucher HB, Termin A, Kelly AM, Pette D. Antagonistic effects of chronic low frequency stimulation and thyroid hormone on myosin expression in rat fast-twitch muscle. *J Biol Chem* 265: 13974–13980, 1990.
48. Kochakian CD, Tillotson C. Influence of several C19 steroids on the growth of individual muscles of the guinea pig. *Endocrinology* 60: 607–618, 1957.
49. Krege JH, Hodgin JB, Couse JF, Enmark E, Warner M, Mahler JF, Sar M, Korach KS, Gustafsson JA, Smithies O. Generation and reproductive phenotypes of mice lacking estrogen receptor beta. *Proc Natl Acad Sci USA* 95: 15677–15682, 1998.
50. Labarge S, McDonald M, Smith-Powell L, Auwerx J, Huss JM. Estrogen-related receptor-alpha (ERRalpha) deficiency in skeletal muscle impairs regeneration in response to injury. *FASEB J* 28: 1082–1097, 2014.
51. Laghi F, Langbein WE, Antonescu-Turcu A, Jubran A, Bammert C, Tobin MJ. Respiratory and skeletal muscles in hypogonadal men with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 171: 598–605, 2005.
52. Larsson L, Yu F. Gender-related differences in the regulatory influence of thyroid hormone on the expression of myosin isoforms in young and old rats. *Acta Physiol Scand* 159: 81–89, 1997.

53. Lee AV, Jackson JG, Gooch JL, Hilsenbeck SG, Coronado-Heinsohn E, Osborne CK, Yee D. Enhancement of insulin-like growth factor signaling in human breast cancer: estrogen regulation of insulin receptor substrate-1 expression in vitro and in vivo. *Mol Endocrinol* 13: 787–796, 1999.

54. Lee AV, Weng CN, Jackson JG, Yee D. Activation of estrogen receptor-mediated gene transcription by IGF-I in human breast cancer cells. *J Endocrinol* 152: 39–47, 1997.

55. Lee SJ, Reed LA, Davies MV, Girgenrath S, Goad ME, Tomkinson KN, Wright JF, Barker C, Ehrmantraut G, Holmstrom J, Trowell B, Gertz B, Jiang MS, Sebald SM, Matzuk M, Li E, Liang LF, Quattlebaum E, Stotish RL, Wolfman NM. Regulation of muscle growth by multiple ligands signaling through activin type II receptors. *Proc Natl Acad Sci USA* 102: 18117–18122, 2005.

56. Lemoine S, Granier P, Tiffoche C, Rannou-Bekono F, Thieulant ML, Delamarche P. Estrogen receptor alpha mRNA in human skeletal muscles. *Med Sci Sports Exerc* 35: 439–443, 2003.

57. Li X, Larsson L. Contractility and myosin isoform compositions of skeletal muscles and muscle cells from rats treated with thyroid hormone for 0, 4 and 8 weeks. *J Muscle Res Cell Motil* 18: 335–344, 1997.

58. MacIntosh BR. Role of calcium sensitivity modulation in skeletal muscle performance. *News Physiol Sci* 18: 222–225, 2003.

59. MacLean HE, Chiu WS, Notini AJ, Axell AM, Davey RA, McManus JF, Ma C, Plant DR, Lynch GS, Zajac JD. Impaired skeletal muscle development and function in male, but not female, genomic androgen receptor knockout mice. *FASEB J* 22: 2676–2689, 2008.

60. Maggiolini M, Donze O, Jeannin E, Ando S, Picard D. Adrenal androgens stimulate the proliferation of breast cancer cells as direct activators of estrogen receptor alpha. *Cancer Res* 59: 4864–4869, 1999.

61. Maher AC, Akhtar M, Tarnopolsky MA. Men supplemented with 17beta-estradiol have increased beta-oxidation capacity in skeletal muscle. *Physiol Genom* 42: 342–347, 2010.

62. McCormick KM, Burns KL, Piccone CM, Gosselin LE, Brazeau GA. Effects of ovariectomy and estrogen on skeletal muscle function in growing rats. *J Muscle Res Cell Motil* 25: 21–27, 2004.

63. McKeran RO, Slavin G, Andrews TM, Ward P, Mair WG. Muscle fibre type changes in hypothyroid myopathy. *J Clin Pathol* 28: 659–663, 1975.

64. Meeuwsen IB, Samson MM, Verhaar HJ. Evaluation of the applicability of HRT as a preservative of muscle strength in women. *Maturitas* 36: 49–61, 2000.

65. Miyashita A, Suzuki S, Suzuki M, Numata H, Suzuki J, Akahori T, Okubo T. Effect of thyroid hormone on in vivo contractility of the canine diaphragm. *Am Rev Respir Dis* 145: 1452–1462, 1992.

66. Mooradian AD, Morley JE, Korenman SG. Biological actions of androgens. *Endocr Rev* 8: 1–28, 1987.

67. Moran AL, Nelson SA, Landisch RM, Warren GL, Lowe DA. Estradiol replacement reverses ovariectomy-induced muscle contractile and myosin dysfunction in mature female mice. *J Appl Physiol* 102: 1387–1393, 2007.

68. Moran AL, Warren GL, Lowe DA. Removal of ovarian hormones from mature mice detrimentally affects muscle contractile function and myosin structural distribution. *J Appl Physiol* 100: 548–559, 2006.

69. Petrofsky JS, Burse RL, Lind AR. Comparison of physiological responses of women and men to isometric exercise. *J Appl Physiol* 38: 863–868, 1975.

70. Piccone CM, Brazeau GA, McCormick KM. Effect of oestrogen on myofibre size and myosin expression in growing rats. *Exp Physiol* 90: 87–93, 2005.

71. Pollanen E, Ronkainen PH, Horttanainen M, Takala T, Puolakka J, Suominen H, Sipila S, Kovanen V. Effects of combined hormone replacement therapy or its effective agents on the IGF-1 pathway in skeletal muscle. *Growth Hormone IGF Res* 20: 372–379, 2010.

72. Ranatunga KW, Thomas PE. Correlation between shortening velocity, force-velocity relation and histochemical fibre-type composition in rat muscles. *J Muscle Res Cell Motil* 11: 240–250, 1990.

73. Reiser PJ, Moss RL, Giulian GG, Greaser ML. Shortening velocity in single fibers from adult rabbit soleus muscles is correlated with myosin heavy chain composition. *J Biol Chem* 260: 9077–9080, 1985.

74. Resnicow DI, Deacon JC, Warrick HM, Spudich JA, Leinwand LA. Functional diversity among a family of human skeletal muscle myosin motors. *Proc Natl Acad Sci USA* 107: 1053–1058, 2010.

75. Ronda AC, Buitrago C, Colicheo A, de Boland AR, Roldan E, Boland R. Activation of MAPKs by 1alpha,25(OH)2-Vitamin D3 and 17beta-estradiol in skeletal muscle cells leads to phosphorylation of Elk-1 and CREB transcription factors. *J Steroid Biochem Mol Biol* 103: 462–466, 2007.

76. Ronkainen PH, Kovanen V, Alen M, Pollanen E, Palonen EM, Ankarberg-Lindgren C, Hamalainen E, Turpeinen U, Kujala UM, Puolakka J, Kaprio J, Sipila S. Postmenopausal hormone replacement therapy modifies skeletal muscle composition and function: a study with monozygotic twin pairs. *J Appl Physiol* 107: 25–33, 2009.

77. Schiaffino S, Reggiani C. Fiber types in mammalian skeletal muscles. *Physiol Rev* 91: 1447–1531, 2011.

78. Schiaffino S, Reggiani C. Molecular diversity of myofibrillar proteins: gene regulation and functional significance. *Physiol Rev* 76: 371–423, 1996.

79. Sciote JJ, Horton MJ, Zyman Y, Pascoe G. Differential effects of diminished oestrogen and androgen levels on development of skeletal muscle fibres in hypogonadal mice. *Acta Physiol Scand* 172: 179–187, 2001.

80. Sheffield-Moore M, Urban RJ. An overview of the endocrinology of skeletal muscle. *Trends Endocrinol Metab* 15: 110–115, 2004.

81. Simental JA, Sar M, Lane MV, French FS, Wilson EM. Transcriptional activation and nuclear targeting signals of the human androgen receptor. *J Biol Chem* 266: 510–518, 1991.

82. Singh R, Artaza JN, Taylor WE, Gonzalez-Cadavid NF, Bhasin S. Androgens stimulate myogenic differentiation and inhibit adipogenesis in C3H 10T1/2 pluripotent cells through an androgen receptor-mediated pathway. *Endocrinology* 144: 5081–5088, 2003.

83. Sinha-Hikim I, Cornford M, Gaytan H, Lee ML, Bhasin S. Effects of testosterone supplementation on skeletal muscle fiber hypertrophy and satellite cells in community-dwelling older men. *J Clin Endocrinol Metab* 91: 3024–3033, 2006.

84. Sipila S, Poutamo J. Muscle performance, sex hormones and training in peri-menopausal and post-menopausal women. *Scand J Med Sci Sports* 13: 19–25, 2003.

85. Sipila S, Taaffe DR, Cheng S, Puolakka J, Toivanen J, Suominen H. Effects of hormone replacement therapy and high-impact physical exercise on skeletal muscle in post-menopausal women: a randomized placebo-controlled study. *Clin Sci (Lond)* 101: 147–157, 2001.

86. Smith FM, Holt LJ, Garfield AS, Charalambous M, Koumanov F, Perry M, Bazzani R, Sheardown SA, Hegarty BD, Lyons RJ, Cooney GJ, Daly RJ, Ward A. Mice with a disruption of the imprinted Grb10 gene exhibit altered body composition, glucose homeostasis, and insulin signaling during postnatal life. *Mol Cell Biol* 27: 5871–5886, 2007.

87. Smith GI, Yoshino J, Reeds DN, Bradley D, Burrows RE, Heisey HD, Moseley AC, Mittendorfer B. Testosterone and progesterone, but not estradiol, stimulate muscle protein synthesis in postmenopausal women. *J Clin Endocrinol Metab* 99: 256–265, 2014.

88. Spangenburg EE, Geiger PC, Leinwand LA, Lowe DA. Regulation of physiological and metabolic function of muscle by female sex steroids. *Med Sci Sports Exerc* 44: 1653–1662, 2012.

89. Staron RS, Hagerman FC, Hikida RS, Murray TF, Hostler DP, Crill MT, Ragg KE, Toma K. Fiber type composition of the vastus lateralis muscle of young men and women. *J Histochem Cytochem* 48: 623–629, 2000.

90. Suzuki S, Yamamuro T. Long-term effects of estrogen on rat skeletal muscle. *Exp Neurol* 87: 291–299, 1985.

91. Thornton AM, Zhao X, Weisleder N, Brotto LS, Bougoin S, Nosek TM, Reid M, Hardin B, Pan Z, Ma J, Parness J, Brotto M. Store-operated Ca^{2+} entry (SOCE) contributes to normal skeletal muscle contractility in young but not in aged skeletal muscle. *Aging (Milano)* 3: 621–634, 2011.

92. Tiidus PM. Influence of estrogen on skeletal muscle damage, inflammation, and repair. *Exerc Sport Sci Rev* 31: 40–44, 2003.

93. Vadaszova A, Hudecova S, Krizanova O, Soukup T. Levels of myosin heavy chain mRNA transcripts and protein isoforms in the fast extensor digitorum longus muscle of 7-month-old rats with chronic thyroid status alterations. *Physiol Res* 55: 707–710, 2006.

94. Velders M, Schleipen B, Fritzemeier KH, Zierau O, Diel P. Selective estrogen receptor-beta activation stimulates skeletal muscle growth and regeneration. *FASEB J* 26: 1909–1920, 2012.

95. Wattanapermpool J, Reiser PJ. Differential effects of ovariectomy on calcium activation of cardiac and soleus myofilaments. *Am J Physiol Heart Circ Physiol* 277: H467–H473, 1999.

96. Welle S, Tawil R, Thornton CA. Sex-related differences in gene expression in human skeletal muscle. *PLos One* 3: e1385, 2008.

97. Wiik A, Ekman M, Johansson O, Jansson E, Esbjornsson M. Expression of both oestrogen receptor alpha and beta in human skeletal muscle tissue. *Histochem Cell Biol* 131: 181–189, 2009.

98. Wiik A, Glenmark B, Ekman M, Esbjornsson-Liljedahl M, Johansson O, Bodin K, Enmark E, Jansson E. Oestrogen receptor beta is expressed in adult human skeletal muscle both at the mRNA and protein level. *Acta Physiol Scand* 179: 381–387, 2003.

99. Wiles CM, Young A, Jones DA, Edwards RH. Muscle relaxation rate, fibre-type composition and energy turnover in hyper- and hypo-thyroid patients. *Clin Sci (Lond)* 57: 375–384, 1979.

100. Yang LY, Arnold AP. Interaction of BDNF and testosterone in the regulation of adult perineal motoneurons. *J Neurobiol* 44: 308–319, 2000.

101. Yu F, Degens H, Li X, Larsson L. Gender- and age-related differences in the regulatory influence of thyroid hormone on the contractility and myosin composition of single rat soleus muscle fibres. *Pflügers Arch* 437: 21–30, 1998.

102. Zhu YS, Yen PM, Chin WW, Pfaff DW. Estrogen and thyroid hormone interaction on regulation of gene expression. *Proc Natl Acad Sci USA* 93: 12587–12592, 1996.

Sports Medicine (2021) 51:1401–1415
https://doi.org/10.1007/s40279-021-01451-8

REVIEW ARTICLE

Integrating Transwomen and Female Athletes with Differences of Sex Development (DSD) into Elite Competition: The FIMS 2021 Consensus Statement

Blair R. Hamilton[2,3] · Giscard Lima[1,4] · James Barrett[3] · Leighton Seal[3] · Alexander Kolliari-Turner[2] · Guan Wang[65] · Antonia Karanikolou[2] · Xavier Bigard[5,6,7] · Herbert Löllgen[6] · Petra Zupet[6] · Anca Ionescu[6] · Andre Debruyne[6,7] · Nigel Jones[8,9] · Karin Vonbank[10] · Federica Fagnani[4] · Chiara Fossati[4,11] · Maurizio Casasco[6,7,12] · Demitri Constantinou[7,13] · Bernd Wolfarth[7,14] · David Niederseer[15] · Andrew Bosch[16] · Borja Muniz-Pardos[17] · José Antonio Casajus[17] · Christian Schneider[7,18] · Sigmund Loland[19] · Michele Verroken[20,21] · Pedro Manonelles Marqueta[7,22] · Francisco Arroyo[7,23] · André Pedrinelli[7,24] · Konstantinos Natsis[6,7,25,26] · Evert Verhagen[27] · William O. Roberts[7,28] · José Kawazoe Lazzoli[7,29] · Rogerio Friedman[30] · Ali Erdogan[7,31] · Ana V. Cintron[7,32] · Shu-Hang Patrick Yung[7,33] · Dina C. Janse van Rensburg[34] · Dimakatso A. Ramagole[34] · Sandra Rozenstoka[6,7,35] · Felix Drummond[6,7,36] · Theodora Papadopoulou[6,7,37] · Paulette Y. O. Kumi[38] · Richard Twycross-Lewis[39] · Joanna Harper[40] · Vasileios Skiadas[41] · Jonathan Shurlock[42] · Kumpei Tanisawa[43] · Jane Seto[44,45] · Kathryn North[44,45] · Siddhartha S. Angadi[46] · Maria Jose Martinez-Patiño[47] · Mats Borjesson[7,48,49] · Luigi Di Luigi[7,50] · Michiko Dohi[7,51] · Jeroen Swart[7,52] · James Lee John Bilzon[7,53] · Victoriya Badtieva[7,54,55] · Irina Zelenkova[17] · Juergen M. Steinacker[6,7,56] · Norbert Bachl[6,7,57,58] · Fabio Pigozzi[4,6,7,11] · Michael Geistlinger[7,59] · Dimitrios G. Goulis[60] · Fergus Guppy[2,61] · Nick Webborn[62] · Bulent O. Yildiz[63] · Mike Miller[64] · Patrick Singleton[64] · Yannis P. Pitsiladis[1,2,4,6,7]

Accepted: 10 March 2021 / Published online: 24 March 2021

Abstract
Sport is historically designated by the binary categorization of male and female that conflicts with modern society. Sport's governing bodies should consider reviewing rules determining the eligibility of athletes in the female category as there may be lasting advantages of previously high testosterone concentrations for transwomen athletes and currently high testosterone concentrations in differences in sex development (DSD) athletes. The use of serum testosterone concentrations to regulate the inclusion of such athletes into the elite female category is currently the objective biomarker that is supported by most available scientific literature, but it has limitations due to the lack of sports performance data before, during or after testosterone suppression. Innovative research studies are needed to identify other biomarkers of testosterone sensitivity/responsiveness, including molecular tools to determine the functional status of androgen receptors. The scientific community also needs to conduct longitudinal studies with specific control groups to generate the biological and sports performance data for individual sports to inform the fair inclusion or exclusion of these athletes. Eligibility of each athlete to a sport-specific policy needs to be based on peer-reviewed scientific evidence made available to policymakers from all scientific communities. However, even the most evidence-based regulations are unlikely to eliminate all differences in performance between cisgender women with and without DSD and transwomen athletes. Any remaining advantage held by transwomen or DSD women could be considered as part of the athlete's unique makeup.

1 Introduction

Since antiquity, athletic and Olympic competitions have been separated according to the traditional binary concept of male/female to promote fairness and equity, as well as being divided by criteria such as weight, age, affiliation,

Extended author information available on the last page of the article

Key Points

The use of testosterone concentration limits of 5 nmol/L in transwomen and DSD women athletes is a justifiable threshold based on the best available scientific evidence.

There is a distinct lack of sports performance data to inform and update sports policy for DSD women and transwomen athletes.

Fair integration or exclusion of transwomen and DSD women athletes needs to be based on peer-reviewed experimental sporting performance evidence when such evidence becomes available.

amateur or professional status, and level of competition [1]. The binary classification of male and female was based on different methods, including physical examination (1966), Barr bodies (1968), Y chromosome (1991), and sex-determining region Y (*SRY*) gene (1996) [2]. A female athlete, when suspected to be male, could have been classified as either male or female depending on the previous methodology applied. For example, an individual with androgen insensitivity syndrome, with a 46, XY karyotype, would be classified as a female in 1966 and as a male in 1968, whereas an individual with congenital adrenal hyperplasia with a 46, XX karyotype, would be classified as male in 1966 and as a female in 1968. These examples illustrate such methods were unreliable, discriminatory, not fit for purpose, and that the integration of athletes outside of the binary of male and female is not a new problem (Table 1).

Integrating athletes who previously experienced male puberty into elite female sport is far from straight forward and remains highly contentious. For this reason, the concept of "*athletic gender*" was recently proposed [3, 4] which involves designating athletes to a gender for sports performance only and not social identity using quantitative criteria based on performance [3]. This concept speaks to a "*start over*" notion put forward by Maayan Sudai [5], who proposes the introduction of a classification system based on physiological parameters for athletes, regardless of gender. This would be analogous to the classification system used to assess eligibility to compete in Paralympic events [5]; however, the application of this would be very difficult for

Table 1 Summary of what is already known in this area, and future considerations in integrating transwomen and DSD women into elite women's sport

What is already known	Future considerations
The binary classification of athletes fails to consider differences in sex development (DSD) women and transwomen athletes	The use of testosterone concentration limits of 5 nmol/L in transwomen and DSD women athletes is a justifiable threshold. This level could be refined for specific events with the emergence of new supporting evidence
Testosterone production and action are the primary factors used in determining differences in performance between cis men and cis women	Any treatment is a purely personal and private decision and no sports body should provide recommendations on treatment
Only observational data showing the sporting performance of transwomen and DSD athletes exist	Fair integration of transwomen and DSD women athletes into elite sport needs to be based on peer-reviewed experimental evidence
Recent additions in the scientific literature including original studies provide the necessary impetus for the development of more evidence-based integration of DSD women and transwomen into elite competition	Any safety risks to cisgender female athletes due to the inclusion of transwomen in female elite sport must be evidence-based to justify exclusion
	The assumption that the physiology of elite DSD women and transwomen athletes is the same as elite male athletes is an oversimplified view
	New innovative scientific approaches are needed to guide new sports-specific policy (e.g., quantifying bioactive testosterone and individual sensitivity to testosterone, the role of sex chromosomes in athletic performance, and the extent to which muscle memory is retained after prolonged high testosterone exposure)
	There is a distinct lack of sports performance data to inform and update sports policy, in part due to the lack of funding and lack of elite athletic participants in this research area
	The participation of transwomen and DSD women elite athletes in research will be hindered by their low numbers in elite competition. Recruitment for research may have to be targeted also at the sub-elite level with the specific requirement of being an athlete at higher than grassroots level
	The need to develop approaches to distinguish between predisposition to outstanding performances (e.g., haematological and anatomical features) and any unfair advantages held by transwomen or DSD women

sport's governing bodies due to its complexity and financial commitment to implement at all levels of sport.

The concept of athletic gender could help safeguard fair competition and prevent an unfair advantage, principles which underpin the true essence of sport [6], and would be in line with the fundamental principles of the Olympic Charter which emphasizes the need to respect the freedom and rights of athletes, as well as the importance of competing without any form of discrimination. The Olympic Charter states that "*The enjoyment of the rights and freedoms set forth in this Olympic Charter shall be secured without discrimination of any kind, such as race, colour, sex, sexual orientation, language, religion, political or other opinion, national or social origin, property, birth or other status*" [7] and importantly refers to sex and not gender. Sex is considered in Olympic sports only when it could determine the outcome of a competition. Some sports do not use a sex classification, e.g. shooting, sailing, or horse riding.

The terms "*sex*" and "*gender*" have different meanings and their overlap is conceptually complex. *Sex* refers to any individual's biology, such as anatomical or chromosomal differences, which are used to categorize an individual as male or female, whereas gender refers to socially constructed roles related to sex distinctions [8]. While gender identity is a self-defined social construct that shapes how an individual chooses to live, gender identity alone will not be enough to determine the appropriate sports category for each individual that allows fair competition, especially in the case of elite sport.

The current article aims to highlight the main issues to be considered surrounding the participation of female athletes with previously high testosterone concentrations (transwomen) and female athletes with naturally high testosterone concentrations [differences in sex development (DSD)] in elite female sport. The two cases, cisgender women athletes with DSD (DSD women, for short) and transwomen athletes, will be presented separately to enhance reader understanding, while future research considerations will be discussed together in Sect. 5, because the considerations for both groups of athletes are similar. It is important to note that the fluidity of gender identity does include non-binary and transmen athletes. However, in this article, the authors wish to focus on the integration of DSD women and transwomen athletes into the elite female category of sports. The reasoning for this is that transmen (birth-assigned female transitioned to male) athletes are perceived to not have the same magnitude of competitive advantage as transwomen or DSD women athletes when integrated into male elite sports [9] and that non-binary individuals are less likely to undertake gender-affirming treatment and are predominantly female sex assigned at birth [10], forgoing the effects of male puberty.

2 Methods

Here, we present the International Federation of Sports Medicine (FIMS) consensus on integrating DSD women and transwomen athletes into elite female sport based on identifying, selecting, and critically appraising the very limited relevant primary research. An added objective of this consensus was to provide a roadmap for future research direction. The review of the evidence was performed by the first and second author (BH and GL) using the following keywords: "transgender" or "transwomen", "intersex" or "DSD", "gender identity", "testosterone", "competition", and "sport". The first draft of the manuscript was written by the first and last authors (BH and YP). Of all 78 invited authors, 1 author declined the invitation and 7 authors elected to withdraw their names during one of the draft rounds. These names are not included on the authorship list above. All remaining 70 authors reviewed, commented on and approved the final draft. The drafting of the consensus statement was initiated by the last author (YP) via email for ease of verification and process during the unprecedented constraints due to the COVID-19 pandemic. Voting on the consensus statements was performed remotely using Google Forms (Google™, California, USA). The voting result was collated by the first author (BH) along with dissenting opinions and discussions which were manifested and reported in the manuscript. All statements received unanimous approval by all named authors except for the statement on the testosterone limit of 5 nmol/L, which received majority approval and the voting result is included in the article. The authors consider it essential to declare the extent of agreement, as well as dissenting views.

3 DSD Women Athletes

3.1 Background

DSD is a group of rare conditions involving genes, hormones and reproductive organs [11, 12]. This article will focus on the integration of DSD women athletes in the elite female category of sports who currently have high testosterone concentrations which the binary classification of sports fails to consider. The German Federal Parliament approved a law that came into effect in December 2018 that permits children with DSD born with ambiguous sexual anatomy who are not distinctly male or female to indicate a third gender category on their birth certificate [13]. This action follows a court ruling by the Federal Constitutional Court of Germany in October 2017 that ruled the existing regulations discriminated against people with DSD, the principle being that the gender identity of an individual must be protected as a fundamental human right [14].

The views of the Court of Arbitration for Sport (CAS) have evolved concerning legal sex being a factor to determine the eligibility of an athlete to compete in a male or female category. In the Dutee Chand vs. the Athletics Federation of India and the International Association of Athletics Federations (IAAF) arbitration tribunal in 2014, CAS stated in their decision that "*The distinction between male and female is a matter of legal recognition*" [15]. In contrast, in the Caster Semenya and Athletics South Africa vs. IAAF tribunal in 2018, CAS stated that "*a person's legal sex alone may not always constitute a fair and effective means of making that determination*" [16]. The Human Rights Council under the United Nations recently released a statement on discrimination against women in sport [17]. While not limited to discrimination concerning DSD women and androgen sensitivity, the position taken is that both member and non-member states of the United Nations should work in unison to recognize protected characteristics and eliminate discrimination.

3.2 The Challenge

Conditions such as DSD are rare and primarily of genetic origin [18] and are presented concomitantly with ambiguous genitalia at birth which can occur phenotypically in undervirilized genotypic males or virilized genotypic females. These features can result in individuals assigned female at birth possessing testosterone concentrations comparable to cisgender males and, therefore, much higher than non-DSD women, including those with polycystic ovary syndrome [19]. Hyperandrogenic 46, XY DSD female athletes in the 2011 IAAF World Championships were 140 times more prevalent (0.7% of athletes had testosterone concentrations of > 15.6 nmol/L [20]) compared to 0.005% [20, 21] reported in the general population [22, 23], which could be an indicator of performance advantage [20]. A possible indicator of fair integration of DSD women athletes into competitive sport would be a similar prevalence of DSD women and non-DSD women athletes in the championships as in the general population.

The DSD condition is a natural attribute as opposed to a doping issue, such as the misuse of anabolic steroids. However, observational data have shown a clear difference in performance in DSD women athletes depending on whether testosterone concentrations were suppressed or not. For example, there was an average performance reduction of approximately 5.7% in the best performances of three female distance runners who had their testosterone concentrations suppressed from 21–25 to 2 nmol/L over 2 years [23]. Although a notable finding, no firm conclusions can be reached due to the reliance on a small number of athletes. Within DSD women athletes, there are individuals with 46, XY karyotype, and androgen insensitivity, which can be either complete androgen insensitivity (CAIS) or partial androgen insensitivity (PAIS). Therefore, testosterone concentrations in such individuals will not have the same functional effect as those with normal androgen receptors. This complexity needs to be considered if testosterone concentration, either as a single parameter or more likely as one of several parameters, will evolve into a viable solution.

3.3 The Present Rulings in Elite Sport

Following an observational study by Bermon and Garnier describing the serum androgen levels of male and female athletes and their relationship to performance in track and field events [24], the eligibility regulations for the female classification were created and published by the IAAF (now World Athletics) in April 2018. Implementation of the policy was planned for November 2018 [25]. However, this study [24] and the subsequent regulations have been subject to much debate [26–30]. The IAAF regulations permitted female athletes with specific DSD's (i.e., testosterone concentrations ≥ 5 nmol/L and sufficient sensitivity to androgens) to compete in international competitions in the female category from 400 up to 1500 m if they reduced testosterone concentrations to < 5 nmol/L for at least 6 continuous months. These requirements needed to be maintained for the athlete to continue to be eligible for the female category of the events described in the regulation.

Considering a challenge brought by Caster Semenya against these regulations, the IAAF agreed to delay the implementation and await the decision from CAS. The panel's decision was released in May 2019, with the statement that the "*Panel has dismissed the requests for arbitration considering that the Claimants were unable to establish that the DSD Regulations were invalid*" [31]. Semenya appealed to Switzerland's Federal Supreme Court, which suspended the implementation of the eligibility regulation in June 2019. However, Semenya ultimately lost her appeal [32] in August 2020 and the eligibility regulations were reinstated with the court citing that "*fairness in sport is a legitimate concern and forms a central principle of sporting competition*" [33].

In addition to the media frenzy both for and against the inclusion of DSD women athletes in the female category of sports [34], editorials have been published sparking subsequent critiques and rebuttals in response [32, 35]. This fervour has also sparked academic and general community outrage at the IAAF ruling, which has been declared as discriminatory against Semenya. Idiosyncratically, the emotional and legal argument is that Semenya is being victimized and unfairly treated as a female athlete, yet her sex is not biologically clearly defined in the male/female binary definition. This case is an inevitable consequence of the antithesis between the binary concept of gender applied to

sport and the new realm of gender fluidity, as illustrated by DSD women athletes.

World Athletics in their most recent version of the eligibility regulations for the female classification (athletes with DSD), state that not all DSD women athletes who wish to compete in the female classification should need to reduce their testosterone levels to < 5 nmol/L. They state that: "*A woman who has androgen insensitivity syndrome (AIS) is completely (CAIS) or partially (PAIS) insensitive to testosterone, thereby eliminating (CAIS) or reducing (PAIS) the physiological effect of that testosterone. An athlete with CAIS is not a Relevant Athlete. An athlete with PAIS will only be a Relevant Athlete if she is sufficiently androgen-sensitive for her elevated testosterone concentrations to have a material androgenising effect. The benefit of any doubt on this issue will be resolved in favour of the athlete*" [31].

4 Transwomen Athletes

4.1 Background

Transgender refers to a gender expression that is different from the sex that is assigned at birth. In this article, a specific focus will be placed on transwomen, assigned male at birth who have transitioned to female both socially and legally and have had previous exposure to high testosterone concentrations during puberty. Recently, a controversial bill (i.e., 2019 Tennessee SB2077) prohibiting the participation of transwomen athletes in school sports was introduced in the U.S. legislature. Should this bill pass into law, a burden would be placed on education providers to ensure pupils participate according to the biological sex indicated on their birth certificate. Additionally, the bill seeks to impose a civil penalty of US $10,000 as well as the revocation of public funds for any school that acts contrary to the bill [36].

In March 2020, a second similarly controversial bill (i.e., the 2020 State of Idaho HB500) known as the "*Fairness in Women's Sports Act*", was passed into state law making Idaho the first state to ban transwomen from participating in girls and women's sports [37]. The bill states that "*Athletic teams or sports designated for females, women, or girls shall not be open to students of the male sex*" [38] and that if a student's sex is disputed, "*a student may establish sex by presenting a signed physician's statement that shall indicate the student's sex solely on: (a) the student's internal and external reproductive anatomy; (b) the student's normal endogenously produced levels of testosterone; and (c) an analysis of the student's genetic makeup*" [38].

The bill received praise from the Senior Vice President of U.S. Legal Division, citing that "*Allowing males to compete in girls' sports destroys fair competition and women's athletic opportunities*" [39] but has drawn criticism given the Act conflicts with the right to privacy provision within the 4th Amendment to the American Constitution. Indeed, this has formed the basis of a claim brought by the American Civil Liberties Union against the State of Idaho regarding the legitimacy of the Act [40], alleging that the legislation could violate the federal law known as Title IX which prohibits sex discrimination, not gender discrimination, in educational institutions that receive federal financing [41]. This kind of legislation will inevitably result in tension between domestic law and international treaties developed to promote inclusivity and protect individuals from discrimination based on protected characteristics.

4.2 The Challenge

Although permitted by the IOC since 2004, no recognized transgender athlete has participated in the Olympic Games [42]. The main argument opposing the integration of transwomen athletes into the female category for future Olympics is the perceived sporting advantages that transwomen have over cisgender women, such as lever length or height advantages conferred by skeletal size and bone density despite testosterone reductions [43]. Prior athletic training with high testosterone concentrations may potentially result in advantages such as muscle memory [44], which may persist for some time post testosterone suppression. This is a concern for sports highly dependent on muscle mass, strength, and aerobic capacity. This will be expanded on in Sect. 5.5.

Despite these concerns, evidence on transwomen's sporting performance is scarce (Table 1) and in the case of aerobic performance, non-existent. Couple this with the data already showing that oxygen-carrying haemoglobin levels are reduced in transwomen to female norms levels [45], it is a sports performance proxy that is urgently needing investigation due to the importance of the cardiovascular system during aerobic exercise. Low testosterone concentrations have been reported in transwomen undergoing hormone replacement therapy (HRT) [46] and in a recent meta-analysis, HRT was found not to affect the motor coordination or visuospatial abilities of transwomen [47]. In a study of 50 non-athlete transwomen who had undergone gender-affirming surgery (GAS) coupled with HRT, a reduction in muscle mass and bone mineral density was reported together with an increase in fat mass following HRT initially and 1 year after GAS [48]. These data on non-athletic transwomen and non-sports performance measures make it difficult to suggest that the athletic capabilities of transwomen individuals undergoing HRT or GAS are comparable to those of cisgender women and because of this, the recording of data describing transwomen's sporting performance should be of the highest importance to sporting governing bodies and researchers.

While data on transwomen's athletic performance remain to be experimentally determined, a first retrospective study did evaluate the performances of eight non-elite transwoman masters athletes who had participated in running competitions, first as males and then as females [49]. Running performance was compared using a standard age grading methodology [age grade (%) = age standard × 100/race time] for comparing groups of athletes of any age and gender in track-and-field and distance running [49, 50]. Overall, the group of athletes obtained similar "age-graded" scores in both categories. However, the design of the study may limit its relevance given the small sample size, no reporting of testosterone levels, self-reported run times, no reporting of when the participants ran after their transition, the athletes were not elite, and the findings of this study have not been replicated.

A review paper by Hilton and Lundberg [43] addressed the integration of transwomen in the elite female category of sport. The authors concluded that anthropometric and muscle mass advantages are sustained in transwomen after 12 months of gender-affirming treatment based on studies showing the physiological changes caused by HRT in transwomen and chemical castration in men. Conversely, due to these studies being conducted in non-athletic transgender women, they also concluded that "*it is still uncertain how transgender women athletes, perhaps undergoing advanced training regimens to counteract the muscle loss during the therapy, would respond*" [43].

Despite the lack of direct sport-specific studies of transgender athletes in their review, Hilton and Lundberg raised safety as their primary concern and proposed that 12 months of testosterone suppression is insufficient to mitigate their safety concerns [43]. However, the main criticism of this review is the purely biological argument from an elite male versus elite female position, implying that transwomen athletes are the same as elite male athletes (Table 1). Data showing lower baseline isometric torque and muscle volume [51] in transwomen compared to cisgender males highlight the problematic nature of inferring that transwomen and cisgender males are the same, as this ignores the impact of gender-affirming treatments such as HRT and GAS and the psychological effects of gender dysphoria such as low self-esteem, anxiety and/or depression, and becoming socially isolated [52].

Recently, Roberts et al. [53], retrospectively reviewed pre- and post-HRT military fitness test results in transwomen individuals ($n = 46$) of the U.S. Air Force. These authors found that the push-up (31% more than their female counterparts) and sit-up (15% more than their female counterparts) advantages over ciswomen at baseline had been negated after 2 years, but not after 1 year. This finding agrees with previous studies that have shown that baseline muscular strength in transwomen is not significantly diminished after 1 year [51, 53] but is after 2 years of HRT [53]. Roberts et al. also found that running performance in the 1.5 mile run remained 12% faster on average in transwomen after 2 years of HRT [53]. These findings require replication in trained transwomen athletes, although they would suggest a different rate and extent of mitigation of the advantages held by transwomen given that the strength advantages, but not the cardiovascular advantages, of transwomen were mitigated after 2 years of HRT. These observations also question the required testosterone suppression time of 12 months for transwomen to be eligible to compete in women's sport, as most advantages over ciswomen were not negated after 12 months of HRT. How applicable these performance data are from both Harper [49] and Roberts et al. [53] in determining the extent of advantage remaining in transwomen athletes post-gender-affirming treatment remains to be determined. This will require longitudinal transgender athlete case-comparison studies that control for variations in hormonal exposure and involve numerous indices of performance (Table 1).

4.3 The Present Rulings in Elite Sport

The participation of transgender athletes in the Olympic Games was approved following the *2003 Stockholm Consensus on Sex Reassignment in Sports*, which recommended that transwomen athletes undergoing sex reassignment after puberty be eligible for competition 2 years post-gonadectomy, HRT, and legal recognition of assigned sex [42]. The IOC released one update of the recommendations in 2015 [54]. Most sports governing bodies adopted this policy, declaring the eligibility of transwomen athletes with serum testosterone concentrations < 10 nmol/L for at least 12 months before the first competition and throughout the competition period. There was also no requirement for surgical procedures for any anatomical changes. World Athletics [55], World Rowing [56] and Union Cycliste Internationale (UCI) [57] have all adopted the lower serum testosterone concentration limit of 5 nmol/L for transwomen athletes. Some would consider a 5 nmol/L limit high, as healthy premenopausal women typically have a testosterone concentration < 5 nmol/L (e.g., < 1.7 nmol/L) [19]). The support for the < 5 nmol/L limit (Table 1) for transwomen athletes emerges from a study where 24 healthy, physically active women aged 18–35 years underwent 10 weeks of testosterone treatment [22]. This study reported improved running time to exhaustion during an incremental maximal test on a treadmill by 21.17 s (8.5%) and an increase in lean body mass. However, the average testosterone concentrations of these participants did not exceed 5 nmol/L (from 0.9 ± 0.4 to 4.3 ± 2.8 nmol/L) [22], which is considerably below the 10 nmol/L threshold used by the IOC [54].

World Rugby became the first international sports governing body to ban the participation of transwomen in the elite female level of sport in October 2020. They state that "*Transgender women may not currently play women's rugby because of the size, force- and power producing advantages conferred by testosterone during puberty and adolescence, and the resultant player welfare risks this creates*" [58]. The policy, by its admission, is based on a "*hypothetical cross-over scenario in which a typical male tackler mass is involved in a tackle against a ball carrier with a typical female mass*" [58]. The policy itself speaks to the "common sense" view that transwomen athletes are larger and stronger than their cisgender peers, which mischaracterises transwomen athletes as elite male athletes (Table 1) and has been opposed by rugby unions such as the USA and Canada. England Rugby will also not implement the policy stating to the media that it "*believes further scientific evidence is required alongside detailed consideration of less restrictive measures in relation to the eligibility of transgender players*" [59]. World Rugby's ruling is a prominent polarising example of the need for sports-specific performance data for transwomen athletes.

5 Future Research Considerations

5.1 Testosterone as the Primary Biomarker for Eligibility

Despite being imperfect, serum testosterone concentrations are being considered as the primary biomarker to regulate the inclusion of athletes into the female category. At this time, it is the only method based on an objective biomarker supported by most available scientific literature (Table 1), while also accomplishing the integration of DSD women athletes and transwomen athletes into the female category of sports. This is consistent with the fundamental principles of the Olympic Charter and is an attempt to be fair to all participants by ensuring an equitable competitive environment. However, many unresolved issues need clarification before unreservedly adopting testosterone concentration, or any biomarkers, to define "*athletic gender*" [3]. Resolving these issues will require the scientific and sports medicine community to employ innovative research ideas [e.g., a combination of cell, animal, and human research paradigms (Table 1)] to generate the biological data needed to inform the inclusion or exclusion of transwomen and DSD women athletes in elite female sports.

Areas of research focus could include better methods for quantifying bioavailable testosterone, also known as free testosterone, as a potentially better alternative to total circulating testosterone as a criterion for participation in the female category of sports. Bioavailable testosterone is the testosterone that is taken up and used by the body's cells and could be measured in conjunction with an allowance for androgen insensitivity [3]. An increase in bioavailable testosterone over time seems to induce a greater increase in muscle mass and strength [60], although this finding has been recently disputed [61]. In contrast, when bioavailable testosterone was reduced to castrate levels in young men, isometric strength did not increase after resistance exercise training [62]. Assuming these findings are replicated and if extrapolated to elite DSD women athletes and transwomen athletes, they would imply that decreasing bioavailable testosterone concentrations would mitigate to some extent any previous sporting advantage due to the previously high testosterone concentrations. This is a particularly encouraging future avenue of research.

The role of testosterone in muscle anabolism (i.e., tissue growth, substrate restoration, and recovery) and catabolism (i.e., tissue breakdown and metabolic regulation) is well described [63] and, therefore, could be another avenue of research. The hypothesis is that the low testosterone concentrations induced in transwomen or DSD women will impact negatively on muscle performance and recovery. Therefore, it is essential that researchers replicate or determine the precise time frame, individual variability, and mechanism(s) of this drop off in strength with HRT in trained athletes.

5.2 Genetics

Another pertinent issue is genetic factors (i.e., sex chromosome composition) in influencing athletic performance. Boys and girls demonstrate differences in a range of physical characteristics, including body composition and skinfold thickness [64], height, and explosive strength, even before puberty [65], suggesting that sex chromosome composition plays a role in determining differences in adult athletic performance. Consistent with this, different populations of muscle cells may express different phenotypes of androgen sensitivity, raising the possibility that the muscle response to training may be different between men and women at the same testosterone concentrations. Animal model studies are a feasible option to examine the influences of sex chromosomes and pubertal hormones. For example, the four core genotypes mouse model which incorporates mice with four different combinations of gonads and sex chromosomes [66, 67], has helped identify the influence of sex chromosomes on physical traits, such as obesity and food intake [68, 69]. This model represents an ideal opportunity to study muscle function in the present context as the different combinations of gonads and sex chromosomes will result in different testosterone concentrations. This model may highlight the true effect of testosterone on muscle function.

5.3 Androgen Receptor Function

Elucidating further androgen receptor function is another relevant future avenue of research. Androgen receptors can be modulated by specific proteins called coregulators [70–72] or mediated via the activation of membrane-bound protein receptors to initiate intracellular signalling pathways [73], which can occur even in the presence of low levels of androgens [74]. Investigations into the non-genomic actions of the androgen receptor have been limited to in vitro studies [75, 76] rather than in vivo due to the lack of an appropriate animal model that can distinguish between genomic and non-genomic receptor actions [75]. Androgen receptor knockout mice such as DBD-ARKO [40], which has a deletion of the second zinc finger of the DNA-binding domain, has been created for such research purposes. Given the inherent challenges of human studies, investigators need to adopt similar creative approaches if they are to elucidate the role of androgen receptors in elite DSD women and transwomen athletes.

5.4 Athlete Health

It is important to note that the World Medical Association has urged physicians not to implement the World Athletics policy on classifying women athletes, arguing that the policy is not in line with medical ethics and could be harmful to the athlete [77]. This argument is an outdated approach to protect the privacy of patients. If the athlete is fully informed of the consequences of treatment and not coerced into undergoing treatment, the athlete has free choice to do so (Table 1), which is a fundamental human right [32]. However, when the sex of an athlete is challenged or uncertain, eligibility would need to be determined for women's events. Such a concept to request eligibility is currently being implemented by World Rowing [56]. The justification is that it is ethical and may be necessary for a medical doctor to assist an athlete in determining their eligibility for a sex-restricted event. This requirement is not about treatment and treatment choices, which are always private and not relevant to the sports community. This process is essential to ensure all athletes, including transwomen and DSD women athletes, can compete on an even playing field with cisgender athletes, and currently, as the best proxy, transgender athletes have to demonstrate testosterone concentrations in a similar range to those athletes they wish to compete against. The eligibility of DSD women athletes must not only follow the same principles based on testosterone concentrations, but also needs to consider testosterone receptor function.

The health of athletes should be the number one priority of any sport, and it is clear that World Rugby's new transwomen exclusion policy [58] has the health of athletes at the heart of its policy. However, such exclusion policies should be based on generally accepted scientific consensus, including results from studies conducted in transwomen athletes. The authors of the World Rugby guidelines may be correct in their assumptions using hypothetical modelling of elite male versus elite female athletes [58]; however, until relevant transwomen athletic performance data become available, there is just as much circumstantial evidence to support this policy by World Rugby as there is to oppose it. For example, a study of young untrained women with polycystic ovary syndrome found greater muscle mass did not equate to greater peak muscle force [78]. There is an urgent need, therefore, for well-designed longitudinal studies throughout a transwomen's transition that assesses at regular intervals the main indices of performance relevant to all sports. Such data will prove invaluable to directly evaluate the true safety risks inherent in transwomen playing in the elite female category of sport.

5.5 Muscle Memory

Muscle memory refers to the persistence of cellular phenotype related to previous testosterone exposure [79]. Research shows that in addition to hormone concentrations, the number of myonuclei can affect skeletal muscle training [79, 80]. Indeed, muscle cells have multiple nuclei and their number increases with muscle hypertrophy [81, 82]. In female mice, short-term treatment with testosterone increased both muscle fibre cross-sectional area (CSA) and myonuclei number [79]. After cessation of exposure, muscle fibre CSA reverted to that of the control arm, but the number of myonuclei remained 42% higher than controls for at least 3 months. These resident myonuclei facilitated enhanced muscle hypertrophy during 6-day resistance training overload (31% increase in the fibre CSA vs. 6% in controls); this increase remained 20% higher compared with controls after 14-day overload [79]. The number of myonuclei not only reflects the current size of the fibre, but also the history of the fibre. Current data might fit a "*peak pegging*" hypothesis, where the number of myonuclei found in the fibre represents the largest size the fibre has achieved, and new myonuclei are only added if the fibre grows beyond that size. However, this "*peak pegging*" hypothesis found in female mice does not transfer to young healthy, physically active women. Horwath et al. showed no change in the myonuclei content following a 10-week testosterone administration of 10 mg daily protocol [83] coupled with an interesting finding of a 31% increase in satellite cells associated with type II fibres in the testosterone group. Satellite cells exit quiescence by extrinsic mechanical stretch to the fibre, generating differentiated cells and self-renewing stem cells by asymmetric division

[84], meaning that the myofibres could feasibly repair more quickly with exogenous testosterone administration.

Testosterone has been shown to increase the myonuclear number in men in a dose-dependent manner alongside muscle fibre CSA being well correlated with the myonuclear number [81, 82]. Nevertheless, further data are needed to confirm the extent to which myonuclei are retained over time after human muscle fibres have been exposed to a high testosterone environment. If high numbers of myonuclei are confirmed to be retained in transwomen or DSD women athletes, these results could imply that an advantage of previously high testosterone concentrations remains even after testosterone suppression. The relevant question would remain whether this potential effect is relevant to regulations that seek to prohibit individuals who have this potential advantage from competition.

5.6 Previous Failings Present Opportunity

Finally, it is important to stress that the current physiological data are insufficient to adequately inform policy and result from both a distinct lack of research funding and a limited number of elite athletes available to participate in this research area. For eligibility to be determined in the fairest manner possible, more funding and subsequent research are required to allow specialists in biological sciences and sports medicine to conduct experiments to determine the best solutions for integrating DSD women and transwomen athletes into the elite level of female sport.

5.7 FIMS Consensus Statements for the Integration of DSD Women and Transwomen Athletes into Elite Female Sport

Although serum testosterone concentrations constitute an indicator of androgen production and availability, a reliable biological index of androgen action is still lacking. Promising new developments in sport and exercise science are destined to contribute to the fair inclusion of DSD women and transwomen athletes. A well-coordinated multidisciplinary international research approach should include well-designed, controlled studies on the effect of testosterone on training and sports performance. Providing scientific evidence to use a system of biology multi-omics adequately and ethically (i.e., genomics, transcriptomics, metabolomics, and proteomics) to generate the necessary data and downstream biomarkers will be needed to address all open issues. There must be a transparent roadmap for the scientific community to focus on the best possible outcome of such new research. The authors, therefore, propose the following FIMS consensus statements and roadmap to facilitate the integration of DSD women and transwomen athletes into elite female sport:

- The inclusion of a third category in elite sport is not currently plausible, as the numbers of elite DSD women and transwomen athletes are relatively small.
- The prevalence of transwomen athletes in elite competition is likely to increase in the future, due to the increased visibility of transgender individuals in society [85, 86], which in turn may drive more people to consider expressing their chosen gender identity [87]. Research into transwomen sporting performance is highly relevant for leading scientists, leading clinicians, sport's governing bodies, and the World Anti-Doping Agency and is already a priority for the IOC [88].
- Transwomen have the right to compete in sports. However, cisgender women have the right to compete in a protected category.
- Any inclusion or exclusion policies on DSD women and/or transwomen athletes should be free of any social and/or religious prejudice, bias, or discrimination and should be based solely on the governance of fair competition.
- As each sport can vary greatly in terms of physiological demands, we support the view held also by others [43] stating that individual sport's governing bodies should develop their own individual policies based on broader guidelines developed on the best available scientific evidence, determined experimentally from a variety of sources with a particular preference for studies on transwomen and DSD women athletes.
- With data showing reductions in haemoglobin following testosterone suppression [45], obtaining data on DSD women and transwomen athletes' cardiovascular performance, such as maximal oxygen uptake, should be a priority for researchers due to the importance of the cardiovascular system in numerous sports performance contexts.
- The use of serum testosterone concentrations as the primary biomarker to regulate the inclusion of athletes into male and female categories is currently the most justified solution as it is supported by the available scientific literature (Table 1) and should be implemented at the elite level, where there is an emphasis on performance enhancement.
- DSD women or transwomen athletes should be fully informed by medical personnel of the risks and consequences of testosterone suppression treatment and must never be coerced or forced into testosterone suppression. The athletes must be free to make the decision that is best for them (Table 1).
- No sport's governing body should provide recommendations on treatment; this should be done by medical personnel (Table 1).
- If DSD women and transwomen athletes choose not to have suppressed testosterone, as is their right, they cannot compete in the restricted female category with high

testosterone concentrations above the policy threshold. Instead, they should be offered the chance to compete in the male category.

- A testosterone concentration threshold of 5 nmol/L in DSD women and transwomen athletes should be used as a global recommendation for sport's governing bodies at this present time and may be modified as new evidence arises for an event or sport-specific concentrations (Table 1).
- The statement on the testosterone concentration threshold for transwomen and DSD women athletes was the only point of contention for the FIMS Panel. All 70 authors voted, of whom 87% were in favour of the 5 nmol/L threshold, 2% of authors were in favour of a threshold of 8 nmol/L, 2% were in favour of a threshold around the upper testosterone concentration of normal healthy females of 0.2–1.7 nmol/L [89], and 8% of authors were in favour of no change to the limit until further evidence was acquired. This large but not unanimous majority consensus highlights the area most in need of research, i.e., altered bioavailability of testosterone and performance indices in DSD women and transwomen athletes.
- New innovative avenues must be explored to guide improved, up to date policy (Table 1), for example, quantifying bioactive testosterone and individual sensitivity to testosterone, the role of sex chromosomes in athletic performance, the role of androgen receptors, and the extent to which muscle memory is retained after high testosterone exposure. In addition, identification of other biomarkers (e.g., metabolomics, proteomics) that may better differentiate individual sensitivity to testosterone is needed. Liquid chromatography–mass spectrometry is well accepted as the preferred technique for the analysis of testosterone [90, 91].
- The best available scientific methods, such as well-designed, controlled studies, must be utilised to acquire new scientific evidence on sporting performance measures to derive policies on DSD women and/or transwomen sporting participation. This should be on a sport-by-sport basis when the evidence arises, rather than the universal approach to sports regulations at present due to the lack of individual sports data.

5.8 Dissenting Opinions During Consensus Discussions

During the consensus discussions, there was a constructive debate on the testosterone limit in the elite category of female sports. One author agreed that the concentration of 5 nmol/L was a median value between the upper and lower ranges of female and male testosterone. However, the 5 nmol/L level adopted by World Athletics is based on the inference that there is a relationship between performance and testosterone concentrations and is meant to represent the value above which a performance advantage is no longer within the bounds of healthy cisgender females. This assumption is likely false due to the multifactorial nature of different sports. Although there is evidence to suggest that performance of female athletes with high testosterone levels may be enhanced, it is still a contentious issue that requires research before and after testosterone suppression to identify where the testosterone threshold should be set for such athletes, and the limit may have to follow a sport-by-sport evidenced basis instead of a holistic approach.

The authors also discussed the issue of athletes' health, which is timely given the announcement of World Rugby's transgender guideline which excludes transwomen players to safeguard cisgender female players at the international level. One author opposed the "one size does not fit all" notion of World Rugby's policy due to its assumption that all transwomen are larger in stature and heavier than their cisgender counterparts. This assumption is due to studies like Roberts et al. showing that transwomen are heavier when presented as a pre-treatment average [53]. However, some cisgender women athletes are taller than transwomen or have greater muscle mass than transwomen and anthropometric variation is a part of sport. If the modelling scenario in World Rugby's policy of a "*typical male tackler mass*" involved in a rugby tackle with a "*typical female tackler mass*" [58] is confirmed, an exclusion policy could be implemented on an individual basis and resolving all the practical challenges that this would entail. Safety in sport is of great importance and exclusion based on safety is a justifiable cause but exclusion needs to be evidenced-based and include some consideration of transwomen athletic performance metrics.

Another author strongly affirmed that all cut-offs for hormones that are out of normal ranges for age and/or gender are pathological, not physiological, and are associated with different side effects, some of them increasing health risks and some potentially useful at different levels for physiological performance. The author stated, "*that as sports physicians we have to decide if firstly, we protect athlete's health issues or social issues*" and that sport physicians should mimic society's physicians and be "*a cornerstone for athletes health*".

6 Conclusions

Ultimately, even the most evidence-based policies will not eliminate differences in sporting performance between athletes in the elite category of female sports. However, any advantage held by a person belonging to an athlete in this category could be considered part of the athlete's unique individuality. Whatever the solution, there is an urgent need for a well-coordinated multidisciplinary international

research program, backed by appropriate research grant funding and athlete participation, to generate the evidence to inform future objective policy decisions. Such decisions should be based on the best available scientific evidence from the best available scientific practice and the decisions made will also require a firm political resolve to fairly integrate transwomen and DSD women athletes into elite female sport.

Declarations

Funding This manuscript was not funded.

Conflict of interest/competing interst Blair Hamilton, Gicard Lima, James Barrett, Leighton Seal, Alexander Kolliari-Turner, Guan Wang, Antonia Karanikolou, Xavier Bigard, Herbert Lollgen, Petra Zupet, Anca Ionescu, Andre Debruyne, Nigel Jones, Karin Vonbank, Federica Fagnani, Chiara Fossati, Maurizio Casasco, Demitri Constantinou, Bernd Wolfarth, David Niederseer, Andrew Bosch, Borja Muniz, Jose Casajus, Christian Schneider, Sigmund Loland, Michele Verroken, Pedro Manonelles Marqueta, Francisco Arroyo, André Pedrinelli, Kostas Natsis, Evert Verhagen, William Roberts, José Kawazoe Lazzoli, Rogerio Friedman, Ali Erdoğan, Ana Cintron, Patrick Yung, Christa Janse van Rensburg, Dimakatso Ramagole, Sandra Rozenstoka, Felix Drummond, Theodora Papadopoulou, Paulette Kumi, Richard Twycross-Lewis, Joanna Harper, Vasileios Skiadas, Jonathan Shurlock, Kumpei Tanisawa, Jane Seto, Kathryn North, Siddhartha Angadi, Maria Patino, Mats Borjesson, Luigi Di Luigi, Michiko Dohi, Jeroen Swart, James Bilzon, Victoriya Badtieva, Irina Zelenkova, Juergen Steinacker, Norbert Bachl, Fabio Pigozzi, Michael Geistlinger, Dimitrios Goulis, Fergus Guppy, Nick Webborn, Bulent Yildiz, Mike Miller, Patrick Singleton and Yannis Pitsiladis declare that they have no conflicts of interest relevant to the content of this review.

Ethics approval Not applicable

Consent to participate Not applicable

Consent for publication Not applicable

Availability of data and material Not applicable

Code availability Not applicable

Author Contributions Yannis Pitsiladis was responsible for the manuscript conception. The first draft of the manuscript was written by BH and YP and all authors commented on subsequent versions of the manuscript. All authors read and approved the final manuscript.

References

1. Archives TN. Equality Act 2010. http://www.legislation.gov.uk/ukpga/2010/15/section/195. Accessed 4 Mar 2020.
2. Sánchez FJ, Martínez-Patiño MJ, Vilain E. The new policy on hyperandrogenism in elite female athletes is not about "sex testing." J Sex Res. 2013;50(2):112–5.
3. Harper J, Martinez-Patino M-J, Pigozzi F, Pitsiladis Y. Implications of a third gender for elite sports. Curr Sports Med Rep. 2018;17(2):42–4.
4. Harper J, Lima G, Kolliari-Turner A, Malinsky FR, Wang G, Martinez-Patino MJ, et al. The fluidity of gender and implications for the biology of inclusion for transgender and intersex athletes. Curr Sports Med Rep. 2018;17(12):467–72. https://doi.org/10.1249/JSR.0000000000000543.
5. Sudai M. The testosterone rule—constructing fairness in professional sport. J Law Biosci. 2017;4(1):181–93.
6. Jones BA, Arcelus J, Bouman WP, Haycraft E. Sport and transgender people: a systematic review of the literature relating to sport participation and competitive sport policies. Sports Med. 2017;47(4):701–16.
7. Committee IO. Olympic Charter. 2020. www.olympic.org/documents/olympic-charter. Accessed 15 Jun 2020.
8. Walker PL, Cook DC. Brief communication: gender and sex: vive la difference. Am J Phys Anthropol. 1998;106(2):255–9.
9. Vilain E, Betancurt JO, Bueno-Guerra N, Martinez-Patiño MJ. Transgender athletes in elite sport competitions. Transgender Athletes in Competitive Sport. Oxfordshire: Routledge; 2017.
10. Reisner SL, Hughto JM. Comparing the health of non-binary and binary transgender adults in a statewide non-probability sample. PLoS ONE. 2019;14(8):e0221583.
11. NHS. Differences in sex development. 2019. https://www.nhs.uk/conditions/differences-in-sex-development/. Accessed 15 Sep 2020.
12. Jung EJ, Im DH, Park YH, Byun JM, Kim YN, Jeong DH, et al. Female with 46. XY karyotype Obstet Gynecol Sci. 2017;60(4):378–82.
13. Court GFC. Act on amending the act on the civil status of persons (Personenstandsgesetz). In: Article 4 of the Act of 18.12.2018. 2018. http://www.gesetze-im-internet.de/englisch_gg/englisch_gg.pdf. BGBl. 2018 IS. 2369.
14. Court TFC. Order of the First Senate of 10th October. 2017. http://www.bverfg.de/e/rs20171010_1bvr201916en.html. Accessed 25 Jun 2020.
15. Sport CoAf. CAS 2014/A/3759 Dutee Chand v. Athletics Federation of India (AFI) & The International Association of Athletics Federations (IAAF). 2014. https://www.doping.nl/media/kb/3317/CAS%202014_A_3759%20Dutee%20Chand%20vs.%20AFI%20%26%20IAAF%20%28S%29.pdf. Accessed 26 Feb 2020.
16. Sport CoAf. CAS 2018/0/5794 Mokgadi Caster Semenya v. International Association of Athletics Federations and CAS 2018/0/5798 Athletics South Africa v. International Association of Athletics Federations. 2018. https://www.tas-cas.org/fileadmin/user_upload/CAS_Award_-_redacted_-_Semenya_ASA_IAAF.pdf. Accessed 26 Feb 2020.
17. Council UNHR. Promotion of the cultural rights of everyone and respect for cultural diversity. 2019. https://documents-dds-ny.un.org/doc/UNDOC/GEN/G19/098/67/PDF/G1909867.pdf?OpenElement. Accessed 19 April 2020.
18. Ellaithi M, Kamel A, Saber O, Hiort O. Consanguinity and disorders of sexual developments in the Sudan. Sudan JMS. 2011;6(4):267–70.
19. Handelsman DJ, Hirschberg AL, Bermon S. Circulating testosterone as the hormonal basis of sex differences in athletic performance. Endocr Rev. 2018;39(5):803–29. https://doi.org/10.1210/er.2018-00020.
20. Bermon S, Garnier PY, Hirschberg AL, Robinson N, Giraud S, Nicoli R, et al. Serum androgen levels in elite female athletes. J Clin Endocrinol Metab. 2014;99(11):4328–35.
21. Morel Y, Rey R, Teinturier C, Nicolino M, Michel-Calemard L, Mowszowicz I, et al. Aetiological diagnosis of male sex ambiguity: a collaborative study. Eur J Pediatr. 2002;161(1):49–59.
22. Hirschberg AL, Knutsson JE, Helge T, Godhe M, Ekblom M, Bermon S, et al. Effects of moderately increased testosterone concentration on physical performance in young women: a double

blind, randomised, placebo controlled study. Br J Sports Med. 2019;54:599–604.
23. Bermon S. Androgens and athletic performance of elite female athletes. Curr Opin Endocrinol Diabetes Obes. 2017;24(3):246–51.
24. Bermon S, Garnier P-Y. Serum androgen levels and their relation to performance in track and field: mass spectrometry results from 2127 observations in male and female elite athletes. Br J Sports Med. 2017;51(17):1309–14.
25. Federations. IAoA. Eligibility Regulations for the Female Classification (Athletes with Differences of Sex Development). Version 1.0. 2018. www.iaaf.org/about-iaaf/documents/rules-regulations. Accessed 6 March 2020.
26. Pielke R, Tucker R, Boye E. Scientific integrity and the IAAF testosterone regulations. Int Sports Law J. 2019;19(1–2):18–26.
27. Menier A. Use of event-specific tertiles to analyse the relationship between serum androgens and athletic performance in women. Br J Sports Med. 2018;52(23):1540.
28. Koh B, Adair D, Elphick L. Not by gender, not by sex, but by testosterone saith the IAAF: International Athletics and the new female eligibility regulations. Law Sport J. Available at SSRN: https://ssrn.com/abstract=31846642018.
29. Bermon S, Hirschberg AL, Kowalski J, Eklund E. Serum androgen levels are positively correlated with athletic performance and competition results in elite female athletes. Br J Sports Med. 2018;52(23):1531–2.
30. Loland S. Caster Semenya, athlete classification, and fair equality of opportunity in sport. J Med Ethics. 2020;46(9):584–90.
31. Federations. IAoA. Eligibility Regulations for the female classification (athletes with differences of sex development). Version 2.0. 2019. www.iaaf.org/about-iaaf/documents/rules-regulations. Accessed 1 Jun 2020.
32. Hamilton BR, Martinez-Patiño MJ, Barrett J, Seal L, Tucker R, Papadopoulou T, et al. Response to the United Nations Human Rights Council's report on race and gender discrimination in sport: an expression of concern and a call to prioritise research. Sports Med. 2020. https://doi.org/10.1007/s40279-020-01380-y.
33. Court SFS. DSD Regulations: Caster Semenya's appeal against the decision of the Court of Arbitration for Sport dismissed. 2020. https://www.bger.ch/files/live/sites/bger/files/pdf/en/4A_248_2019_yyyy_mm_dd_T_e_18_18_10.pdf. Accessed 10 Oct 2020.
34. Johnson J, Sotherton K. Caster Semenya ruling: The case for and against IAAF's testosterone regulations. The Telegraph, Online. 2019. https://www.telegraph.co.uk/athletics/2019/05/02/caster-semenya-ruling-case-against-iaafs-testosterone-regulations/. Accessed 8 Apr 2020.
35. Tannenbaum C, Bekker S. Sex, gender, and sports. BMJ. 2019;364:l1120.
36. Committee TGAFR. Fiscal Note HB 1572—2077. 2020. http://www.capitol.tn.gov/Bills/111/Fiscal/HB1572.pdf. Accessed 1 Apr 2020.
37. Luevano A. Fairness In Women's Sports Act bans some transgender athletes. Local News 8 and KIDK Eyewitness News 3, Online 2020. https://localnews8.com/sports/2020/04/02/fairness-in-womens-sports-act-bans-some-transgender-athletes/. Accessed 15 Jun 2020.
38. Idaho TSo. Fairness in Women's Sports Act. 2020. https://legislature.idaho.gov/sessioninfo/billbookmark/?yr=2020&bn=H0500. Accessed 17 Apr 2020.
39. Freedom AD. Idaho governor signs Fairness for Women in Sports Act into law. 2020. http://www.adfmedia.org/News/PRDetail/10948. Accessed 17 Apr 2020.
40. Press A. Civil rights groups file suit against Idaho ban on trans athletes in women's sports. The Guardian. 2020. https://www.theguardian.com/sport/2020/apr/15/idaho-transgender-ban-womens-sports-lawsuit. Accessed 17 Apr 2020.
41. Minsberg T. 'Boys are boys and girls are girls': Idaho Is first state to bar some transgender athletes. The New York Times, Online. 2020. https://www.nytimes.com/2020/04/01/sports/transgender-idaho-ban-sports.html. Accessed 17 Apr 2020.
42. Cavanagh SL, Sykes H. Transsexual bodies at the Olympics: the international Olympic Committee's policy on transsexual athletes at the 2004 Athens Summer Games. Body Soc. 2006;12(3):75–102.
43. Hilton EN, Lundberg TR. Transgender women in the female category of sport: perspectives on testosterone suppression and performance advantage. Sports Med. 2020. https://doi.org/10.1007/s40279-020-01389-3.
44. Pitsiladis Y, Harper J, Betancurt JO, Martinez-Patino M-J, Parisi A, Wang G, et al. Beyond fairness: the biology of inclusion for transgender and intersex athletes. Curr Sports Med Rep. 2016;15(6):386–8.
45. Gooren LJ, Bunck MC. Transsexuals and competitive sports. Eur J Endocrinol. 2004;151(4):425–30.
46. Gooren LJ, Giltay EJ, Bunck MC. Long-term treatment of transsexuals with cross-sex hormones: extensive personal experience. J Clin Endocrinol Metab. 2008;93(1):19–25.
47. Karalexi MA, Georgakis MK, Dimitriou NG, Vichos T, Katsimpris A, Petridou ET, et al. Gender-affirming hormone treatment and cognitive function in transgender young adults: a systematic review and meta-analysis. Psychoneuroendocrinology. 2020;119:104721.
48. T'Sjoen G, Weyers S, Taes Y, Lapauw B, Toye K, Goemaere S, et al. Prevalence of low bone mass in relation to estrogen treatment and body composition in male-to-female transsexual persons. J Clin Densitom. 2009;12(3):306–13.
49. Harper J. Race times for transgender athletes. J Sport Cult Ident. 2015;6(1):1–9.
50. Grubb H. Models for comparing athletic performances. J R Stat Soc Ser D Stat. 1998;47(3):509–21.
51. Wiik A, Lundberg T, Rullman E, Andersson D, Holmberg M, Mandić M, et al. Muscle strength, size, and composition following 12 months of gender-affirming treatment in transgender individuals. J Clin Endocrinol. 2020;105(3):e805–13.
52. NHS. Gender dysphoria. The National Health Service. 2020. https://www.nhs.uk/conditions/gender-dysphoria/. Accessed 29 Jul 2020.
53. Roberts TA, Smalley J, Ahrendt D. Effect of gender affirming hormones on athletic performance in transwomen and transmen: implications for sporting organisations and legislators. Br J Sports Med. 2020. https://doi.org/10.1136/bjsports-2020-102329.
54. Commitee IO. Consensus Meeting on Sex Reassignment and Hyperandrogenism. 2015. https://stillmed.olympic.org/Documents/Commissions_PDFfiles/Medical_commission/2015-11_ioc_consensus_meeting_on_sex_reassignment_and_hyperandrogenism-en.pdf. Accessed 20 Feb 2020.
55. Athletics W. Eligibility Regulations Transgender Athletes. 2019. https://www.worldathletics.org/about-iaaf/documents/book-of-rules. Accessed 22 Oct 2020.
56. Rowing W. Rule 13 and Bye law to rule 13– Men's and Women's Events. 2020. www.worldrowing.com. Accessed 2 Dec 2020.
57. Internationale UC. Rule 13.5.028. 2020. https://www.uci.org/docs/default-source/rules-and-regulations/part-xiii---medical-rules---01.03.2020.pdf. Accessed 5 Dec 2020.
58. Rugby W. World Rugby Transgender Guideline. Online. 2020. https://playerwelfare.worldrugby.org/?documentid=231. Accessed 2 Nov 2020.
59. Ingle S. RFU clears trans women to keep playing domestic women's rugby in England. The Guardian. 2020. https://www.theguardian.com/sport/2020/oct/14/rfu-clears-trans-women-to-play-womens-rugby-at-all-levels-in-england?fbclid=IwAR359QKh

JiCxKr8a_u8yX1vGNN7opyvA52uJVykCfTR7B9ehb2EJvc8_kOM. Accessed 2 Nov 2020.
60. Rønnestad BR, Nygaard H, Raastad T. Physiological elevation of endogenous hormones results in superior strength training adaptation. Eur J Appl Physiol. 2011;111(9):2249–59.
61. Schwanbeck SR, Cornish SM, Barss T, Chilibeck PD. Effects of training with free weights versus machines on muscle mass, strength, free testosterone, and free cortisol levels. J Strength Cond Res. 2020;34(7):1851–9.
62. Kvorning T, Andersen M, Brixen K, Madsen K. Suppression of endogenous testosterone production attenuates the response to strength training: a randomized, placebo-controlled, and blinded intervention study. Am J Physiol Endocrinol Metab. 2006;291(6):E1325–32.
63. Kraemer WJ, Ratamess NA, Nindl BC. Recovery responses of testosterone, growth hormone, and IGF-1 after resistance exercise. J Appl Physiol. 2017;122(3):549–58.
64. Kelso A, Vogel K, Steinacker JM. Ultrasound measurements of subcutaneous adipose tissue thickness show sexual dimorphism in children of three to five years of age. Acta Paediatr. 2019;108(3):514–21.
65. Marta CC, Marinho DA, Barbosa TM, Izquierdo M, Marques MC. Physical fitness differences between prepubescent boys and girls. J Strength Cond Res. 2012;26(7):1756–66.
66. Arnold AP, Chen X. What does the "four core genotypes" mouse model tell us about sex differences in the brain and other tissues? Front Neuroendocrinol. 2009;30(1):1–9.
67. Burgoyne PS, Arnold AP. A primer on the use of mouse models for identifying direct sex chromosome effects that cause sex differences in non-gonadal tissues. Biol Sex Differ. 2016;7(1):68.
68. Chen X, McClusky R, Chen J, Beaven SW, Tontonoz P, Arnold AP, et al. The number of x chromosomes causes sex differences in adiposity in mice. PLoS Genet. 2012;8(5):e1002709.
69. Chen X, Wang L, Loh DH, Colwell CS, Taché Y, Reue K, et al. Sex differences in diurnal rhythms of food intake in mice caused by gonadal hormones and complement of sex chromosomes. Horm Behav. 2015;75:55–63.
70. Heinlein CA, Chang C. Androgen receptor (AR) coregulators: an overview. Endocr Rev. 2002;23(2):175–200.
71. Heemers HV, Tindall DJ. Androgen receptor (AR) coregulators: a diversity of functions converging on and regulating the AR transcriptional complex. Endocr Rev. 2007;28(7):778–808.
72. van de Wijngaart DJ, Dubbink HJ, van Royen ME, Trapman J, Jenster G. Androgen receptor coregulators: recruitment via the coactivator binding groove. Mol Cell Endocrinol. 2012;352(1–2):57–69.
73. Papakonstanti EA, Kampa M, Castanas E, Stournaras C. A rapid, nongenomic, signaling pathway regulates the actin reorganization induced by activation of membrane testosterone receptors. J Mol Endocrinol. 2003;17(5):870–81.
74. Lamont KR, Tindall DJ. Minireview: Alternative activation pathways for the androgen receptor in prostate cancer. Mol Endocrinol. 2011;25(6):897–907.
75. Davey RA, Grossmann M. Androgen receptor structure, function and biology: from bench to bedside. Clin Biochem Rev. 2016;37(1):3.
76. Heinlein CA, Chang C. The roles of androgen receptors and androgen-binding proteins in nongenomic androgen actions. J Mol Endocrinol. 2002;16(10):2181–7.
77. Association WM. World Medical Association urges physicians not to implement IAAF rules on classifying women athletes. 2020. https://www.wma.net/news-post/wma-urges-physicians-not-to-implement-iaaf-rules-on-classifying-women-athletes/. Accessed 8 Jul 2020.
78. Caliskan Guzelce E, Eyupoglu N, Torgutalp S, Aktoz F, Portakal O, Demirel H, et al. Is muscle mechanical function altered In polycystic ovary syndrome? J Endocr Soc. 2019;3(Supplement_1):MON-205.
79. Egner IM, Bruusgaard JC, Eftestøl E, Gundersen K. A cellular memory mechanism aids overload hypertrophy in muscle long after an episodic exposure to anabolic steroids. J Physiol. 2013;591(24):6221–30.
80. Bruusgaard JC, Johansen I, Egner I, Rana Z, Gundersen K. Myonuclei acquired by overload exercise precede hypertrophy and are not lost on detraining. Proc Natl Acad Sci USA. 2010;107(34):15111–6.
81. Sinha-Hikim I, Roth SM, Lee MI, Bhasin S. Testosterone-induced muscle hypertrophy is associated with an increase in satellite cell number in healthy, young men. Am J Physiol Endocrinol Metab. 2003;285(1):E197–205.
82. Sinha-Hikim I, Cornford M, Gaytan H, Lee ML, Bhasin S. Effects of testosterone supplementation on skeletal muscle fiber hypertrophy and satellite cells in community-dwelling older men. J Clin Endocrinol Metab. 2006;91(8):3024–33.
83. Horwath O, Apró W, Moberg M, Godhe M, Helge T, Ekblom M, et al. Fiber type-specific hypertrophy and increased capillarization in skeletal muscle following testosterone administration in young women. J Appl Physiol. 2020;128(5):1240–50.
84. Le Grand F, Rudnicki MA. Skeletal muscle satellite cells and adult myogenesis. Curr Opin Cell Biol. 2007;19(6):628–33.
85. Aitken M, Steensma TD, Blanchard R, VanderLaan DP, Wood H, Fuentes A, et al. Evidence for an altered sex ratio in clinic-referred adolescents with gender dysphoria. J Sex Med. 2015;12(3):756–63.
86. De Vries A, Kreukels B, T'sjoen G, Ålgars M, Mattila A. Increase of referrals to gender identity clinics: a European trend. Transgender Healthc Europe. Book of abstracts. 2015.
87. Bouman WP, de Vries AL, T'Sjoen G. Gender dysphoria and gender incongruence: an evolving inter-disciplinary field. Milton Park: Taylor & Francis; 2016.
88. Committee. IO. Call for Applications. 2019. https://www.olympic.org/news/2019-call-for-applications. Accessed 16 Apr 2020.
89. NHS. Testosterone. 2020. https://www.nbt.nhs.uk/severn-pathology/requesting/test-information/testosterone. Accessed 2 Aug 2020.
90. Stanczyk FZ, Clarke NJ. Advantages and challenges of mass spectrometry assays for steroid hormones. J Steroid Biochem Mol Bio. 2010;121(3–5):491–5.
91. Martínez-Escribano A, Maroto-García J, Ruiz-Galdón M, Barrios-Rodríguez R, Álvarez-Millán JJ, Cabezas-Sánchez P, et al. Measurement of serum testosterone in nondiabetic young obese men: comparison of direct immunoassay to liquid chromatography-tandem mass spectrometry. Biomolecules. 2020;10(12):1697.

Authors and Affiliations

Blair R. Hamilton[2,3] · Giscard Lima[1,4] · James Barrett[3] · Leighton Seal[3] · Alexander Kolliari-Turner[2] · Guan Wang[65] · Antonia Karanikolou[2] · Xavier Bigard[5,6,7] · Herbert Löllgen[6] · Petra Zupet[6] · Anca Ionescu[6] · Andre Debruyne[6,7] · Nigel Jones[8,9] · Karin Vonbank[10] · Federica Fagnani[4] · Chiara Fossati[4,11] · Maurizio Casasco[6,7,12] · Demitri Constantinou[7,13] · Bernd Wolfarth[7,14] · David Niederseer[15] · Andrew Bosch[16] · Borja Muniz-Pardos[17] · José Antonio Casajus[17] · Christian Schneider[7,18] · Sigmund Loland[19] · Michele Verroken[20,21] · Pedro Manonelles Marqueta[7,22] · Francisco Arroyo[7,23] · André Pedrinelli[7,24] · Konstantinos Natsis[6,7,25,26] · Evert Verhagen[27] · William O. Roberts[7,28] · José Kawazoe Lazzoli[7,29] · Rogerio Friedman[30] · Ali Erdogan[7,31] · Ana V. Cintron[7,32] · Shu-Hang Patrick Yung[7,33] · Dina C. Janse van Rensburg[34] · Dimakatso A. Ramagole[34] · Sandra Rozenstoka[6,7,35] · Felix Drummond[6,7,36] · Theodora Papadopoulou[6,7,37] · Paulette Y. O. Kumi[38] · Richard Twycross-Lewis[39] · Joanna Harper[40] · Vasileios Skiadas[41] · Jonathan Shurlock[42] · Kumpei Tanisawa[43] · Jane Seto[44,45] · Kathryn North[44,45] · Siddhartha S. Angadi[46] · Maria Jose Martinez-Patiño[47] · Mats Borjesson[7,48,49] · Luigi Di Luigi[7,50] · Michiko Dohi[7,51] · Jeroen Swart[7,52] · James Lee John Bilzon[7,53] · Victoriya Badtieva[7,54,55] · Irina Zelenkova[17] · Juergen M. Steinacker[6,7,56] · Norbert Bachl[6,7,57,58] · Fabio Pigozzi[4,6,7,11] · Michael Geistlinger[7,59] · Dimitrios G. Goulis[60] · Fergus Guppy[2,61] · Nick Webborn[62] · Bulent O. Yildiz[63] · Mike Miller[64] · Patrick Singleton[64] · Yannis P. Pitsiladis[1,2,4,6,7]

✉ Yannis P. Pitsiladis
Y.Pitsiladis@brighton.ac.uk

1 Centre for Exercise Sciences and Sports Medicine, FIMS Collaborating Centre of Sports Medicine, Rome, Italy

2 Centre for Stress and Age-Related Disease, University of Brighton, Brighton, UK

3 The Gender Identity Clinic Tavistock and Portman NHS Foundation Trust, London, UK

4 Department of Movement, Human and Health Sciences, University of Rome "Foro Italico", Rome, Italy

5 Union Cycliste Internationale (UCI), Aigle, Switzerland

6 European Federation of Sports Medicine Associations (EFSMA), Lausanne, Switzerland

7 International Federation of Sports Medicine (FIMS), Lausanne, Switzerland

8 British Association Sport and Exercise Medicine, Doncaster, UK

9 British Cycling and University of Liverpool, Liverpool, UK

10 Department of Pneumology, Pulmonary Function Laboratory, Medicine Clinic (KIMII), University of Vienna, Vienna, Austria

11 Villa Stuart Sport Clinic, FIFA Medical Center of Excellence, Rome, Italy

12 Italian Federation of Sports Medicine (FMSI), Rome, Italy

13 Centre for Exercise Science and Sports Medicine, University of the Witwatersrand, Johannesburg, South Africa

14 Department of Sports Medicine, Humboldt University and Charité University School of Medicine, Berlin, Germany

15 Department of Cardiology, University Hospital Zurich, University Heart Centre, University of Zurich, Zurich, Switzerland

16 Division of Exercise Science and Sports Medicine, University of Cape Town, Cape Town, South Africa

17 GENUD Research Group, FIMS Collaborating Center of Sports Medicine, Department of Physiatry and Nursing, University of Zaragoza, Zaragoza, Spain

18 Orthopaedic Center Theresie, Munich, Germany

19 Department of Sport and Social Sciences, Norwegian School of Sport Sciences, Oslo, Norway

20 Centre of Research and Innovation for Sport, Technology and Law (CRISTAL), De Montfort University, Leicester, UK

21 Sporting Integrity Ltd, Stoke Mandeville, UK

22 Department of Sports Medicine, San Antonio Catholic University of Murcia, Murcia, Spain

23 FIMS Collaborating Center of Sports Medicine, Guadalajara, Mexico

24 Department of Orthopaedics, University of São Paulo Medical School, São Paulo, Brazil

25 Interbalkan Medical Center, FIMS Collaborating Center of Sports Medicine, Thessaloniki, Greece

26 Department of Anatomy and Surgical Anatomy, Faculty of Health Sciences, School of Medicine, Aristotle University of Thessaloniki, Thessaloniki, Greece

27 Amsterdam Collaboration on Health and Safety in Sports, Department of Public and Occupational Health, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam Movement Sciences, Amsterdam, The Netherlands

28 Department of Family Medicine and Community Health, University of Minnesota, Minneapolis, USA

29 Biomedical Institute, Fluminense Federal University Medical School, Niterói, Brazil

30 Universidade Federal do Rio Grande do Sul, Endocrine Unit, Hospital de Clinicas de Porto Alegre, Porto Alegre, Brazil

31 Gloria Sports Arena, FIMS Collaborating Centre of Sports Medicine, Antalya, Turkey

32 Puerto Rico Sports Medicine Federation, San Juan, Puerto Rico

[33] Asian Federation of Sports Medicine (AFSM), Hong Kong Center of Sports Medicine and Sports Science, Hong Kong, China

[34] Section Sports Medicine, Faculty of Health Sciences, University of Pretoria, Pretoria, South Africa

[35] FIMS Collaboration Centre of Sports Medicine, Sports laboratory, Riga, Latvia

[36] FIMS Collaboration Centre of Sports Medicine, Instituto de Medicina do Esporte, Porto Alegre, Brazil

[37] Defence Medical Rehabilitation Centre, Stanford Hall, Loughborough, UK

[38] Centre for Sports and Exercise Medicine, Queen Mary University of London, London, UK

[39] School of Engineering and Materials Science, Queen Mary University of London, London, UK

[40] School of Sport, Exercise and Health Sciences, Loughborough University, Loughborough, UK

[41] University Hospital Southampton, Southampton, UK

[42] Brighton and Sussex University Hospitals, Brighton, UK

[43] Faculty of Sport Sciences, Waseda University, Tokorozawa, Japan

[44] Murdoch Children's Research Institute, Melbourne, VIC, Australia

[45] Department of Paediatrics, University of Melbourne, The Royal Children's Hospital, Melbourne, VIC, Australia

[46] Department of Kinesiology, School of Education and Human Development, University of Virginia, Charlottesville, VA, USA

[47] Faculty of Educational Sciences and Sports, University of Vigo, Galicia, Spain

[48] Department of Molecular and Clinical Medicine, Sahlgrenska Academy, Center for Health and Performance, Goteborg University, Göteborg, Sweden

[49] Sahlgrenska University Hospital/Ostra, Region of Western Sweden, Göteborg, Sweden

[50] Unit of Endocrinology, Department of Movement, Human and Health Sciences, University of Rome "Foro Italico", Rome, Italy

[51] Sport Medical Center, Japan Institute of Sports Sciences, Tokyo, Japan

[52] UCT Research Unit for Exercise Science and Sports Medicine, Cape Town, South Africa

[53] Department for Health, University of Bath, Bath, UK

[54] I.M. Sechenov First Moscow State Medical University (Sechenov University), Ministry of Health of Russia, Moscow, Russian Federation

[55] Moscow Research and Practical Center for Medical Rehabilitation, Restorative and Sports Medicine, Moscow Healthcare Department, Moscow, Russian Federation

[56] Division of Sports and Rehabilitation Medicine, Ulm University Hospital, Ulm, Germany

[57] Institute of Sports Science, University of Vienna, Vienna, Austria

[58] Austrian Institute of Sports Medicine, Vienna, Austria

[59] Unit of International Law, Department of Constitutional, International and European Law, University of Salzburg, Salzburg, Salzburg, Austria

[60] Unit of Reproductive Endocrinology, 1st Department of Obstetrics and Gynecology, Medical School, Aristotle University of Thessaloniki, Thessaloniki, Greece

[61] School of Pharmacy and Biomolecular Sciences, University of Brighton, Brighton, UK

[62] School of Sport and Service Management, University of Brighton, Eastbourne, UK

[63] Division of Endocrinology and Metabolism, Department of Internal Medicine, Hacettepe University School of Medicine, 06100 Ankara, Turkey

[64] World Olympian Association, Lausanne, Switzerland

[65] Sport and Exercise Science and Sports Medicine Research and Enterprise Group, University of Brighton, Brighton, UK

Received: 26 January 2017 | Revised: 4 April 2017 | Accepted: 6 April 2017

DOI: 10.1111/cen.13350

ORIGINAL ARTICLE

WILEY

Sex differences in athletic performance emerge coinciding with the onset of male puberty

David J Handelsman

ANZAC Research Institute, University of Sydney, Sydney, NSW, Australia

Correspondence
David J. Handelsman,
ANZAC Research Institute, Concord Hospital 2139
Sydney, NSW, Australia.
Email: djh@anzac.edu.au

Summary

Background: Male performance in athletic events begins to exceed that of age-matched females during early adolescence, but the timing of this divergence relative to the onset of male puberty and the rise in circulating testosterone remains poorly defined.

Design: This study is a secondary quantitative analysis of four published sources which aimed to define the timing of the gender divergence in athletic performance and relating it to the rise in circulating testosterone due to male puberty.

Data: Four data sources reflecting elite swimming and running and jumping track and field events as well as hand-grip strength in nonathletes were analysed to define the age-specific gender differences through adolescence and their relationship to the rising circulating testosterone during male puberty.

Results: The onset and tempo of gender divergence were very similar for swimming, running and jumping events as well as the hand-grip strength in nonathletes, and all closely paralleled the rise in circulating testosterone in adolescent boys.

Conclusions: The gender divergence in athletic performance begins at the age of 12-13 years and reaches adult plateau in the late teenage years with the timing and tempo closely parallel to the rise in circulating testosterone in boys during puberty.

KEYWORDS

age group, performance, puberty, swimming, testosterone, track and field

1 | INTRODUCTION

It is well known that men's athletic performance exceeds that of women especially in power sports because of men's greater strength, speed and endurance. This biological physical advantage of mature males forms the basis for gender segregation in many competitive sports to allow females a realistic chance of winning events. This physical advantage in performance arises during early adolescence when male puberty commences after which men acquire larger muscle mass and greater strength, larger and stronger bones, higher circulating haemoglobin as well as mental and/or psychological differences. After completion of male puberty, circulating testosterone levels in men are consistently 10-15 times higher than in children or women at any age.[1] The age at which sex differences emerge is reported as around the age of 12 from a study of individual Norwegian athletes in two running and two jumping events[2] and at 13-14 years in four track and field skills in Polish athletes[3]; however, the relationship to male puberty and circulating testosterone is not clear. This study investigates the age of the gender divergence in performance in elite swimming and a wider range of elite athletic events as well as a community-based study of grip strength among nonathletes to deduce the onset and progression of the gender divergence in performance of athletes and relates this to the timing and tempo of male puberty and the rise in circulating testosterone into adult male levels.

2 | MATERIAL AND METHODS

Four sources of published data were used in this study for which no ethics approval was required. The first was the US Age Group Swimming time standards which lists the prevailing time standard for entry to the top level (AAAA long course criteria) of all boys and girls events for individual years from 1981 to 2016 (accessed Oct 2016).

http://www.usaswimming.org/DesktopDefault.aspx?TabId=2628&Alias=Rainbow&Lang=en

Age groups were classified into five categories 10 and under, 11-12 years, 13-14 year, 15-16 years and 17-18 years. The seven events in common to all age groups were freestyle (50 m, 100 m, 200 m), backstroke, breaststroke and butterfly (all 100 m) and individual medley (200 m).

A second data source was the current world records for boys and girls between the ages of 5 and 19 years available at http://age-records.125mb.com/ (curated by Dominique Eisold, accessed Oct 2016). This included sufficient data to cover the timing of puberty onset with some pre- and postpuberty ages (ages 9-19 years) for a wide range of boys and girls track and field events. For this study, the running events included were 50 m, 60 m, 100 m, 200 m, 300 m, 400 m, 500 m, 600 m, 800 m, 1000 m, 1500 m, 1 mile, 2000 m, 3000 m and 2 miles. Only records recorded by fully automatic timing devices were included whether set indoor or outdoor or at altitude (>1000 m), but wind-assisted records were excluded from this analysis. The jumping events included were high jump, pole vault, long jump, triple jump, standing long jump.

The third data source was from a published study[1] in which serum testosterone was measured in over 100 000 consecutive serum samples processed over 7 years from a single pathology laboratory which was analysed to estimate male and female age-specific reference ranges across the full lifespan.

The fourth was a meta-analysis of secular changes in hand-grip strength in nonathletic children and adolescents from Canada and United States[4] using the data provided on 5676 males and 5489 females in 19 studies conducted between 1966 and 2009.

Data analysis was performed by analysis of variance and nonlinear curve fitting using NCSS 11 Statistical Software (NCSS LLC. Kaysville, Utah, USA). For each event used in this analysis, the age-specific record or age-group time standard was defined for boys (Tb) and girls (Tg) so the difference (expressed as a percentage) between boys and girls for any event was defined as D=(Tg−Tb)*100/Tg. For athletic jumping events, an analogous definition for record length was used (Lb for boys, Lg for girls) with the male advantage defined as D=(Lb−Lg)*100/Lg. For the athletic events where individual year age records were available across the age of puberty, the age-specific difference (as a percentage) for each year of age were pooled into running or jumping categories. For track and field performance, the pooled data were fitted to a four-parameter sigmoidal curve which allowed for asymptotic estimation of the lower (prepubertal) and upper (postpubertal) plateaus from the four parameters. In addition, the timing and tempo of the pubertal increase were defined by the start of puberty, defined as the time when 20% of the ultimate increase due to puberty had occurred (ED_{20}), and mid-puberty as the time when half the ultimate increase had occurred (ED_{50}). For swimming, the pooled gender differences for all strokes and distances were fitted by a smoothed spline curve. For hand-grip strength, the differences were fitted to a piecewise linear-quadratic curve with a single inflexion point.

3 | RESULTS

In swimming performance, the overall gender differences were highly significant with age group ($F_{4,360}$=1481, P<.0001) and stroke ($F_{4,360}$=11.9, P<.0001) as main (between) effects (Figure 1). There was no significant difference according to year (as a within factor, P=.99) so that for further analysis, years were taken as replicates. Using a sigmoidal curve fit for the overall gender differences pooling all strokes and distances, the ED_{20} was 11.4 years and the ED_{50} was 12.8 years.

Within a single stroke (freestyle), in addition to expected age-group effects ($F_{4,525}$=2174, P<.0001), there were also significant effects according to distance ($F_{2,525}$=231.5, P<.0001) whereby the age-group effects was significantly greater the shorter the event distance (Figure 2, 50 m>100 m>200 m, age group x distance interaction, $F_{8,525}$=55.9, P<.0001) (Figure 1). Similarly, for a fixed length of events (100 m) and after taking age-group effects into account, the four form strokes did differ significantly ($F_{3,700}$=12.9, P<.0001) producing significant

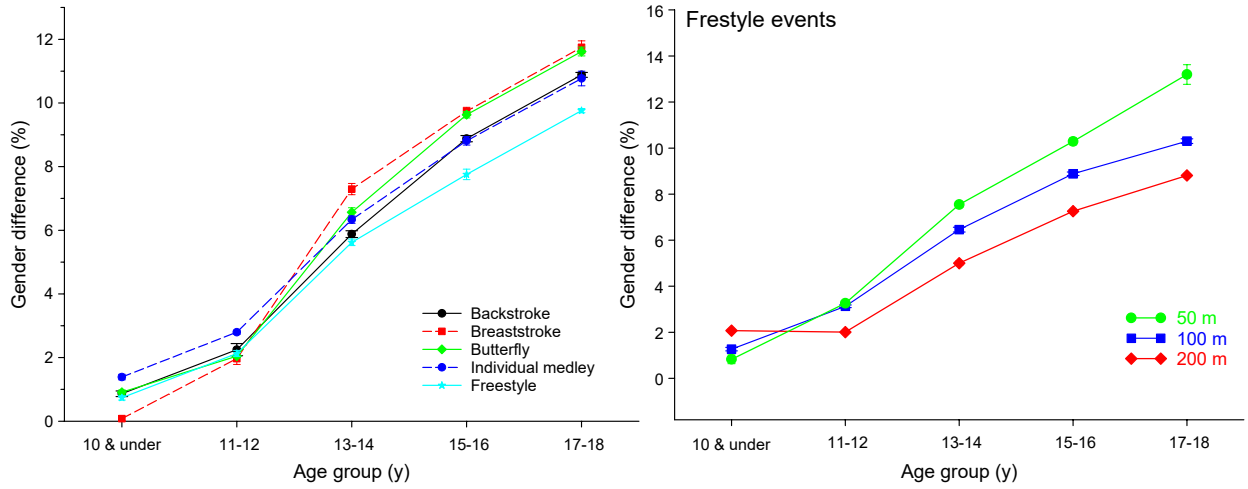


FIGURE 1 Gender differences in performance (in percentage) according to age group and stroke (left panel) or distance in freestyle events (right) in swimming events. Data shown as mean and standard error of the mean. Note greatest increase after the age of 12 years by age in breaststroke and least in freestyle and magnitude of increases are 50 m>100 m>200 m in freestyle events. [Colour figure can be viewed at wileyonlinelibrary.com]

differences between strokes (interaction $F_{12,700}$=23.4, P<.0001), the most prominent being for breaststroke, which displayed the greatest age-group effect, and butterfly followed by backstroke and then freestyle, which showed the least age-group effect (Figure 1).

In track and field athletics, the effects of age on running performance (Figure 2 upper left panel) showed that the prepubertal differences of 3.0% increased to a plateau of 10.1% with an onset (ED_{20}) at 12.4 years and reaching midway (ED_{50}) at 13.9 years. For jumping (Figure 2 upper right panel), the prepubertal difference of 5.8% increased to 19.4% starting at 12.4 years and reaching midway at 13.9 years. The timing of the male advantage in running, jumping and swimming was similar and corresponded to the increases in serum testosterone in males (Figure 2 lower panel).

To examine age of gender divergence in strength in an analogous data set from a nonathletic population (Canadian and US children and adolescents), the age trends in hand-grip strength showed a difference in hand-grip strength commencing from the age of 12.8 years onwards (Figure 3). Prior to the age of 13 years, boys had a marginally significant greater grip strength than girls (n=45, t=2.0, P=.026), but after the age of 13 years, there was a strong significant relationship between age and difference in grip strength (n=18, r=.89, P<.001).

4 | DISCUSSION

The present study shows that the gender divergence in performance for swimming and for running and jumping track and field events is very closely aligned to the timing of the onset of male puberty, which typically has onset at around 12 years of age.[5,6] These findings are consistent with reports on the timing of the gender differences in performance observed among Norwegian athletes in two running and two jumping events[2] and for track and field skills among Polish athletes.[3] This study extends the findings to swimming and a wider range of running and jumping track and field events. This timing is also consistent with the start of the gender divergence in fat-free (muscle) mass[7] and strength increases.[8,9]

In this study, the timing and tempo of male puberty effects on running and jumping performance were virtually identical and very similar

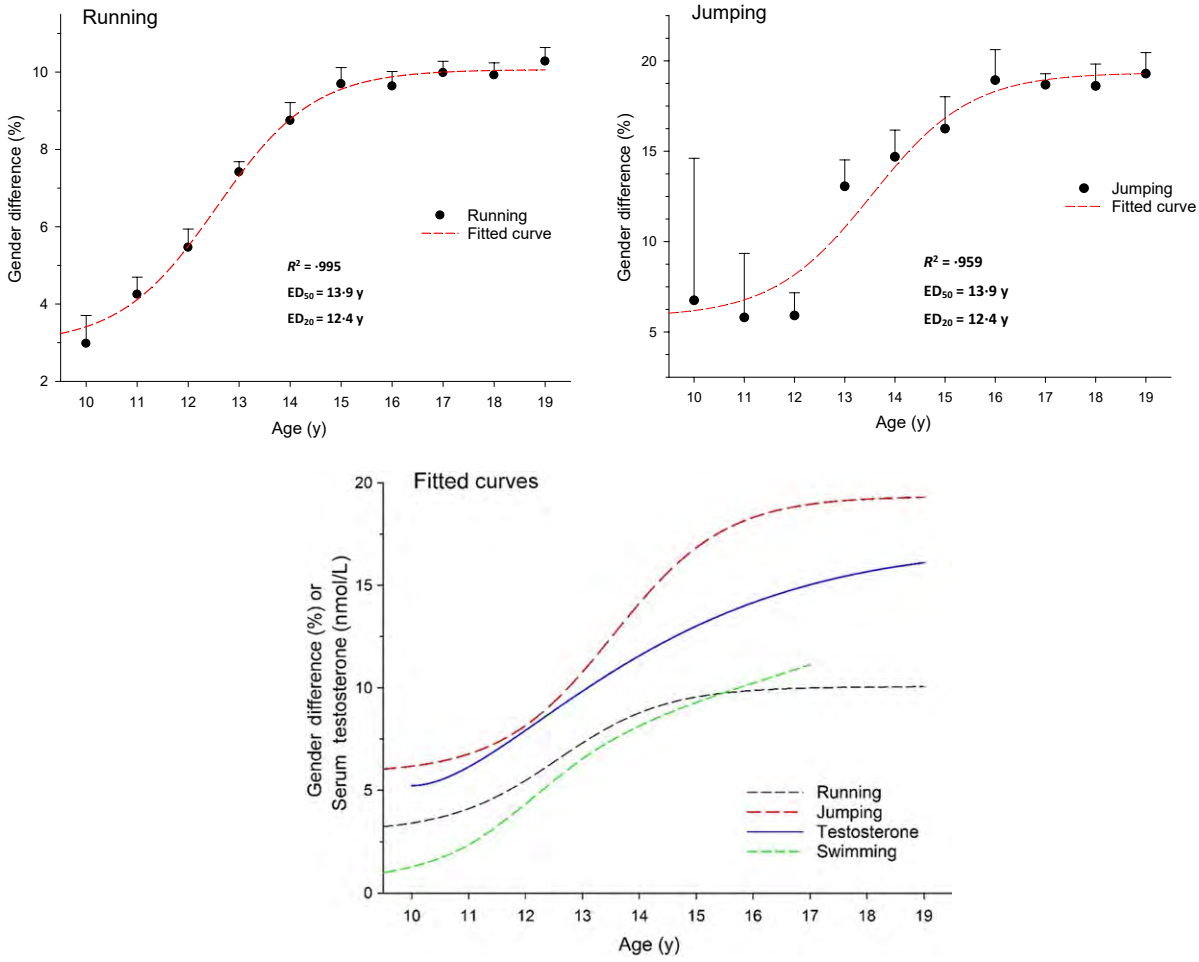


FIGURE 2 Gender differences in performance (in percentage) according to age (in years) in running events including 50 m, 60 m, 100 m, 200 m, 300 m, 400 m, 500 m, 600 m, 800 m, 1000 m, 1500 m, 1 mile, 2000 m, 3000 m and 2 miles (upper left panel) and in jumping events including high jump, pole vault, triple jump, long jump and standing long jump (upper right panel). Fitted sigmoidal curve plot of gender differences in performance (in percentage) according to age (in years) in running, jumping and swimming events as well as serum testosterone (lower panel). Data shown as mean and standard error of the mean of the pooled gender differences by age. [Colour figure can be viewed at wileyonlinelibrary.com]

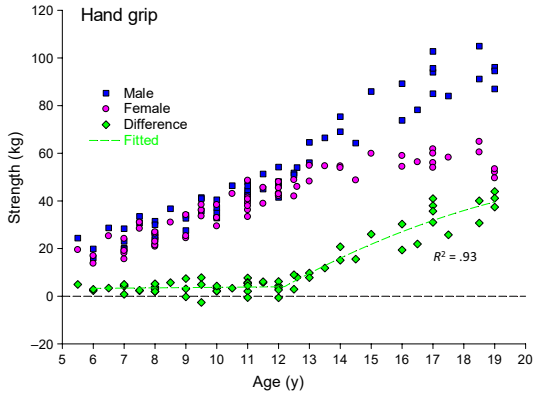


FIGURE 3 Hand-grip strength in children and adolescents from 19 studies including 5676 males (square) and 5489 females (circles) and the differences between male and females (diamonds) conducted between 1966 and 2009. The dotted line represents the fitted curve using a piecewise linear-quadratic curve fit with an automatically defined inflexion point at 12.8 years. [Colour figure can be viewed at wileyonlinelibrary.com]

to those in swimming events. Furthermore, these coincided with the timing of the rise in circulating testosterone due to male puberty. In addition to the strikingly similar timing and tempo, the magnitude of the effects on performance by the end of this study was 10.0% for running and 19.3% for jumping, both consistent with the gender differences in performance of adult athletes previously reported to be 10%-12% for running[10,11,12] and 19% for jumping.[12] The similar magnitude of the plateau effects observed for the oldest (postpubertal) stages in this study with mature adult gender differences suggests there are likely minimal if any further divergences in gender performance among athletes after the age of 20 years.

In the swimming events, despite the continued progressive improvements in individual male and female event records, the stability of the gender difference over 35 years shown in this study suggests that the gender differences in performance are stable and robust. These findings are consistent with a previous report of no narrowing of the gender gap in swimming event performance over more than three decades.[12] These findings contribute to discounting previous suggestions that the gender gap in performance of athletes was narrowing and might even disappear,[13] interpretations which were confounded by the increasing participation of females in elite sports through the 20th century that led to short-term accelerating improvement until women approached closer to contemporary female performance plateau.[12] The greater effect of male puberty on shorter freestyle events is consistent with the greater power demands of short sprint events than for longer freestyle events that involve more endurance. The consistent differences between form strokes over 100-m events, even after accounting for the very dominant age-group effect, suggest that the power demands on performance were most prominent in breaststroke and least in freestyle, presumably due to the different mechanical demands of the different strokes.

The gender divergence in hand-grip strength among nonathletic children and adolescents strengthens the view that these gender divergences are a feature of normal male puberty rather than being a feature that manifests only in elite athletes.

The similar time course of the rise in circulating testosterone with that of the gender divergences in swimming and track and field sports is strongly suggestive that these effects arise from the increase in circulating testosterone from the start of male puberty.[1] Somatic effects of male puberty differ in responsiveness to the postpubertal increase in serum testosterone. Muscle effects of testosterone have been established in well-controlled, interventional clinical experiments in healthy young[14,15] and older[16] men. Testosterone increases muscle mass and strength over weeks to months with a strong dose-response evident from below to above physiological testosterone doses and concentrations. Analogous findings are reported in androgen-deficient (hypogonadal) men administered testosterone replacement therapy[17] and in women receiving appropriately lower testosterone doses,[18] and observational dose-effect relationship between endogenous testosterone and upper or lower body muscle mass is reported in healthy men.[19] Most if not all sex differences in maximal oxygen uptake are explained by differences in muscle mass.[20-22]

Adult male circulating testosterone also has marked effects on bone development leading to longer, stronger and denser bone than in age-matched females.[23] However, testosterone effects on bone are slower in onset and probably less reversible than effects on muscle. For example, men achieve peak bone mass at the end of skeletal maturation only in the early 1920s, about a decade after the start of sustained exposure to adult male testosterone levels. Furthermore, while testosterone deficiency may lead to loss of bone density,[23] the overall structural framework of the skeleton is likely to change slowly if at all. Hence, the extent to which testosterone-induced bone changes contribute to the male advantage in adolescent athletic performance is unclear but is probably at least not maximal until the third decade of life by which time the gender differences are already stabilized.

A further biological advantage of adult male circulating testosterone concentrations is the increased circulating haemoglobin. Men have ~10 g/L greater haemoglobin than women[24] with the gender differences also evident from the age of 13-14 years.[25] Testosterone effects on haemoglobin are replicated by administration of exogenous testosterone in a dose-dependent fashion[26] within 1-3 months.[27] Like the effects on muscle, the erythropoietic effect of testosterone is relatively rapid and reversible in contrast to the slower effects on bone. Although a higher haemoglobin is likely to provide advantages in endurance rather than power events, it is unclear how much the relatively modest magnitude of this gender difference contributes to the male advantage in athletic performance.

Finally, exposure to adult male testosterone concentrations is likely to produce some mental or psychological effects.[28] However, the precise nature of these remains controversial and it is not clear whether, or to what extent, this contributes to the superior elite sporting performance of men in power sports compared with the predominant effects on muscle mass and function.

The strength of the present study is that it includes a wide range of swimming as well as track and field running and jumping events as well as strength for nonathletes for males and females across the ages spanning the onset of male puberty. The similar timing of the gender divergence in each of these settings to that of the rise in circulating

testosterone to adult male levels strongly suggests that they all reflect the increase in muscular size and strength although the impact of other androgen-dependent effects on bone, haemoglobin and psychology may also contribute. Limitations of this study include that it could not extend to all swimming or track and field events due to the restricted participation of younger age groups in more gruelling events. Furthermore, the testosterone measurements were not from the individual athletes included in the analysis of available published data so that the comparisons are cohort-wise rather than based on individuals.

It is concluded that the gender divergence in athletic performance begins at the age of 12-13 years and reaches adult plateau in the late teenage years. Although the magnitude of the divergence varies between athletic skills, the timing and tempo are closely parallel with each other and with the rise in circulating testosterone in boys during puberty to reach adult male levels.

ACKNOWLEDGEMENTS

The author is grateful to Professor Ken Fitch for helpful advice.

CONFLICT OF INTERESTS

Nothing to declare.

REFERENCES

1. Handelsman DJ, Sikaris K, Ly LP. Estimating age-specific trends in circulating testosterone and sex hormone-binding globulin in males and females across the lifespan. *Annu Clin Biochem*. 2016;53:377-384.
2. Tonnessen E, Svendsen IS, Olsen IC, Guttormsen A, Haugen T. Performance development in adolescent track and field athletes according to age, sex and sport discipline. *PLoS ONE*. 2015;10:e0129014.
3. Malina RM, Slawinska T, Ignasiak Z, et al. Sex differences in growth and performance of track and field athletes 11-15 years. *J Hum Kinet*. 2010;24:79-85.
4. Silverman IW. The secular trend for grip strength in Canada and the United States. *J Sports Sci*. 2011;29:599-606.
5. Beccuti G, Ghizzoni L. Normal and abnormal puberty. In: De Groot LJ, Beck-Peccoz P, Chrousos G, et al., eds. *Endotext*. MDText.com, Inc.: South Dartmouth, MA; 2000.
6. Day FR, Bulik-Sullivan B, Hinds DA, et al. Shared genetic aetiology of puberty timing between sexes and with health-related outcomes. *Nat Commun*. 2015;6:8842.
7. Malina RM, Bouchard C, Beunen G. Human growth: selected aspects of current research on well-nourished children. *Ann Rev Anthropol*. 1988;17:187-219.
8. Sartorio A, Lafortuna CL, Pogliaghi S, Trecate L. The impact of gender, body dimension and body composition on hand-grip strength in healthy children. *J Endocrinol Invest*. 2002;25:431-435.
9. Henneberg M, Brush G, Harrison GA. Growth of specific muscle strength between 6 and 18 years in contrasting socioeconomic conditions. *Am J Phys Anthropol*. 2001;115:62-70.
10. Cheuvront SN, Carter R, Deruisseau KC, Moffatt RJ. Running performance differences between men and women:an update. *Sports Med*. 2005;35:1017-1024.
11. Seiler S, De Koning JJ, Foster C. The fall and rise of the gender difference in elite anaerobic performance 1952-2006. *Med Sci Sports Exerc*. 2007;39:534-540.
12. Thibault V, Guillaume M, Berthelot G, et al. Women and men in sport performance: the gender gap has not evolved since 1983. *J Sports Sci Med*. 2010;9:214-223.
13. Beneke R, Leithauser RM, Doppelmayr M. Women will do it in the long run. *Br J Sports Med*. 2005;39:410.
14. Bhasin S, Storer TW, Berman N, et al. The effects of supraphysiologic doses of testosterone on muscle size and strength in normal men. *N Engl J Med*. 1996;335:1-7.
15. Finkelstein JS, Lee H, Burnett-Bowie SA, et al. Gonadal steroids and body composition, strength, and sexual function in men. *N Engl J Med*. 2013;369:1011-1022.
16. Bhasin S, Woodhouse L, Casaburi R, et al. Older men are as responsive as young men to the anabolic effects of graded doses of testosterone on the skeletal muscle. *J Clin Endocrinol Metab*. 2005;90:678-688.
17. Bhasin S, Storer TW, Berman N, et al. Testosterone replacement increases fat-free mass and muscle size in hypogonadal men. *J Clin Endocrinol Metab*. 1997;82:407-413.
18. Huang G, Basaria S, Travison TG, et al. Testosterone dose-response relationships in hysterectomized women with or without oophorectomy: effects on sexual function, body composition, muscle performance and physical function in a randomized trial. *Menopause*. 2014;21:612-623.
19. Mouser JG, Loprinzi PD, Loenneke JP. The association between physiologic testosterone levels, lean mass, and fat mass in a nationally representative sample of men in the United States. *Steroids*. 2016;115:62-66.
20. Jones NL, Makrides L, Hitchcock C, Chypchar T, McCartney N. Normal standards for an incremental progressive cycle ergometer test. *Am Rev Respir Dis*. 1985;131:700-708.
21. Svedenhag J. Maximal and submaximal oxygen uptake during running: how should body mass be accounted for? *Scand J Med Sci Sports*. 1995;5:175-180.
22. Genberg M, Andren B, Lind L, Hedenstrom H, Malinovschi A. Commonly used reference values underestimate oxygen uptake in healthy, 50-year-old Swedish women. *Clin Physiol Funct Imaging*. 2016; doi: 10.1111/cpf.12377. [Epub ahead of print]
23. Vanderschueren D, Laurent MR, Claessens F, et al. Sex steroid actions in male bone. *Endocr Rev*. 2014;35:906-960.
24. Murphy WG. The sex difference in haemoglobin levels in adults - mechanisms, causes, and consequences. *Blood Rev*. 2014;28:41-47.
25. Krabbe S, Christensen T, Worm J, Christiansen C, Transbol I. Relationship between haemoglobin and serum testosterone in normal children and adolescents and in boys with delayed puberty. *Acta Paediatr Scand*. 1978;67:655-658.
26. Coviello AD, Kaplan B, Lakshman KM, Chen T, Singh AB, Bhasin S. Effects of graded doses of testosterone on erythropoiesis in healthy young and older men. *J Clin Endocrinol Metab*. 2008;93:914-919.
27. Bachman E, Travison TG, Basaria S, et al. Testosterone induces erythrocytosis via increased erythropoietin and suppressed hepcidin: evidence for a new erythropoietin/hemoglobin set point. *J Gerontol A Biol Sci Med Sci*. 2014;69:725-735.
28. Celec P, Ostatnikova D, Hodosy J. On the effects of testosterone on brain behavioral functions. *Front Neurosci*. 2015;9:12.

How to cite this article: Handelsman DJ. Sex differences in athletic performance emerge coinciding with the onset of male puberty. *Clin Endocrinol (Oxf)*. 2017;87:68–72. https://doi.org/10.1111/cen.13350

UNITED STATES SPORTS ACADEMY

ASSESSING THE POTENTIAL TRANSGENDER IMPACT ON GIRL CHAMPIONS IN AMERICAN HIGH SCHOOL TRACK AND FIELD

A dissertation submitted to
the faculty of the United States Sports Academy
in partial fulfillment of the requirements
for the degree of

Doctor of Education

in

Sports Management

by

Gabriel A. Higerd

chair: Dr. Brandon Spradley

Daphne, Alabama

December 2020

UNITED STATES SPORTS ACADEMY

Approval
of a dissertation submitted by

Gabriel Higerd

Brandon Spradley, Ed. D.
Chair, Dissertation Committee

December 2020
Date

Fred J. Cromartie, Ed. D.
Member, Dissertation Committee

December 2020
Date

Bonnie Tiell, Ed. D.
Member, Dissertation Committee

December2020
Date

Tomi Wahlström, D. M.
Provost/CAO

December 2020
Date

DEDICATION

To my five beautiful and wonderful girls. May you grow in stature, knowledge and wisdom; and fulfill the purposes for which you were made. I love you forever and unconditionally.

ACKNOWLEDGEMENTS

The researcher is deeply grateful for all those who have helped me along the way in this journey; the researcher could not do it on my own.

Above all, to my wife Mari, whose contributions to my success in every way cannot be measured nor appropriately accounted for. To my children, who shared workspaces, bandwidth, and their dad with this project. To my parents and grandparents, who placed high value and priority on higher education and supported me all along the way.

To the football coaches who dramatically shaped who the researcher is, in particular Victor Santa Cruz, Brian Willmer, Peter Shinnick, Dr. Sean Rochelle, Tom Gault, plus the myriad of assistant coaches who the researcher has had the pleasure to work with and be coached by. Among them are Bo Beatty, Rudy Carlton, Ben Buys, Spencer Danielson, Justin Riddle, Manako Tuafua and so many more that the researcher loves and appreciates. To my colleague and NFL Hall of Famer, Jackie Slater, for his friendship and mentorship over the years.

The researcher is grateful to the Army leaders who built into me during a pivotal period in my life, and the soldiers and cadets that this researcher was honored to lead and mentor. thankful for the faculty at Concorida University Irvine, who guided my master's studies and instilled a desire to keep learning and growing. To LTC Mark Goeller, who helped me secure my first teaching position at Claremont McKenna College. To the staff at Azusa Pacific University, for their care for others. To Dr. Sharron Lehman who gave me my first teaching position in the kinesiology department and who recommended her alma mater, the United States Sports Academy, as a good fit for me to pursue my

doctorate. To Dr. Eric Sorenson for his mentorship as the researcher began teaching in sport and exercise science. To the thousands of students and players over the years, who allowed me to live out my passion for teaching, training, and competing, it is them that the researcher misses most when not on the field or in the classroom.

The researcher is grateful to the faculty and staff at the United States Sports Academy. In particular, the researcher would like to thank my committee members for their encouragement, advice, and guidance, Dr. Fred Cromartie, Dr. Bonnie Tiell, and my committee chair Dr. Brandon Spradley. Their enthusiasm and support of my work went a long way in buoying my spirit and mindset. To Dr. Liz Weaver of the United States Military Academy at West Point, the researcher is grateful for her careful reading of the manuscript, including edits and suggestions.

Finally, to my grandfather Andrew Kubica "Gramps," a farm boy, WWII veteran, rocket scientist literally responsible for putting a man on the moon, father of 5 girls, and grandfather of 11. May his oft-repeated refrain echo in my ears and remain true, "promise me this, that you will never become an educated idiot."

TABLE OF CONTENTS

LIST OF TABELS

LIST OF FIGURES

ABSTRACT

Higerd, G. A., Doctor of Education in Sports Management at the United States Sports Academy: presented December 2020. Title: *Assessing the potential transgender impact on girl champions in American high school track and field*. Chair: Dr. Brandon Spradley.

The intersection of the transgender movement and sport has caught the attention of average Americans as well as sporting bodies and regulators. The rise in the numbers, as well as the acceptance of transgender individuals, have accelerated the need to create modern transgender sport policies. The goal of this research was to assist those seeking to make informed, evidence-based transgender policy decisions. Accordingly, the purpose of the study was to:

1. Investigate the underlying basis for post-pubertal sex segregation in sport.
2. Assess the effect of event distance on the performance differences between the sexes.
3. Assess the probability of a girls' champion being biologically male (46, XY).

The research included a three-question quantitative design investigating the scope and scale of sex differences in high-school track and field, and the implications of sex differences on the probability of transgender disruption of the female classification. The study focused on biologically driven performance differences and the prevalence of *potential female champions* (PFCs) (males better than the best female) that may be *male to female* (MTF) transgender athletes.

The study investigated roughly one million American high school track and field performances (N = 920,115) available through the track and field database Athletic.net. In the sample, 400,929 were female (46, XX) and 519,186 male (46, XY), which included five states (CA, FL, MN, NY, WA), over three years (2017 – 2019), in eight events; high jump, long jump, 100M, 200M, 400M, 800M, 1600M, and 3200M.

The research first addressed the question: Is there a statistically significant relationship in the performances of female and male high school track and field athletes? Second, is there a statistically significant relationship between event distance and the percentage of males that are superior performers to the best female? Third and finally, is there a statistically significant probability of one or more 46, XY MTF transgender individuals being a girls' champion in an event?

A z-test was used on data from each event to analyze the relationship between sex and performance. Correlation and regression assessments analyzed the relationship of event distance and sex, as represented by the percentage of PFCs. A Monte Carlo random number generation simulation, consisting of 1,110,000 trials through 111 simulations, compared transgender population estimates and known PFCs in the selected events to project the theoretical MTF transgender density at the top of the female field.

Results for the first question indicate in each of the eight events the null is rejected in favor of the alternative hypothesis: There is a statistically significant positive relationship between performance and being 46, XY ($p < .001$), with mean differences in performance by sex ranging from 14% at the low end in the 100M, to 24% at the high end in the long jump, and the mean difference of all the events is 18% in favor of males.

Additional evaluation of the performance distributions reveals the average male performance is better than 94%-98% of female performances (top 2%-6% of the female field). The average female performance is worse than 93%-97% of male performances (bottom 3%-7% of the male field). Approximately one-third or more (32%-43%) of male performances fit within the top 1% of female performances.

Overall the participation was 44% female and 56% male, but the participation gap varied from 14% to 50% in favor of boys' dependent on the event, with a strong correlation ($r = .93$, $p < .001$) between participation percentage and distance, showing girls have higher participation rates in comparison to boys in events that are more dependent on power and speed, and less participation in a percentage comparison in events that rely on endurance.

The findings for the second question reveal a moderately positive relationship when comparing percent PFC and distance ($r = .31$, $p < .001$). However, post hoc analysis of performance alone suggests that there is not a statistically significant relationship between distance and mean difference in performance ($r = -.19$, $p = .652$). The smallest gap in performance, and average percentage of male PFCs, occurred in the 100M, with larger gaps occurring at 400M and beyond.

The results for the third question indicate that if transgender population density estimates were true and representative of high school track and field athletes, and if being transgender was independent, uniformly distributed attribute among the 46, XY sample, there is a simulated 81%-98% probability of transgender dominance occurring in the female track and field events. Additionally, in the simulation trials where there was at least one transgender PFC, there was an average of two to three MTF individuals. Thus, in the majority of cases, the entire podium (top performers in the state) would be MTF transgender athletes.

The data provides sufficient and strong evidence to support post-pubertal sex segregation in sport. It presents insufficient evidence that policies should be tailored by event distance. Finally, since female sport is an invaluable asset and societal good, the

findings provide critical data for policymakers to make informed, evidence-based decisions that protect and promote competitive female sport.

CHAPTER I

INTRODUCTION

Female sports have developed into a vibrant and empowering outlet for millions of girls, and the rise in opportunities given to girls in America has become a point of national pride. Recently, however, the intersection of the transgender movement and female athletics has spawned controversy. The conflict between the transgender movement and competitive female sport has become a topic on the minds of sporting bodies big and small.

Presently, transgender sport policy is gaining national attention. Lawsuits and countersuits are working their way through the courts; championships are being won and lost by transgender individuals (Randnofsky, 2020). Some states are restricting their laws, whereas others are liberalizing to become more transgender-inclusive (Tamerler, 2020). Weighing in on the issue have been diverse voices such as presidential candidates, prominent feminists, Lesbian Gay Bisexual Transgender Questioning Intersex Asexual (LGBTQIA+) leaders, evangelical pastors, news outlets, and sports governing bodies such as the International Olympic Committee (IOC), the National Collegiate Athletic Association (NCAA), and the National Federation of State High School Associations (NFHS) (Raff, 2020).

The massive attention on transgender policies in sport is warranted. Transgender sports policy, at least in theory, has the potential to fundamentally transform female athletics. Academics and cultural observers suggest that beyond just disrupting the status

quo, that the transgender moment could, and some contend, "should," eliminate the binary of male and female sports (Knox et al., 2019).

A thorough investigation is warranted into the performance aspects of transgender athletes competing in the female classification. Although the body of evidence documenting sex differences in the human species is vast, the research applying differences to sport performance in light of the transgender movement is scant. Throughout the mass of controversy, the call for more research has come from all sides. This topic is evolving at a rapid pace, often with policy outpacing the body of evidence (Coleman, 2017). Credible research is needed that will provide an evidence-based framework for transgender sport policy—one where the interests of all parties are evaluated in light of the data. This study is a quantitative evaluation of an estimated one million high school track and field performances to probe questions related to biologically driven performance differences and potential consequences of particular transgender sport policies.

Statement of the Problem

Sports governing bodies exist to foster competition and establish rules of the contest (Keating, 1973). Competition is possible only if there is at least the appearance of universal rules applied fairly, theoretically allowing all participants an "opportunity" to prevail and achieve meaningful competition. Without consistently applied rules of the game, the resulting institutionalized unfairness becomes an assault on the heart of sport. The main issue at hand is whether transgender inclusion in female athletics constitutes an incursion on fair competition.

Segregation by Sex

Who gets to play is a fundamental question sporting organizations must answer. Youth sports often restrict participation based on criteria such as age or size. Golf restricts participation based on earning one's player card, combat sports and weightlifting restrict participation based on body weight, and most restrict participation based on sex. Of the myriad of restrictions, the one that has arisen as controversial, and only as of late, is sex. The rise of the transgender movement has the sports world, and certainly sporting organizations, on edge.

Separating the sexes for purposes of sport has a long history. Sex segregated sport via institutionalized policy, or by natural kinds, or a combination of both, is pervasive. This binary stretches throughout virtually every sporting endeavor, in every culture on planet earth, with few exceptions (Harper, 2019). Perhaps it is because of the width and depth of this categorical distinction that the transgender movement in sport is seen as a topic of monumental consequence.

From ancient times, the dichotomy of male and female has been ever-present, if not ever important. Sex distinctions go back to the B.C.E. era and can be seen, to one extent or another, in every culture that has ever existed (Costa & Guthrie, 1994). Since the rise of sport in the modern world, highlighted by the reintroduction of the Olympic Games in 1896, male and female divisions in sport have been on the global stage ("The women of Sparta," 2012). Unlike the vast global variance in the interaction between the sexes seen on an interpersonal and societal basis, sports participation has been surprisingly uniform in its application of sex distinctions. If anything, the differences in sex are highlighted and magnified in sports.

The separation of participants by sex, for purposes of sport competition, was never controversial in toto. To be sure, there were, and are, controversies over the status of the separate sex-based category (Jones et al., 2017). For example, valid claims of females being separate and not equal exist, but few voices have been decrying the very existence of separate sex-based categories.

Sex has been shown to be such a highly correlated variable to sport performance that virtually all meaningful competition, everywhere on the globe, has divided participants on sex (International Olympic Committee, 2017). Assuming that the existence of a binary male and female sports is a settled issue, and it is but for the most committed deconstructionists such at Buzuvis (2016) who view "the separation of men's and women's sport itself contributes to the stereotype of female athletic inferiority" (p. 48); the present controversy resides "upon which basis" organizations separate participants into the two categories of male and female. Sporting bodies seek to align natural kinds or "like-with-like" to the best of their reason and ability when assigning categories of competition, and thus they face the question: What should be done, and how to accommodate transgender individuals who do not identify with the biological sex that they were embodied with?

The Need for Research

Significant social and societal level structural changes to sports are being contemplated and undertaken to accommodate transgender athletes. Sports governing bodies are considering and making policies to address the new and rising occurrence of transgender athletes in sport, with dramatic discrepancies between institutions. Some American states have adopted firm interscholastic stances that reserve participation on

the basis of biological sex, while others with equal conviction and fervor, are basing participation on gender identity (Raff, 2020).

As the transgender movement is gaining wider acceptance in society, its collision with the male and female athletic binary is escalating tensions. Passionate advocates are lined up on all sides of the issue. It seems the one thing all can agree on is that there is a lack of research on the topic (Anderson, 2018; Jones et al., 2017). Investigations into the underlying performance differences in the sexes and the implications for transgender sport policies are desperately needed. This dissertation addresses this research gap.

The study investigates the underlying basis for post-pubertal sex segregation in sport, the effect of event distance on the performance differences between the sexes, and assesses the probability of a girls' champion being biologically 46,XY. This research provides unique and original findings for a particular population that has not been fully studied. The resultant data from the hypotheses testing, points to the extent and degree of the impact of transgender participation on female sport. It can inform governing bodies who are considering policies that may reshape the sporting world. The study addresses whether transgender athletes will rise to dominate female athletics, or if their inclusion is immaterial to the competitive framework of girls' sports.

Research Questions and Hypotheses

The study is paramountly concerned with biologically driven performance differences and the prevalence of *potential female champions* (PFCs) that may be *male to female* (MTF) transgender athletes. A PFC is an athlete competing in the male sport classification whose performance is better than the top female in the event in the state. For example, if the 20 fastest coed runners in an event were evaluated and the fourth

fastest athlete is a 46,XX (female) runner, the three 46,XY (male) runners in front of the 46,XX athlete are PFCs.

The primary aim of the study is to investigate the statistical probability of one or more athletes being both an MTF transgender person and a PFC, in eight selected high school track and field events, among representative states. The findings are likely to carry implications for sports policy development and legislative efforts.

Research Questions

1. Is there a statistically significant relationship in the performances of 46,XX (female) and 46,XY (male) high school track and field athletes, in selected events?
2. Is there a statistically significant relationship between event distance and percentage of potential female champions?
3. What is the probability of one or more 46,XY potential female champions also being an MTF transgender individual? P(n[PFC and MTF] $\geq$ 1).

Hypotheses

Question 1 Hypothesis

$H1_0$: There is not a statistically significant relationship in the performance of the 46,XY and 46,XX high school track and field athletes.

$H1_a$: There is a statistically significant positive relationship between performance and being 46,XY.

Question 2 Hypothesis

$H2_0$: There is not a statistically significant relationship between the percentage of PFCs and event distance.

$H2_a$: There is a statistically significant relationship between the percentage of PFCs and the event distance.

Question 1 utilizes a z-test, with the z-score as the test statistic and a significance α of .05. This question will help in assessing whether it can be established in the events selected at the high school level that there is, or is not, a statistically significant difference in performance between male and female persons. The results will show if the performance distributions are normally distributed or skewed right or left, along with the kurtosis for the two sex classes. It will show the standard deviation within female and male performances and the possible difference between groups. These results assist in confirming the existence of a statistically significant relationship and bimodal distribution of sex and performance among the sample.

Question 2 assesses the correlation between event distance and percent PFC with the test statistic being Pearson's coefficient *r* and significance established at *r* =.5 to 1.0, or *r* =.-0.5 to -1.0. The test is non-directional in case there are surprising results that go against the theory of endurance leveling the playing field. The charts will show how many 46,XY in each event are PFCs. It will be of interest to know if this PFC percentage size similar across events.

Question 2 will help in assessing whether shorter events, relying more predominantly on maximal strength and power (adenosine triphosphate phosphocreatine [ATP-PC] bioenergetic pathway), have a larger percentage of PFCs than events that rely on more anaerobic endurance or aerobic capabilities. This question will help illuminate if there are certain sporting events that are less impacted by a 46,XX and 46,XY difference. The implications could point to a graded response to transgender inclusion based on

event distance, rather than blanket application statewide, similar to the International Association of Athletics Federations (IAAF) (2019) policy governing testosterone concentrations more vigorously at certain event lengths.

Question 3 will consist of a Monte Carlo simulation asking what is the probability of one or more 46,XY PFCs in a state, in an event, also being an MTF transgender individual. The researcher conducted 10,000 random number simulations per event, per state (111 total) to assess the probability that 46,XY athletes could be the state female champions in the events selected in a given year: P(n[PFC and MTF] $\geq$ 1).

Presumably, Americans may differ significantly on what they deem an acceptable probability on the question of transgender champions. However, this study's strength will not be in its satisfaction or non-satisfaction of an arbitrary significance of statistical probability; its strength and value to the scientific community will be in its illumination of the probability by using such a vast dataset. A statistically high probability has implications for restricting participation in the female category to 46,XX persons only; ensuring females win and are able to take part in meaningful competition. An insignificant statistical result could be instructive in that it would support keeping and expanding more transgender-inclusive policies that have unrestricted access to the female classification.

Definitions of Terms

Writing for Accuracy and Clarity in a Changing Field

In light of the transgender movement, extensive elaboration on definitional terms is needed. Rapid cultural changes have progressed at an exponential pace. The shift in culture has accompanied a linguistic revolution. Linguistics intersecting with the

transgender movement is so much at the forefront of the cultural conversation that Merriam-Webster declared the singular "they" as its 2019 word of the year (Word of the year 2019, n.d.). The acronym device LGBTQIA+ continues to grow increasingly long. In light of the rapid changes, further explanation is needed in the researcher's attempt to be both non-antagonistic and accurate.

One could be excused if they are struggling to keep up with the latest LGBTQIA+ vernacular. For example, if one were to embrace the full logic and language of the transgender movement, the following are true as of 2020: NCAA policy allows for a women's soccer team to be made up of one hundred percent persons labeled men (NCAA, 2011); an athlete that can be simultaneously labeled a man, and yet the women's national champion; the women's Division III hammer throw record holder is identified as a man (Harper, 2019). To many, a rising number of women need to be screened for testicular cancer, and an increasing amount of men have birthed a child (Hattenstone, 2019). To any other generation in human history, the reader likely would have significant difficulty comprehending preceding sentences such as these. Some such as Anderson (2018) reject this progression from historical norms.

Given the current environment, the researcher must make some linguistic decisions for the sake of scientific accuracy and clarity; for the sake of effective communication and understanding. Navigating the progressive and sometimes contradictory taxonomy and definitions is not an easy task. In exploring this topic, the researcher will attempt to definitively and effectively identify persons while trying to reasonably avoid controversy and scorn from either side of the transgender movement.

Operational Definitions

46,XY– A human being with the genetic karyotype of 46 paired chromosomes with XY distinctions. Normal function characterized by testicular gonads, if not removed, that contribute to male development, especially in puberty—historically known as a boy, man, and male.

46,XX– A human being with the genetic karyotype of 46 paired chromosomes with XX distinctions. Normal function characterized by ovaries that contribute to female development, especially in puberty—historically known as a girl, woman, and female.

Cross Sex Hormone Therapy (CSHT) – hormonal interventions that are intended to inhibit the natural hormonal production and response of an individual, and artificially replicate, to some extent, the hormonal status of the opposite/desired sex.

Disorders of sexual development (DSD) – A variety of chromosomal and developmental conditions that affect the normal sexual and/or hormonal expression and development. Often referred to as "differences" of sexual development to avoid inferring the potentially negative connotation that the individual is disordered.

Female champion – The top official 46,XX performance in a given event, in a given state, in a given year. This performance need not have been accomplished in the finals of the state championship, hence they may not be officially recognized in the record books as state champion.

Female to male (FTM) – A 46,XX individual who has a transgender man/boy gender identity. Understanding that the term could more accurately be "woman to man" rather than "female to male" for coherence; but due to the acceptance in the literature, the researcher will sacrifice accuracy and coherence for familiarity.

Gender – "The behavioral, cultural, or psychological traits typically associated with one sex…the external expression of sexual status as male or female" (Merriam-Webster, n.d.). "The attitudes, feelings, and behaviors that a given culture associates with a person's biological sex" (APA, 2019, p. 138).

Gender identity – The subjective identification of one's gender.

Intersex – A catchall term typically associated with disorders of sexual development (DSD). Sometimes presents in an individual possessing an external phenotype that is incongruous to their hormonal production. A prominent population that is usually conflated into the transgender discussion of sport, especially given prominent decisions by the Court of Arbitration for Sport (CAS) and the International Olympic Committee (IOC). Most notably different than a transgender person in that they are typically raised as a gender from birth that may be in opposition to their hormonal function, rather than the typical self-identification of transgender individuals.

Male to female (MTF) – A 46,XY individual who has a transgender woman/girl gender identity.

Misgender – Using vocabulary, especially pronouns, to describe someone that does not reflect the gender identity of the individual. Misgendering is technically about gender, but like most things in the transgender movement, in practice, it conflates sex and gender. For example, it would typically be frowned upon to say someone is a female man or a male woman, even if the statement was definitionally true.

Normal – Conforming to the standard or common type; usual; free from disease or malformation, or from experimental therapy or manipulation.

Non-binary – A variety of terms to identify an individual whose gender identity fluctuates between male and female, both, or neither. A spectrum of gender identities that are outside the gender binary and not exclusively male or female (e.g., *Genderfluid, Bigender, Agender, Pangender, Third Gender, Genderqueer, Two-Spirit).*

Potential female champion (PFC) – A 46,XY athlete who had a performance in a given event, in a given state, in a given year, that was greater than the female (girls' classification) champion.

Sex – The status of being biologically male or female; normally expressed with congruent genitalia, hormone, and reproductive capacity.

Transgender – An individual who ascribes to a gender identity that is counter to their biological sex.

+ – A symbol to indicate an ever-expansive list of gender, sexuality, lifestyles, and identities.

Conflating Gender and Sex

The researcher will seek to avoid conflating "gender" and "sex." Most definition sections in the field of gender and sexuality begin by declaring that gender is the sense that one has of being a man or woman, and that sex is a status assigned at birth, annotated as male or female, and that the two are not exclusive. Paradoxically, it is exceedingly common in academia, especially in the hard sciences, to conflate the two throughout the literature. When data describes "men" on average being taller, "women" going through menopause, "boys" beginning puberty typically after "girls," researchers are conflating gender and sex (Round et al., 1999). Frequently, articles that begin with defining gender as a social construct, go on to violate that definition and conflate sex in gender in their

work. As gender typically is defined as a social construct or internal identification of being male or female, if the standard definitions hold, it implies that there could be male women and female men.

Notably, the American Psychological Association (2020) definitionally separates sex and gender, but it then seems to endorse conflating the terms. E.g., "use specific nouns to identify people or groups of people…use 'male' and 'female' as adjectives," but "in general, avoid using 'males' and 'females' as nouns; instead use 'men' and 'women'" (p. 139).

Difficulty with Non-binary and Gender-fluidity

Gender-fluidity is the notion that gender is a spectrum rather than fixed positions, and that individuals reside anywhere along the spectrum from fully man/masculine to fully woman/feminine or anywhere in between, and with no fixed position. In effect, the individual alone is the only one competent to know where they are on the spectrum at any one time. Someone could be far along the man/male spectrum at one moment, walk into another room, and be fully woman/female. If this logic is embraced, it is problematic to even use the terms cisgender, transgender, etc., as they only apply definitively to the past, rather than the present or the future.

The rise of transgender and non-binary identification poses questions as to when it is ever appropriate to assume one's gender. Headlines such as "police looking for man," or "women's body found" are presumptuous, as gender requires direct interrogation of the individual. The man who won the race last weekend could be the woman who won the race today. Adding to the difficulty is the fact that bodily appearance is not a relevant factor in gender. According to this worldview, a woman may be a lumberjack with a

flowing beard, and the Victoria's Secret model may be a man; accordingly, following this logic, activists claim, "the best and only fair way to determine if someone is male or female is to ask them" (Buzuvis, 2016, p. 47). Philosophically and taken to its logical conclusion, the non-binary worldview disarms from ever making pronouncements on man or women, unless it was about oneself (personal pronoun) or if it had been immediately preceded by a first-hand proclamation of the identity by another. Even this would run the risk of misgendering someone who changed identities since the last known proclamation.

Difficulties with Chronological Reflections of Gender

The researcher, when using pronouns, will refer to a person in the pronoun of the time period of reference, rather than retroactive application. For example, the researcher will not suggest "she," Caitlyn Jenner, was the world's greatest athlete in 1976, setting a world record in the decathlon. Because by linguistic extension, if Jenner were a woman in the year 1976, Jenner would be, athletically speaking, the greatest women's athlete in world history (Litsky, 1976).

Difficulty with Personal Pronouns

Personal pronouns such as they, ze, zir, xe, hir, per, ve, ey, hen, etc., have recently found their way into academic literature. The APA Publication Manual (2020) states, "use 'they' as a generic third-person singular pronoun to refer to a person whose gender is unknown" (p. 120). (Yet the APA has a separate standard for animals, one in which sex determines pronoun usage). To avoid misgendering someone, a very real threat in the era of gender fluidity, as shown above, the researcher seeks to avoid using personal pronouns altogether where possible. When gender, sex, and pronouns are used, it will be used to

refer to the classical historical understanding of gender and sex as interchangeable or in terms that the sports organizations themselves use, such as “girls” and “boys” for American high school track and field. All other uses of he or she, man or woman, boy or girl, male or female, are good-faith assumptions of the researcher and not intended to offend or misgender.

Delimitations on Taxonomy

Since gender for many has become an objectively meaningless and biologically undefinable reality in many circles, the researcher will attempt to abide by sex, biological sex, or physiology, for the identification of subjects to the greatest extent possible. The researcher will use the standard of forensic anthropologists, that is, experts who are called upon to make post-mortem determinations of sex. These experienced experts can quickly and routinely assess sex by examining anthropometric traits alone (Langley & Dudzik, 2016). In such an evaluation, the gender identity of the individual under examination is inconsequential to the determination of male or female. However, although it would be scientifically accurate to make distinctions between gender and sex, in practice, it has the potential to agitate animosity. Therefore, in light of the controversy and with accuracy in mind, for the most part, the researcher will use 46,XY and 46,XX, where historically one would place male/man/boy and female/woman/girl.

Scope of the Study

The study examines an estimated one million American high school track and field athlete performances ($N = 920{,}115$). Slightly more performances from the 46,XY category are observed ($n = 519{,}186$ vs. 400,929), similar to the greater male participation observed nationwide as shown in available participation data (National Federation of

State High School Associations, 2019). The performances are assessed from five states, using eight events, over a three-year period (2017, 2018, and 2019).

Data was extracted from official results tracked through a paid subscription to athletic.net®. Official results posted from athletic.net require that they are uploaded from the meet host, and must report all places, all events, all participants, and all marks of that meet. Results were filtered by fully automatic time (FAT) only, over the course of the three outdoor seasons, 2017, 2018, and 2019, in the following eight high school track and field events: high jump; long jump; 100 meters; 200 meters; 400 meters; 800 meters; 1600 meters; 3200 meters. The subjects originate from five states representative of five regions of the United States: Northwest; West; Midwest; Northeast; Southeast. The states selected are: California; Florida; Minnesota; New York; Washington.

Each state selected has both a high number of participants and are "fully inclusive" (i.e., not requiring any hormonal or biological interventions for 46,XY persons to compete in the female classification). The states selected have a high number of track and field participants (among the largest in their region).

Delimitations

The researcher made multiple delimitations as to the level, event type, state, and other pertinent characteristics of the population under consideration.

Level Selection

The researcher chose subjects in American high schools, with the rationale that high school sports are integral to the fabric of American life. It is at the high school level that serves a point of contention for policymakers throughout the nation. Many watch the Olympic/elite level of sport and appreciate the controversies with transgender, intersex,

and doping athletes. However, whereas the elite level is followed in a voyeuristic manner, high school sports occur in the local neighborhood. High school sports are the competition ground for the masses and have implications for millions of young people and their families; therefore, an important population to study.

Event Selection

Measurements of validity and reliability are critical to the scientific method. This is why the researcher chose track and field events. Their standardization in measurements, timing, and results are well suited to post-hoc data analysis. The researcher selected the eight events because they progress from high-force, short-work durations, in a linear direction toward lower-force and longer-work durations of performance. It has been demonstrated that certain events have a lower divergence of performance in the sexes, and the IOC has adopted hormone rules that are restrictive for certain events and not for others (IAAF, 2019). The researcher wanted to know whether such policies would be supported by the data at the high school level.

The researcher chose the specific events because of their status as “meter” and “seconds” sports. As such, they are subject to a direct comparison between the sexes. Sports such as basketball, soccer, and wrestling are examples of sports where sex certainly contributes to performance, yet the nature of those sports limits direct scientific comparison between the sexes. Some meter sports such as shot put, javelin, and discus would be exceptionally valid to determine upper body differences in the sexes, but the throwing instrument weight difference makes direct comparisons difficult. Other sports such as weightlifting and powerlifting are similarly difficult to compare because of their

differential weight classes. A comparison of these events requires advanced weight classification modeling that is beyond the scope of this study.

State Selection

The researcher is deciding to investigate states that are important to the discussion of transgender sport policy. The selected states contain a valid population of tens, and even hundreds of thousands, of participants in both female and male categories. The researcher chose to select states that currently had very progressive transgender policies that have little or no transgender restrictions.

Regional differences may influence sports participation. The Williams study reflects a state and regional difference in terms of transgender identification (Herman et al., 2017). Therefore, trying to capture regional performance differences is important, as it will strengthen the study and hold national implications for the findings. Not all states have the same impact. Some are more significant than others to the national climate. The researcher selected the states in their respective regions that arguably have the greatest impact on high school athletics. The states selected are regional heavyweights.

Year Selection

Another decision the researcher is making is to investigate three years. Outliers, by definition, can confound the data analysis. The researcher chose multiple years to strengthen the study's findings. By taking the mean of three years, there is less of a risk of an outlier affecting the data set. If one of the states studied happened to have an exceptionally dominant 46,XX performer, it is unlikely that the athlete's dominance would persist over the course of three years; the inclusion of that time dilutes her outlier performances.

Additional Delimitations

Accessing valid data is critical to any research project, and this is no exception. The researcher chose the data resource company athletic.net for the data because of its status as a trusted authority, has national standards for data reporting, and processes for minimizing and correcting data errors. This source has the additional benefit as an easily accessible data source and shall allow for future researchers seeking to replicate or expand upon the study's findings.

The researcher chose Question 1, in regards to a significant difference between 46,XY and 46,XX persons athletically related physiological attributes, because it will establish if there is a scientifically verifiable difference between 46,XY and 46,XX persons. The existing evidence would suggest that there will be a difference, but this study will be enlightening by showing the extent of the difference in this particular population.

FTM transgender athletes are not the focal population in this study. Without hormonal intervention, there is limited controversy or disruption of sporting competitiveness of boys' athletics. FTM transgender persons on CSHT who are experiencing and anabolic steroid performance boost are still deemed as not acquiring a benefit that would make their participation as unfair (NCAA, 2011). In other words, the fastest 46,XX runners could be utilizing CSHT to acquire an anabolic steroid effect, yet they would still not be deemed as possessing an unfair advantage over 46,XY athletes.

The research avoids a particular focus on persons with disorders of sexual development. Recent high-profile cases at the elite level have involved DSD persons, rather than transgender athletes. However, the rules emplaced to address DSD persons

with hormonal restrictions are applied uniformly and, therefore, relevant to both populations. The question of DSD athlete participation is important and relevant, yet more nuanced and complex, as there are dozens of DSDs that manifest, some affecting XX and other XY persons. Some XY persons could have ambiguous genitalia or female phenotype so that, by no decision of their own, are raised female. The fairness of these persons should be considered elsewhere as a recent women's world championship 800 meters podium consisted of three athletes with DSDs (Harper, 2019).

Finally, the researcher will choose to largely avoid the question of hormonal interventions to blunt 46,XY advantage. Elite and collegiate competition requires hormonal interventions for a period of time in order for a 46,XY person could join the female classification. Other researchers are studying if these are valid criteria to protect fairness in female sports (Coleman, 2017). This study is not focused on whether hormone rules appropriately eliminate the advantage that 46,XY persons, who have gone through male puberty, possess.

Use of the Monte Carlo Simulation

The selection of the Monte Carlo simulation was deemed a good fit for estimating the probability of the hypothetical population of MTF PFCs.

> The Monte Carlo method is defined as representing the solution of a problem as a parameter of a hypothetical population, and using a random sequence of numbers to construct a sample of the population, from which statistical estimates of the parameter can be obtained. (Lund, 1981, p. 1)

Limitations

There are several limitations of this study. One limitation is that it is dependent on the Herman et al. (2017) Williams Institute at the University of California at Los Angeles report on transgender population numbers being accurate (or understated). If the true transgender population is less than estimated by the Williams Institute, then the probability of transgender champions in the findings would be reduced. Similarly, if transgender numbers prognosticated in the Williams report are unrepresentative of the transgender athlete populations, the strength of the findings will be diminished.

Another limitation of the study is that it fails to consider the male upper body performance advantage. Most team sports, which includes that vast majority of high school athletes, are dependent to some degree on the contribution of the upper body. This study's exclusive reliance on lower body dominant events likely will mitigate the difference in 46,XX and 46,XY performance, since the gap in upper body performance in the sexes is greater than the gap in lower body performances (Sandbakk et al., 2018).

A potential limitation is that some girl champions may, unbeknownst to the researcher, be a 46,XY MTF transgender person or have a DSD that confers a material athletic advantage. This limitation is mitigated by no known public controversy in the events in each of the states regarding DSD or transgender persons, and the three-year scope has a buffering effect on such occurrences.

Assumptions

The researcher assumes several things about the data. The first is that the official data inputted into the data site are accurate and correct and that the results were from track meets that adhered to all applicable rules, regulations, and measurement protocols.

The researcher assumes a number of notions about the transgender population. First, the research assumes that the number of transgender athletes statistically mirrors the number of similarly aged transgender young adults in the given state. Second, the researcher assumes that the transgender population distribution between accomplished athletes, represented as PFCs, and poorly performing athletes are the same. The researcher also assumes that being a PFC is independent of being MTF. Finally, the researcher assumes that being transgender has some sense of permanency, and hence the athlete would not be a transgender athlete one day and a non-transgender participant the next. This assumption, by nature, fails to account for gender diverse persons (i.e., non-binary, two-spirit, gender non-conforming, questioning, etc.).

Significance of the Study

At a critical time where local, state, and national governing bodies are considering transgender policies and the cultural conversation over the issue is high and contentious, the need for objective data on the subject is great. Even as emotional and philosophical argumentation play a role in policymaking, the establishment of facts through rigorous scientific inquiry is critical to informed, evidence-based decisions. This study will be significant in that it seeks to answer if and to what extent 46,XY persons differ from 46,XX persons at the high-school level, in lower body dominant track and field events. There has not been a study of similar scope and scale regarding the transgender movement that has such a direct relationship to American high-school athletics.

This study is significant to the world of sports management if it both does, or does not, affirm the alternate hypotheses. Findings that produce a failure to reject the null, are just as important as findings of statistical significance. Significant findings would have

critical implications for protecting the girls' sporting classification for the sake of competitive fairness and meaningful competition. Conversely, if the alternative hypotheses are not supported by the evidence, sports governing bodies would not be able to use performance as the sole rationality for limiting the girls' category to 46,XX persons only.

CHAPTER II

REVIEW OF LITERATURE

Virtually all competitive sport is segregated based on sex. This literature review will explore the body of evidence and arguments for having such a system. The transgender movement presents a challenge to the sports organizations and legislators who oversee the rules, regulations, and laws that keep the sexes in a segregated posture.

It is critical to evaluate the claims of those in the exercise physiology field as to the nature and scope of sex differences. The central justification for segregation is to provide an equal opportunity to those in the female classification to experience meaningful competition and opportunities for success. The review will evaluate anatomical and physiological evidence of sex differences that are pertinent to athletic performance.

Since the equality statute Title IX has been so monumental in the American sports world, it is important to review its origin and effect. The review traces the origin of the law through the subsequent interpretations and to the current status of Title IX.

A work on transgender policy, would not be complete without reviewing the transgender movement in general and its relation to sport in particular. This review includes the philosophical positions of the various parties and questions of fairness and inclusion. These aspects will affect how policy will shape the future of sport.

The review is arranged in five parts. Part I will review the literature regarding innate differences in biology and sex differentiation. It begins with human development and the emerging osteological, body mass, cardiorespiratory, metabolic, and nervous

system differences following the onset of puberty; and documents strength, power and athletic performance differences along with the role of testosterone. Part II will review the equality statute of Title IX and its relation to American sport, including a review of the law's origins, regulations, and subsequent interpretations. Part III will review the transgender movement in America, to include several philosophical views on gender and sex. The section looks at key historical developments in the movement, along with the recent assessments of the status of the transgender population in America. Part IV will review issues related to policy-making regarding the transgender movement. It reviews the various arguments for and against transgender inclusion in the female classification, along with the varying approaches to policy development, and it will look at current policies and recent developments that will influence future policymakers. Finally, Part V will briefly review the Monte Carlo simulation as a method for determining probabilities.

Part I: Innate Differences in Biology and Sex Differentiation

Advances in the fields of genetics, endocrinology, urology, epidemiology, taxonomy, physiology, and others, have produced an unprecedented body of evidence regarding sex and sex differentiation. A thorough discussion of transgender athlete policy must include the observation of biology and physiology.

Early Human Development

As Carlson (2018) notes, upon the conjoining of the sperm from the male and the egg from the female, new life has been created. Every new human life is a genetically unique being with differences in the sexes appearing at the earliest stage of human development. The human being, at this point, has a DNA genetic composition that is unique amongst the billions of other humans that are living and have ever lived. Normal

chromosomal destiny at this point is either 46, XX or 46, XY. In atypical cases, genetic mutations influence the normal expression of this dichotomous destiny, but these are rare (Kousta et al., 2010).

The human body consists of cells with the nucleus in the center of each cell. Within cells are the chromosomes that serve as structures for human genes. These genes determine human traits. For example, over 3,000 genes affect how skeletal muscle is expressed (Haizlip et al., 2015). Furthermore, differences between male and female humans are driven by gene expression. Carlson (2018) explains that the typical number of chromosomes in each cell of the body is 23 pairs for a total of 46 chromosomes; half inherited from the father and half from the mother. Chromosome pairs are numbered from 1 to 22, and dependent on its structure, the 23rd pair is labeled X or Y. These are also known as the sex chromosomes, because of their determination of male or female. Males have one X and one Y chromosome, whereas females have two X chromosomes. A full image of all 46 paired chromosomes is known as a karyotype. Normal karyotype for a female is 46, XX, and for a male is 46, XY.

Disorders of Sex Development

Variations of the 46, XY and 46, XX binary appear through gene mutation. These are natural, extremely rare, conditions of sexual development, labeled disorders of sex development (DSD), that add a layer of complexity to the male and female binary. DSDs are now understood to indicate a condition of disharmony between chromosomal, gonadal, and anatomical sex (Wisniewski & Mazur, 2009). Recently revised nomenclature reflect a better understanding of genetic advances and reverse gender-based diagnostic labels. Since 2006, and driven in part by the Lawson Wilkins Pediatric

Endocrine Society and the European Society for Pediatric Endocrinology, the term intersex has become DSD; hermaphrodite has become ovotesticular DSD; XX Male is now 46, XX testicular DSD; female psudohermaphrodite is now 46, XX DSD (Kousta et al., 2010).

Many DSDs have negative effects on health and athletic performance potential. Klinefelter's syndrome involves two X and one Y chromosome. Turner's syndrome involves one X, and Mosaicism is a mixture of cells with XX/XY or X/XY chromosomes. These DSDs of a chromosomal nature, for the most part, are not seen as producing a sports performance advantage. Conversely, androgen excess DSDs, or ovotesticular DSDs, can result in external female genital phenotype, but male virilization, a response to endogenous testosterone; a likely athletic advantage over normal females (Tucker & Collins, 2010).

46, XY DSDs are rare conditions where genetic abnormalities affects testosterone and dihydrotestosterone (DHT) or tissue and cellular responses to androgens. The two most notable conditions are 5-alpha reductase deficiency, type 2 (5ARD2) and androgen insensitivity syndrome (AIS) (Wisniewski & Mazur, 2009; Clark et al., 2019). Often these DSDs result in ambiguous genitalia and misidentification at birth. 5ARD2, for example, occurs when a 46, XY individual fails to develop initial male characteristics such as a penis, but does develop male analogous musculature because of normal male levels of testosterone.

Abnormal conditions that can occur in both 46, XX and 46, XY persons. The conditions polycystic ovary syndrome (PCOS) and congenital adrenal hyperplasia (CAH) occur in 46, XX persons with the most common CAH being 21-hydroxylase deficiency

(21OHD) (Clark et al., 2019). CAH involves the adrenal glands producing excessive amounts of testosterone in females. These persons are genetically female and lack male sex organs but develop secondary male sex characteristics, which may help performance (Tucker & Collins, 2010). Bermon et al. (2014) found the incidence of elite female competitors with a DSD were 7.1 in 1,000 versus 1 in 20,000 in the general public; a 140 times difference/overrepresentation, suggesting some DSDs have an ergogenic effect that may impact the competitive landscape of female athletics.

Prepubescent Sex Differences

Athletic performance of males and females begins to diverge in childhood and eventually reaches a vast bimodal divergence through puberty. The general consensus is that there is no prepubescent performance advantage for either sex (Dore et al., 2005; Round et al., 1999; Sandbakk et al., 2018). However, there are some reported prepubescent differences.

Dore et al., (2005) report observing coordination superiority in cycling efficiency as early as age 10; before the onset of testosterone. Temfemo et al. (2009), who found that boys had significantly less body fat than girls prior to puberty. In evaluating swimming performance, Senefeld et al., (2019) found in their sample, prior to age 10, the top five 46, XX children were faster (3%), than 46, XY children. The top five 46, XX children demonstrated faster swimming velocities, and the 10^{th}-50^{th} place 46, XX children demonstrated similar swimming velocities as 46,XY youth until around age 10. After age 10, however, 46, XY children demonstrated increasingly faster swimming velocities than 46, XX children.

In investigating world records from age 5 to 19 in boys and girls, Handelsman et al. (2017) report 46, XY children were found to have a relatively small 3% prepubescent advantage; despite having the same testosterone counts. Between 11 and 13-years old, 46, XY children began experiencing an influx of circulating testosterone, correlating to a widening performance gap. The onset of sex differences in swimming events was at 11.4 years old and reached mid-puberty, the time when half the ultimate difference had occurred, at 12.8 years old. In running, the start of difference occurred at 12.4 years old and reached midway at 13.9 years. The authors also report that the shorter the event, the greater and earlier the divergence. This phenomenon seems to affirm the fact that 46, XY persons, with the assistance of testosterone, develop more Type II fibers that are invaluable in short, powerful events.

A classic longitudinal study from Round et al. (1999) shows linear increases that were similar in both sexes up till puberty, at which time the 46, XY children increase in strength quickly outpaces their 46, XX counterparts. Their study had virtually zero overlap in the biceps strength scores; the weakest 46, XY child in the sample at 16, was stronger than the strongest 46, XX child at that age. This finding is similar to those of Sandbakk et al. (2018), when evaluating world-record performances in children.

Researchers of the timing of the gender divergence attribute it to the rise in testosterone with the onset of male puberty (Clark et al., 2019; Handelsman, 2017; Handelsman et al., 2018; Sandbakk et al., 2018). Round et al. (1999) summarizes the prepubescent performance development and difference; “for the girls, development of strength is proportional to the general increase in height and weight, while for the boys there is an additional factor that can be fully explained by the term *testosterone”* (p 56).

Anthropometric Differences Following the Onset of Puberty

The onset of puberty, with testosterone acting as a stimulator, human secondary sex characteristics pertinent to athletic performance diverge between 46, XY and 46, XX persons significantly. This section will explore the underlying sex differences in factors such as osteological, mass and body composition, muscle type and size, and metabolic function and capacity.

Sex-specific changes are induced by circulating levels of hormones, including; testosterone, estrogen, progesterone, luteinizing hormone, follicle-stimulating hormone, and growth hormone (Sandbakk et al., 2018). Through the process of puberty, 46, XY individuals on average possess larger size, lower body fat, increased aerobic and anaerobic metabolism, more hemoglobin, more blood, larger hearts that produce higher stroke volume, greater cardiac output and VO_2 max, and superior neuromuscular adaptation (Dore et al., 2005).

Osteological Differences

There is a difference in skeletal system structure of 46, XY and 46, XX persons that manifest in puberty. 46, XY individuals have a distinctively greater bone size and density than do 46, XX persons of the same age. Sex differences in bone are absent prior to puberty, but following the onset of puberty diverge markedly (Almeida et al., 2017; Dore et al., 2005; Handelsman et al., 2018; Jones et al., 2016; Staron et al., 2000). Temfemo et al. (2009) and Dore et al. (2005) report significant gender differences in height from 14 years old on in favor of 46, XY adolescents.

Handelsman et al. (2018) report 46, XY persons have stronger, longer, and denser bones. 46, XX individuals have an earlier onset of puberty, triggering a growth spurt and

an earlier growth plate fusion than 46, XY persons, contributing to a smaller stature. As a consequence, on average, 46, XY persons are 7% to 8% taller, and 46,XX humerus cross-sectional areas are 65% to 75% and femur 85% of those of 46,XY persons. Bone size and mass present a performance advantage when longer and stronger bones contribute to superior fulcrum force and power. In addition to presenting a force advantage in explosive, power dominant activities such as jumping, sprinting, throwing, and striking; stronger bones help in injury susceptibility in lower force activities, with 46, XX athletes having more serious and frequent stress fractures (Handelsman et al., 2018; Moreira & Bilezikian, 2017).

Certain sex-differentiated bone structures affect kinematic abilities. The pelvic Q-angle is the vector that represents the direction of force placed on the patellar tendon with the quadriceps muscles. 46, XX individuals have a greater Q-angle stemming from a widening of the 46, XX pelvis in puberty. This wider angle negatively affects athletic performance as it reduces the force of knee extension. Therefore, 46, XY and 46, XX individuals with identical musculature and height will not be able to generate the same performance results in this kinematic function (Fischer & Mitteroecker, 2017). Additionally, 46, XX and 46, XY persons have different elbow carrying angles, a product of the ulna and humerus not being in a straight line, contributing to yet another kinematic advantage for 46, XY persons (Handelsman et al., 2018).

Body Mass Differences

Reporting on anthropomorphic characteristics shows that 46, XY persons are heavier, have more fat-free mass (FFM), and less body fat (14.8% vs. 23.3%) when compared to females (Jones et al., 2016). 46, XX muscle mass is generally 25% to 40%

less than 46, XY individuals, and 46, XY persons have relatively more muscle mass located in the upper body (Sandbakk et al., 2018). Janssen et al. (2000) reported that the skeletal-muscle mass of 46, XY persons is 36% greater than 46, XX persons with 40% and 33% differences in the upper and lower body, respectively. Temfemo et al. (2009) report no sex difference was found for lean body mass from 11 and 12 years, but from 13-16 years, there was a higher value in 46, XY children (43.6kg versus 40.0kg or 9% at age 13; 51.2kg versus 44.1kg or 15% at age 16).

Muscle Type and Size

Haizlip et al. (2015) report key differences between 46, XY and 46, XX muscle are driven by gene expression. Muscles are distinguished by fiber type. Fibers are generally divided into two types; Type I, (slow-twitch), and Type II (fast-twitch). Type II contributes to superior explosive power and speed, whereas Type I contribute to greater endurance capabilities and are slower to fatigue, smaller, and less strong.

Welle et al. (2008), Staron et al. (2000), and Haizlip et al. (2015) report 46, XX persons have a greater ratio of Type I muscle fiber mass to Type II. 46, XY persons have a higher rate of protein synthesis per muscle fiber leading to greater protein mass per fiber. 46, XY individuals had greater lean muscle mass, less body fat, and a larger cross-sectional area of every fiber type. The 46, XY's distribution of fiber type showed far more of the powerful fast-twitch Type II fibers and a higher percentage of these fibers. They further conclude, even for the slow-twitch Type I fibers, the 46, XY person's cross-sectional areas of each fiber were bigger than 46,XX persons, and therefore, more effective for work output. In agreement with these findings, Haizlip et al. (2015) showed

46, XY Type I fibers were 19% bigger and Type IIA and Type IIX fibers were 59% and 66% bigger than their 46, XX counterparts.

Cardiorespiratory System Differences

Blood and Hemoglobin. Blood accomplishes the two critical functions of transporting oxygen from the lungs to the tissues of the body and removing carbon dioxide from the tissues. 46, XY persons have more blood volume on average than 46, XX individuals (Lundgren et al., 2015). Hemoglobin, carried by red blood cells, is an iron-protein molecule responsible for oxygen transport (Haff & Triplett, 2015). Circulating hemoglobin levels are, on average, are 12% higher in 46, XY than 46, XX persons (Murphy, 2014). Handelsman et al. (2018) estimate that this advantage will produce an average maximal oxygen transfer of ~10% greater in 46, XY persons, directly impacting athletic performance: "To put this into context, any drug that achieved such increases in hemoglobin would be prohibited in sports for blood doping" (p. 817).

Sandbakk et al. (2018) report that while most of the sex differences in cardiac and blood parameters are equalized when normalized for body size, the concentration of hemoglobin in remains markedly lower in 46, XX individuals than in 46, XY persons (12–16 versus 14–18 g/100 mL).

Heart, Stroke Volume, and Cardiac Output. On average 46, XY persons have greater heart size and left ventricular mass critical to blood volume circulation than 46, XX persons; contributing to greater cardiac-output and higher VO2max (Foryst-Ludwig & Kintscher, 2013). Cardiac-output being a product of stroke rate and stroke volume. Lundgren et al. (2015) indicate no sex difference in maximal heart rate (stroke rate);

therefore, the cardiac-output discrepancies are explained by greater stroke volume of larger 46, XY hearts.

Lung Capacity. 46, XY persons have, on average, superior pulmonary function to 46,XX persons. Sex difference in pulmonary function is largely explained due to 46, XY persons on average being bigger and taller, which is a strong predictor of lung capacity and function (Handelsman et al., 2018). Townsend et al. (2012) report the difference in lungs appears early during adolescence; where 46, XY individuals generate higher respiratory pressures than 46, XX persons at all lung volumes, and even 1-year old to 10-year old 46,XY persons have higher peak expiratory flow (maximum speed of expiration) than 46,XX persons, and higher airway conductance (amount of air reaching the alveoli per unit of time per unit of pressure).

VO_2 max. VO_2 max (maximal oxygen uptake) is considered the most valid measurement of aerobic endurance. In the general population, VO_2 max differs between 46, XX and 46, XY persons by ~50% in absolute terms (L/min) and by 20% to 30% when normalized to body mass (mL · min–1 · kg–1) (Pate & Kriska, 1984). In trained athletes, a 15% to 20% difference in this value relative to body mass among equally talented and well-trained 46, XY and 46, XX athletes appear. The highest body-mass-normalized VO_2 max values reported (among cross country skiers) are ~90 and ~75 mL ·min–1 · kg–1 for 46, XY and 46, XX persons, respectively (Sandbakk et al., 2016). For 46, XY distance runners and cyclists, values in the range of 70 to 85 mL ·min–1 · kg–1 are frequently seen, in 46, XY persons and 46, XX values ~10% to 15% lower are typically reported (Joyner & Coyle, 2008; Sandbakk, 2018).

Nervous System Differences

There appear to be early neural advantages of 46, XY persons. Dore et al. (2005) report optimal velocity in cycling efficiency favoring 46, XY over 46, XX persons, seen as early as age 10. Sandbakk et al. (2018) report a sex difference in children running events; there was a pre-pubertal difference of 3%, suggesting additional factors apart from testosterone contribute to a 46, XY athletic advantage.

Haizlip et al. (2015) reported that muscle contractility, that is, the speed of muscle contractions, was significantly higher in maximal rate of force generation in 46, XY versus 46, XX persons. Jones et al. (2016) found that men produced a significantly higher average power per kg of fat-free mass (FFM). Meaning, FFM by itself fails to explain all the sex differences in power. The study found that when normalized to FFM, power was higher across all lifts in 46, XY persons.

Central nervous system differences in the sexes are not yet fully explained, but researchers such as Levine et al. (2016) report significant differences in spatial thinking and spatial skills favoring 46, XY individuals, researchers are exploring the underlying mechanism of such a difference.

Psychological Differences

There appear to be sex-related psychological differences that affect athlete performance. 46, XX runners have been shown to pace themselves more evenly than 46, XY persons (Sandbakk et al., 2018). Speechley et al. (1996) found that 46, XX persons could sustain higher speeds during the second half of a race, when comparing 46, XY runners of equal levels of running performance. Trubee et al. (2014) report as race temperatures increase, differences in pacing widens with 46, XX runners being more

conservative under such conditions. Sandbakk et al. (2018) further reports, sex difference in pacing appears to involve psychological factors such as decision making and level of competitiveness. Oglas and Masters (2003) also found that 46, XY distance runners are more competitive than their 46, XX peers. This finding was in agreement with Deaner et al. (2015) who found in addition to a competitive difference, 46, XY runners take more risks.

Genetics Versus Training

Performance-related genes have been extensively studied, attempting to quantify contributions to performance and the performance gap between nature (intrinsic) and nurture (extrinsic) factors. Rankinen et al. (2010) found the most significant known genetic traits that contribute to elite performance are sex, height, skeletal muscle, and VO_2max. Researchers Tucker and Carlson (2012) determined sex was the number one genetic contributor to elite performance, "is a key predictor of absolute levels of performance, and is the most fundamental biological characteristic where genes influence performance" (p. 557).

The 30-year review from the HERITAGE (health, risk factors, exercise training, genetics) study by MacArthur and North (2005) suggest that extrinsic factors such as training and nutrition are indispensable to success, but insufficient to account for elite performance. Genetic heritability (including sex) contributed from 20% to 70% of the conditions linked to elite performance.

Strength and Power Development Differences

It is well established that 46, XY persons on average have significantly greater strength and power in comparison to their 46, XX counterparts (Handlesman, 2017;

Handelsman et al., 2018; Sandbakk et al., 2018). 46, XY and 46, XX persons develop strength in adolescence noticeably different. Temfemo et al. (2009) suggest a rise in testosterone levels in 46, XY persons induces a selective hypertrophy of Type II muscle fibers. Jones et al. (2016), controlling for lean muscle mass, found males had greater strength (1.3 times) in the deadlift and more power per kg of FFM. In agreement, Perez-Gomez et al. (2008) report males have greater absolute and relative muscle mass, strength, power, and greater peak acceleration. Even after normalization for fat mass, controlling for FFM, the differences in strength and power remain.

In trained individuals in the deadlift exercise, findings from Jones et al. (2016) showed that 46, XY lifters produced higher average and peak power across all tests, in absolute terms. Their numbers were roughly twice what the 46, XX persons were (197.8$\pm$ 46.3 vs. 100.0 $\pm$ 18.4). Likewise, 46, XY lifters produced more than twice the power of 46, XX counterparts (698 $\pm$ 36 Watts vs 336 $\pm$ 36 Watts). The study's main conclusion was that 46, XY persons produce greater peak and average power that is driven by superior muscle mass and the strength that it produces. Agreeing with these findings, Thomas et al. (2007) found significantly greater power outputs in 46, XY persons performing both upper and lower body exercises across a variety of loads.

Temfemo et al. (2009) studied over 500 boys and girls aged 11 to 16 years and found that for all jumping/power events, the measures of 46, XY children were significantly higher. Significant differences are observed from age 14 on that the researchers attribute to the increase in leg length and leg muscle volume. Similarly, Dore et al. (2005) found for girls, power increased from age 10 through 16-18 years old and then plateaued, and there was a significant sex-related difference in power from 13 years

onward in power. The authors attribute the difference to muscle dimensions, muscle fiber type, anaerobic metabolism, and neuromuscular adaptation.

Testosterone Differences

Testosterone is a powerful anabolic androgen. Sex difference in athletic performance coincides with rising testosterone concentrations in male puberty, resulting in 46, XY persons having 15- to 20-times greater circulating testosterone than 46, XX persons. This wide, bimodal testosterone difference and the clear dose-response relationships between testosterone and mass, strength, hemoglobin etc., largely account for the sex differences in athletic performance (Handelsman et al., 2018).

Round et al. (1999) conclude that for females, quadriceps strength increases in proportion to their general increase in height and weight, whereas the 46, XY's stark gains are due to the rapid rise in the testosterone concentrations. Clark et al. (2019) report healthy/normal male and female testosterone levels are dimorphic/distinct with zero overlap. The normal range for 46, XY individuals was 8.8 to 30.9 nmol/L, versus 0.4 to 2.0 nmol/L in females, meaning the low end of the 46, XY range, was four or five times as great as the 46, XX range. Persons with DSDs, such as 46, XY DSD have a median/mean range of 13.4 to 31.2 nmol/L, those with AIS have 11.9 to 55.7 nmol/L, and testosterone in 46, XX females with PCOS is 0.34-5.5 nmol/L; higher than normal 46, XX persons, but less than normal 46, XY levels.

Testosterone's Correlation with Performance

Recent developments in doping testing at the elite levels has led to the creation of a longitudinal individualized biomarker profile called the Athlete Biological Passport (ABP). In addition to anti-doping evaluation, the ABP provides one of the best study

samples of elite athletes for scientific purposes. Bermon et al. (2014) used the ABP blood samples from the 2011 World Championships to study serum testosterone levels in elite female athletes. They found that power athletes (throwers, sprinters, and jumpers) showed a significantly higher testosterone concentration than long-distance athletes. They established that the upper end of the normal range (99th percentile) for females was 3.09 nmol/L, significantly lower than the previously arbitrarily established 10 nmol/L allowed by the International Association of Athletics Federations (IAAF) (now known as World Athletics). Based in part on this data, World Athletics lowered the threshold to 5 nmol/L in 2019 (Harper, 2019).

A timely and consequential study by Bermon and Garnier (2017) was used to uphold the World Athletics policy on hyperandrogenism before the Court of Arbitration for Sports (CAS). Their study sample of over 2,100 female athletes at the 2011 and 2013 track and field IAAF World Championships showed a significant correlation with testosterone and performance in multiple female events. Their conclusions show a significant advantage to having higher testosterone in the 400m, 400m hurdles, 800m, hammer, and pole vault. Further research is needed to replicate the findings on a longitudinal basis, particularly because it fails to explain why certain events would not be included, such as power-dominant events, like the shot put, or the 200m and 100m sprints.

Additionally, Bermon and Garnier (2017) report androgen concentrations did affect performance in females, but not in males. Male sprinters showed higher testosterone concentrations than the other male athletes. Interestingly and counterintuitively, throwers showed lower testosterone than other male athletes. Within

males, testosterone concentrations failed to have a statistically significant impact on performance on the whole.

Following an extensive systematic review, researchers Handelsman et al. (2018) report the appropriateness of lowering of the IAAF testosterone threshold to 5nmol/L. The nonoverlapping, bimodal testosterone distribution and its dose-response relationship to muscle mass, strength, and athletic performance suggest that circulating testosterone is a valid and appropriate eligibility criterion for female sports participation. They argue the limit at 5 nmol/L is high enough that it gives allowance to 46, XX persons with PCOH to compete without suppressing their testosterone. However, given that the incidence of PCOH in elite female athletes is greater than 100 times higher than the general population, further monitoring and study are needed.

Bermon et al. (2014) report 46, XX athletes using oral contraceptives (OC) had significantly lower androgen numbers, confirming the proposed usage of OC to bring testosterone numbers down to the normal female range. An elite athlete sample reports 14.5% of 46, XX persons using OC compared to a much higher number of around 47% in the general population, suggests that female athletes seek to keep maximum androgen concentrations.

Testosterone Supplementation Via CSHT

Cross-sex hormone therapy (CSHT) involves the introduction of exogenous hormones such as testosterone into the body. Haizlip et al. (2015) report testosterone supplementation causes increases in muscle mass, fiber cross-sectional area, increased strength, and decreased body fat in human subjects. Handelsman et al. (2018) report

exogenous testosterone replicates the effects of endogenous testosterone on every reproductive and nonreproductive organ or tissue, with the sole exception of the testis.

Establishing a dose-response relationship with testosterone and performance is important and critical if testosterone concentration is to be the criterion for exclusion in transgender sport policy. A review by Velho et al. (2017) found that the non-athlete FTM transgender patients increased body mass index (BMI) by 1.3% to 11.4%, FFM by 8.5% to 12.3%, hemoglobin by 4.5% to 12.5%, and hematocrit by 4.4% to 17.6% with CSHT. However, they report anabolic steroids, such as the commonly used testosterone undecanoate, do not come without risks; adverse effects such as hypertension, erythrocytosis, joint and muscle pain, and hair loss occur in 50% of patients.

In a large population review, Klaver et al. (2017) documented CSHT in both FTM and MTF persons and compared body weight, body fat percentage, and lean body mass. Their findings show an increase in body weight for both MTF and FTM persons undergoing treatment: 1.8kg and 1.7kg, respectively. The non-athlete MTF persons had an increase in body fat 3.5kg or 25%, whereas the FTM individuals lost body fat of 2.6kg or 10.5%. FFM decreased by 2.4kg in MTF persons and increased 3.9kg in FTM persons. When comparing FFM, following 12 months of MTF CSHT (primarily cyproterone acetate) and FTM CSHT, there was still zero overlap in lean body mass. In other words, a hypothetical male taking testosterone blockers will still on average, have an advantage in lean muscle over 46, XX persons; even 46, XX persons that have been taking anabolic steroids for a year. The findings indicate that FTM athletes are likely no risk to the competitive makeup of the male category of sport; even while taking anabolic steroids.

Joanna Harper (2015), a transgender athlete and researcher who has been an expert witness in front of the CAS on behalf of limiting access to the female category, believes that CSHT for one year is sufficient to limit the 46, XY performance advantage to an acceptable level of competitiveness in sport. In support of this hypothesis is Harper's (2015) study that included a very small sample (N=8) of fellow MTF recreational runners who self-reported their recollection of times pre- and post-treatment in recreational running races. Results showed a similar age group comparison post-treatment, now in the female category. The findings have yet to be supported by corroborating controlled, longitudinal trials.

Gooren and Bunck (2004) findings confirm significantly higher muscle area in 46, XY individuals than in 46, XX individuals prior to CSHT, even though some overlap in the distributions do occur. They report that even after one year of hormone treatment, a 46, XY MTF person still had significantly more muscle area than an untreated 46, XX individual, and although androgen suppression in MTFs increases the muscular overlap with 46, XX persons, the difference remains significantly greater following three years of CSHT. The treated MTF muscle area was 271cm^2 versus a 46, XX area of 238.8cm^2, which amounts to substantial potential performance advantage.

In Wiik et al. (2020) in one of the few longitudinal assessments of CSHT on transgender persons found that following 12 months of CSHT, MTF individuals maintained their strength levels, and only decreased their muscle volume by 5% and CSA by 4%. This study had a N = 12 of untrained MTF individuals.

Advantage Independent of Testosterone. Klaver et al. (2017) suggest FFM still remains vastly greater than 46, XX persons following CSHT. Temfemo et al. (2009)

found power differences in 46, XY and 46, XX persons were due to the superior in leg length and leg muscle volume of males; and Jones et al. (2016) showed that 46, XY persons produced significantly greater averages of power per kg of lean muscle mass (10 watts/ kg of FFM compared to 7 watts/kg of FFM for 46,XX); and Perez-Gomez et al. (2008) in agreement report males have greater absolute and relative muscle mass, strength, power, and greater peak acceleration; even after controlling for fat-free mass, the differences in strength and power remain. MTF athletes will, on average, have a significant advantage in power and strength; regardless of treatment. Thus, the question remains whether or not an MTF transgender person could fairly compete with a 46, XX person in athletic competition.

The evidence suggests that post-pubertal MTF transgender persons do not sacrifice their athletic advantage simply by lowering their testosterone concentrations. Unexplained benefits of male virilization need to be further researched, but what is known builds a strong case that there is advantage independent of testosterone once an individual has gone through 46, XY puberty.

Post-Pubescent Athletic Performance Differences

Physiological differences in the sexes matter in so much as they contribute to athletic performance differences in the competitive framework of sport. The following is a summary of the differences in performance terms.

Following the unfolding of Title IX, there were, and are, those who view sex differences as a result of disparate treatment and sex stereotypes, and would eventually be eradicated (Jones et al., 2017). This view theorized that, like the classroom, eventually sports could also be gender-blind (Coleman et al., 2020).

In an investigation into performance gaps and their historical context, Millard-Stafford et al. (2018) investigated performance differences in individual sports since 1972, when the Title IX legislation transformed female participation rates. Data from U.S. Olympic trials from 1968 to 2016 show an initial narrowing of performance in both swimming and running. However, following the steep curve of the 1970s, the performance gaps have effectively stabilized from the 1980s on. From 1972 to 1980, the male/female performance gap closed by 2% in swimming and 5% in running, but failed to close from that point onward to 2016. Currently, they report performance gaps of 13% existing in running and 8% in swimming. The steady plateau in the performance gaps over the most recent 40-years suggests a stable model for evaluating sex differences in sport.

Kinematic factors and event distance are contributing factors to the performance difference in 46, XY versus 46, XX athletics. Sandbakk et al. (2018) in a comprehensive review of world record performances, report a gap of 8% to 12% with significant variation due to kinematic factors and event distance.

Handelsman (2017) reports that towards the end of adolescence, males have a quite massive 10% advantage in running and 19.3% in jumping. Sandbakk et al. (2018) report sprint events that last 10-60 seconds have been stable since 2005 with a difference of approximately 10-12%. Senefeld et al., (2019) report the sex difference in performance at age 18 was larger for swimming sprint events (9.6%; 50-200m) than endurance events (7.1%; 400-1500m), but others such as Sandbakk et al. (2018) and Millard-Stafford et al. (2018) report the smallest performance gap in running occurs in the shortest distance, the

100m sprint. Similar to this trend in running, Sandbakk et al. (2018) report speed-skating and short track cycling sex differences increase with distance.

Upper body dominant events of kayaking and canoeing and double-poling cross-country skiing report the greatest performance gaps of up to 23% (Handelsman, 2017; Sandbakk et al., 2018). At short distances, these events show the biggest gaps of any sports, suggesting that the more reliant the event is on upper body power, the greater the gap. Hegge et al. (2015) report after normalization of power differences for lean upper-body mass during poling by cross-country skiers, performance gaps up to 30% remained.

The narrowest gaps in sex performance, in some cases less than 5%, are in ultra-endurance swimming (Handelsman, 2017; Sandbakk et al., 2018). Sandbakk et al. (2018) and Millard-Stafford et al. (2018) find that the largest performance gap between the sexes in swimming occurs in the shortest distance, the 50m sprint.

The 9% to 13% sex difference in track and field running events, may not seem like a vast difference to some, but as Tucker and Collins (2010) note, the female world record holder in the 400m would not place in the top 400 male performances in that event in any given year. As an additional illustration, the female world record holder in the 100-meter sprint (10.49 seconds) would not even make the 93-man qualifying field (10.39 cutoff) in America's 2019 NCAA Division I outdoor track and field championships (NCAA Division I Outdoor qualifying, n.d.).

To put the performance gaps into context amongst the competitive spaces, the winning margin (the difference in performance by which a competitor misses a gold medal, any medal, or making the final), in elite track or swimming events during the last three Olympics is less than 1% in both male and female events (Handelsman et al., 2018).

Summary of Sex Differences

To summarize the points in Part I, 46, XX and 46, XY persons are different from the moment of conception. This difference manifests in two distinct versions of the same species. The versions are so different in terms of athletic performance, that 46, XX persons would have virtually no chance to compete, much less win, against a field of competitive 46, XY athletes. Most of the difference occurs in conjunction with male puberty and the production of testosterone. The effects of 46, XY maturation during these years results in being taller, bigger, stronger, faster, amongst other anthropomorphic, metabolic, and physiological advantages. Once the gains are made through puberty, many advantages such as denser and larger bones, larger heart and lungs, and greater numbers of Type II muscle fibers, they are a permanent fixture of the person. Even vastly more lean muscle mass, which can be reduced (but not reversed) through CSHT, will still be far more than 46, XX persons, on average. Transgender persons who developed into mature 46, XY persons, and later seek to participate with 46, XX athletes are, on average, at a scientifically verifiable advantage.

Part II: Title IX and American Sport

Historical Review of Title IX

In the last 40 years, female involvement has been the area with perhaps the most profound and visible impact in American sports. Title IX is the most significant piece of legislation regarding American sports (Anderson et al., 2006). It is credited for changing the landscape of female opportunities in athletics. The central theme of equality in The Declaration of Independence instigated the revolution that formed our nation, and Title IX was written and adopted in keeping with that ideal of equality (McAffee, 2001).

The figures show more than 3,000,000 girls now participate in high school athletics, and over 250,000 college women play sports. In 1972 there were only roughly 300,000 high school girl athletes and less than 32,000 college women athletes (National Federation of State High School Associations, 2019). Girls today have many opportunities to compete in a vast array of sports at the high school and collegiate level; before 1972 that was not the case. From 1981 to 1999 alone, the total number of women's intercollegiate sports teams increased from 5,695 to 9,479 (Kennedy, 2010).

Origins of Title IX

Title IX was brought about in the aftermath of the 1960s and early 1970s social and moral revolutions that were taking place. Along with social changes, came an awareness of female disparities in educational settings and a recognition that inequities in the education system were detrimental to females in education. Title IX was an attempt to end discrimination based on sex in education and ensure that the government promoted the Constitution's ethic of equality to all persons.

In 1972, President Richard Nixon signed into law a series of amendments to the United States Code, and in doing so filled a gap that was left by Title VII of the 1964 Civil Rights Act (Coleman et al., 2020). In 37 words, Title IX of the Education Amendments, Section 901 (1972) states: "No person in the United States shall on the basis of sex, be excluded from participation in, be denied the benefits of, or be subjected to discrimination under any education program or activity receiving federal financial assistance." Because federal funds are linked to student aid programs, federal grants, and the like, nearly every educational institution would be subject to its demands. In effect, this was a mandate placed on the entire nation.

The language of Title IX in its 37 words is and was non-controversial. Title IX was passed with unanimous congressional approval. It is simple and lacks language typically viewed as contentious or divisive. However, those 37 words seem to be about the only thing lawmakers have been able to agree on (Kennedy, 2010). The simplicity was a unifying aspect but perhaps one of its enduring weaknesses. In the midst of rising confusion in the wake of its passing, Congress tried to act, but as more specificity emerged, it could not reach political consensus to pass any significant clarifying legislation. What was intended to be simple, has been complex and controversial. In turn, the courts, institutions, and individuals have struggled to accomplish the ideals set out by Title IX. The goal of equal opportunities, when placed into the real world of budgets, personnel, equipment, admissions, etc. quickly was shown to be lacking in details and guidance (Anderson et al., 2006).

Given the lack of guidance, Title IX has been strongly shaped by the executive branch. When the 1975 regulations, intended to clarify issues, came to a vote, Congress failed to act, leading to a precedent of Title IX regulation and policy via executive action (Leahy, 1997). Since the 1972 vote, there have been hundreds of pages of regulations, multiple clarifications, interpretations, and even commissions by administration officials who have effectively cultivated the 37 words into an expansive labyrinth of guidance encompassing athletics, sexual misconduct, staffing, and more (Anderson & Cheslock, 2004).

Subsequent Interpretations and Regulations for Compliance

Under the Title IX umbrella are a litany of items including education, athletics, sexual orientation, gender identity, and sexual misconduct; items that have been deemed

law, yet do not carry the weight of the legislative process. The original language of the statute made no mention of athletics or sexual assault, arguably the two most current citations of the law. It seems clear from the floor and chamber discussions that these issues were not on the minds of the original signators, and only mentioned twice in passing by Indiana Senator Birch Bayh; once regarding privacy and athletic facilities in 1972, and once noting that Title IX will not require gender-blended football teams (Cozzillio et al., 2007).

The 1975 regulations were the first time that athletics were mentioned explicitly. Part 106 of the OCR regulations stated that a recipient of federal funds shall provide "equal athletic opportunities for both sexes" (OCR, 1975, para. 106.41). The regulation gave a number of factors to help determine compliance. In a first, it asked whether the availability of sports and levels of completion effectively accommodate the interests and abilities of members of both sexes (Leahy, 1997). This interests and abilities factor was ambiguous in its application, which led to confusion and lawsuits in the years immediately following the guidance. In the three years alone following 1975, there were over 100 reported claims of discrimination delivered to the HEW (Hatlevig, 2005).

Needing to clarify guidance for institutions and with a Congress unwilling to take action, the administration looked to the newly created Office of Civil Rights (OCR) for answers. In response, the OCR published its December 11, 1979, guidance that became known as the *Intercollegiate Athletics Policy Interpretation* (OCR, 1979). This interpretation attempted to resolve many of the questions and uncertainty faced by schools. The policy was intended to be a manual by which Title IX in intercollegiate athletics could be more objectively measured for compliance. The three areas of

emphasis that the policy covered were; (a) scholarship opportunities in college athletics; (b) benefits and opportunity equality; and (c) fulfillment of interests and abilities of students. Midway down the 10,000-word policy was the heart of the guidance to address interest and abilities. What was described was a three-pronged criterion that eventually would become known as the three-part test. The three-part test involved the following:

> 1. Whether intercollegiate level participation opportunities for male and female students are provided in numbers substantially proportionate to their respective enrollments; or
>
> 2. Where the members of one sex have been and are underrepresented among intercollegiate athletes, whether the institution can show a history and continuing practice of program expansion which is demonstrably responsive to the developing interest and abilities of the members of that sex; or
>
> 3. Where the members of one sex are underrepresented among intercollegiate athletes, and the institution cannot show a continuing practice of program expansion such as that cited above, whether it can be demonstrated that the interests and abilities of the members of that sex have been fully and effectively accommodated by the present program.
>
> (OCR, 1979, p. 12)

The three-part test has become the most prominent enforcement mechanism and guidance for the equality of the sexes in athletics. Part one of the test, known as the substantial proportionality part, motivates most institutional attempts at compliance. The second and third part are used to prove compliance should the first be insufficient. Part

one is known to be the preferred option due to its "safe harbor" provision. Due to the subjective nature of part two and part three, there is an incentive to satisfy part one outright (Hatlevig, 2005). This predominant view, has had an unintended negative effect on non-revenue mens' sports by transforming the anti-discrimination statute into "an affirmative action law that mandates gender quotas" (Ganzi, 2004, p. 543).

Given ongoing confusion, especially concerning satisfaction of part two and three of the three-part test, further guidance continued to shape the application of the law. The OCR published further guidance in 1990 to attempt to address some of the issues in what it titled, *The Title IX Athletics Investigator's Manual* (Bonnette & Daniel, 1990). The manual was another way to trying to clarify compliance in an age of compliance by the devastating contraction of male non-revenue sports (Hatlevig, 2005; Ganzi, 2004; Anderson & Cheslock, 2004; Anderson et al., 2006). The manual seemed to try to stop the losses in male athletics by disapproving of quotas, but it paradoxically offered mirror percentages of male and female students as the ideal (Bonnette & Daniel, 1990). Many were left uncertain and unsatisfied with the ambiguous subjective nature of the policy that remained.

With continued uncertainty over the years, the OCR published further guidance in 1996 titled, *Clarification of intercollegiate athletics policy guidance: The three-part test* (1996). This policy guidance sought to eliminate quotas in the application of the three-part test. It specified that eliminating mens' teams were not advised and provided some choices for the schools in compliance. At this stage, the imbalance of women and mens' non-revenue teams had swung dramatically in favor of women over men (NCAA, 2013).

The new century brought with continued uncertainty about the subjective nature of the three-pronged test, prompting the publishing of *Further clarification of intercollegiate athletics policy guidance regarding the Title IX compliance* (2003). This guidance was prompted by a bipartisan 2002 Congressional commission that was formed to identify Title IX problems and find solutions. Despite the exhaustive work of the commission, which produced a number of constructive reforms, the findings were rejected before they were even officially presented. In an election year, the Secretary of Education preempted the report saying that he would reject any finding not unanimously supported. Therefore, what was produced was continued commitment to the three-part test and the subjectivity of the second two parts (Anderson & Cheslock, 2006).

Since Title IX is more a functionary of executive action rather than signed legislation, it is highly dependent on the political climate at the time. The Bush administration issued guidance in 2005 that emphasized the use of part-three. It did so in support of interests who were concerned about the status of male non-revenue sports. In a more counter move, the Obama administration reversed and rescinded the part three guidance in 2010 in favor of part one (Kimmel, 2015).

Perhaps in the most consequential move for the transgender movement, President Obama's Department of Education issued transformative Title IX guidance known as a *Dear Colleague* letter. The guidance, for the first time in this presidential context, established that the government should interpret "sex discrimination" to include claims based on gender identity. However, the landmark change was short-lived, as President Trump in 2017, withdrew and rescinded the Dear Colleague guidance (American Bar Association, 2018).

The Sports Exception to Title IX

Even though Title IX is structurally a non-discrimination rule that is neutral, the regulatory structure has presented sex-affirmative exceptions. The Supreme Court currently "distinguishes sex from stereotype, but also race from sex on the ground that the latter but not the former involve inherent differences . . . these differences are properly considered when doing so serves to empower rather than to subordinate" (Coleman et al., 2020, p. 79). In short, segregation by sex is currently admissible, even though it, by nature, discriminates. In the following Part V, the researcher will review why recent developments in the courts place this understanding in question.

Part III: Transgender Developments in America

Transgender Basics and Population Estimates

Although there have been gender-diverse persons throughout history, dating back even to ancient times, as documented by Harper (2019), the transgender movement reached a *tipping point* of visibility and cultural acceptance only very recently. Evidence of mainstream adoption of transgender logic change can be seen throughout society. At its most basic level:

> transgender is an umbrella term that incorporates differences in gender identity wherein one's assigned biological sex doesn't match their felt identity. This umbrella term includes persons who do not feel they fit into a dichotomous sex structure through which they are identified as male or female. Individuals in this category may feel as if they are in the wrong gender, but this perception may or may not correlate with a desire for surgical or hormonal reassignment. (Meier & Labuski, 2013)

In the most widely cited attempt to quantify the population of transgender persons, The Williams Institute at UCLA's School of Law estimates that 1.4 million transgender persons live in the United States (0.58% of the population) (Flores et al., 2016). This estimate is twice the amount that Gates (2011) reported. This number is expected to double again by 2030 because of the dramatic rise in persons identifying as transgender. Young Americans are increasingly comfortable being transgender, and "the number of 'out' trans kids is growing beyond the small percentages described in earlier population surveys. This increase is not yet well understood, but it appears to be an upward trajectory" (Coleman et al., 2020, p. 115).

Findings from Flores et al. (2016) and Herman et al. (2017) show that younger age groups are more likely to identify as transgender. Wilson and Kastanis (2015) report an estimated 1.3% to 3.2% of youth are transgender. Herman et al. (2017) reports 0.7% of 13 to 17-year olds are transgender. Johns et al. (2019 reports 1.8% high school students identify as transgender when sampling 10 states and 9 large urban areas. The researcher will use the Herman et al. (2017) numbers when investigating research question three of this dissertation. For the states considered in the study the population the estimates are: California, 0.85%; Florida, 0.78%; Minnesota, 0.85%; New York, 0.79%; Washington, 0.70% (pp. 4-5).

Competing Gender Worldviews

There is an ontological/metaphysical debate as to what it means to be a gendered human being. Generally, there are two different worldviews on the subject. The first is an ontological, objective view of being, and the second is an autonomous, subjective worldview.

The Binary/Objective View

The binary/objective view generally holds that gender is an objective assignment that it is inseparable and complementarian with sex. Those that hold to this view, see sex as integral to whether one is a man or woman. A prominent proponent of this view is one of the greatest tennis stars of all time, Martina Navratilova. As a gay rights activist, Navratilova has been a recipient of the National Equality Award from the Human Rights Campaign. Navratilova ignited controversy with a 2019 op-ed in The Sunday Times of London, among which was said:

> a man can decide to be female, take hormones if required by whatever sporting organization is concerned, win everything in sight and perhaps earn a small fortune, and then reverse his decision and go back to making babies if he so desires. It's insane and it's cheating. (Navratilova, 2019, p. 23)

Famed author of the Harry Potter series, J.K. Rowling likewise drew controversy by supporting the objective philosophical belief, among which has said, "If sex isn't real, there's no same-sex attraction. If sex isn't real, the lived reality of women globally is erased" (Miller & Yasharoff, 2020, para. 3). In an extensive essay, Rowling (2020) writes, "[Woman] is not a costume. 'Woman' is not an idea in a man's head. 'Woman' is not a pink brain, a liking for [high heels] or any of the other sexist ideas now somehow touted as progressive" (para. 29).

Rawling and Navratilova have been labeled Trans-Exclusionary Radical Feminists (TERFs), a term embraced by some, but interpreted as derogatory or a slur by others such as Rowling (Miller & Yasharoff, 2020).

The objective view has brought together some unusual coalitions of otherwise ideological enemies. Self-described "radical feminists," such as *Fair Play for Women* in the United Kingdom, and *The Women's Liberation Front* in the U.S., who among other things, seek to "dismantle the gender-caste system" have been partnering with conservatives, including Christian pro-family organizations such as the *Family Policy Alliance*, to promote the objective view in policy (Woman's Liberation Front, n.d.; Schmidt, 2020).

The Autonomous/Subjective View

The second view is that gender is a subjective state of mind or "self-identification." This autonomous view suggests that what matters is the subjective internal conviction. An individual's subjective convictions should be seen as objective facts. Harry Potter actor Daniel Radcliffe rebuked Potter creator Rowling following this line of thinking when he stated emphatically, "trans women, are women" (Miller & Yasharoff, 2020, "Who has come out against Rowling" section).

According to this subjective worldview, it is objectional to question the status of anyone's beliefs on the issue of their gender. Cases are frequently cited of persons not being affirmed for who they "are" which to them is fully male or fully female. Any obstructions from allowing them to live as fully one or the other are seen as unlawfully discriminatory, unjust, and hurtful. This could include persons being forced to play on the "wrong" sex's team (James et al., 2016).

According to the National Center for Transgender Equality, sex and gender cannot and should not be differentiated in society. They call for full affirmation of one's gender, as well as the conflation of sex. *A transgender man is a full man, and a*

transgender woman is a full woman according to their worldview. This belief is already displayed in pronoun usage in media and academia (American Psychological Association, 2015). It is now taboo, and in some cases, illegal, to point to biological sex when describing or interacting with certain persons (Herman et al., 2017).

A less common, yet increasingly vocal group of advocates would like the male and female binary to be erased altogether. Adherents such as Burdge (2007) seek to fundamentally alter the way society is ordered and views gender:

> Gender is a ubiquitous social construct that wields power over every individual in our society. The traditional dichotomous gender paradigm is oppressive, especially for transgendered people whose sense of themselves as gendered people is incongruent with the gender they were assigned at birth…we must target society's traditional gender dichotomy for change. (p. 243)

Key Transgender Developments in American Culture

Cultural developments in the transgender movement have progressed rapidly in recent times. The following is a review of some of the cultural developments in American culture: in 1952 Christine Jorgensen became the first American to have sex reassignment surgery; clashes between police and transgender persons happened in California in 1959 and 1966; in 1967, *The Transsexual Phenomenon* was published by Harry Benjamin M.D.; in June of 1969 the Stonewall Riots in New York are largely considered to spark the LGBTQIA+ revolution; in 1987, the Diagnostic and Statistical Manual (DSM) of mental disorders designated transgender persons with *gender identity disorder*; in 2013,

the DSM replaced gender identity disorder with gender dysphoria (The Associated Press, 2017).

Anderson (2020) further documents how in 2014, Lavern Cox became the first transgender person to be featured on the cover of Time Magazine in a year hailed as the "transgender tipping point" (para. 1). Similar front-page attention was given to; the Vanity Fair's Caitlin Jenner cover (July 2015), National Geographic cover of a 9-year transgender girl (Dec 2016), Vogue's (Feb 2017) cover and GQ's (April 2017) cover story, and 2016 full-length article in *Time* of Evan Hempel, described "the man who gave birth."

Other cultural signals that momentum unquestionably is on the side of the transgender movement, are developments such as: Nearly every state having mechanisms for changing one's gender on state identification (Transgender Law Center, n.d.); and the Common Application and the Universal College Application, which are used by more than 600 colleges and universities, now allowing students to self-identify their gender when applying to college (Prinster, n.d.).

Key Transgender Developments in Sports

The sporting context has the greatest implications for gender identity because of the segregated space and performance gaps. Some of the developments in the sports have been: In 1977 Renee Richards became the first MTF transgender person to be eligible to play professionally as a woman; November 2010, Kye Allums, George Washington University basketball player competes on a women's team as a FTM transgender individual; in 2012, Gabrielle Ludwig made news as a 50-year-old 6-foot 8 inch, MTF star of a women's junior college basketball team in California (Harper, 2019).

In 2012, Fighter Fallon Fox exploded on the mixed martial arts (MMA) scene with several dominating performances, including Fox's last fight, where Fox gave Tamikka Brents a concussion, orbital bone fracture, and seven staples to the head in the first round. When asked about the experience, Brents replied:

> I can only say, I've never felt so overpowered ever in my life and I am an abnormally strong female in my own right. Her grip was different; I could usually move around in the clinch against other females but couldn't move at all in Fox's clinch. (Murphy, 2014, para. 5)

Following that performance, UFC superstar and champion fighter Ronda Rousey commented publicly that she refused to fight Fox. Fox has yet to fight again (Harper, 2019).

Perhaps the central moment in transgender sport, and perhaps of the entire transgender movement, came in June 2015 when former Olympic track and field champion and current celebrity, Bruce Jenner, declared to the world through the magazine Vanity Fair, that Jenner now identified as a woman. Multiple accolades and awards followed, culminating in receiving the Arthur Ashe Courage Award at the ESPY Awards in July 2015. According to Brockes (2017), time, in regards to the transgender movement, may now be appropriately labeled "pre-Caitlin" and "post-Caitlin."

In the post-Caitlin era, one of the highest visibility transgender issues in sport does not deal with competition, but rather fan usage of private places such as restrooms and locker rooms in government-owned buildings. In North Carolina, House Bill 2 (HB2) or the "bathroom bill" as it became known, prompted the NCAA and others to bring all

of their political and financial clout to bear on the state, eventually extracting a compromise bill (Tracy, 2017).

More significant questions involve performance-related transgender issues. In 2017, controversy ensued when Texas FTM wrestler Mack Beggs won the girls wrestling title in undefeated fashion after being allowed to take testosterone for a "valid medical purpose," yet not allowed to compete with boys. This situation created an unsatisfactory result for all sides of the issue (Barnes, 2019).

Internationally, in March 2017, 39-year-old Laurel Hubbard won a female competition in weightlifting in dominating fashion, at an age that is abnormally high to be in prime condition beat the entire, much younger field. Hubbard beat the second-place competitor by massive 19 kilos in a sport frequently differentiated by 1 or 2-kilo differences. Leading one opponent of Hubbard's competitors to opine, "It's difficult when you believe that you're not the same. If its not even, why are we doing the sport?" (Payne, 2017, para. 3). Hubbard is perhaps the most competitive MTF transgender individual in position to qualify for the rescheduled 2021 Olympic games. Hubbard's inclusion would be a first.

In collegiate competition, in June 2019, CeCe Telfer won the NCAA Division II national championship in the 400-meter hurdles, becoming the first MTF transgender champion at that level. Telfer's title came after competing as an unaccomplished male athlete in 2016 and 2017 known as Craig Telfer (ranked 6th on team, 433rd in DII, 1,744th in the NCAA in the 400m hurdles), and sitting out 2018 while undergoing the NCAA required 12 months of suppressed testosterone. Telfer had a personal record (PR) of 57.01 as a male sophomore and a 56.50 as a senior in the female category of the indoor

season 400m; and a 24.03 as a male and a 24.33 in the female category in the indoor 200m (CeCe Telfer, n.d.; Craig Telfer, n.d.).

In 2019 June Eastwood became the first consequential MTF transgender athlete at the NCAA Division I level. In October 2019, Eastwood was honored by The Big Sky Conference as its "Women's Cross Country Athlete of the Week" following Eastwood's second-place finish out of a field of 204 46, XX runners (Morton, 2019). Eastwood previously competed as a male for three years at the University of Montana (2015-2017) as Jonathan Eastwood, before taking the required 12 months of testosterone suppression in 2018. Eastwood had 6,000m marks of 19:03.9 as a male sophomore in 2016, 20:04.0 as a male junior in 2017, and 20:18.7 in the female classification in 2019. As a male in 2017, Eastwood ranked 4th on the team, 31st in the Big Sky Conference, 648th in Division I, and 835th in the NCAA. In the female classification in 2019, Eastwood ranked 1st on the team, 3rd in the Big Sky Conference, 120th in Division I, and 126th in the NCAA (Jonathan Eastwood, n.d.; June Eastwood, n.d.).

Finally, in the most consequential MTF transgender case at the high school level; in the state of Connecticut, over the three-year period (2016-2019), two MTF athletes combined to win fifteen individual state championships in the girls' division. This case has served as the basis for a federal Title IX complaint, on behalf of 46, XX competitors, that the United States Department of Education's Office of Civil Rights is currently investigating (Coleman et al., 2020).

The MTF athletes who were dominant Connecticut State Champions, would not have been successful, had they competed in the boys' division. Andraya Yearwood, one of the MTF Connecticut athletes, has continuously competed in the female classification.

Yearwood's 55m PR of 7.02 ranks 1st in the state and 11th nationally as a girl (Andraya Yearwood, n.d.). If Yearwood were in the boys' category, Yearwood would rank 207th in the state and 3,870th nationally (High school - men's 55-meter dash rankings, n.d.).

Connecticut does not have a hormonal or surgical requirement for transitioning. Therefore, there is no sit-out time similar to the NCAA. One can theoretically compete as a male one week and as a female the next. Terry Miller, the other of the MTF transgender athletes, competed as a male in the 2017 outdoor and 2018 indoor seasons and then mid-year transitioned to compete in the female category for the 2018 outdoor, then 2019 indoor and outdoor seasons. Miller got faster in the female classification. In the indoor 55m as a male, Miller had a PR of 7.27 in 2018 (ranked 337th in the state), and in the female classification had a PR of 6.91 in 2019 (and was a state champion), with 21 of 23 sprints that year being faster than the male PR. In the indoor 300m as a male, Miller had a PR of 41.94, versus 38.90 in 2019, with 8 of 9 races being faster than the male PR. In the outdoor 100m as a male, Miller had a PR of 12.02, versus 11.72 PR in the female classification. In the outdoor 200m as a male, Miller had PR of 25.12 versus 24.17 in the female classification (Terry Miller: Bulkeley HS Track and field bio n.d.; Terry Miller: Bloomfield HS Track and field bio, n.d.).

Key Transgender Developments in Law

The legal status of transgender persons in America continues to evolve. The key developments include: in 1975, Minneapolis became the first city and in 1993 Minnesota the first state to pass a law prohibiting transgender discrimination; in 2005, California became the first state to mandate transgender health care; in 2010 the first transgender judge, Phyllis Frye was placed into office; in 2011, the Office of Personnel Management

issued a memo on federal agencies and support for transgender employees; in 2012, the Equal Employment Opportunity Commission ruled that Title VII of the 1964 Civil Rights Act applied to transgender employees; in 2014, the Department of Justice assumed the position that Title VII of the Civil Rights Act of 1964 applied to gender identity; on January 20, 2015, President Obama mentions transgender in the state of the union speech; and perhaps one of the most significant legal signals was the 2016 Dear Colleague letter that put institutions on notice that they may be in violation of the law if they do not have in place policies and practices to fully affirm a person's gender identity (Taylor et al., 2018). At the national level, the Trump Administration has withdrawn the Obama Guidance of 2016, and the OCR is investigating a complaint alleging that the Connecticut policy allowing unconditional inclusion of MTF athletes in competition violates Title IX (Eaton-Robb, 2019).

Recently there have been 20 bills filed, in 17 states, seeking to regulate transgender participation in athletics, with only one becoming law (Barnes, 2020). In April, 2020, Idaho's governor Brad Little signed into law HB 500 which limits the female classification of sport to 46, XX persons. This action, considered to be the most restrictive in the country, has triggered negative responses from a wide range of organizations and individuals who want less restrictions on the female classification (Ennis, 2020).

In March, 2020 Arizona's House of Representatives passed a restrictive bill limiting the female classification to 46, XX persons, it states:

> Athletic teams or sports designated for females, women or girls may not be open to students of the male sex. If disputed, a student may establish the student's sex

> by presenting a signed physician's statement that indicates the student's sex based on:
>
> 1. internal and external reproductive anatomy, 2. the student's normal endogenously produced levels of testosterone, 3. an analysis of the student's genetic makeup. (HB 2706, 2020, p. 1)

In 2019, The 116th Congress passed H.R. 5 "The Equality Act" (2019), which redefines "sex" in federal civil rights law to include "gender identity." The legislation would make it unlawful to discriminate against individuals based on their gender identity, and practically speaking, it would prevent distinguishing between people on the basis of sex, such as male and female-only sports. With Republican control of the Senate and White House, The Equality Act is unlikely to pass, but the 2019 vote in the House of Representative is an important social signal and massive implication if elections swing power to favor the Democrat party (The Equality Act, H.R. 5, 116th Cong., 2019).

On June 15, 2020, the Supreme Court, in the landmark ruling Bostock v. Clayton County, Georgia (2020), declared discrimination against transgender persons was discrimination on the basis of sex and thus violated Title VII. Title VII states it is, "unlawful . . . for an employer to fail or refuse to hire or to discharge any individual, or otherwise to discriminate against any individual . . . because of such individual's race, color, religion, sex, or national origin" (42 U. S. C. §2000e–2[a]). In making this ruling, the court defined transgender discrimination as sex discrimination. In Bostock, the court definitively ruled that sex in the classical sense has been replaced by, or at least must include, gender identity. The looming question is whether these findings will next be applied to Title IX in the context of education-based sport (Coleman et al., 2020). Many

experts and observers note, including the dissenting opinion by Justice Alito, that this new analysis of Title VII may apply to Title IX, and in doing so, fundamentally transform sports in America (Barnes, 2020).

Emerging Conflict Regarding Transgender Policies

Fully affirming state laws that privilege gender identity over biological sex, could be subject to a federal challenge in court. Jurist Ray Hacke (2018) notes that the legal status of transgender policy could fall under the Equal Protection Clause (EPC), Title IX, and state female equality statutes. Hacke argues that the EPC should apply because physiological differences between males and females place females at a competitive disadvantage, and it creates a potential safety risk. Additionally, he argues, Title IX can be invoked to privilege 46, XX athletes, because of the plain reading of the text, and state of mind and intent of the signatories. This assertion is also the position of the Trump Administration's Justice Department. In a *Statement of Interest* brief, the Attorney General of the United States, William Barr, said:

> Allowing biological males to compete in all-female sports deprives women of the opportunity to participate fully and fairly in sports and is fundamentally unfair to female athletes. Sports are an important part of education and character development and provide an arena where individual discipline can result in achievement and recognition. The purpose of all-female athletics is to ensure that women have an equal opportunity to participate, compete and excel in this important part of life. Title IX has been a major step forward in the long fight to achieve this equality. As reflected in Title IX, the basis for single-sex athletics, is rooted in the reality of biological differences between the sexes. Clearly then,

> eligibility to participate on a single-sex team must be based on objective biological fact. Girls should not be forced, through the dismantling of Title IX, to be sidelined in their own sports. (U.S. Department of Justice, 2020, para. 2)

Part IV: Aspects of Transgender Sport Policy

Fairness

Thoughtful consideration of transgender sport policy requires reflection on the philosophical nature of fairness. Considerations of fairness must distinguish how fairness affects one segment of the population versus another. Often, what is fair for one is not for the other. Sometimes an intersection of fairness between the two can be established, but often achieving fairness for both groups is illusory. MTF transgender individuals may have an unfair advantage if allowed to compete with 46, XX females, especially if without restrictions. However, it can also be viewed as unfair to discriminate against transgender athletes solely based on their sex (Chen, 2018).

Many such as Handelsman et al. (2018) have discussed that achieving fairness is unworkable due to it being a subjective concept, changing over time as societal views change. What once was considered unfair, like getting paid to play, or having organized training, could evolve over time. The desire for distributive justice and a fondness for anti-discrimination and human rights legislation is present on both sides of the issue.

The Mission of Sports Institutions: Competition Versus Inclusion

The two prominent positions that are in intellectual and philosophical opposition in developing transgender policy, and can be broadly summarized as being *competition-focused* versus *inclusion-focused*. For the transgender advocate, what matters most is an MTF athlete's liberty to self-identify, and in doing such, be treated equally as female in

all of society. On the other side, of highest value is the ability of 46, XX persons to participate in fair and meaningful competition.

Arguments for a Competition Focus

While proponents of both views esteem the virtue building nature of sport on the individual, the reasons that sporting organizations exist is to regulate and be discriminatory as to the rules and participants and to ensure that there is a realistic possibility of competing for the wins. If regulating fairness in the pursuit of winning is not central to their mission, there is no need for such sporting bodies (Vamplew, 2007).

The idea of "natural kinds" is a philosophical view that living beings can be rationally sorted based on their form, function, appearance, and ability to reproduce with another member of the same kind (Bird, 2010; Chen, 2018; Dupré, 1981). Because of the vast differences between 46, XY and 46, XX persons, as noted in Part I, applying the concept of "natural kinds" to sports is a logical progression.

Sex-segregated opportunities for participation in competitive athletes promote values of fitness and athleticism for lifelong health benefits for 46, XX individuals. Competitive sport, regardless of the level, imparts socially valuable traits, including teamwork, sportsmanship, and leadership, as well as success-critical traits like goal setting, time management, perseverance, discipline, and grit (Coleman et al., 2020).

Equality for females is a widely accepted and embraced philosophy in society. This commitment to female equality extends to sport. Sports are good for girls. "Girls who play sport stay in school longer, suffer fewer health problems, enter the labor force at higher rates, and are more likely to land better jobs. They are also more likely to lead" (Kotschwar, 2014, p. 1). Additionally, although participation contributes to equality,

competitive equity and in victories matter, as well as scholarships, accolades, finals, and podiums (Coleman, 2017).

Because of their different secondary sex characteristics, 46, XX persons, in general, have different physical capacities and experiences than 46, XY persons, regardless of how they identify. Thus, segregation on the basis of sex or at a minimum, sex traits, is necessary to promote female equality. If there were not a female classification, competitive sport would be exclusively 46, XY:

> …..even at their absolute best (female Olympic Champions), the women would lose to literally thousands of boys and men, including to thousands who would be considered second tier in the men's category. And because it only takes three male-bodied athletes to preclude the best females from the medal stand, and eight to exclude them from the track, it doesn't matter if only a handful turn out to be gender nonconforming. (Coleman, 2019, p. 4)

Thus, "if we want females to win some competitions, we need to separate them" (Coleman et al., 2020, p. 99).

Arguments for an Inclusion Focus

The World Anti-doping Agency (2015) argues that sport is a critical social structure in which character is built and people thrive. The inclusion focus argument suggests that policies should be fully affirming, to include as many people, especially young people, to help them thrive. The National Transgender Discrimination Survey (NTDS) documented the negative health effects caused by stigma and gender identity regulations, including reporting 41% of respondents said to have attempted suicide in their lifetime (Wilson & Kastanis, 2015). Given this alarming statistic, full inclusion and

affirmation is seen by many as the way forward towards improving these results. Proponents of this view believe because competitions become a conduit for physical activity, teamwork, life lessons, etc., "inclusion, equal opportunity, and acceptance should be the goals when establishing such standards" (Zeigler & Huntley, 2013).

Historian Vamplew (2007) notes, "there is nothing in the nature of sport itself which determines who can and cannot play. In the purest form of sport only self-exclusion should apply" (p. 851). Knox et al. (2019) and Jones et al. (2017) advocate that there is a fundamental right for everyone to be recognized in the gender in which they identify; including sports. Dr. Karissa Niehoff, the executive director of the National Federation of State High School Associations, summarizes the position in:

> [This] is not about the winning and losing. It's about the successful development of these kids, . . .And if we don't treat them [transgender student-athletes] respectfully, their development is going to lose. That's a much bigger issue than someone not getting a medal or a place in a race. Much bigger issue. (Albl, 2019, para. 8).

The position for inclusion can be summarized as; valuing equal protection for all of its citizens, especially for the most vulnerable, and because of the positive benefits of sports in the educational setting, society should afford those values to the transgender youth, regardless of competitive factors.

Frequent Transgender Activists' Arguments and Opponent Responses

Transgender activists do not typically embrace a position that acknowledges transgender participation negatively affecting the competitive female classification framework. Most often, it is argued by these proponents that inclusion is paramount, but

also that MTF inclusion is not as disruptive as opponents claim. The two most frequently used arguments from a transgender activist are an appeal to the *absence of evidence* on transgender individuals and the *myth of the level playing field*.

The absence of evidence argument suggests that no direct or consistent evidence exists that shows MTF athletes have an athletic advantage over 46, XX persons (Jones et al., 2017). Such activists would point to the lack of controlled longitudinal studies with transgender MTF athlete subjects. Certain activists such as Buzuvis (2016) are even in denial of an objective male advantage, and that policies are based on "biased assumption of [female] athletic inferiority" (p. 35). Making their point, Jones et al. (2017) suggests "there is no direct or consistent research suggesting transgender female individuals have an athletic advantage at any stage of their transition and, therefore, competitive sport policies that place restrictions on transgender people need to be considered and potentially revised" (p. 701).

Opponents, however, would note; the absence of evidence is not evidence of absence. This "appeal to ignorance" or *argumentum ad ignorantiam* is logical fallacy first attributed to philosopher John Lock (1690) as cited in Walton (1999). The tactic involves a reversal of the burden of proof: "if proposition A is not known (proved) to be true (false), therefore A is false (true)" (p. 368). Opponents would concede that there is little athlete performance data on athletes transitioning; however, this absence can be explained because randomized placebo-controlled studies would be unethical in most cases (Bermon et al., 2014). For example, data extracted from doping programs in female athletics in the former German Democratic Republic are helpful to determine the effect of

testosterone on FTM persons, but replicating the results is impracticable (Franke & Berendonk, 1997).

The myth of the level playing field argument proposed by proponents such as Buzuvis (2016) and Jones et al. (2017), posits that there are always going to be unfair advantages, and thus, it is a "myth" that there is ever fairness in sport. For example, longer limbs of Michael Phelps, Lebron James, and Usain Bolt, and the lower center of gravity of Simone Biles put them at an advantage. Geography, socioeconomic status, training facilities, equipment, support systems, nutrition, etc., will always favor some over others; therefore, biological sex is simply another variable of potential inequality amongst many. This argument is most frequently refuted by opponents by the extensive evidence reviewed in Part I. Coleman, as cited in Coleman et al. (2020), presents the central opponent response to the "myth" claim, "[t]here is no characteristic that matters more than testes and testosterone. Pick your body part, your geography, and your socioeconomic status and do your comparative homework" (p. 98).

Transgender Sports Policies

How to effectively police restrictions on the female classification presents a number of difficulties. Americans are decidedly private. Policing restrictions inevitably intrudes into that private space, and using the least-invasive evidence-based methods are the challenge of organizations with such policies. Nevertheless, the necessity of maintaining fair play in female events will remain as long as separate female competitions exist (Handelsman et al., 2018).

Many past policies have struggled with how to enforce their protections on the female classification. Controversy over sex verification, in particular, has existed for over

75 years (Tucker & Collins, 2010; Vamplew, 2007). Simpson et al. (2000) report a troubling history of sex verification, including "nude parades" of the 1960s, that were invasive, embarrassing, inaccurate with ambiguous genitalia, and ultimately unable to account for certain disorders of sex development (DSD). Barr body tests were later implemented to test for the presence of two X chromosomes; a positive test would equal a certificate of femininity. The world record holder in the women's 100m was disqualified using this test in 1965. They further report that false-negatives and false-positives were a common problem with the Barr body test, with androgen insensitivity syndrome (AIS) being one the most common DSD that was misread by the test.

The IAAF banned compulsory gender testing in 1992 and emplaced a suspicion-based protocol. The Barr body test was replaced by identifying the polymerase chain reaction for the Y-linked SRY gene in 1992. At the 1996 Atlanta Olympics, 8 of over 3,000 samples tested positive, but all were ruled to have naturally occurring AIS and allowed to compete. The International Olympic Committee (IOC) formally abolished compulsory testing in 1999 (Tucker & Collins, 2010).

A concern for forced medical treatment in order to be "fully affirmed" seems well found. Medical freedom of choice is a longstanding American value and has been honored by the courts (Monahan et al., 1995). Forced medical treatment is rare in America and generally unlawful. Longstanding American legal precedent and practice establishes the right for persons to reject invasive medical procedures. Many transgender persons are uncomfortable with any sort of medical treatment. The surging numbers of young transgender people is taking place at a time where many young people even view eating food with hormones as unthinkable (Brewer, 2008). Jones et al. (2017) and other

activists believe CSHT and surgical interventions should not be a requirement at any sporting level. With this mindset, transgender persons may or may not want any cosmetic or hormonal treatment and are likely to have a legal case to challenge an American policy stipulating invasive medical treatment as a pre-condition for participation.

Finding Middle Ground

Middle ground between competition and inclusion is going to be a significant challenge. Harper (2019) and others are actively researching at Loughborough University in the U.K. to present evidence-based policies that are both fair to 46, XX athletes, and inclusive to MTF transgender persons. Ziegler and Huntley (2013) and Chen (2018) believe that policies should first value inclusion and opportunity to participate for all, while balancing fairness of competition for all. Yet to be determined is how valuing inclusion and opportunity, and competitive fairness, is to be accomplished through policy creation. It has yet to be shown how one can have a commitment to full affirmation, medical freedom, and fair and meaningful female sports; simultaneously.

International Transgender Policies

International transgender policies have undergone recent changes. The Athlete Biological Passport (ABP) has developed into an alternate means of sex testing at the international level. A longitudinal profile of ABP for elite athletes can hopefully detect abnormalities (cheating) better than drug tests, but it can also be used to determine serum androgen levels of MTF athletes or hyperandrogenism (DSDs) in elite female populations.

The international cutoff for testosterone formerly was 10nmol/L, chosen arbitrarily by the IAAF (now World Athletics) in the absence of normative androgen

statistics in elite female athletes. Recent findings of normative elite female ranges, show the 99th percentile for testosterone is 3.08 nmol/L, prompting the recent tightening of the female threshold from 10 nmol/L to 5 nmol/L in 2019 (Bermon et al., 2014; Clark et al., 2019).

The World Athletics eligibility regulations for transgender athletes (2019) reads as follows:

> 1. [MTF athletes] must provide a written and signed declaration, in a form satisfactory to the Medical Manager, that [they] gender identity is female.
>
> 2. [MTF athletes] must demonstrate to the satisfaction of the Expert Panel (on the balance of probabilities), in accordance with clause 4, that the concentration of testosterone in [they] serum has been less than 5 nmol/L continuously for a period of at least 12 months.
>
> 3. [MTF athletes] must keep her serum testosterone concentration below 5 nmol/L for so long as [they] wishes to maintain [their] eligibility to compete in the female category of competition. (p. 5)

There is no longer a requirement to have a legal recognition of the athlete's gender identity or any surgical anatomical changes at the international level.

In July 2020, World Rugby announced that it may be the first international sports federation to prevent MTF athletes from competing in their sport, citing safety due to injury risk and performance differential remaining, even after testosterone suppression (Orchard, 2020). World Rugby report at least a 20% to 30% greater risk of injury to a 46, XX player when tackled by a 46,XY person who has gone through puberty (Ingle, 2020). World Rugby is scheduled to vote on the restrictive policy in November 2020.

Intercollegiate Transgender Policies

In America, collegiate sport is primarily overseen by the NCAA. The latest policy guidance regarding transgender athletes at the collegiate level comes from the dated handbook titled, NCAA Inclusion of Transgender Student-Athletes (2011) that requires:

> A trans female (MTF) student-athlete being treated with testosterone suppression medication for Gender Identity Disorder or gender dysphoria and/or Transsexualism, for the purposes of NCAA competition may continue to compete on a men's team but may not compete on a women's team without changing it to a mixed team status until completing one calendar year of testosterone suppression treatment. (p. 13)

The NCAA policy establishes medical privacy as a guiding principle, and yet mandates medical treatment to fully affirm MTF persons (p. 14). The policy cites "damage" caused by forcing someone to participate on a team other than their gender identity, and then subjects MTF individuals to that very outcome unless they are willing to undergo invasive medical treatment (Lenzi, 2017).

Interscholastic Transgender Policies

According to Lenzi (2017), high school educational institutions promote competitions for young people and identify and celebrate the champions. At the high school level of competition, there is no national sporting authority, and thus, each state independently sets its transgender athlete policies. Presently, 17 states allow transgender students to compete exclusively on the basis of gender identity, regardless of treatment or legal status. Ten states allow competition based on birth certificate, and 16 states operate

on a case-by-case basis, using a review process, or requiring MTF individuals to complete CSHT. Six states have no policy at all (Barnes, 2020; Eaton-Robb, 2019).

The five states included in this dissertation's study are California, Florida, Minnesota, New York and Washington: All are considered as part of the 17 that are fully affirming; that is, competition classification based exclusively on gender identity.

Part VI: History, Theory, and Use of Monte Carlo Simulations

The researcher feels it necessary to briefly discuss the Monte Carlo method due to the methodology of question three of this study. The Monte Carlo method is a means of approximately solving problems by simulating the problem by using random numbers. Lund (1981) defines it as "representing the solution of a problem as a parameter of a hypothetical population and using random sequence of numbers to construct a sample population, from which statistical estimates of the parameter can be obtained" (p. 4). As such, it has been widely utilized and valid means of probabilistic inference.

The Monte Carlo method utilizes repeated random sampling to obtain numerical results. It is one of the main applications of computerized random number generators. The Monte Carlo method goes back to the B.C.E. ancient world; it rose in use in the 18th and 19th centuries and reached wide use in America starting in the 1950s (Lund, 1981). Today, the Monte Carlo simulation is used in a wide range of fields, from nuclear physics to finance to sports.

Monte Carlo simulations rely on the *law of large numbers* and uniform random number distributions (Lund, 1981). The law of large numbers states that as sample sizes (simulation numbers) grow, its mean gets closer to the average of the whole population (Hsu & Robbins, 1947).

Some of the notable uses of Monte Carlo simulations in the sports world are: Freeze (1974) determining the effect of batting order on wins in baseball; Baumer (2009) estimating the magnitude of the effect of baserunning skills upon a team's run-scoring ability; Silva et al. (2002) developed a final classification ranking model in soccer; and Newton and Aslam (2006) using a Monte Carlo simulation to predict the probability of winning in tennis. Given the long history, wide use, and validity of the Monte Carlo method, its use in this study seems justified.

Summary of the Literature

In summary, human beings have a chromosomal destiny of 46, XX or 46, XY that fundamentally affects their physical development. Variations occur on rare occasions, and these are known as DSDs. Prepubescent differences are marginal, if existing at all. With the onset of puberty, corresponding to a rise in endogenous testosterone in 46, XY persons, a large divergence in the human species takes place.

46, XY persons are 7% to 8% taller than 46, XX. In addition to longer bones, 46, XX individuals have only 65% to 75% of the cross-sectional area of the humerus and 85% of the 46, XY femur, leading to higher injury risk, lower force generation capability, and reduced kinetic mass for 46, XX persons. 46, XY people have roughly 36% greater body mass, and a difference of 40% in the upper body alone. 46, XY persons have a greater ratio of fast-twitch Type II muscle fibers in relation to slow-twitch Type I ones. Additionally, the fiber cross-sectional area of both fiber types is significantly larger than 46, XX counterparts: Type I are roughly 19% bigger and Type II 60% larger in 46, XY persons.

Due to their bigger size, 46, XY people have a larger blood volume, and critically, around 12% greater hemoglobin, bringing more oxygen to tissues. The 46, XY heart size, stroke volume, and cardiac output are all larger. These factors, as well as bigger lungs, lead to up to a 20% to 30% greater VO_2 max for 46, XY persons. Nervous system differences translate to superior speed and maximal rate of force development for 46, XY persons. Changes are even seen in the brain, with 46, XY and 46, XX persons approach pacing, competition, and risk-taking different, with 46, XX persons doing better at pacing, but 46, XY persons taking more risk and being more competitive. The evidence suggests that the cause of sex differences occurs via male puberty, with testosterone being the key difference. 46, XY persons frequently have 15 to 20 times the testosterone concentrations found in 46, XX persons.

The physiological differences translate to between an estimated 5% and 30% performance difference in sports, depending on the event. 46, XX persons have roughly 50% of the upper body strength and 60% to 70% lower body strength of 46, XY counterparts. Power differences in 46, XY individuals are frequently reported to be twice that of 46, XX persons. Gaps in performance are the smallest in endurance swimming and largest in events dependent on the upper body, such as cross country skiing and rowing. In track and field, there is a difference of around 9% for the short sprint distances of 100m, and an estimated 19% gap in jump performance.

Part II reviewed how Title IX is a critical piece of equality legislation that has assisted millions of girls to participate and thrive in sports. Its legislative history has not been without issues, but the overall good that it brings to girls is unquestioned. Part III reviewed the rise of the transgender movement and its subsequent impact on the culture;

awareness and the embrace of the transgender movement is widespread. Part IV reviewed practical and philosophical considerations in making transgender policy. Competing interests are facing off in ideological warfare on the subject. Supreme Court and Congressional signals point to the potential for a decisive nationwide legal decision on the topic, but until then, athletic organizations and state governments are trying to balance the concerns of competitive fairness for girls' sport, and inclusion of MTF transgender persons.

CHAPTER III

METHODOLOGY

Statement of the Problem

Biological differences in 46, XY and 46, XX persons affect the structure and function of the human body (Carlson, 2018). This function and structure translate to differences in sports performance (Handelsman, 2017). However, despite the differences, many state high school governing bodies favor gender identity as the determining factor for separation into boys' and girls' categories (Tamerler, 2020). In light of the growing transgender movement, further research is needed regarding biological performance considerations for sports competition. The main issue at hand is whether transgender inclusion in female athletics constitutes an incursion on fair competition.

Is it reasonable to allow high school participants to self-select if they are eligible to participate in the protected girls' classification? This question is important and culturally relevant. To inform an evidence-based answer, it is reasonable and prudent to produce research on the nature of sex differences in high school athletes, as well as the possibility for disruption of the girls' classification. The research to date on the subject is scant.

Purpose of the Study

The purpose of the study was to evaluate the potential existence, nature, and extent of an advantage for 46, XY persons competing in the girls' category as MTF transgender track and field athletes, and it investigates the probability of MTF transgender persons outperforming the top state 46, XX competitors. The study seeks to

estimate the probability that an MTF athlete could be better than the best 46, XX, in states with self-identification and no hormone intervention requirements, and estimates of transgender teens are true and representative of 46, XY high school athletes. In short, if all transgender youth came "out of the closet," how would the best 46, XX persons do?

Nature of the Study

The study is concerned with the differences in athletic performance of the sexes, performance differences as it relates to sex and event distance, and the prevalence of athletes who are PFCs that may be transgender.

In sport, those that are the best (champions) are prominent. In general, society is not concerned about the doping habits of the persons in last place, but great efforts are taken, and governing bodies go to extremes to ensure the performers at the top, the champions, are rightly established. Cheating, blood doping, lying about age, and the like, create controversy with strong feelings of disdain that a champion, attained such status unjustly (World Anti-Doping Agency, 2015). Because of the unavoidable attention to the champions, and potential for transgender athletes to upset the competitive balance of girls' sport, this is a needed research project.

Descriptive statistics describe the nature, scale, and scope of the differences in 46, XY and 46, XX performance. Measures of central tendency were assessed. Histograms visually describe the presence or lack of a bimodal distribution, along with the shape and spread of the distributions. Z-tests for differences were done. Correlations for sex and performance were tested. Correlations between the percentage of PFCs and type of event were conducted. Best fit lines via linear regression assess the effect that endurance would have on the difference in sex performance. Finally, a Monte Carlo simulation involving

random number generation provides repeated sampling to obtain the statistical probability of transgender disruption of the girls sporting classification.

Selection of Subjects

The study's subjects were an approximately one million American high school track and field performances (N = 920,115). The subjects came from eight events, in five states, over three years, in both sexes (46, XX n = 400,929, 46XY n = 519,186). Official results data proceeds from the following events: high jump; long jump; 100m; 200m; 400m; 800m; 1600m; and 3200m. A representative state comes from five regions of the United States: Northwest, West, Midwest, Northeast, and Southeast. Utilizing purposeful sampling, the states selected were California, Florida, Minnesota, New York, and Washington. The years under consideration are 2017, 2018, and 2019.

Three years are selected to increase statistical power and control for outlier performances. The five states are representative of the region from which they are selected. A criterion for state selection was the state policy regarding transgender athletes being "fully inclusive," (e.g., not requiring any hormonal or biological interventions for 46, XY persons to compete in the girls classification). An additional criterion required states in each region to have a high number of participants. The eight events were selected for reliability, validity, and comparison purposes. Both the boys and girls in the events compete by the same rules, distance, and equipment, leading to "apples to apples" comparison. The study uses the outdoor season because of its broader participation.

Instrumentation

Data was collected from official results available through the now ubiquitous track and field result database athletic.net®. The stat tracking site athletic.net® 2020, a

subsidiary of Deker Net LLC., is the premier source of publicly available track and field data in the United States. Official results posted from athletic.net® have operational controls that require that they are uploaded from the meet host, and must report all places, all events, all participants, all marks of that meet. Its wide use and accepted validity make it an ideal medium for research.

Parameters of "official" and "fully automatic time" (FAT) were used. FAT mandates that the clock is automatically started by the starting device (gun), and the finish is automatically recorded or analyzed via a photo finish.

Statistical analysis was done through the data analysis add-in tool pack of Microsoft® Excel 2016 MSO (16.0.12624.20422). The data was extracted into Excel, .csv format through the use of ParseHub® version 54.0.1 data conversion and parsing software. The applications used in the data analysis application were "correlation," "descriptive statistics," "z-test: two-sample for means," "regression," and "random number generation."

Procedures

The flow of the procedures was extraction, trimming, formatting and cleaning, combining and sorting, describing and charting, running, analyzing, and finally reporting the data.

Extraction

Following committee and IRB approval for the study, the researcher established an account with athletic.net and ParseHub. A new project was created in ParseHub by which marks (times and distances) are extracted, page-by-page in a copy and paste method. Data filters were; "high school level," "outdoor" season, "FAT," "official,"

"English" standards (except for the jumps, as they were measured utilizing the "metric" system).

The collection began with California, high jump, boys, 2017; progressed through 2019; then girls 2017, 2018, 2019. The collection then replicated the process through the events from high jump, long jump, 100m, 200m, 400m, 800m, 1600m, 3200m. Each year in each event was saved as its own .csv file, later to be copied into a spreadsheet with all states and years included.

The collection extracted only times (distances for jumps) posted and no other personally identifiable information. Therefore, the anonymity of the human subjects was maintained and the utmost regard for the ethical treatment of human subjects in line with IRB guidance.

Trimming

Each year's dataset was scrubbed for defects in the count of participants to ensure page completeness. Data from 2017 Minnesota in the events of long jump, 100M, and 200M were partially available and therefore rejected from the study for both sexes. The data excludes wind adjustment figures due to the season best reporting in the list. If the filtering box was checked for wind adjustments, many participants would be completely eliminated from the study, despite having competed many wind legal results in the year.

The data scrub was done utilizing the "text to columns" tab under data, then "delimited" function in Excel. Next, the researcher removed the right tail 0.5% bottom outliers in each year/event. This data trimming was done to increase the representativeness of the data. The 0.5% is sufficient to account for outlier occurrences, i.e., subjects that received times that are atypical of the generalized data because of such

things as falls, injuries, or disability. This decision increase histogram representativeness due to the x-axis scale and the right-tailed nature of continuous data.

To make the 0.5% systematic data trim, the researcher identified the original count in the data, multiplied the count number by .995 in Excel using the formula function, and rounded to the nearest integer. Once identified, the bottom 0.5% of numbers were deleted from the dataset.

Formatting and Cleaning

The researcher next formatted the data for sorting and direct comparison. The data cleaning converted all times that are "mm:ss.00" to the seconds and hundredths "sss.00" format. This formatting is essential for processing descriptive statistics, sorting, and visualization in Excel. Because Excel stores time as date-time, a fraction of a 24-hour period, all "mm:ss.00" times must be multiplied by 86400 in order to be converted to seconds.

To complete this task, the researcher placed 86400 into a blank cell, then copied 86400, highlight the times to convert, right-clicked "paste special," "multiply," and clicked enter. After complete, the researcher checked that formatting is numbered to two decimal places and sorted best to worst to verify. As a final step, there was an additional visual review for anomalies and ensure times and distances were posted in the proper format.

Combining and Sorting

The researcher created a spreadsheet per event that has all years included, with years still in separate columns. This information was used to build the percent PFC chart on which questions 2 and 3 depend. Next, the researcher created a spreadsheet that

combines all years into one column per sex. This spreadsheet creation was done by copying each year for each sex into one column, then using the sort function for the column from best performance to worst. This spreadsheet was the source data for question 1.

Question 1

Once the data was collected, formatted, and sorted into the appropriate spreadsheets, the researcher moved on to the questions. Question 1 asks if there a statistically significant, bimodal distribution in the performances of 46, XX and 46, XY persons. The researcher took the three years combined, separated by sex, created histograms, and ran descriptive statistics on each. The researcher created bins; whose size well reflects the distribution curve of the two sexes. Next, a histogram per sex was run in Excel. Then distributions of the sexes in frequency bins were placed side by side to create a cluster columns (histogram) chart showing both distributions on one axis. There were eight charts developed for the eight events.

The researcher then ran descriptive statistics on the separate performances distinguished by sex. This process was done through the “data” tab and “data analysis” add-in tool pack in Excel. Following the descriptive statistics, the researcher ran a z-test: two sample for means.

Following the z-test, the researcher run a correlation between sex and time. This analysis was done by creating an additional column on the summary spreadsheets for each event. To do this code, “2” for 46, XY and “1” for 46, XX, then combine into a single column and sort from best to worst. In the data analysis tool pack run correlation on “time” (“distance” for jumps) and “sex.”

Question 2

Question 2 asks if there is a statistically significant relationship between event distance and PFC size. The Pearson product-moment correlation was used to determine the strength and direction of a linear relationship between the two continuous variables. To do this analysis, the researcher created a spreadsheet combining all years into one tab per event, keeping years in separate columns. Then, the number of PFC for each year were identified by dividing PFC *n* by total 46, XY to get the percentage that is PFC (percent PFC). Following this step, total 46XY was multiplied by the Williams Institute transgender percentage for each state's 13 to 17-year-olds, to get the expected number of MTF transgender athletes in each year. The numbers were rounded to the nearest integer.

Once a table of distance and percent PFC is developed, the researcher ran a correlation with "event distance" and "percent PFC" in the data analysis application. Regression analysis produced the associated *p* value. Next, a scatterplot with "percent PFC" on the y-axis and "event distance" on the x-axis was created. The researcher converted the high jump and long jump from nominal levels of measurement to ratio levels of measurement in order to unify the distance measurements. To accomplish this analysis, the high jump, for presentation and analysis purposes, was estimated to be a 15-meter event and the long jump to be a 30-meter event. These distances are deemed an approximation of the critical physiological work output time that is the determinant of success in the events.

Following the percent PFC scatterplot, the researcher used the simple linear regression function in the Excel chart. This function was accomplished by right-clicking

the chart, clicking the plus button, clicking the trendline selection, selecting more options, checking the box for "display equation on chart" and "display *r* squared on chart."

Question 3

For Question 3, the researcher ran 111 random number Monte Carlo simulations of 10,000 trials each. A fresh Excel tab displayed the data from one year. In each simulation, "*n* MTF Trans" signifies the number of random numbers to draw in the trial and is a product of percent transgender estimate and *n* 46, XY. The number ranges from which to draw the random numbers will be between "1" and the total *n* 46, XY for that event and year.

The assumption is that being MTF is independent, uniformly distributed random variables. Using the "data analysis" application and the "random number generation" function, the number of variables were filled to equal 10,000 to trigger 10,000 trials per simulation. The number of random numbers to be filled equaled "*n* MTF Trans." The uniform distribution between "1" and the total 46, XY count was selected and the researcher clicked "enter" to run the trials.

Next, the researcher counted the number of simulated MTF female champions in each of the 10,000 trials. Using the "=COUNTIF" formula function, the column to be counted was selected, then a comma placed, followed by a less than or equal (<=) sign and the PFC number for that trial. The autofill pasted the "=COUNTIF" function to the right across all 10,000 trials by using the copy function, highlighting the entire row of the 10,000 trials, and paste.

The number of trials that have $\geq$ 1 MTF female champion were then counted. The researcher did this count by using the "=COUNTIF" function, highlighting the row to be

counted and use a comma and then ">=1". The total MTF female champion number was then be converted into a percentage. Finally, the researcher ran a post-hoc analysis of the descriptive statistics of the MTF female champion row.

Design and Preparation for Data Analysis

The research included a three-question quantitative design investigating the scope and scale of sex differences in high-school track and field and the implications of sex differences on the probability of transgender disruption of the female classification. Descriptive statistics, z-test, correlation, regression, and a Monte Carlo simulation were utilized. Inferential statistics were used to test the hypothesis of whether PFCs have a statistically significant probability of being transgender.

Question 1

Question 1 had variables sex and performance time (distance for jumping events). Sex is the independent variable, nominal level of measurement (dichotomous), and possible outcomes are 46, XY or 46, XX. Performance time (distance jumped) is the dependent variable, ratio level of measurement (continuous) with possible outcomes in the sss.00 format (0.00m for jumps).

Is there a statistically significant relationship in the performances of 46, XX and 46, XY high school track and field athletes, in selected events?

$H1_0$: There is not a statistically significant relationship in the performance of the 46, XY and 46, XX high school track and field athletes.

$H1_a$: There is a statistically significant positive relationship between performance and being 46, XY.

Test used: z-test. Test Statistic: z-score. Significance: $\alpha = .05$.

Question 1 allowed the researcher to create histograms, run descriptive statistics, conduct a z-test and correlation analysis. Histograms were created for the distributions of performances based on sex. Along the x-axis is time for the running events and distance for the jumping events. The y-axis displays the frequency of performances for the 46, XY and 46, XX samples. Separate histograms were created for each event and sex. A combination overlay histogram with both sexes on the same chart shows a visual representation of the two samples.

Descriptive statistics were run on both 46, XY and 46, XX samples independently. The statistics considered were mean, mode, standard deviation, variance, kurtosis, range, and count. The most important descriptive statistics to the researcher are mean, standard deviation, and kurtosis.

A z-test for the relationship between 46, XY and 46, XX population was done. This result tests if a difference in the performance of sexes is likely due to chance alone. The test statistic is the z-score with a significance $\alpha = .05$.

Finally, a test of correlation between sex and performance was done with the Pearson correlation coefficient r being the test statistic. The researcher reports on the strength and direction of the relationship for all eight events. Question 1 was expected to confirm existing literature on the existence of a performance difference in the sexes, but it makes new discoveries as to the extent of the difference in track and field high school athletes.

Question 2

Question 2 has the variables event distance in meters and percent PFC. Event distance is the independent variable, ratio level of measurement, and possible outcomes

are 15m; 30m ; 100m ; 200m ; 400m ; 800m ; 1600m ; 3200m. Percent PFC is the dependent variable, ratio level of measurement (continuous) with possible outcomes in the 0.00% format.

Question 2 asks, Is there a statistically significant relationship between event type and percentage of potential female champions? I.e., do shorter events relying on maximal strength and power (short-time/bioenergetic pathway) have a larger amount of PFCs than events that rely on more aerobic/endurance capabilities?

$H2_0$: There is not a statistically significant relationship between the percentage of PFCs and the event distance.

$H2_a$: There is a statistically significant relationship between the percentage of PFCs and the event distance.

Test: linear regression. Test Statistic: Pearson's coefficient *r*. Non-directional.

Significance: $\alpha = .05$

Using the data analysis Excel application, simple linear regression tests for significance. The linear equation $y = mx + b$ was produced. A scatterplot was created with distance on the x-axis and percent PFC on the y-axis. A best-fit line accompanies the chart with the equation, and *r* squared visible on the chart. The summary output for the regression includes the correlation coefficient, Pearson's *r*, the correlation of determination, R^2, and *p*-value for significance.

Question 3

Question 3 has variables PFC, MTF, and (PFC and MTF). PFC and MTF are the independent variables, PFC is ordinal level of measurement, with integers the possible outcomes. MTF is ratio level of measurement with whole numbers (non-negative

integers) being possible outcomes. (PFC and MTF) is the dependent variable, ratio level of measurement (count) with whole numbers (non-negative integers) being possible outcomes.

Question 3 asks, what is the probability of one or more 46, XY potential female champions also being an MTF transgender individual? P(*n*[PFC and MTF] $\geq$ 1). This question was explored via a Monte Carlo simulation. There were 111 simulations (eight events, three years, five states) in which each simulation involved 10,000 trials of random number generation based on the assumed rate of transgender athletes in each event, each year. Each trial used the transgender population estimates from the Williams Institute to suggest how many of the 46, XY participants are theoretically transgender. The estimated percentage of transgender persons age 13 to 17 in the given states are: California, 0.85%; Florida, 0.78%; Minnesota, 0.85%; New York, 0.79%; Washington 0.70% (Herman et al., 2017).

Each simulation began with taking the data generated in spreadsheets that have the total 46, XY number for an event, and the *n* PFC in the event. Using these numbers, the researcher multiplied the Williams' estimate by the total *n* 46, XY event sample to get the estimated *n* MTF transgender total for the event. This total *n* MTF number was rounded to the nearest integer and dictated the number of random numbers to be selected in the 10,000 trials.

In each trial, random numbers, based on *n* MTF, were generated from "1" to the total 46, XY for the event; assuming a uniform random distribution. Any number randomly generated from "1" to *n* PFC, simulates a PFC being an MTF transgender athlete, and thus would generate evidence towards rejecting the null. Any trial that failed

to have a random number from "1" to *n* PFC, simulates no PFCs being MTF transgender athletes, and thus would generate evidence towards failing to reject the null.

To run the simulation, in the "data analysis" application in Excel, the researcher clicked "random number generation"; filled in the number of variables, "10,000"; filled in the number of random numbers, *n* MTF; selected "uniform" distribution; populated the draw range between "1" and total 46, XY. Next, using the "=COUNTIF" function, populated a row below each trial that counts the occurrence of "1" to *n* PFC being selected in each of the 10,000 trials (e.g., =COUNTIF(A3:A39, "<=354"). Then, using the "=COUNTIF" function, counted the number of trials that have one or more MTF PFCs (e.g., =COUNTIF(A40:NTR40, ">=1"). Finally, the researcher turned that number into a percentage based on 10,000 trials to get the probability of one or more 46, XY athletes being both a PFC and an MTF transgender person.

Post-hoc descriptive statistics of the *n* (PFC and MTF) could shed further light on the scale of potential PFC and MTF individuals. For example, if a track and field event had a mean of three or greater, with relatively small standard deviation, it could suggest the likelihood that all three spots on the podium in girls' events could be theoretically filled by 46, XY persons. Alternatively, if the mean is lower than one it could suggest that the probability of 46, XY MTF dominance of the girl category is unlikely, thus lending support to policies contingent on gender identity, rather than sex.

CHAPTER IV

RESULTS

Statement of the Problem

Biological differences in 46, XY and 46, XX persons affect the structure and function of the human body (Carlson, 2018). This function and structure translate to differences in sports performance (Handelsman, 2017). However, despite the differences, many state high school governing bodies favor gender identity as the determining factor for separation into boys' and girls' categories (Tamerler, 2020). In light of the growing transgender movement, further research is needed regarding biological performance considerations for sports competition. The main issue at hand is whether transgender inclusion in female athletics constitutes an incursion on fair competition.

Is it reasonable to allow high school participants to self-select if they are eligible to participate in the protected girls' classification? This question is important and culturally relevant. To inform an evidence-based answer, it is reasonable and prudent to produce research on the nature of sex differences in high school athletes, as well as the possibility for disruption of the girls' classification.

The research included a three-question quantitative design, $N = 920{,}115$, of which $n = 400{,}929$ were female (46, XX) and $n = 519{,}186$ male (46, XY) among 5 states (CA, FL, MN, NY, WA) over 3 years (2017 – 2019). The researcher excluded 1 year from the analysis in the Minnesota 2017 long jump, 100 meter, and 200 meter events because of incomplete dataset in the 46, XY sample (only top 599, 399, and 899 available

respectively). The researcher excluded data from New York 1600 meter and 3200 meter events because of marginal participation by females in those events (n = 75 - 264).

Purpose of Study

1. Investigate the underlying basis for post-pubertal sex segregation in sport.
2. Assess the effect of event distance on the performance differences between the sexes.
3. Assess the probability of a girls' champion being biologically male (46, XY).

Research Question 1

Is there a statistically significant relationship in the performances of 46, XX (female) and 46, XY (male) high school track and field athletes, in selected events?

- $H1_0$: There is not a statistically significant relationship in the performance of the 46, XY and 46, XX high school track and field athletes.
- $H1_a$: There is a statistically significant positive relationship between performance and being 46, XY.

A z-test: two sample for means between male and female performance was done to test if a difference in the performance of sexes is likely due to chance alone and a test of correlation between sex and performance was be done with the Pearson correlation coefficient r being the test statistic.

High Jump Results

In the comparison of 46, XX (n = 23,390) and 46, XY (n = 26,843) high jump distances, there was a statistically significant relationship between sex and jump distance (z = -221.3 ; $p < .001$). The means for 46, XX and 46, XY were 1.35 m (SD = 0.124) and 1.62 m (SD = 0.148) respectively. Findings reveal a strong positive correlation between

sex (46, XY) and distance jumped (r = .698). The mean advantage in favor of male performances was 18.18% (see Table 4.1).

Table 4.1

High Jump Comparison of Distance (Meters) and Sex

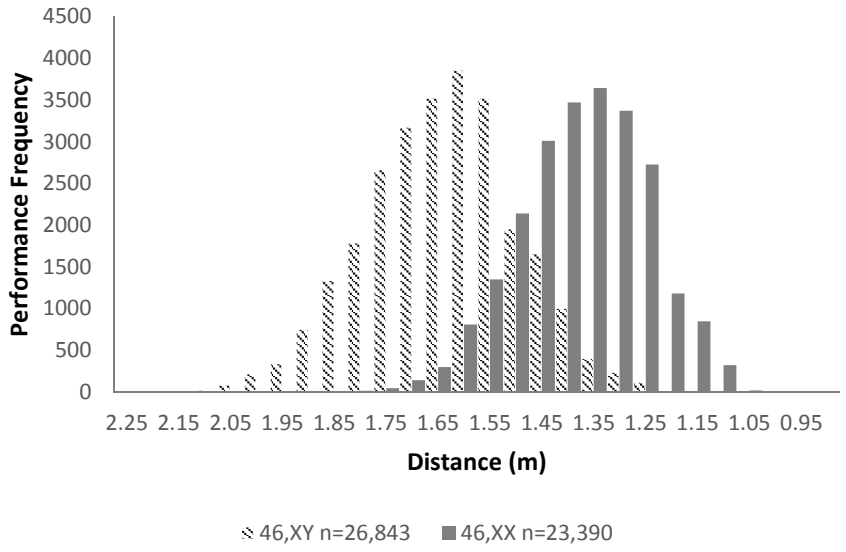
	46, XX	46, XY		
n	23,390	26,843		
M	1.35	1.62	M_{diff}	18.18%
SD	0.124	0.148	*Z*	-221.3
Variance	0.015	0.022	*p*	< .001
Kurtosis	-0.124	-0.077	*r*	0.698

Note. M_{diff} = mean difference in the samples.

As shown in Figure 4.1, the distributions of the high jump performances are of similar shape (kurtosis 46, XY= -0.124, 46, XX =-0.077) and slightly skewed left (negative).

Figure 4.1

High Jump Performance Distribution by Sex



Note. Mean difference = 18.18%. Mode 46, XX = 1.32 46, XY = 1.52; Range 46, XX = .94, 46, XY = 1.09.

Long Jump Results

In the comparison of 46, XX (n = 45,705) and 46, XY (n = 54,506) long jump distances, there was a statistically significant relationship between sex and jump distance (z = -266.61 ; $p < .001$). The means for 46, XX and 46, XY were 4.08 m (SD = 0.607) and 5.20 m (SD = 0.730) respectively. Findings reveal a strong positive correlation between sex (46, XY) and distance jumped (r = .638). The mean advantage in favor of male performances was 24.14% (see Table 4.2).

Table 4.2

Long Jump Comparison of Distance (Meters) and Sex

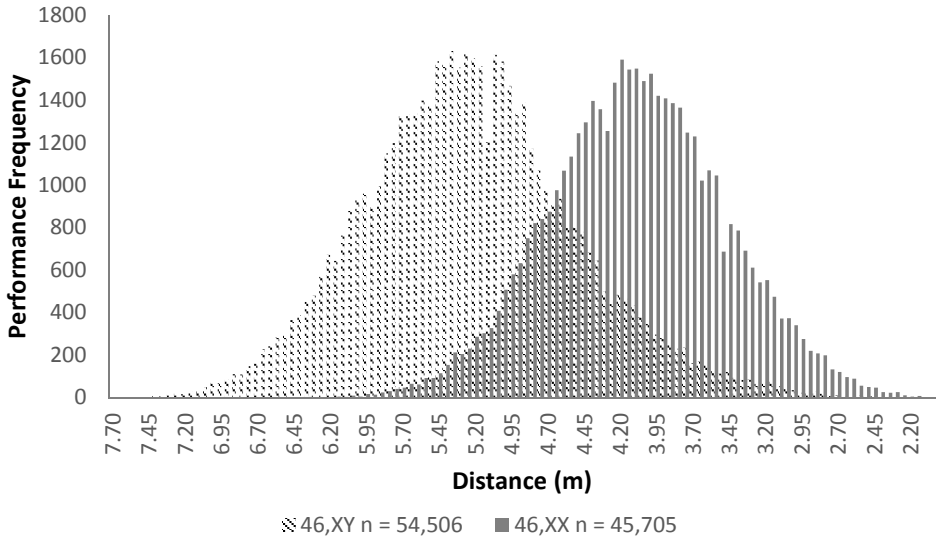
	46, XX	46, XY		
n	45,705	54,506		
M	4.08	5.20	M_{diff}	24.14%
SD	0.607	0.730	*Z*	-266.61
Variance	0.369	0.533	*p*	< .001
Kurtosis	-0.076	0.198	*r*	0.638

Note. M_{diff} = mean difference in the samples.

As shown in Figure 4.2, the distributions of the long jump performances are of similar shape (kurtosis 46, XX = -0.076, 46, XY = 0.198) and normally distributed.

Figure 4.2

Long Jump Performance Distribution by Sex



Note. Mean difference = 24.14%. Mode 46, XX = 3.96 46, XY = 5.18; Range 46, XX = 4.72, 46, XY = 5.29.

100 Meter Results

In the comparison of 46, XX (n = 79,663) and 46, XY (n = 94,447) 100M run times, there was a statistically significant relationship between sex and time (z = 342; p < .001). The means for 46, XX and 46, XY were 14.76 seconds (SD = 1.286) and 12.78 seconds (SD = 1.098) respectively. Findings reveal a strong negative correlation between sex (46, XY) and time (r = -0.639). The mean advantage in favor of male performances was 14.38% (see Table 4.3)

Table 4.3

100 Meter Comparison of Time (Seconds) and Sex

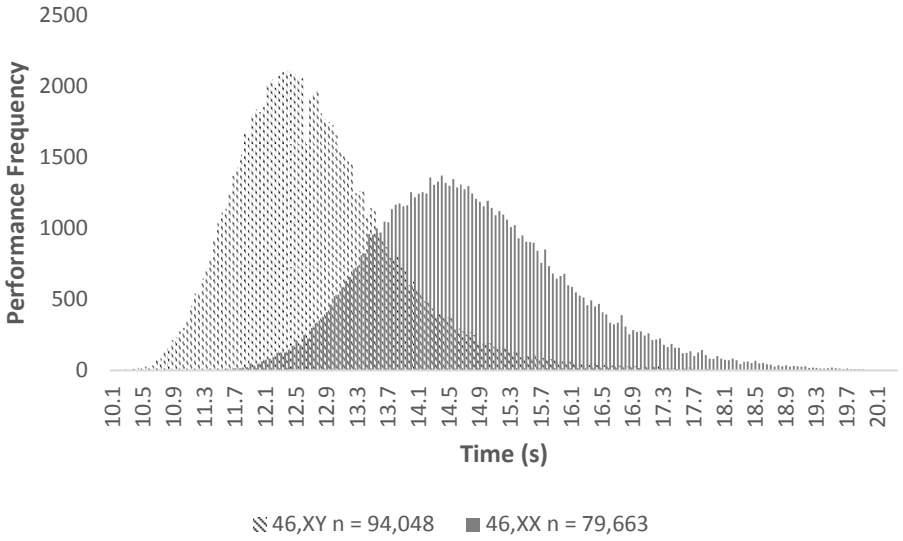
	46, XX	46, XY		
n	79,663	94,447		
M	14.76	12.78	M_{diff}	14.38%
SD	1.286	1.098	*Z*	342
Variance	1.653	1.205	*p*	< .001
Kurtosis	0.509	2.629	*r*	-0.639

Note. M_{diff} = mean difference in the samples.

As shown in Figure 4.3, the distributions of the 100M performances are of different shape (kurtosis 46, XX = 0.509, 46, XY = 2.629) and slightly skewed right (positive).

Figure 4.3

100 Meter Performance Distribution by Sex



Note. Mean difference = 14.38%. Mode 46, XX = 14.50 46, XY = 12.43; Range 46, XX = 9.14, 46, XY = 8.57.

200 Meter Results

In the comparison of 46, XX (n = 75,192) and 46, XY (n = 88,045) 200M run times, there was a statistically significant relationship between sex and time (z = 356.2; p < .001). The means for 46, XX and 46, XY were 30.75 seconds (SD = 2.856) and 26.15 seconds (SD = 2.269) respectively. Findings reveal a strong negative correlation between sex (46, XY) and time (r = -0.668). The mean advantage in favor of male performances was 16.17% (see Table 4.4).

Table 4.4

200 Meter Comparison of Time (Seconds) and Sex

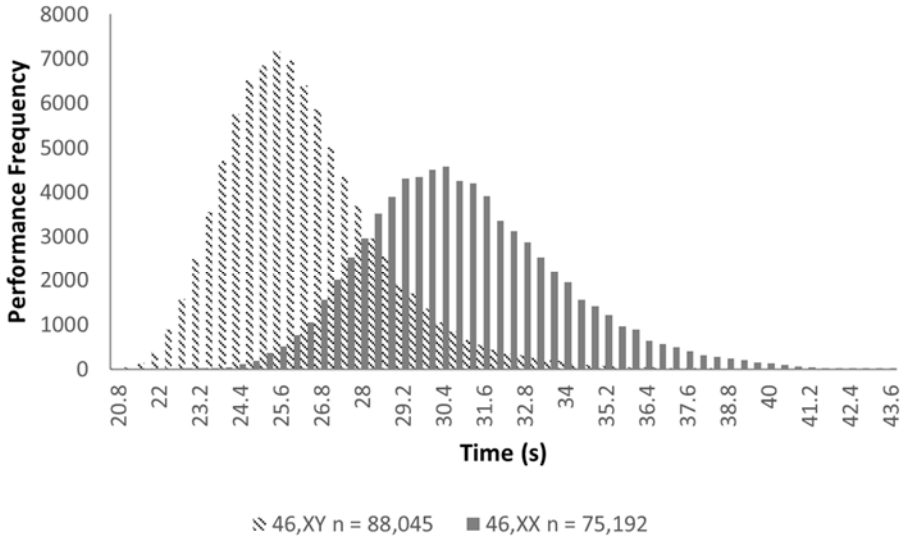
	46, XX	46, XY		
n	75,192	88,045		
M	30.75	26.15	M_{diff}	16.17%
SD	2.856	2.268	*Z*	356.2
Variance	8.159	5.151	*p*	< .001
Kurtosis	0.45	1.488	*r*	-0.638

Note. M_{diff} = mean difference in the samples.

As shown in Figure 4.4, the distributions of the 200M performances are of different shape (kurtosis 46, XX= 0.45, 46, XY = 1.488) and slightly skewed right (positive).

Figure 4.4

200 Meter Performance Distribution by Sex



Note. Mean difference = 16.17%. Mode 46, XX = 29.57, 46, XY = 25.10; Range 46, XX = 20.81, 46, XY = 17.83.

400 Meter Results

In the comparison of 46, XX (n = 52,050) and 46, XY (n = 69,517) 400M run times, there was a statistically significant relationship between sex and time (z = 293.5; p < .001). The means for 46, XX and 46, XY were 70.77 seconds (SD = 7.41) and 59.31 seconds (SD = 5.713) respectively. Findings reveal a strong negative correlation between sex (46, XY) and time (r = -0.658). The mean advantage in favor of male performances was 17.62% (see Table 4.5).

Table 4.5

400 Meter Comparison of Time (Seconds) and Sex

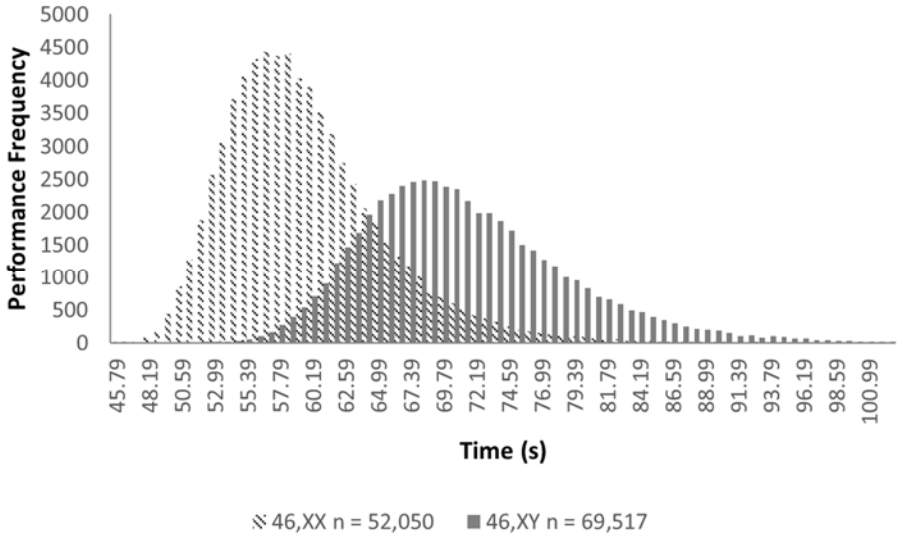
	46, XX	46, XY		
n	52,050	69,517		
M	70.77	59.31	M_{diff}	17.62%
SD	7.41	5.713	*Z*	293.5
Variance	54.91	32.64	*p*	< .001
Kurtosis	0.696	1.498	*r*	-0.658

Note. M_{diff} = mean difference in the samples.

As shown in Figure 4.5, the distributions of the 400M performances are of different shape (kurtosis 46, XX = 0.696, 46, XY = 1.498) and slightly skewed right (positive).

Figure 4.5

400 Meter Performance Distribution by Sex



Note. Mean difference = 17.62%. Mode 46, XX = 70 46, XY = 57.54; Range 46, XX = 50.68, 46, XY = 47.56.

800 Meter Results

In the comparison of 46, XX (n = 56,670) and 46, XY (n = 76,599) 800M run times, there was a statistically significant relationship between sex and time (z = 288; p < .001). The means for 46, XX and 46, XY were 172.03 seconds (SD = 7.41) and 143.67 seconds (SD = 5.713) respectively. Findings reveal a strong negative correlation between sex (46, XY) and time (r = -0.632). The mean advantage in favor of male performances was 17.96% (see Table 6).

Table 4.6

800 Meter Comparison of Time (Seconds) and Sex

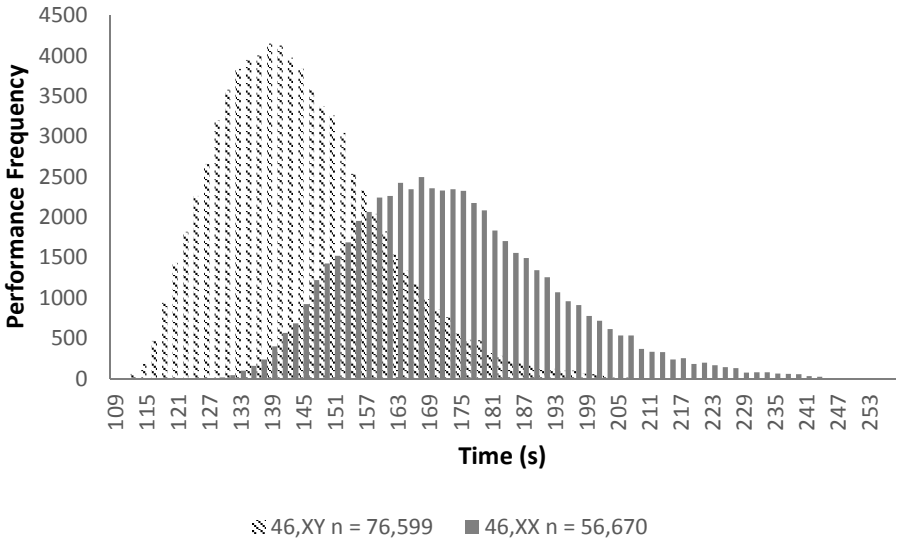
	46, XX	46, XY		
n	56,670	76,599		
M	172.03	143.67	M_{diff}	17.96%
SD	19.32	15.42	*Z*	288
Variance	373.33	237.82	*p*	< .001
Kurtosis	0.352	0.608	*r*	-0.632

Note. M_{diff} = mean difference in the samples.

As shown in Figure 6, the distributions of the 800M performances are of different shape (kurtosis 46, XX= 0.352, 46, XY = 0.608) and slightly skewed right (positive).

Figure 4.6

800 Meter Performance Distribution by Sex



Note. Mean difference = 17.96%. Mode 46, XX = 161.8 46, XY = 134; Range 46, XX = 144.22, 46, XY = 110.37.

1600 Meter Results

In the comparison of 46, XX (n = 43,914) and 46, XY (n = 68,787) 1600M run times, there was a statistically significant relationship between sex and time (z = 257.7; p < .001). The means for 46, XX and 46, XY were 382.99 seconds (SD = 43.47) and 320.35 seconds (SD = 33.24) respectively. Findings reveal a strong negative correlation between sex (46, XY) and time (r = -0.631). The mean advantage in favor of male performances was 17.81% (see Table 4.7).

Table 4.7

1600 Meter Comparison of Time (Seconds) and Sex

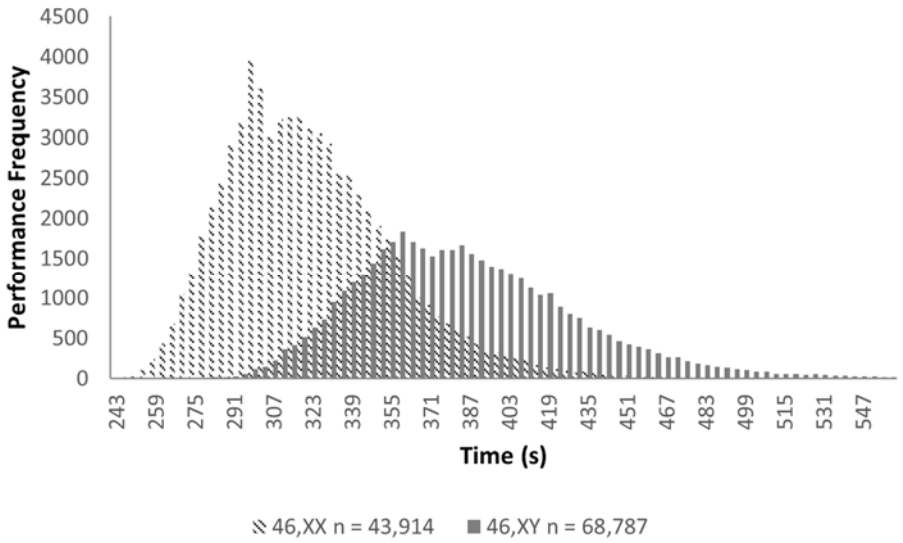
	46, XX	46, XY		
n	43,914	68,787		
M	382.99	320.35	M_{diff}	17.81%
SD	43.47	33.24	*Z*	257.7
Variance	1,889.49	1,104.83	*p*	< .001
Kurtosis	0.487	0.544	*r*	-0.631

Note. M_{diff} = mean difference in the samples.

As shown in Figure 7, the distributions of the 1600M performances are of similar shape (kurtosis 46, XX= 0.487, 46, XY = 0.544) and skewed right (positive).

Figure 4.7

1600 Meter Performance Distribution by Sex



Note. Mean difference = 17.81%. Mode 46, XX = 354.72, 46, XY = 287.79; Range 46, XX = 284.01, 46, XY = 225.38.

3200 Meter Results

In the comparison of 46, XX (n = 24,345) and 46, XY (n = 40,442) 3200M run times, there was a statistically significant relationship between sex and time (z = 183.7; p < .001). The means for 46, XX and 46, XY were 818.81 seconds (SD = 92.06) and 691.73 seconds (SD = 72.59) respectively. Findings reveal a strong negative correlation between sex (46, XY) and time (r = -0.608). The mean advantage in favor of male performances was 16.83% (see Table 4.8).

Table 4.8

3200 Meter Comparison of Time (Seconds) and Sex

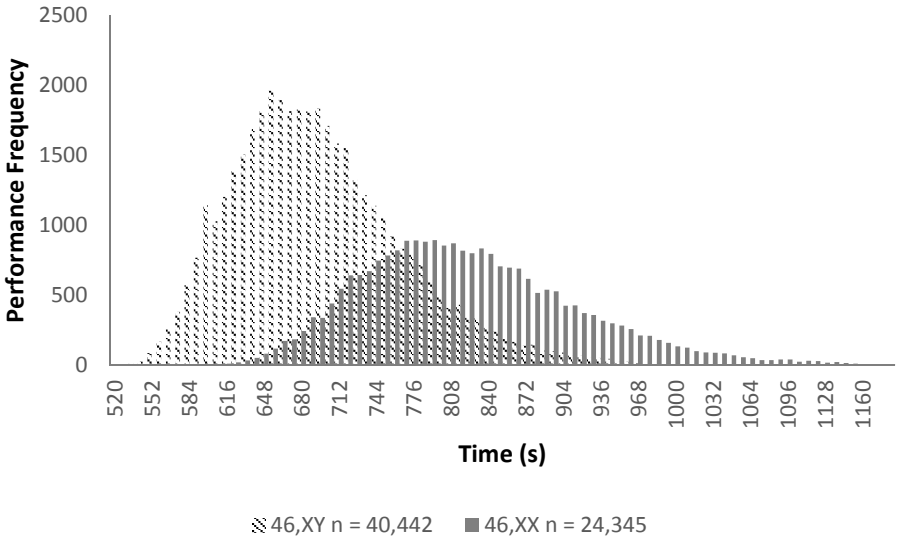
	46, XX	46, XY		
n	24,345	40,442		
M	818.81	691.73	M_{diff}	16.83%
SD	92.06	72.59	*Z*	183.7
Variance	8,475.55	5,268.60	*p*	< .001
Kurtosis	0.351	0.492	*r*	-0.608

Note. M_{diff} = mean difference in the samples.

As shown in Figure 4.8, the distributions of the 3200M performances are of similar shape (kurtosis 46, XX= 0.351, 46, XY = 0.492) and skewed right (positive).

Figure 4.8

3200 Meter Performance Distribution by Sex



Note. Mean difference = 16.83%. Mode 46, XX = 805 46, XY = 662.17; Range 46, XX = 586.06, 46, XY = 500.99.

Table 4.9 displays the summary statistics for question 1.

Table 4.9

Summary Differences Comparing Sex and Performance in High School Track and Field Events

	n		*M*			*SD*		Variance		Kurtosis				
	46, XX	46, XY	46, XX	46, XY	M_{diff}	46, XX	46, XY	46, XX	46, XY	46, XX	46, XY	*Z-score*	*p*	*r*
High Jump	23,390	26,843	1.35	1.62	18.2%	0.12	0.15	0.02	0.02	-0.12	-0.08	-221.3	< .001	0.70
Long* Jump	45,705	54,506	4.08	5.20	24.1%	0.61	0.73	0.37	0.53	-0.08	0.20	-266.6	< .001	0.64
100M*	79,663	94,447	14.76	12.78	14.4%	1.29	1.10	1.65	1.20	0.51	2.63	342.0	< .001	-0.64
200M*	75,192	88,045	30.75	26.15	16.2%	2.86	2.27	8.16	5.15	0.45	1.49	356.2	< .001	-0.67
400M	52,050	69,517	70.77	59.31	17.6%	7.41	5.71	54.91	32.64	0.70	1.50	293.5	< .001	-0.66
800M	56,670	76,599	172.03	143.67	18.0%	19.32	15.42	373.33	237.82	0.35	0.61	288.0	< .001	-0.63
1600M**	43,914	68,787	382.99	320.35	17.8%	43.47	33.24	1889.49	1104.83	0.49	0.54	257.7	< .001	-0.63
3200M**	24,345	40,442	818.81	691.73	16.8%	92.06	72.59	8475.55	5268.60	0.35	0.49	183.7	< .001	-0.61
Total	400,929	519,186			17.9%									

Note. N = 920,115. States = CA, FL, MN, NY, WA. 2017, 2018, 2019 outdoor seasons. *Excludes 2017 MN. **Excludes NY.

Post hoc analysis of M_{diff} and distance reveal non-statistically significant relationship (p = .652) between mean percentage difference and event distance, r = - .19, R^2 = .036, $F(1, 6) = 0.225$.

Post hoc analysis of participation differences by sex and event distance reveal a statistically significant ($p < .001$) strong correlation between sex participation percentage difference and event distance r = .93, R^2 = .87, $F(1, 6) = 38.58$ (see Table 4.10 and Figure 4.9).

Table 4.10

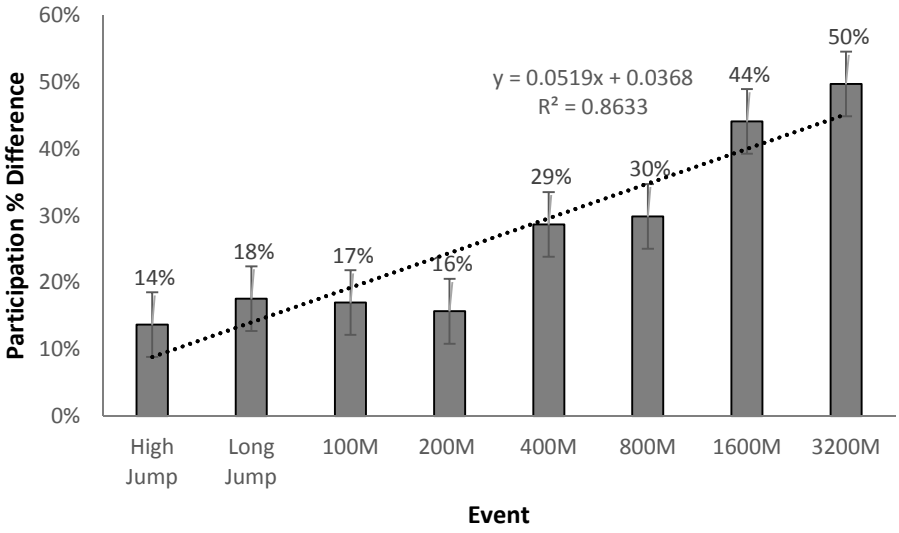
Difference in Participation by Sex

	n 46, XX	*n* 46, XY	% Difference by Sex
High Jump	23,390	26,843	13.7%
Long Jump*	45,705	54,506	17.6%
100M*	79,663	94,447	17.0%
200M*	75,192	88,045	15.7%
400M	52,050	69,517	28.7%
800M	56,670	76,599	29.9%
1600M**	43,914	68,787	44.1%
3200M**	24,345	40,442	49.7%
Total	400,929	519,186	25.7%

Note. N = 920,115. 44% Female and 56% Male. States = CA, FL, MN, NY, WA. 2017, 2018, 2019 outdoor seasons. *Excludes 2017 MN. **Excludes NY.

Figure 4.9

Difference in Participation Between Boys and Girls and Event Distance



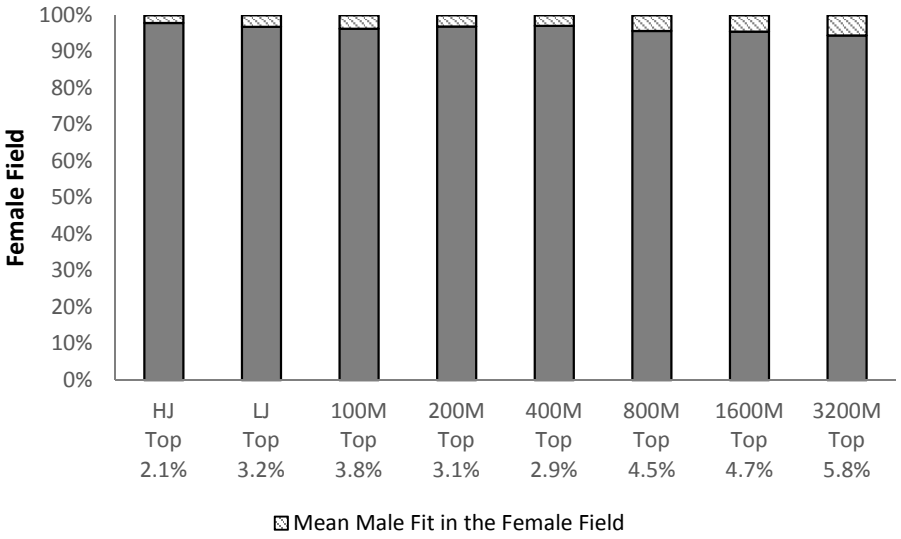
Note. N = 920,115. 400,929 female, 519,186 male. High jump = 15M and long jump = 30M for regression and correlation. r = .93, R^2 = .87, $F(1, 6) = 38.58$, p <.001.

Mean Fit Comparisons of Male and Female Performances

As shown in Figure 4.10, the mean of the male performances fit within the top 2.1%-5.8% of the 46, XX distribution.

Figure 4.10

Mean Male Performance Fit Among the Female Field

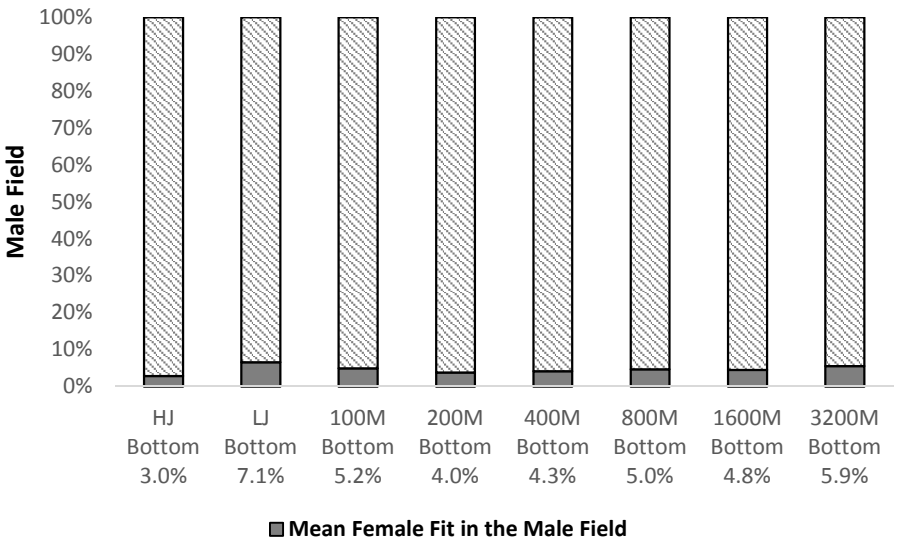


Note. The average male performance is better than 94.2%-97.9% of female performances. High Jump *M*(46, XY = 1.62), mean placement = 491 of 23,390. Long Jump *M*(46, XY = 5.20), mean placement 1,457 of 45,705. 100 meters *M*(46, XY = 12.78), mean placement = 3,046 of 79,663. 200 meters *M*(46, XY = 26.15), mean placement = 2,330 of 75,192. 400 meters *M*(46, XY = 59.31), mean placement = 1,515 of 52,050. 800 meters *M*(46, XY = 143.67), mean placement = 2,527 of 56,670. 1600 meters *M*(46, XY = 320.35), mean placement = 2,061 of 43,914. 3200 meters *M*(46, XY = 691.73), mean placement = 1,412 of 24,345.

As shown in Figure 4.11, the mean of the female performances fit within the bottom 3.0%-7.1% of the 46, XY distribution.

Figure 4.11

Mean Female Performance Fit Among the Male Field

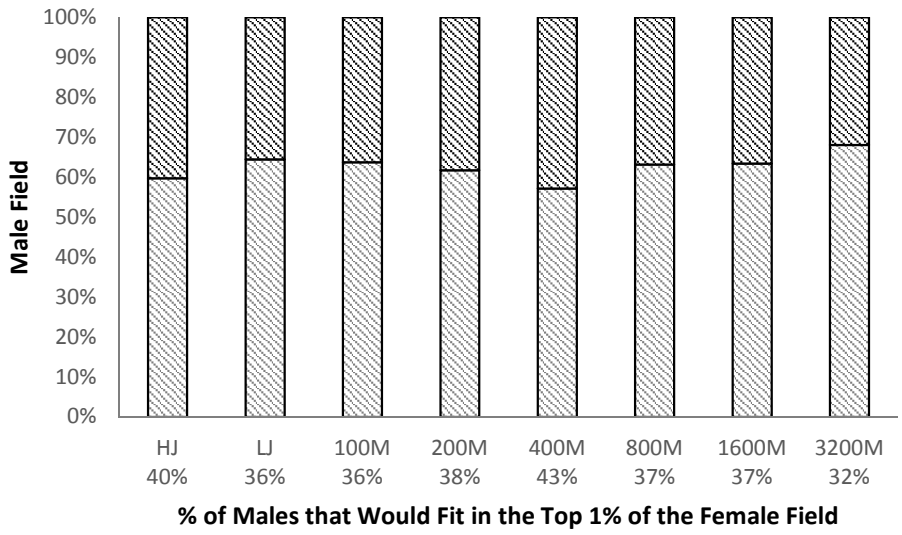


Note. The average female performance is worse than 92.9%-97% of male performances. High Jump *M*(46, XX = 1.35), mean placement = bottom 799 of 26,843. Long Jump *M*(46, XX = 4.08), mean placement = bottom 3,854 of 54,506. 100 meters *M*(46, XX = 14.76), mean placement = bottom 4,875 of 94,447. 200 meters *M*(46, XX = 30.75), mean placement = bottom 3,539 of 88,045. 400 meters *M*(46, XX = 70.77), mean placement = bottom 3,001 of 69,517. 800 meters *M*(46, XX = 172.03), mean placement = bottom 3,795 of 76,599. 1600 meters *M*(46, XX = 382.99), mean placement = bottom 3,268 of 68,787. 3200 meters *M*(46, XX = 818.81), mean placement = bottom 2,374 of 40,442.

As shown in Figure 4.12, the number of 46, XY performances that would fit in the top 1% of the female field are between 32-43%.

Figure 4.12

Percentage of Male Performances That Fit in the Top 1% of the Female Field



Note. 32%-43% of male performances fit within the top 1% of female performances. High Jump 46, XX top 1% mark = 1.65, top 46, XX 1% mark among 46, XY field = 10,811 of 26,843. Long Jump 46, XX top 1% mark = 5.49, top 46, XX 1% mark among 46, XY field = 19,379 of 54,506. 100 meters 46, XX top 1% mark = 12.30, top 46, XX 1% mark among 46, XY field = 34,266 of 94,447. 200 meters 46, XX top 1% mark = 25.25, top 46, XX 1% mark among 46, XY field = 33,720 of 88,045. 400 meters 46, XX top 1% mark = 57.53, top 46, XX 1% mark among 46, XY field = 29,776 of 69,517. 800 meters 46, XX top 1% mark = 136.91, top 46, XX 1% mark among 46, XY field = 28,227 of 76,599. 1600 meters 46, XX top 1% mark = 304.73, top 46, XX 1% mark among 46, XY field = 25,191 of 68,787. 3200 meters 46, XX top 1% mark = 651.52, top 46, XX 1% mark among 46, XY field = 12,917 of 40,442.

Research Question 2

Is there a statistically significant relationship between event distance and the percentage of potential female champions (PFCs)?

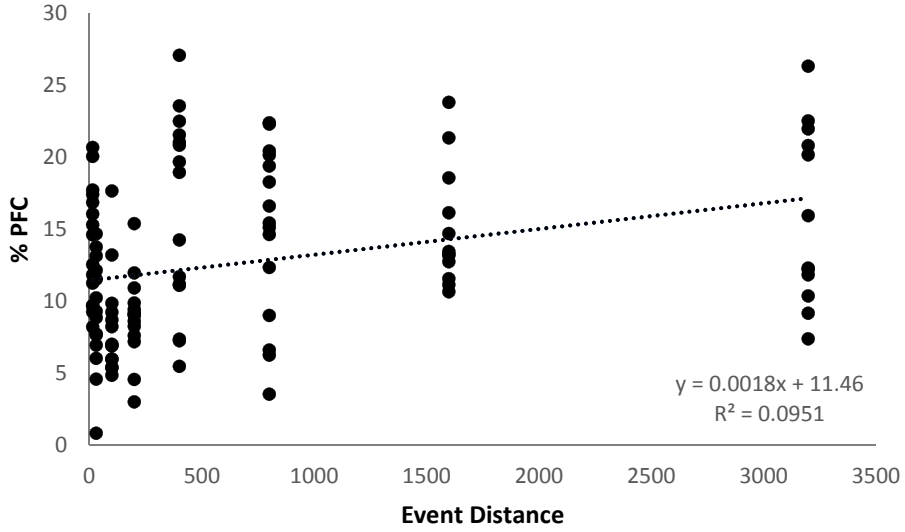
- $H2_0$: There is not a statistically significant relationship between the percentage of PFCs and the event distance.
- $H2_a$: There is a statistically significant relationship between the percentage of PFCs and the event distance.

The researcher converted the high jump and long jump from nominal levels of measurement to ratio levels of measurement in order to unify the distance measurements. To accomplish this analysis, the high jump, for presentation and analysis purposes, was estimated to be a 15-meter event and the long jump to be a 30-meter event.

The Pearson product-moment correlation was used to determine the strength and direction of a linear relationship between the two continuous variables. Regression was used to assess the relationship between distance and percentage of PFCs in an event. The Pearson correlation resulted in $r = .31$ suggesting a moderately-sized positive correlation between distance and percent PFC (see Figure 4.13). The results of the regression suggested that distance explained 10% of the variance, $R^2 = .10$, $F(1, 109) = 11.46$, $p < .001$.

Figure 4.13

Scatterplot of Percentage of 46, XY PFCs Per Event Distance

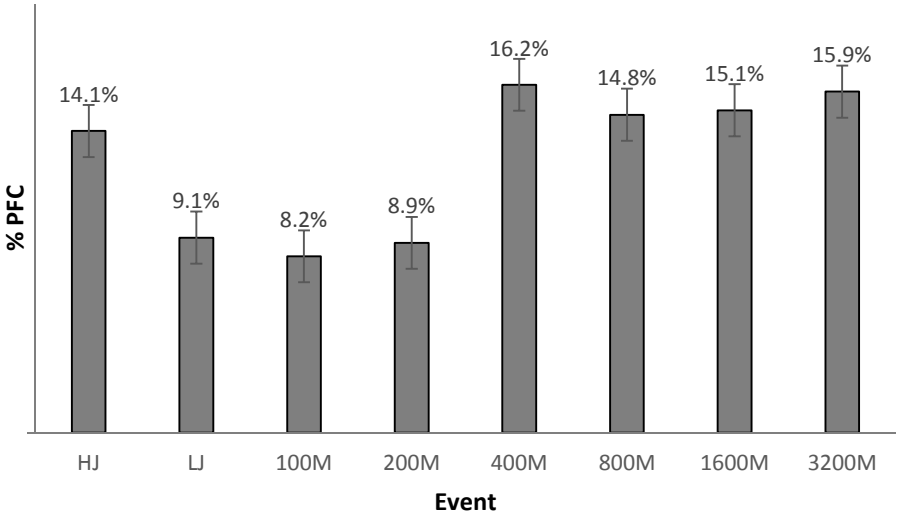


Note. PFC = potential female champion. Results from 3 years, 5 states, and 8 events (n = 111). States = CA, FL, MN, NY, WA. 2017, 2018, 2019 outdoor seasons. One year removed in Long Jump, 100M, and 200M for incomplete dataset in Minnesota for 2017. Three years removed in 1600M and 3200M for lack of 46, XX participation in the events in New York. High Jump coded as a 15-meter distance and long jump coded as a 30-meter distance. $r = .31$, $p < .001$.

As shown in Figure 4.14, the average percent PFC totals varied from 8.23%-16.24%, with the smallest occurring in the sprint events and the largest in the 400M.

Figure 4.14

Percentage of Male Performances Better than the Best Female



Note. HJ = high jump, LJ = long jump. PFC = potential female champion.

Research Question 3

What is the probability of one or more 46, XY potential female champions also being an MTF transgender individual? P(n[PFC and MTF] $\geq$ 1).

The researcher conducted 111 simulations in which each simulation consisted of 10,000 trials of random number generation based on the assumed rate of transgender athletes in each event, each year. The assumption is that being MTF is an independent, uniformly distributed random variable. Each trial used the transgender population estimates from the Williams Institute at UCLA to suggest how many of the 46, XY participants are theoretically transgender. The estimated percentage of transgender persons age 13 to 17 in the given states are: California, 0.85%; Florida, 0.78%; Minnesota, 0.85%; New York, 0.79%; Washington 0.70% (Herman et al., 2017). The results of the simulations are displayed in Table 4.11.

Table 4.11

Monte Carlo Simulation Results Inferring the Probability of a 46,XY Female Champion

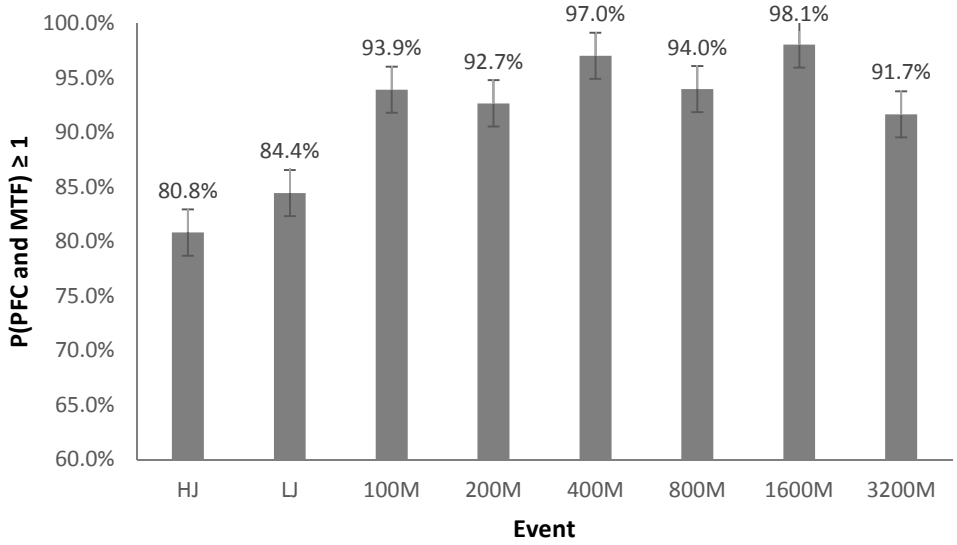
	HJ	LJ	100M	200M	400M	800M	1600M	3200M	Σ / x̄
Simulations	15	14	14	14	15	15	12	12	111
P(*n*[PFC and MTF] ≥ 1)	0.808	0.844	0.939	0.927	0.970	0.940	0.981	0.917	0.916
Mean *n* (PFC and MTF) ≥ 1	1.78	2.38	3.92	4.44	5.06	5.57	6.24	3.48	4.11
Ave. Mode *n* (PFC and MTF) ≥ 1	1.40	2.00	3.57	4.00	4.53	5.33	5.92	3.17	3.74
Ave. *SD*	1.21	1.41	1.82	1.88	2.00	2.03	2.17	1.63	1.77
Ave. Range	7.40	9.43	12.29	12.71	13.87	14.53	14.92	10.67	11.98

Note. Total simulations = 111, trials *n* = 1,110,000. One year removed in Long Jump, 100M, and 200M for an incomplete dataset in Minnesota for 2017. Three years removed in 1600M and 3200M for lack of female participation in the events in New York. The estimated percentage of transgender persons age 13 to 17 in the given states are: California, 0.85%; Florida, 0.78%; Minnesota, 0.85%; New York, 0.79%; Washington 0.70% (Herman et al., 2017).

If the assumptions are met, the results indicate a high probability that the female champion will be a 46,XY MTF athlete. Following 111 simulations of 1,110,000 trials, the P(*n*[PFC and MTF] ≥ 1) is between 81%-98% in the selected American high school track and field events (see Figure 4.15).

Figure 4.15

Probability of One or More 46,XY Potential Female Champions Being MTF Transgender

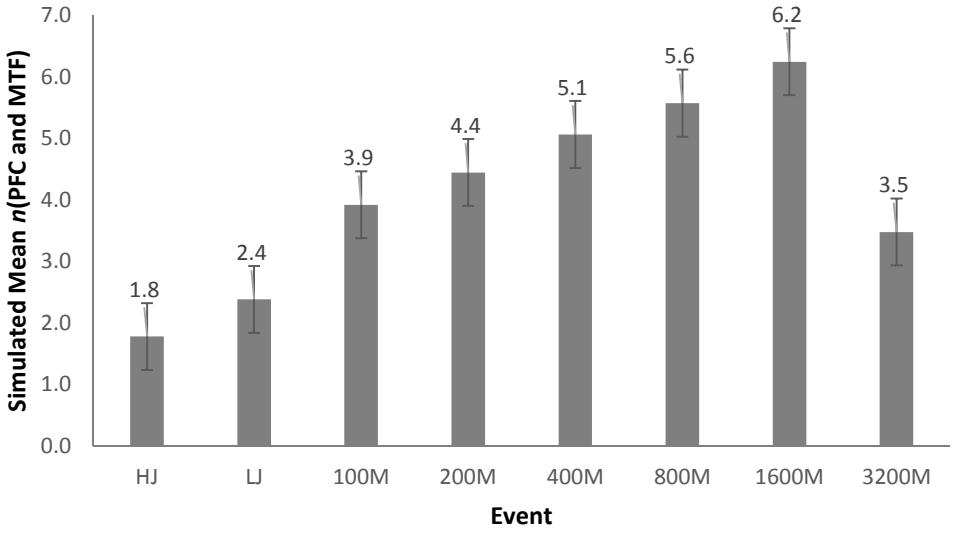


Note. Post hoc regression analysis of the results result in a non-significant relationship between P(*n*[PFC and MTF] ≥ 1) and distance ($p = .44$). $R^2 = .10$, $F(1, 6) = .68$, $r = .32$.

In the simulation when P(*n*[PFC and MTF] ≥ 1), the average size of the (PFC and MTF) group in each event varied on average between 2 and 6 individuals (1.8-6.2). The size of the *n*(PFC and MTF) ≥ 1 simulated in each is intended to show the depth of the (PFC and MTF) group in each event (see Figure 4.16).

Figure 4.16

Average Size of Simulated MTF Transgender Athletes atop the Female Standings



Note. Total simulations = 111, trials n = 1,110,000. Mean Size of n(PFC and MTF) $\geq$ 1 when P(n[PFC and MTF] $\geq$ 1). Post hoc regression analysis of the results reveal a non-significant relationship between mean size of n(PFC and MTF) $\geq$ 1 when P(n[PFC and MTF] $\geq$ 1*)* and distance (p = .56). R^2 = .06, $F(1, 6)$ = .38, r = .24.

Question 3 Results Summary

The Monte Carlo simulations n = 111, P(n[PFC and MTF] $\geq$ 1) is between 81%-98% in the selected American high school track and field events selected, assumptions being met. When a simulated trial resulted in P(n[PFC and MTF] $\geq$ 1) the average n(PFC and MTF) $\geq$ 1 was between 2 and 6 MTF athletes (1.8-6.2). The range for n(PFC and MTF) $\geq$ 1 was between 7.4-14.9 MTF individuals. The average mode for n(PFC and MTF) $\geq$ 1 was between 1.4 and 5.9 MTF individuals.

CHAPTER V

CONCLUSION, DISCUSSION, AND RECOMMENDATIONS

Statement of the Problem

Biological differences in 46, XY and 46, XX persons affect the structure and function of the human body (Carlson, 2018). This function and structure translate to differences in sports performance (Handelsman, 2017). However, despite the differences, many state high school governing bodies favor gender identity as the determining factor for separation into boys' and girls' categories (Tamerler, 2020). In light of the growing transgender movement, further research is needed regarding biological performance considerations for sports competition. The main issue at hand is whether transgender inclusion in female athletics constitutes an incursion on fair competition.

Is it reasonable to allow high school participants to self-select if they are eligible to participate in the protected girls' classification? This question is important and culturally relevant. To inform an evidence-based answer, it is reasonable and prudent to produce research on the nature of sex differences in high school athletes, as well as the possibility for disruption of the girls' classification.

Conclusions

As shown in Table 5.1, research questions 1 and 2 resulted in rejecting the null, and research question 3 resulted in a high probability of a MTF transgender individual being the female champion in each event (n[PFC and MTF] $\geq$ 1).

Table 5.1

Summary Research Question Results

Research Questions	Results
Question 1: Is there a statistically significant relationship in the performances of 46, XX (female) and 46, XY (male) high school track and field athletes, in selected events?	Reject the null in each event ($p < .001$)
Question 2: Is there a statistically significant relationship between event distance and percentage of potential female champions (PFCs)?	Reject the null (moderately positive) $r = .31$
Question 3: What is the probability of one or more 46, XY potential female champions also being an MTF transgender individual? P(n[PFC and MTF] ≥ 1).	P = 81%-98% event dependent

Note. $N = 920{,}115$. 46, XX $n = 400{,}929$. 46, XY $n = 519{,}186$. PFC = An 46, XY performance that is better than the best 46, XX per state, per year.

Research Question 1

Is there a statistically significant relationship in the performances of 46, XX (female) and 46, XY (male) high school track and field athletes, in selected events?

- $H1_0$: There is not a statistically significant relationship in the performance of the 46, XY and 46, XX high school track and field athletes.
- $H1_a$: There is a statistically significant positive relationship between performance and being 46, XY.

In each of the eight events the null is rejected in favor of the alternative hypothesis: There is a statistically significant positive relationship between performance and being 46, XY.

The mean difference between sex and distance in the high jump was statistically significant ($p < .001$), with an advantage in favor of males' performances being 18.2%. The mean difference between sex and distance in the long jump was statistically significant ($p < .001$), with an advantage in favor of 46, XY performances being 24.1%. The mean difference between sex and time in the 100M was statistically significant ($p < .001$), with an advantage in favor of male performances being 14.4%. The mean difference between sex and time in the 200M was statistically significant ($p < .001$), with an advantage in favor of male performances being 16.2%. The mean difference between sex and time in the 400M was statistically significant ($p < .001$), with an advantage in favor of male performances being 17.6%. The mean difference between sex and time in the 800M was statistically significant ($p < .001$), with an advantage in favor of male performances being 18%. The mean difference between sex and time in the 1600M was statistically significant ($p < .001$), with an advantage in favor of male performances being 17.8%. The mean difference between sex and time in the 3200M was statistically significant ($p < .001$), with an advantage in favor of male performances being 16.8%. Correlation between mean percentage difference and distance were non-significant ($p < .652$, $r = -.19$).

The mean male performance is better than 94%-98% of female performances (top 2%-6%). The mean female performance is worse than 93%-97% of male performances (bottom 3%-7%). Finally, approximately one-third or more (32%-43%) of male performances fit within the top 1% of female performances.

Research Question 2

Is there a statistically significant relationship between event distance and the percentage of potential female champions (PFCs)?

- $H2_0$: There is not a statistically significant relationship between the percentage of PFCs and the event distance.
- $H2_a$: There is a statistically significant relationship between the percentage of PFCs and the event distance.

The null is rejected in favor of the alternative hypothesis: There is a statistically significant relationship between the percentage of PFCs and the event distance ($p < .001$), as $r = .31$ indicates a moderate positive correlation between the two variables.

Research Question 3

What is the probability of one or more 46, XY potential female champions also being an MTF transgender individual? P(n[PFC and MTF] ≥ 1).

If the assumptions are met, the results indicate a high probability that the female champion will be a 46, XY MTF athlete. The P(n[PFC and MTF] ≥ 1) is between 81%-98% in the selected American high school track and field events. High jump = 80.8%. Long jump = 84.4%. 100M = 93.9%. 200M = 92.7%. 400M = 97%. 800M = 94%. 1600M =98.1%. 3200M = 91.7%.

Discussion

The first purpose of this study was to investigate the underlying basis for post-pubertal sex segregation in sport. If there were not categories based on sex, the top half of performances would be overwhelmingly male, and the top hundreds, and even thousands of positions in some events, would be exclusively males. A female would never get remotely close to winning at the state level without sex segregation.

The data provides sufficient and strong evidence to support post-pubertal sex segregation. Evidence in support of this point is the statistically significant difference in favor of males on the average of 17.9% across the events investigated. The gap between the sexes is large and persistent across all events (14.4%-24.1%). The average male performance fits at the very top of the female field (top 2.1%-5.8%), and the average female performance fits at the very bottom of the male field (bottom 3.0%-7.1%). A massive number of males (32%-43%) are able to perform within the top 1% of female performances. Put another way, a female that was better than 99% of girls would only be better than 57%-68% of boys. The best of the best females under a non-segregated format would be slightly better than average, and the average female would place at the very bottom of the field. Such an outcome would likely be discouraging and suppress sports participation by females.

In participation, boys showed higher participation in every event in every state. Overall the participation was 44% female and 56% male, but the participation gap varied from 14% to 50%. The exceptionally strong correlation ($r = .93$, $p < .001$) between participation percentage and distance is extremely interesting to the researcher because of its contrarian nature. Girls are participating at rates that more closely mirrors the male population in the shortest events. As the distance increases, so does the participation rate discrepancy. Girls have higher participation rates in comparison to boys in events that are more dependent on power and speed, and less participation in events that rely on endurance.

In terms of performance, boys have a slightly higher standard deviation in the jumps, but a smaller standard deviation in the runs. Boys have a smaller variance in every

event except for high jump. Boys have a greater kurtosis in all events. This distribution suggests that the boys' field is more centrally located. The girls have a greater spread and thus greater diversity in performance in comparison to the boys. If competitiveness is described as times that are closer to each other, boys are more competitive. If diversity is described as having a greater variety of performance, girls are more diverse.

The second purpose of this study was to assess the effect of event distance on performance differences between the sexes. The findings suggest a moderately positive relationship when comparing percent PFC and distance ($r = .31$, $p < .001$), but post hoc analysis of performance alone suggests that there is not a statistically significant relationship between distance and mean performance difference. Due to the mixed findings, it seems imprudent to make policies that are event dependent. For illustration, the greatest mean difference was seen in the long jump (24.1%), the 400M had the highest number of PFCs (16.2%), and the 1600M had the highest probability of a simulated PFC being a MTF transgender athlete (98.1%). It is not reasonable to have certain rules for certain events and not for others, as seen at the international (World Athletics) level.

It was interesting that the findings confirm a counterintuitive smaller performance gap in the 100M despite the large male advantage in maximal power, ability to generate a greater stride rate, and naturally possessing a greater stride length on average. These findings agree with those by Millard-Stafford et al. (2018) in their investigation of world record performances in that there seems to be a smaller sprint advantage and a larger endurance advantage for males.

These findings have consequential implications for team and contact sports. If the gaps between female and male performance are as large and significant as indicated by this studies data in non-contact, lower body dominant events, surely they are more significant and impactful in contact and upper body contribution sports, since the lower bodies of males versus females are more similar than the upper bodies of males versus females. The performance difference between the sexes in other sports is likely considerably higher than in this study.

The third purpose of the study was to assess the probability of a girls' champion being biologically male (46, XY). It is probable that the girls' champion in each state and in each event would be male. If the assumptions about transgender persons were true, there is an overwhelming probability (81%-98%) that a MTF transgender athlete would be at the top of the female ranking list in each state, each year. In the simulation trials where there was at least 1 transgender PFC, there was an average of 2-6 MTF individuals. Thus, in the majority of cases, the entire podium (top of the state) would be MTF.

Presently, transgender dominance of female sport is not occurring on anything resembling a large scale. However, the goal of many, including transgender activists and allies, is to have full acceptance and full integration of transgender persons at every level, in every institution. The data suggest, that full acceptance and full integration, without restrictions, would lead to transgender dominance of the female category of high school track and field and would likely apply to other sports as well.

It is apparent that currently transgender participation in athletics does not mirror the approximated 0.7%-0.85% rate that the Williams Institute projects. Nevertheless, if

transgender athlete participation were at that rate, and if being transgender did not affect performance, the probability found identified in the Monte Carlo Simulation could become reality.

Why is there presently not widespread transgender dominance? Perhaps the estimates of transgender youth are wildly overstated. Perhaps the distribution of MTF and FTM persons not equal; the Williams report does not distinguish between the two. Perhaps MTF persons are less interested in sports, or maybe they are interested, but some other barrier, whether physical, logistical, psychological, or sociological exists to suppress participation as suggested by Jones et al. (2017) and Herrick and Duncan (2018). Perhaps being MTF somehow inhibits performance independent of any medical intervention? Answers to these questions are unknown.

Recommendations for Further Study

There are multiple areas of further study and questions that should be addressed related to this study. Future research should investigate participation rates: Why do girls participate less in track and field than boys in general, and some events in particular? Why are girls participating less in endurance events in comparison to boys than in power and speed events? A multi-variate analysis of participation percentage differences, performance, and event distance would be valuable. Are there certain factors that drive participation or are effective at expanding participation rates?

Future studies should investigate performance differences further: Why is the girls' field more diverse and the boys more competitive? Are poor performing males quitting, whereas poor performing girls are sticking it out? Are the girls more mentally

resilient when facing the prospect of being at the bottom? Does the motivation to perform well drive participation rates?

Further research should explore the relationship of performance and event type: What is the physiological basis for the smaller performance difference in power and speed-dependent sprints and greatest in the endurance-dependent distance races? Why does a much bigger difference in Type II muscle fibers, power, and maximal strength physiology, and a narrower gap in Type I fibers, and endurance capabilities, produce the inverse results in terms of performance and event distance?

Future research needs to explore direct evidence applied to team and contact sports. Because this study only dealt with non-contact, individualistic, lower body dominant events, and the vast majority of high school sports participation is in contact, team and upper body inclusive sports, a detailed assessment of the difference in additional areas of key performance indicators is needed. Such analysis would include each aspect of human performance to include; strength, power, speed, agility, anaerobic endurance, and aerobic endurance. Some interesting sex difference events to study would be; bench press for upper body maximal strength, back squat for lower body maximal strength, shot put for upper and lower body power, 55M for speed, three-cone and pro-shuttle for agility, and half marathon for aerobic endurance.

Further study of the transgender population is needed. The ongoing efforts to document and assess the transgender population in American teenagers and those efforts should continue. Further studies could illuminate if the Williams numbers are accurate. Some other questions are: Should they differentiate between MTF and FTM percentages?

Why are transgender sports participation unremarkable and unrepresentative of the estimates, even in affirming and supportive communities?

Research should continue on how medical interventions influence performance. To what exact extent does cross sex hormone therapy (CSHT) level the playing field? Beyond the physiology, are requirements to have medical intervention, even ethical or legal? Do treatment first policies incentivize life-altering treatment that may otherwise not be desired, but for the ability to play sports? Investigating if MTF individuals forfeit a 46, XY athletic advantage would best be done by pursuing direct evidence. Such an analysis could involve conducting a pre-treatment physiological performance assessment on a substantial sample of trained MTF athletes; the participants should be followed in a longitudinal assessment tracking treatment while maintaining a supervised strength and conditioning program with follow-up post-transition performance tests.

Competitive sport is an interwoven, important, cultural force, and female sport is an invaluable asset and societal good. Sound research is needed to provide the data necessary for policymakers to make informed, evidence-based decisions that protect and promote competitive female sport.

REFERENCES

Albl, M. (2019, June 24). CIAC's transgender policy faces test with new lawsuit. Runner space. https://www.runnerspace.com/gprofile.php?mgroup_id=44531&do=news&news_id=580238

Almeida, M., Laurent, M. R., Dubois, V., Claessens, F., O'Brien, C. A., Bouillon, R., Vanderschueren, D. & Manolagas, S. C. (2017). Estrogens and androgens in skeletal physiology and pathophysiology. *Physiological Reviews*, *97*(1), 135-187.

American Bar Association. (2018, February 6). Title IX enforcement: "Rules" on gender identity, college-campus sexual assault shift under Trump. https://www.americanbar.org/news/abanews/aba-news-archives/2018/02/title_ix_enforcement/

American Psychological Association (2015). Guidelines for psychological practice with transgender and gender nonconforming people. *American Psychologist*, 70(9), 832-864. doi.org/10.1037/a0039906

American Psychological Association. (2020). *Publication Manual of the American Psychological Association (7th Ed.).*

Anderson, D. J., & Cheslock, J. J. (2004). Institutional strategies to achieve gender equity in intercollegiate athletics: Does Title IX harm male athletes? *American Economic Review*, 307-311.

Anderson, D. J., Cheslock, J. J., & Ehrenberg, R. G. (2006). Gender equity in intercollegiate athletics: Determinants of Title IX compliance. *The Journal of Higher Education, 77*(2), 225-250.

Anderson, R. T. (2018). A brave new world of transgender policy. *Harvard Journal of Law & Public Policy*, *41*(1), 309-353.

Anderson, T. (2020, February 18). It's time for a new tipping point for transgender folks in Hollywood. Out.com. https://www.out.com/print/2020/2/18/its-time-new-tipping-point-transgender-folks-hollywood

Andraya Yearwood: Track and field bio (n.d.) Athletic.net. https://www.athletic.net/TrackAndField/Athlete.aspx?AID=14519891

Barnes, K. (2019). What does the journey of transgender wrestler Mack Beggs teach us? *ESPN.* https://www.espn.com/espnw/voices/story/_/id/27652214/what-does-journey-transgender-wrestler-mack-beggs-teach-us

Barnes, K. (2020, June 23). The battle over Title IX and who gets to be a woman in sports: Inside the raging national debate. *ESPN.* https://www.espn.com/espnw/story/_/id/29347507/the-battle-title-ix-gets-woman-sports-raging-national-debate

Baumer, B. S. (2009). Using simulation to estimate the impact of baserunning ability in baseball. *Journal of Quantitative Analysis in Sports*, *5*(2), 1-16.

Bermon, S., Garnier, P. Y., Hirschberg, A. L., Robinson, N., Giraud, S., Nicoli, R., Baume, N., Saugy, M., Fenichel, P., Bruce, S. J., Henry, H., Dolle, G., & Ritzen, M. (2014). Serum androgen levels in elite female athletes. *The Journal of Clinical Endocrinology & Metabolism*, *99*(11), 4328-4335.

Bermon, S., & Garnier, P.Y. (2017). Serum androgen levels and their relation to performance in track and field: Mass spectrometry results from 2127 observations

in male and female elite athletes. *British Journal of Sports Medicine*, (17), 1309-1316.

Bird, A. (2010). Discovering the essences of natural kinds. In *The Semantics and Metaphysics of Natural Kinds* (pp. 133-144). Routledge.

Bonnette, V. M., & Daniel, L. (1990). *Title IX athletics investigator's manual.* US Department of Education, Office of Educational Research and Improvement, Educational Resources Information Center.

Bostock v. Clayton County, Georgia (2020). *U.S. Supreme Court.* https://www.supremecourt.gov/opinions/19pdf/17-1618_hfci.pdf

Brewer, M. S., & Rojas, M. (2008). Consumer attitudes toward issues in food safety. *Journal of Food Safety*, *28*(1), 1-22.

Brockes, E. (2017, May 8). Caitlyn Jenner on transitioning: 'It was hard giving old Bruce up. He still lives inside me.' *The Guardian.* https://www.theguardian.com/tv-and-radio/2017/may/08/caitlyn-jenner-bruce-transitioning-kardashians-reality-tv-star

Burdge, B. J. (2007). Bending gender, ending gender: Theoretical foundations for social work practice with the transgender community. *Social work*, *52*(3), 243-250.

Buzuvis, E. (2016). Hormone check: Critique of Olympic rules on sex and gender. *Wisconsin Journal of Law, Gender & Society*, *31*(1), 29-56.

Carlson, B. M. (2018). *Human Embryology and Developmental biology (6th ed.)*. Elsevier Health Sciences.

CeCe Telfer: Track and field bio (n.d.). Athletic.net. https://www.athletic.net/TrackAndField/Athlete.aspx?AID=14308523

Chen, V. (2018). Ethical Issues Concerning Transgender Athletes. *Penn Bioethics Journal*, *14*(1). 15-18.

Clark, R. V., Wald, J. A., Swerdloff, R. S., Wang, C., Wu, F. C., Bowers, L. D., & Matsumoto, A. M. (2019). Large divergence in testosterone concentrations between men and women: Frame of reference for elite athletes in sex$_2$ specific competition in sports, a narrative review. *Clinical Endocrinology*, *90*(1), 15-22.

Coleman, D. L. (2017). Sex in sport. *Law & Contemporary Problems, 80*(1), 63-126. https://scholarship.law.duke.edu/cgi/viewcontent.cgi?article=4849&context=lcp

Coleman, D. L. (2019, April 2). *Oral Testimony re: H.R. 5.* U.S. House of Representatives, Committee on the Judiciary.

Coleman, D. L., Joyner, M. J., & Lopiano, D. (2020). Re-affirming the value of the sports exception to Title IX's general non-discrimination rule. *Duke Journal of Gender and Law & Policy, 27*(1), 69-134.

Connecticut High School Men's 55 Meter Dash Rankings: 2019 (n.d.). Athletic.net https://www.athletic.net/TrackAndField/Division/Event.aspx?DivID=103101&Event=41&page=2

Costa, D. M., & Guthrie, S. R. (1994). *Women and sport: Interdisciplinary perspectives*. Human Kinetics.

Cozzillio, M.J., Levinstein, M.S., Dimino, Sr., M.R., & Feldman, G. (2007). *Sports law, cases and materials* (2nd ed.). Carolina Academic Press.

Craig Telfer: Track and field bio (n.d.). Athletic.net https://www.athletic.net/TrackAndField/Athlete.aspx?AID=9001456

Deaner, R. O., Lowen, A., Rogers, W., & Saksa, E. (2015). Does the sex difference in competitiveness decrease in selective sub-populations? A test with intercollegiate distance runners. *PeerJ*, *3*, e884. https://peerj.com/articles/884/

Dore, E., Martin, R., Ratel, S., Duché, P., Bedu, M., & Van Praagh, E. (2005). Gender differences in peak muscle performance during growth. *International Journal of Sports Medicine*, *26*(04), 274-280.

Dupré, J. (1981). Natural kinds and biological taxa. *The Philosophical Review*, *90*(1), 66-90.

Eaton-Robb, P. (2019, August 8). Civil rights probe opened into transgender athlete policy. *Associated Press News.* https://apnews.com/f553f2367f964462aab8cabdae38622e

Ennis, D. (2020, March 30). Idaho governor signs the nation's most anti-transgender measures into law. *Forbes.* https://www.forbes.com/sites/dawnstaceyennis/2020/03/30/idaho-governor-signs-the-nations-most-anti-transgender-measures-into-law/#4fee513c7a12

Fischer, B., & Mitteroecker, P. (2017). Allometry and sexual dimorphism in the human pelvis. *The Anatomical Record*, *300*(4), 698-705.

Flores, A. R., Herman, J. L., Gates, G. J., & Brown, N. T. (2016). *How many adults identify as transgender in the United States?* The Williams Institute: UCLA School of Law. https://williamsinstitute.law.ucla.edu/wp-content/uploads/Race-Ethnicity-Trans-Adults-US-Oct-2016.pdf

Foryst-Ludwig, A., & Kintscher, U. (2013). Sex differences in exercise-induced cardiac hypertrophy. *Pflügers Archiv-European Journal of Physiology*, *465*(5), 731-737.

Franke, W. W., & Berendonk, B. (1997). Hormonal doping and androgenization of athletes: A secret program of the German Democratic Republic government. *Clinical Chemistry*, *43*(7), 1262-1279.

Freeze, R. A. (1974). An analysis of baseball batting order by Monte Carlo simulation. *Operations Research*, *22*(4), 728-735.

Ganzi, D. M. (2004). After the commission: The government's inadequate responses to Title IX's negative effect on men's intercollegiate athletics. *Boston University Law Review 84*, 543-570.

Gates, G. J. (2011). *How many people are lesbian, gay, bisexual, and transgender?* The Williams Institute: UCLA School of Law. https://williamsinstitute.law.ucla.edu/wp-content/uploads/Gates-How-Many-People-LGBT-Apr-2011.pdf

Gooren, L. J., & Bunck, M. C. (2004). Transsexuals and competitive sports. *European Journal of Endocrinology*, *151*(4), 425-429.

Hacke, R. (2018). Girls will be boys, and boys will be girls: The emergence of the transgender athlete and a defensive game plan for high schools that want to keep their playing fields level-for athletes of both genders. *Sports Lawyers Journal*, *25*, 57-89.

Haff, G. G., & Triplett, N. T. (Eds.). (2015). *Essentials of strength training and conditioning (4th ed.)*. Human kinetics.

Haizlip, K. M., Harrison, B. C., & Leinwand, L. A. (2015). Sex-based differences in skeletal muscle kinetics and fiber-type composition. *Physiology*, *30*(1), 30-39.

Handelsman, D. J. (2017). Sex differences in athletic performance emerge coinciding with the onset of male puberty. *Clinical Endocrinology*, *87*(1), 68-72.

Handelsman, D. J., Hirschberg, A. L., & Bermon, S. (2018). Circulating testosterone as the hormonal basis of sex differences in athletic performance. *Endocrine Reviews*, *39*(5), 803-829.

Harper, J. (2015). Race times for transgender athletes. *Journal of Sporting Cultures and Identities*, *6*(1), 1-9.

Harper, J. (2019). *Sporting gender: The history, science, and stories of transgender and intersex athletes*. Rowman & Littlefield Publishers.

Hatlevig, E. (2005). Title IX compliance: Looking past the proportionality prong. *Sports Lawyers Journal.*, *12*, 87-122.

Hattenstone, S. (2019, April 20). The dad who gave birth: 'Being pregnant doesn't change me being a trans man.' *The Guardian.* https://www.theguardian.com/society/2019/apr/20/the-dad-who-gave-birth-pregnant-trans-freddy-mcconnell

HB 2706 (2020). State of Arizona: House of Representatives: Fifty -fourth Legislature

Hegge, A. M., Myhre, K., Welde, B., Holmberg, H. C., & Sandbakk, Ø. (2015). Are gender differences in upper-body power generated by elite cross-country skiers augmented by increasing the intensity of exercise? *PloS One*, *10*(5), e0127509. https://doi.org/10.1371/journal.pone.0127509.

Herman, J. L., Flores, A. R., Brown, T. N., Wilson, B. D., & Conron, K. J. (2017) *Age of individuals who identify as transgender in the United States.* The Williams

Institute: UCLA School of Law. http://williamsinstitute.seemysite.us/wp-content/uploads/TransAgeReport.pdf

Herrick, S. S., & Duncan, L. R. (2018). A systematic scoping review of engagement in physical activity among LGBTQ+ adults. *Journal of Physical Activity and Health,* 15(3), 226-232.

High school - men's 55-meter dash rankings: 2019 (n.d.) Athletic.net. https://www.athletic.net/TrackAndField/Division/Event.aspx?DivID=102510&Event=41&page=38

Hsu, P. L., & Robbins, H. (1947). Complete convergence and the law of large numbers. *Proceedings of the National Academy of Sciences of the United States of America, 33*(2), 25-31.

Ingle, S. (2020). Trans women face potential women's rugby ban over safety concerns. *The Guardian.* https://www.theguardian.com/sport/2020/jul/19/transwomen-face-potential-womens-rugby-ban-over-safety-concerns

International Association of Athletics Federations. (2019). *Eligibility regulations for the female classification: Athletes with differences of sex development.* https://www.sportsintegrityinitiative.com/wp-content/uploads/2019/05/IAAF-Eligibility-Regulations-for-the-Female-Classi-2-compressed.pdf

International Olympic Committee. (2017, December 21). *The Olympic games of antiquity: Factsheet.* https://stillmed.olympic.org/media/Document%20Library/OlympicOrg/Factsheets-Reference-Documents/Games/OG-of-Antiquity/Factsheet-The-Olympic-Games-of-Antiquity.pdf#_ga=2.91003091.813709799.1592431932-264464841.1592431932

James, S. E., Herman, J. L., Rankin, S., Keisling, M., Mottet, L., & Anafi, M. (2016). *The Report of the 2015 U.S. Transgender Survey*. National Center for Transgender Equality. http://www.ustranssurvey.org/

Janssen, I., Heymsfield, S. B., Wang, Z., & Ross, R. (2000). Skeletal muscle mass and distribution in 468 men and women aged 18–88 yr. *Journal of Applied Physiology*. *89*(1), 81-88.

Johns, M. M., Lowry, R., Andrzejewski, J., Barrios, L. C., Demissie, Z., McManus, T., Rasberry, C. N., Robin, L., & Underwood, J. M. (2019). Transgender identity and experiences of violence victimization, substance use, suicide risk, and sexual risk behaviors among high school students—19 states and large urban school districts, 2017. *Morbidity and Mortality Weekly Report*, *68*(3), 67-71.

Jonathan Eastwood: Cross country bio (n.d.) Athletic.net. https://www.athletic.net/CrossCountry/Athlete.aspx?AID=8776921

Jones, B. A., Arcelus, J., Bouman, W. P., & Haycraft, E. (2017). Sport and transgender people: A systematic review of the literature relating to sport participation and competitive sport policies. *Sports Medicine, 47*(4), 701-716.

Jones, M., Jagim, A., Haff, G., Carr, P., Martin, J., & Oliver, J. (2016). Greater strength drives difference in power between sexes in the conventional deadlift exercise. *Sports*, *4*(3), 43-53.

Joyner, M. J., & Coyle, E. F. (2008). Endurance exercise performance: The physiology of champions. *The Journal of Physiology*, *586*(1), 35-44.

June Eastwood: Cross country bio (n.d.) Athletic.net https://www.athletic.net/CrossCountry/Athlete.aspx?AID=16118860&L=0

Keating, J. (1973). The ethics of competition and its relation to some moral problems in athletes. *Philosophic Exchange*, *4*(1), 5-20. https://digitalcommons.brockport.edu/phil_ex/vol4/iss1/15

Kennedy, C. L. (2010). A new frontier for women's sports (beyond Title IX). *Gender Issues, 27*, 78-90.

Kimmel, A. P. (2015). Title IX: An imperfect but vital tool to stop bullying of LGBT students. *Yale Law Journal*, *125*(7), 2006-2037.

Klaver, M., Dekker, M. J. H. J., de Mutsert, R., Twisk, J. W. R., & den Heijer, M. (2017). Cross$_2$ sex hormone therapy in transgender persons affects total body weight, body fat and lean body mass: A meta$_2$ analysis. *Andrologia*, *49*(5), e12660.

Knox, T., Anderson, L. C., & Heather, A. (2019). Transwomen in elite sport: Scientific and ethical considerations. *Journal of Medical Ethics*, *45*(6), 395-403.

Kotschwar, B. (2014). Women, sports, and development: Does it pay to let girls play? *Peterson Institute for International Economics*. No. PB14-8. 1-12. https://www.piie.com/publications/pb/pb14-8.pdf

Kousta, E., Papathanasiou, A., & Skordis, N. (2010). Sex determination and disorders of sex development according to the revised nomenclature and classification in 46, XX individuals. *Hormones*, *9*(3), 218-231.

Langley, N., & Dudzik, B. (2016, September 16). Sex estimation. *Oxford Bibliographies.* https://www.oxfordbibliographies.com/view/document/obo-9780199766567/obo-9780199766567-0152.xml

Leahy, C. D. (1997). The title bout: A critical review of the regulation and enforcement of Title IX in intercollegiate athletics. *Journal of College and University Law, 24*(3), 489-544.

Lenzi, M. J. (2017). The trans athlete dilemma: A constitutional analysis of high school transgender student-athlete policies. *American University Law Review*, *67*(3), 841-890.

Levine, S. C., Foley, A., Lourenco, S., Ehrlich, S., & Ratliff, K. (2016). Sex differences in spatial cognition: Advancing the conversation. *Wiley Interdisciplinary Reviews: Cognitive Science*, *7*(2), 127-155.

Litsky, F. (1976, July 31). Jenner triumphs in decathlon, breaks world mark. *The New York Times.* https://www.nytimes.com/1976/07/31/archives/jenner-triumphs-in-decathlon-breaks-world-mark-viren-wins-5000-for.html

Lund, T. (1981). An introduction to the Monte Carlo method. *CERN Libraries, 67*, 1-67. https://cds.cern.ch/record/131030/files/CERN-HS-RP-067.pdf

Lundgren, K. M., Karlsen, T., Sandbakk, Ø., James, P. E., & Tjønna, A. E. (2015). Sport-specific physiological adaptations in highly trained endurance athletes. *Medicine and Science in Sports and Exercise*, *47*(10), 2150-2157.

MacArthur, D. G., & North, K. N. (2005). Genes and human elite athletic performance. *Human Genetics*, *116*(5), 331-339.

McAffee, T. B. (2001). Inalienable rights, legal enforceability, and American constitutions: the fourteenth amendment and the concept of Unenumerated rights. *Wake Forest Law Review*, *36*(3), 747-794.

Meier, S. C., & Labuski, C. M. (2013). The demographics of the transgender population. In *International handbook on the demography of sexuality* (pp. 289-327). Springer.

Merriam-Webster. (n.d.). Gender. Merriam-Webster.com. https://www.merriam-webster.com/dictionary/gender

Millard-Stafford, M., Swanson, A. E., & Wittbrodt, M. T. (2018). Nature versus nurture: Have performance gaps between men and women reached an asymptote? *International Journal of Sports Physiology and Performance*, *13*(4), 530-535.

Miller, R. W. & Yasharoff, H. (2020, June 9) What's a TERF and why is 'Harry Potter' author J.K. Rowling being called one? *USA Today.* https://www.usatoday.com/story/news/nation/2020/06/09/what-terf-definition-trans-activists-includes-j-k-rowling/5326071002/

Monahan, J., Hoge, S. K., Lidz, C., Roth, L. H., Bennett, N., Gardner, W., & Mulvey, E. (1995). Coercion and commitment: Understanding involuntary mental hospital admission. *International Journal of Law and Psychiatry*, *18*(3), 249-263.

Moreira, C. A., & Bilezikian, J. P. (2017). Stress fractures: Concepts and therapeutics. *The Journal of Clinical Endocrinology & Metabolism*, *102*(2), 525-534.

Morton, V. (2019, October 30). Transgender NCAA runner named conference's 'Women's Athlete of the Week.' *The Washington Times.* https://www.washingtontimes.com/news/2019/oct/30/june-eastwood-montana-transgender-runner-named-big/

Murphy, A. (2014, September 17). Exclusive: Fallon Fox's latest opponent opens up to #WHOATV. #woahtv. https://whoatv.com/exclusive-fallon-foxs-latest-opponent-opens-up-to-whoatv/

Murphy, W. G. (2014). The sex difference in hemoglobin levels in adults—mechanisms, causes, and consequences. *Blood Reviews*, *28*(2), 41-47.

National Collegiate Athletic Association. (2011*). NCAA inclusion of transgender student-athletes.* Office of Inclusion. https://www.ncaa.org/sites/default/files/Transgender_Handbook_2011_Final.pdf

National Collegiate Athletic Association. (2013). *NCAA Sports sponsorship and participation rates report 1981-82 – 2012-13.* https://www.ncaapublications.com/p-4334-1981-82-2012-13-ncaa-sports-sponsorship-and-participation-rates-report.aspx

National Federation of State High School Associations. (2019). *High school athletics participation survey.* https://www.nfhs.org/sports-resource-content/high-school-participation-survey-archive/

National Federation of State High School Associations. (2019, September 5). *Participation in high school sports registers first decline in 30 years.* https://www.nfhs.org/articles/participation-in-high-school-sports-registers-first-decline-in-30-years/

Navratilova, M. (2019, February 17). The rules on trans athletes reward cheats and punish the innocent. *The Sunday Times*. https://www.thetimes.co.uk/edition/comment/the-rules-on-trans-athletes-reward-cheats-and-punish-the-innocent-klsrq6h3x

NCAA Division I Outdoor Qualifying: 100 Meters Men. (n.d.). TFRRS Database. https://www.tfrrs.org/lists/2568/2019_NCAA_Division_I_Outdoor_Qualifying

Newton, P. K., & Aslam, K. (2009). Monte Carlo tennis. *SIAM Review*, *48*(4). 722-742.

Office for Civil Rights. (1972). *Title IX of the Education Amendments: Section 901(a).* U.S. Department of Education.

Office for Civil Rights. (1975). *CFR Part 106: Nondiscrimination on the basis of sex in education programs or activities receiving financial assistance.* U.S. Department of Education.

Office for Civil Rights. (1979). *A Policy Interpretation: Title IX and intercollegiate athletics*. U.S. Department of Education.

Office for Civil Rights. (1996). *Clarification of intercollegiate athletics policy guidance: The three-part test.* U.S. Department of Education.

Office for Civil Rights. (2003). *Further clarification of intercollegiate athletics policy guidance regarding the Title IX compliance.* U.S. Department of Education.

Ogles, B. M., & Masters, K. S. (2003). A typology of marathon runners based on cluster analysis of motivations. *Journal of Sport Behavior*, *26*(1), 69-85.

Orchard, S. (2020, July 20). World Rugby could ban transgender women because of safety reasons. BBC Sport. https://www.bbc.com/sport/rugby-union/53476972

Pate, R. R., & Kriska, A. (1984). Physiological basis of the sex difference in cardiorespiratory endurance. *Sports Medicine*, *1*(2), 87-89.

Payne, M. (2017, March 22). Transgender woman wins international weightlifting title amid controversy over fairness. *The Washington Post.* https://www.washingtonpost.com/news/early-lead/wp/2017/03/22/transgender-woman-wins-international-weightlifting-title-amid-controversy-over-fairness/

Perez-Gomez, J., Rodriguez, G. V., Ara, I., Olmedillas, H., Chavarren, J., González-Henriquez, J. J., Dorado, C., & Calbet, J. A. (2008). Role of muscle mass on sprint performance: Gender differences? *European Journal of Applied Physiology*, *102*(6), 685-694. https://link.springer.com/article/10.1007/s00421-007-0648-8

Prinster, R. (n.d.). Words Matter: Affirming gender identity through language. Insight into Diversity. http://www.insightintodiversity.com/words-matter-affirming-gender-identity-through-language/

Raff, K. (2020, March 12). A clash across America over transgender rights. *The New York Times.* https://www.nytimes.com/2020/03/12/us/transgender-youth-legislation.html

Randnofsky, L. (2020, March 9). The race to replace the binary of men's and women's sports. *The Wall Street Journal.* https://www.wsj.com/articles/the-race-to-replace-the-binary-of-mens-and-womens-sports-11583769636

Rankinen, T., Roth, S. M., Bray, M. S., Loos, R., Pérusse, L., Wolfarth, B., & Bouchard, C. (2010). Advances in exercise, fitness, and performance genomics. *Medicine and Science in Sports and Exercise*, *42*(5), 835-46.

Round, J. M., Jones, D. A., Honour, J. W., & Nevill, A. M. (1999). Hormonal factors in the development of differences in strength between boys and girls during adolescence: A longitudinal study. *Annals of Human Biology*, *26*(1), 49-62.

Rowling, J. K. (2020, June 10) J.K. Rowling writes about her reasons for speaking out on sex and gender issues. JKRowling.com. https://www.jkrowling.com/opinions/j-k-rowling-writes-about-her-reasons-for-speaking-out-on-sex-and-gender-issues/

Sandbakk, Ø., Hegge, A. M., Losnegard, T., Skattebo, Ø., Tønnessen, E., & Holmberg, H. C. (2016). The physiological capacity of the world's highest ranked female cross-country skiers. *Medicine and Science in Sports and Exercise*, *48*(6), 1091-1100.

Sandbakk, Ø., Solli, G. S., & Holmberg, H. C. (2018). Sex differences in world-record performance: The influence of sport discipline and competition duration. *International Journal of Sports Physiology and Performance*, *13*(1), 2-8.

Schmidt, S. (2020, February 7) Conservatives find unlikely ally in fighting transgender rights: Radical feminists. *The Washington Post.* https://www.washingtonpost.com/dc-md-va/2020/02/07/radical-feminists-conservatives-transgender-rights/

Senefeld, J.W., Clayburn, A.J., Baker, S. E., Carter, R. E., Johnson, P. W., Joyner, M. J. (2019) Sex differences in youth elite swimming. *PLoS ONE 14*(11): e0225724. https://doi.org/10.1371/journal.pone.0225724

Silva, C. F., Garcia, E. S., & Saliby, E. (2002). Soccer championship analysis using Monte Carlo simulation. *Proceedings of the Winter Simulation Conference, IEEE* (2), 2011-2016.

Simpson, J. L., Ljungqvist, A., Ferguson-Smith, M. A., de la Chapelle, A., Elsas II, L. J., Ehrhardt, A. A., Genel, M., Ferris, E. A., & Carlson, A. (2000). Gender verification in the Olympics. *JAMA*, *284*(12), 1568-1569.

Speechly, D., Taylor, S., & Rogers, G. (1996). Differences in ultra-endurance exercise in performance-matched male and female runners. *Medicine & Science in Sports & Exercise*, *28*(3), 359-365.

Staron, R. S., Hagerman, F. C., Hikida, R. S., Murray, T. F., Hostler, D. P., Crill, M. T., Ragg, K. E., & Toma, K. (2000). Fiber type composition of the vastus lateralis muscle of young men and women. *Journal of Histochemistry & Cytochemistry*, *48*(5), 623-629.

Tamerler, J. (2020). Transgender athletes and Title IX: An uncertain future. *Jeffrey S. Moorad Sports Law Journal*, *27*, 139-179. https://digitalcommons.law.villanova.edu/mslj/vol27/iss1/5

Taylor, J. K., Haider-Markel, D. P., & Lewis, D. C. (2018). *The remarkable rise of transgender rights.* University of Michigan Press.

Temfemo, A., Hugues, J., Chardon, K., Mandengue, S. H., & Ahmaidi, S. (2009). Relationship between vertical jumping performance and anthropometric characteristics during growth in boys and girls. *European Journal of Pediatrics*, *168*(4), 457-464.

Terry Miller: Bloomfield HS Track and field bio (n.d.) Athletic.net. https://www.athletic.net/TrackAndField/Athlete.aspx?AID=14046370

Terry Miller: Bulkeley HS Track and field bio (n.d.) Athletic.net. https://www.athletic.net/TrackAndField/Athlete.aspx?AID=10875325

The Associated Press. (2017, February 23). A look back at the history of US transgender rights. MyNBC5. https://www.mynbc5.com/article/a-look-back-at-the-history-of-us-transgender-rights/8972867#

The Equality Act. House Resolution 5 (2019). U.S. House of Representatives: 116th Congress.

The women of Sparta: Athletic, educated, and outspoken radicals of the Greek world. (2012, January 18). *Ancient History Encyclopedia*. https://www.ancient.eu/article/123/

Thomas, G. A., Kraemer, W. J., Spiering, B. A., & Volek, J. S. (2007). Maximal power at different percentages of one repetition maximum: Influence of resistance and gender. *Journal of Strength and Conditioning Research*, *21*(2), 336-342.

Title VII of the Civil Rights Act (1964). *Volume 42 of the United States Code: section 2000e.*

Townsend, E. A., Miller, V. M., & Prakash, Y. S. (2012). Sex differences and sex steroids in lung health and disease. *Endocrine Reviews*, *33*(1), 1-47.

Tracy, M. (2017, April 4). NCAA ends boycott of North Carolina after so-called bathroom bill is repealed. *The New York Times.* https://www.nytimes.com/2017/04/04/sports/ncaa-hb2-north-carolina-boycott-bathroom-bill.html

Transgender Law Center. (n.d.). State-by-state overview: Changing gender markers on birth certificates. https://transgenderlawcenter.org/resources/id/state-by-state-overview-changing-gender-markers-on-birth-certificates

Trubee, N. W., Vanderburgh, P. M., Diestelkamp, W. S., & Jackson, K. J. (2014). Effects of heat stress and sex on pacing in marathon runners. *The Journal of Strength & Conditioning Research*, *28*(6), 1673-1678.

Tucker, R., & Collins, M. (2010). The science of sex verification and athletic performance. *International Journal of Sports Physiology and Performance*, *5*(2), 127-139.

Tucker, R., & Collins, M. (2012). What makes champions? A review of the relative contribution of genes and training to sporting success. *British Journal of Sports Medicine*, *46*(8), 555-561.

U.S. Department of Education. (2002). *Secretary's Commission for Opportunity in Athletics.* http://www.ed.gov/about/bdscomm/ list/athletics/transcripts.html.

U.S. Department of Justice. (2020, March 24). *The Department of Justice files statement of interest in Title IX womens' equal opportunities case.* https://www.justice.gov/opa/pr/department-justice-files-statement-interest-title-ix-womens-equal-opportunities-case

Vamplew, W. (2007). Playing with the rules: Influences on the development of regulation in sport. *The International Journal of the History of Sport*, *24*(7), 843-871.

Velho, I., Fighera, T. M., Ziegelmann, P. K., & Spritzer, P. M. (2017). Effects of testosterone therapy on BMI, blood pressure, and laboratory profile of transgender men: A systematic review. *Andrology*, *5*(5), 881-888.

Walton, D. (1999). The appeal to ignorance, or argumentum ad ignorantiam. *Argumentation*, *13*, 367-377.

Welle, S., Tawil, R., & Thornton, C. A. (2008). Sex-related differences in gene expression in human skeletal muscle. *PloS ONE*, *3*(1), 1-7.

Wiik, A., Lundberg, T. R., Rullman, E., Andersson, D. P., Holmberg, M., Mandić, M., Brismar, T. B., Leinhard, O. D., Chanpen, S., Flanagan, J. N., Arver, S., & Gustafsson, T. (2020). Muscle strength, size, and composition following 12 months of gender-affirming treatment in transgender individuals. *The Journal of Clinical Endocrinology & Metabolism*, *105*(3), e805-e813.

Wilson, B. D. M., & Kastanis, A. (2015). Sexual and gender minority disproportionality and disparities in child welfare: A population-based study. *Children and Youth Services Review, 58,* 11–17.

Wisniewski, A. B., & Mazur, T. (2009). 46, XY DSD with female or ambiguous external genitalia at birth due to androgen insensitivity syndrome, 5-reductase-2 deficiency, or 17-hydroxysteroid dehydrogenase deficiency: A review of quality of life outcomes. *International Journal of Pediatric Endocrinology*, *2009*(1), e567430.

Woman's Liberation Front. (n.d.) *We believe.* womensliberationfront.org. https://womensliberationfront.org/wolf-key-documents/

Word of the year 2019. (n.d.) *Merriam-Webster*. https://www.merriam-webster.com/words-at-play/word-of-the-year/they

World Anti-Doping Agency. (2015). *World anti-doping code.* https://www.wada-ama.org/en/resources/the-code/world-anti-doping-code

World Athletics. (2019, October 1). *World Athletics eligibility regulations for transgender athletes.* http://www.athletics.org.tw/Upload/Web_Page/WA/Eligibility%20Regulations%20for%20Transgender%20Athletes,%20.pdf

Ziegler, E. M., & Huntley, T. I. (2013). It got too tough to not be me: Accommodating transgender athletes in sport. *Journal of College and University Law*, *39*(2), 467-509.

APPENDICES

Appendix A: IRB Approval

UNITED STATES SPORTS ACADEMY
"America's Sports University®"
One Academy Drive Daphne, Alabama 36526
PH: (251) 626-3303 FX: (251) 626-8829
Email: academy@ussa.edu Website: http://www.ussa.edu/

MEMORANDUM **28 October 2020**

From: Dr. Roch A. King, USSA IRB Chair

To: Gabriel Higerd
Dr. Brandon Spradley

Subject: FINAL REVIEW APPROVAL OF HUMAN SUBJECT RESEARCH

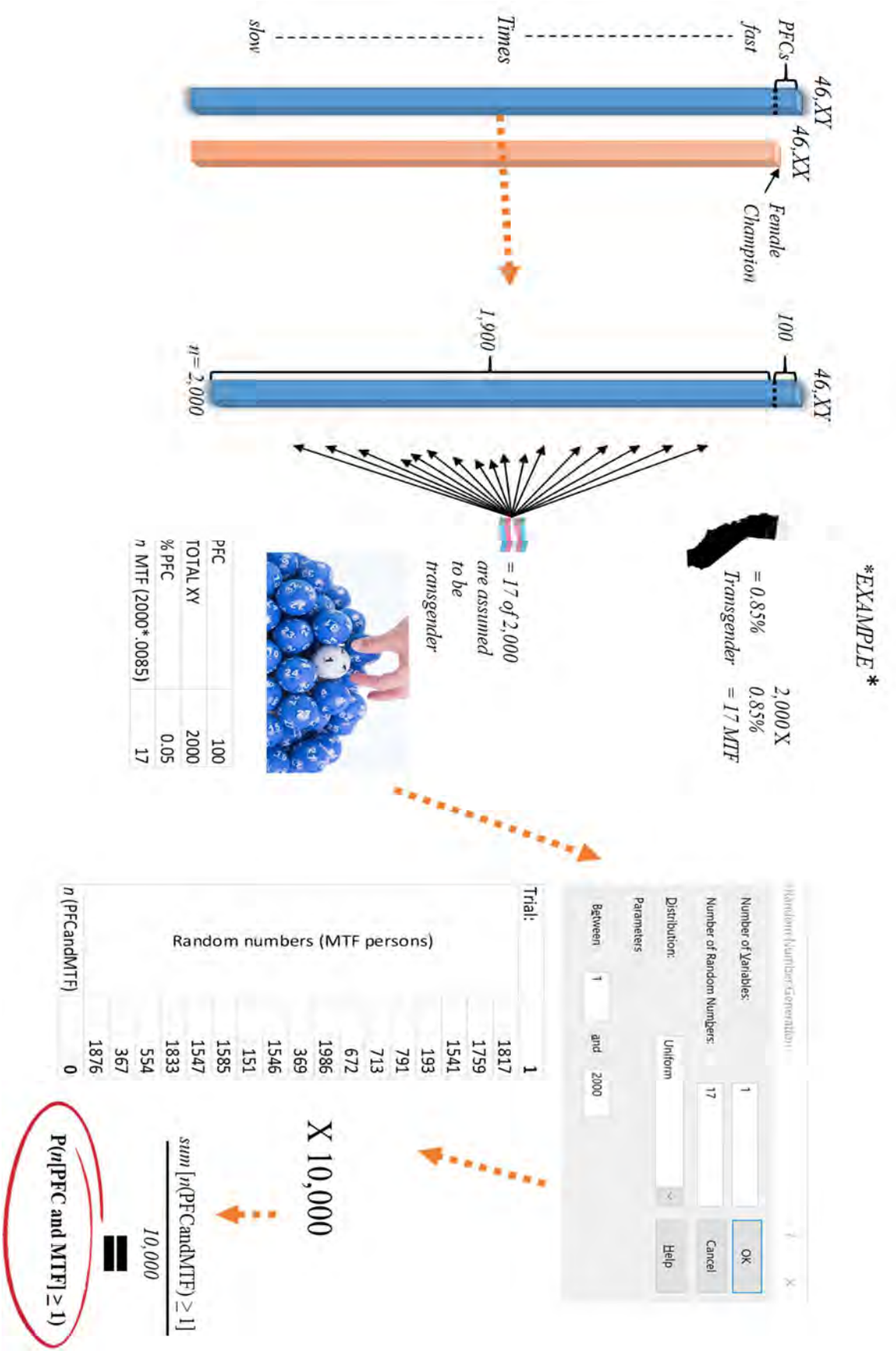
Ref: (a) Appendix C
(b) Appendix D
(c) Appendix E

Encl: None

1. **IRB Identification # USSA.2020.017-IR-EX-C**
2. Per references (a) through (c), the research protocol titled **Assessing the potential transgender impact on girl champions in American high school track & field** has been submitted by Hugh Morrissey from the Ed.D. in Sport Management program. The graduate advisor is Dr. Brandon Spradley
3. The purpose of this study is to **investigate underlying basis for postpubertal sex segregation in sport; assess the effect of event distance on the performance differences between the sexes; and assess the probability of a girl champions being biologically male.**
4. The research is considered **Exempt** as all data is from pre-existing, publicly available sources. No human subjects are involved.
5. This research is **Approved.**
6. Required Modifications:
   a. **None**
7. Date of Approval is 28 October 2020. The Duration of Approval is to be one calendar year (28 October 2021). If research is to be continued beyond 29 October 2021 please submit your renewal application to this office by 28 September 2021 to allow time for adequate processing.
8. You are required to report when the research has concluded and to provide this office with copies of any articles or presentations resulting from this research. Additionally, any presentations or publications must include acknowledgment of IRB approval.

Roch A. King
Roch A. King, Ph.D.
Chair, USSA IRB

Appendix B: Monte Carlo Concept Diagram



Journal of Orthopaedic Surgery 2001, 9(1): 35–38

Sports related fractures: A review of 113 cases

Wan Hazmy Che Hon and Shong Hing Kock
Department of Orthopaedic & Traumatology, Seremban Hospital, Malaysia

ABSTRACT

To establish a profile of fractures related to sporting activities, to determine the extent to which these injuries resulted in morbidity, and to find suitable target areas for injury prevention, a one-year study was undertaken at Seremban Hospital, Negeri Sembilan, Malaysia, from July 15, 1998 to July 14, 1999. All patients presenting to the Department of Orthopaedic & Traumatology with fractures sustained during sporting activities were enrolled. Fractures related to sporting activities are increasing and an entity to be recognized despite the good function outcome of the treatment. Suitable target areas for injury prevention were found to be football (for males), netball (for females), the second decade of life and competitive sporting events.

Key words: *fracture, sports, football*

INTRODUCTION

Injuries sustained during sporting activities are common, the majority being soft-tissue injuries involving ligaments, tendons or muscles. However, with increasing competitiveness in sports even at the recreational level, there are increasing numbers of sports-related fractures noted. We have conducted this study in order to establish a profile of fractures related to sporting activities, to determine the extent to which these injuries resulted in morbidity and to identify suitable target areas for injury prevention.

MATERIALS AND METHODS

This was a one-year prospective observational study done at the Seremban Hospital, a level 1 hospital in Negeri Sembilan, Malaysia. All patients presenting to the Department of Orthopaedic & Traumatology between July 15, 1998 and July 14, 1999 with fractures sustained during sporting activities were enrolled, in total 113. These included both competitive and non-competitive (recreational) sporting activities.

Following appropriate initial management of the fractures, the patients were given a questionnaire in order to identify the demographic characteristics of the patients and the types and levels of sporting activities. We have divided the sporting activities into two levels, competitive and non-competitive (recreational). The competitive sports are played either by professionals or by amateurs at schools, colleges or local clubs. The non-competitive sports are also divided into two groups: regular players and weekend players.

The injuries were classified into contact or non-contact injuries. Contact injuries were defined as injuries sustained following collision with other players or caused by sporting equipment, for example, being hit by a hockey stick or collision with the goalpost. Non-contact injuries were mainly due to falls. Further information, for example, of fracture

Address correspondence and reprint requests to: Dr Wan Hazmy Che Hon, Department of Orthopaedic & Traumatology, Seremban Hospital, 70300 Seremban, Negeri Sembilan, Malaysia. E-mail: whazmy@tm.net.my.

characteristics, whether closed or open fractures, and the clinical diagnosis, was taken, and associated injuries were also noted.

Management of the fractures does not differ from other trauma cases. Surgical treatment was indicated in cases of open fractures, articular disruption and failure of conservative treatment. Patients were followed-up until discharge from the Orthopaedic Sports Clinic. The healing process was monitored by subsequent radiograph and clinical assessment. Any complication was recorded. On the last consultation, the functional status of the patient was assessed based on clinical assessment and return to the pre-morbid sporting activities.

RESULTS

Of the 113 cases enrolled in the study, males sustained 92% of the fractures, and 62.5% of these occurred during football games. Netball caused 33.3% of the female fractures. The median age was 18 years (range 7 to 59) with 70% of the patients in the second decade of life. Malays constituted 62% of the study population followed by Indians (18.6%), Chinese (14.1%) and other ethnic groups (5.3%).

Football contributed 57.5% of the fractures followed by basketball (5.3%) and high jump (4.4%) (Table 1). Sixty-nine cases (61%) occurred during competitive sports and this mainly involving amateur players (98.6%). Only one professional player sustained a fracture during this study. Out of the 44 patients who sustained fractures during recreational sporting activities, 84% were regular players while another 16% were weekend players.

Table 1
Distribution of patients with sports-related fractures

Type of sports	Number of patients
Football	65
Basketball	6
High Jump	5
Rugby	4
Martial Arts	4
Takraw	3
Hockey	3
Athletics	3
Volleyball	3
Netball	3
Skateboarding	3
Motorcross	2
Cycling	2
Jet-skiing	2
Roller-skating	1
Tennis	1
Diving	1
Gymnastics	1
Fishing	1
Total	113

Fifty-seven fractures (50.4%) were due to contact injuries. Of these, 54.4% were caused by collision with other players, while 45.6% were injured by related sporting equipment. Non-contact injuries occurred in 56 of the cases (49.6%).

There were 96.5% closed fractures. Only 4 patients had an open fracture, 2 during football and one each during running and jet-skiing. Seventy-six fractures (67.3%) involved the upper limbs with isolated radius fractures constituting 34.2% of the cases followed by fractures of both radius and ulna (16.8%) and humeral fracture (18.4%).

Fractures of the lower limb occurred in 29.2%, mainly involving the tibia alone (42.4%) and both the tibia and fibula (24.2%) (Table 2). Three cases involved the spine while one had an acetabular fracture. Only four cases were associated with other injuries.

Table 2
Anatomical distribution of fracture sites

Upper limb	
Radius alone	26
Radius and ulna	19
Humerus	14
Ulna alone	6
Phalanges	7
Metacarpal	2
Clavicle	1
Scapula	1
Lower limb	
Tibia alone	14
Tibia and fibula	8
Fibula alone	4
Femur	2
Patella	1
Calcaneum	1
Talus	1
Metatarsal	1
Phalanges	1
Others	
Spine	3
Pelvic	1

Ninety fractures (79.6%) were treated conservatively while 23 patients needed operative management. There was straightforward healing in 85% of the fractures. Seventeen cases were complicated by malunion (17.6%), delayed union (17.6%), stiffness of the adjacent joint (52.9%) and chronic limb pain

(11.9%). The mean hospital stay was 3.6 days (range 2 to 14 days). The mean period of follow-up was 13.3 weeks (range 6 to 32 weeks).

On the last follow-up it was found that 90 (79.6%) of the patients were back to their pre-injury sporting level. Twenty- three patients had residual disabilities, which prevented them from performing to their pre-morbid activity level. Residual disabilities were seen in 43.5% of the patient treated surgically, compared to only 14.4% in patients treated conservatively. The main components of the residual disabilities were stiffness of the adjacent joint (52.1%) and recurrent pain (43.5%). One patient had loss of one phalanx of the finger.

DISCUSSION

The increasing awareness of the value of regular exercise has resulted in millions of individuals participating in sporting activities. The tremendous development of sports has led to an increasing number of sports-related injuries. Even though soft-tissue injuries constitute the major component of the injuries (80% according to Mc Latchie, 1982), some sports have a higher vulnerability for fractures. Some (Choyce MQ et al. 1998) found that fractures constituted 68% of the sports hand injuries admitted to the Accident and Emergency Departments. Others (Emshoff R et al. 1997) found that sports were the most common cause of mandibular fractures, accounting for 31.5% of the entire sample in their study, followed by road traffic accidents (27.2%) and falls (20.8%). The major causative factor in sports-related mandibular fractures was skiing (55.3%) whereas cycling and soccer accounted for 25.4% and 8.9% respectively. Our study highlighted the fact that fractures during sporting activities are not uncommon in Malaysia. As it is fully understood that football and netball are the most commonly played sports, the intensity of fracture occurrences (62.5% in football for males and 33.3% in netball for females) needs further attention. Furthermore, 70% of the cases occurred in the second decade of life, which is an important age group as far as further planning and development of sports in this country is concerned.

As world records continue to be broken and competition assumes ever-increasing intensity, athletes both at the competitive and recreational levels push themselves to the limits of their endurance and performance parameters. This situation may lead to many injuries in elite athletes. This was also obvious in our study where 61% of the fractures occurred during competitive sports, mainly involving amateur players. It should be noted that we do not directly manage professional players since most of them are referred either to private hospitals or to the National Sports Center. Contrary to common belief, weekend players constitute only 16% of fractures sustained during recreational sporting activities.

Detailed analysis and knowledge of the mechanisms of injury allow an understanding of the injury patterns and factors that lead to the risk of fractures. Among the common mechanism of injuries are overuse injuries, collision with other players, collision with objects, infringement of rules, dangerous techniques and vehicular accidents. In ice hockey, for instance, more than 80% of head and neck injuries are due to fighting and high-sticking. This source of injury is unacceptable and could be radically altered.

Some maneuvers used in sport may leave a specific body part vulnerable, such as the low tackle in football. This information enables coaching and medical staffs to modify training and competition situations in order to reduce the risk factors. In our study, we found almost equal risks of fractures due to contact and non-contact injuries. Even though the more disabling injuries caused by collisions and falls are often the most difficult to control by changes in rules or introduction of protective equipment, greater efforts must be taken to minimize the injuries.

The majority of our patients sustained closed fractures, which indicated the low velocity of the injury and explained the conservative management in 79.6% of the cases. The predisposition to radio-ulna and tibia-fibula fractures might need special consideration of the use of protective equipment. Even though there were only 4 cases of open fractures and 3 cases involving the spine, the possible morbidity following these injuries needs serious attention and preventive measurements.

The majority of our patients resumed their pre-injury activity level. The most common residual disabilities were joint stiffness and pain. Joint stiffness needs special consideration and signifies the importance of intensive physiotherapy and well-planned post-injury rehabilitation programs. On the other hand, reaction to pain is individualistic, and some athletes are stoic whereas others seem to feel the slightest change in their physiology. The former may hide their injury and their pain, and the latter may complain at the smallest setback. Functional assessment using both objective and functional tests plays an important role as a guide to safe re-integration of the athlete into training and competition.

CONCLUSION

Sports related fractures are increasing. Suitable target areas for injury prevention are football (for males), netball (for females), the second decade of life and competitive sporting events. Medical input into rule changes, pressure to enforce existing regulations and further research into the injury profiles of different sports in order to develop preventive strategies are valuable contributions to make sports safer.

REFERENCES

1. **Choyce MQ, Potts M, Maitra AK.** A profile of sports hand injuries in an accident and emergency department. *J. Accident Emergency Med* 1998. Jan 15; 1:35–38.
2. **Emshoff R, Schoning H., Rothler G, Waldhart E**. Trends in the incidence and cause of sport related mandibular fractures: A retrospective analysis. *J. Oral Maxillofac Surg* 1997 Jun 55:6, 585–92.
3. **Mc Latchie GR.** Risk factors in sport. Proceedings of Royal Medico-Chirurgical Society. *Scottish Medical Journal* 1982, 27:189.
4. **Reid DC.** Sports injury assessment & rehabilitation. *Sports Medicine and Therapy*. 1st Edition 1992. Churchill Livingstone Publication:1–12.

Journal of Athletic Training 2020;55(11):1174–1180
doi: 10.4085/1062-6050-0536.19

www.natajournals.org

Collision and Contact Sport Participation and Quality of Life Among Adolescent Athletes

David R. Howell, PhD, ATC*‡; Michael W. Kirkwood, PhD†§; Scott Laker, MD†§; Julie C. Wilson, MD*‡||

*Departments of Orthopedics, Sports Medicine Center, and Rehabilitation Medicine, Children's Hospital Colorado, Aurora; Departmens of ‡Orthopaedics, §Physical Medicine and Rehabilitation, and ||Pediatrics, University of Colorado School of Medicine, Aurora

Context: Researchers investigating collision and contact sport participation during high school have found mixed results. Understanding the association between current contact and collision sport participation and quality-of-life outcomes can enhance our knowledge about the risks and benefits of sport participation.

Objective: To examine quality-of-life outcomes among high school athletes who reported participation in collision and contact sports in the year preceding assessment compared with no- or limited-contact sport athletes.

Design: Cross-sectional study.

Setting: Preparticipation physical examination.

Patients or Other Participants: High school athletes 13 to 18 years of age.

Main Outcome Measure(s): We obtained sport participation and Patient-Reported Outcomes Measurement Information System (PROMIS) Pediatric-25 outcomes, which assess self-reported, quality-of-life domains in the preceding 7 days (ie, state assessment). Our grouping variable was collision and contact versus no- or limited-contact sport participation during the year preceding assessment. We used multivariable linear regression models to identify the associations between PROMIS scores and collision and contact sport participation and adjusted for sex; age; history of bone, muscle, ligament, or tendon injury; history of acute fracture or dislocation; and history of concussion.

Results: A total of 143 (51%) athletes reported collision and contact sport participation (24% female, mean age = 15.1 ± 1.7 years) and 138 (49%) reported no- or limited-contact sport participation (66% female, mean age = 15.4 ± 1.2 years). A higher proportion of collision and contact sport athletes reported a history of time loss for bone, muscle, ligament, and tendon injuries (51% versus 29%, $P < .001$) and for acute fracture or dislocation (46% versus 26%, $P < .001$) than did no- or limited-contact athletes. After adjusting for covariates, we found that collision and contact sport participation was significantly associated with lower state anxiety ($\beta = -1.072$, 95% confidence interval = −1.834, −0.310, $P = .006$) and depressive ($\beta = -0.807$, 95% confidence interval = −1.484, −0.130, $P = .020$) symptom scores.

Conclusions: Collision and contact sport athletes reported fewer anxiety and depressive symptoms in the week preceding evaluation than did no- or limited-contact sport athletes, but they had more extensive orthopaedic injury histories. Potential benefits and risks are associated with collision and contact sport participation. These data reinforce the need to examine the assumption that youth collision and contact sports are associated with negative quality of life.

Key Words: youth athletes, pediatric athletes, anxiety, depression, pain

Key Points

- Compared with high school athletes participating in no- or limited-contact sports, those participating in organized contact or collision sports displayed fewer anxiety and depressive symptoms.
- Among boys with no reported history of diagnosed concussion, those involved in contact or collision sports exhibited lower levels of state anxiety than no- or limited-contact sport participants.
- A higher proportion of athletes in contact and collision sports had a history of orthopaedic injury than athletes in no- or limited-contact sports, which may reflect the physical risks of contact and collision sports.
- Potential benefits and risks exist for athletes participating in high school collision sports.

Sport participation among adolescents in the United States has continually risen over the past several decades, with approximately 8 million students reporting current sport participation during high school.[1,2] Physical activity and sport participation are associated with multiple benefits for youth, such as improved physical health, reduced risk of obesity, and decreased rates of smoking and drug use.[3–5] These benefits may be related to the environment in which children participate in sports, and factors such as the type of sport played may influence the relative benefit gained from participation.[3] Sport participation also carries inherent risks for youth athletes. Due to rapid growth and maturation during adolescence, youth athletes may be vulnerable to a variety of traumatic and repetitive stress injuries.[5,6] The cumulative effects of multiple injuries sustained during adolescence may carry into adulthood, resulting in chronic pain, dysfunction, or repeat injury.[5]

Beyond general sport participation, playing collision or contact sports (or both) during high school has received

increased attention due to concerns over exposure to repetitive head impacts. Most researchers have focused on the association between collision sport participation during high school, such as football, and neurologic outcomes during adulthood. Given the importance of brain development during adolescence, examining the effect of sport participation during this time of life is important. In 1 report, researchers[7] observed that beginning football participation at a younger age (ie, <12 years) was associated with increased self-reported neurologic impairments among a sample of former professional football athletes, whereas another group[8] did not observe an association between age of football exposure and neurologic outcomes. Other studies[9–13] of broad samples of current and former athletes (nonprofessional) revealed no significant associations between participation in collision sports and negative neurologic outcomes. Specifically, among older adults, collision sport participation during college was not associated with neurobehavioral quality-of-life (QoL) outcomes.[9] Furthermore, no significant associations between participation in football before the age of 12 and neurocognitive function were found among current high school and collegiate athletes.[10,11] How head-impact exposures in collision and contact sports during adolescence affect QoL remains unclear, as little attention has been paid to the health quality characteristics of youth athletes during their time of active sport participation. In addition, investigations that include equal representation of female and male athletes, as well as different sports and competition levels, are needed.[14] Delineation of the relative benefits and risks of collision and contact sport participation compared with no- or limited-contact sport participation on QoL among a broad sample of athletes may help to inform relevant stakeholders regarding individual decisions on which sports to pursue during high school.

The primary purpose of our work was to examine whether current participation in collision and contact sports during high school was associated with short-term QoL ratings among a sample of healthy adolescent athletes undergoing a preparticipation examination relative to athletes who reported current participation in no- or limited-contact sports. We hypothesized that QoL domain scores would not be different between collision and contact and no- or limited-contact sport athletes. Secondarily, we sought to determine whether the exposure to repetitive head impacts (rather than diagnosed concussions) that are expected to occur during collision sports (ie, boys' football and lacrosse) was associated with short-term QoL outcomes relative to no- or limited-contact sport participation. We proposed that collision sport participation would not be significantly associated with QoL domain scores relative to no- or limited-contact sport participation among high school boys who reported no history of diagnosed concussion.

METHODS

Participants and Study Setting

We investigated high school athletes who were undergoing a preparticipation physical examination (PPE) during May 2018 and May 2019. Participants who completed the PPE were from a single school district and invited to volunteer for the study, which involved answering the study questionnaires during the PPE. If a participant returned for a PPE during 2019, only data from his or her initial visit were included in the analysis. Before the study, the local institutional review board and local school district approved the study protocol. All enrolled participants and their parent or legal guardian provided written informed consent at the time of enrollment. Our inclusion criteria consisted of being between the ages of 13 to 18 years and receiving full clearance to participate in sport at the time of the examination (in the case of a recent injury). Our exclusion criteria consisted of a preexisting neurologic or psychiatric disorder, a recent concussion diagnosis with ongoing symptoms, or not speaking English.

Assessment Protocol

All participants completed a set of standardized questionnaires during their PPE that documented their medical history, demographic characteristics, and QoL. The variables obtained for their history and characteristics included sex, age, average training time in sport per week, level of competition (varsity, junior varsity, C team, or D team), and the different organized sports in which they were involved during the year preceding evaluation. To assess lifetime injury history, participants filled out a questionnaire as a part of their PPE. The questions included in the current study were (1) "Have you ever had an injury to a bone, muscle, ligament, or tendon that caused you to miss a practice or a game?" (2) "Have you ever had any broken or fractured bones or dislocated joints?" and (3) "Have you ever had a head injury or concussion?" If necessary, parents were available to assist with completing the questionnaires. Height and weight were obtained and recorded by trained personnel.

Outcome Variables

To address our primary and secondary aims, we asked participants to complete the Patient-Reported Outcomes Measurement Information System (PROMIS) version 1.1 Pediatric-25.[15,16] This instrument assesses 6 domains: physical function mobility, anxiety, depressive symptoms, fatigue, peer relationships, and pain interference. A single question addresses pain intensity. Each question was answered based on the participant's experience during the past 7 days. Version 1.1 Pediatric-25 is intended for self-reporting by individuals 8 to 17 years of age. Six questions have 5 response options each (ranging from 0 to 4); the pain is rated on a scale from 0 (*no pain*) to 10 (*worst pain you can think of*). The total raw score within each domain is calculated as the sum of all responses (range = 0–24), and a higher score represents more of the concept being measured (ie, zero = *with no trouble* or *never*, 4 = *not able to do* or *almost always*). With the exception of pain interference, each of the 6 domains has demonstrated acceptable levels of reliability using the static pen- and paper- short-form assessment.[16]

Grouping Variable

To address our primary purpose, we divided study participants into 2 groups: collision and contact sport athletes or no- or limited-contact sport athletes. So that we could categorize the participants, we asked them to indicate

Table 1. Demographic and Injury History Characteristics of the Sport Groups[a]

Variable	Sport Group		*P* Value	Effect Size (95% Confidence Interval)
	Collision or Contact (n = 143)	No or Limited Contact (n = 138)		
Female athletes	34 (24%)	91 (66%)	<.001	
Male athletes	109 (76%)	47 (34%)		
Age, y[b]	15.2 (14.3, 16.0)	15.3 (14.5, 16.6)	.18	0.21 (−0.02, 0.45)
Age group	13–15 y: 66 (46%)	13–15 y: 59 (43%)	.63	
	16–18 y: 77 (54%)	16–18 y: 79 (57%)		
Height, cm	170.7 (14.6)	167.9 (8.2)	.05	0.24 (0.00, 0.47)
Weight, kg	64.6 (17.1)	63.2 (15.0)	.46	0.09 (−0.15, 0.32)
Average time training, h/wk	11.6 (4.9)	11.7 (5.0)	.85	0.02 (−0.23, 0.28)
Varsity athlete	52 (36%)	43 (31%)	.36	
History of bone, muscle, ligament, or tendon injury	73 (51%)	40 (29%)	<.001	
History of acute fracture or dislocation	65 (46%)	35 (26%)	<.001	
History of stress fracture	12 (9%)	7 (5%)	.26	
History of diagnosed concussion	35 (25%)	27 (20%)	.29	
Organized sport(s) played in the past year[c]	Basketball: 42	Baseball: 7		
	Cheerleading: 6	Cross-country: 20		
	Diving: 2	Dance: 5		
	Extreme sports: 1	Football (flag/touch): 2		
	Football: 74	Golf: 7		
	Gymnastics: 10	Horseback riding: 1		
	Lacrosse: 10	Marching band: 4		
	Martial arts: 10	Poms: 17		
	Rugby: 1	Racquetball: 1		
	Soccer: 29	Softball: 14		
	Water polo: 1	Swimming: 13		
	Wrestling: 12	Tennis: 10		
		Track & field: 28		
		Volleyball: 15		
Lifetime years playing collision and contact sports	7.8 ± 3.6	4.6 ± 3.2	<.001	0.93 (0.57, 1.30)

[a] Continuous variables are presented as mean ± SD; categorical variables are presented as the number included in each group (%).
[b] Data presented as median (interquartile range).
[c] Participants selected all organized sports they played in the past year, so the total n sums to >100%.

all organized sports in which they had been active over the past year. Consistent with the classification of sports according to contact provided by Rice et al,[17] the following were considered collision and contact sports: basketball, cheerleading, diving, extreme sports, football, gymnastics, lacrosse, martial arts, rugby, soccer, water polo, and wrestling. An athlete who reported playing multiple sports was placed in the collision and contact group if any of these sports was listed. Those who reported involvement in other sports were classified as no or limited contact (Table 1).

To address our secondary purpose, we then grouped participants who reported collision sport participation in the past year and compared them with the no- or limited-contact sport athletes. Consistent with the work of past researchers, *collision sports* were defined as those in which purposeful body-to-body collisions occurred as a part of typical gameplay, and they included football, rugby, boys' ice hockey, and boys' lacrosse.[9] No boys' ice hockey or rugby athletes were enrolled in our study, however, leaving only boys' football and lacrosse in this subanalysis. To determine how exposure to repetitive head impacts, rather than concussions, were associated with QoL, we also removed from this analysis participants who reported a history of a diagnosed concussion. Furthermore, as no female athletes in our study reported collision sport participation, only male athletes were included in this analysis.

Statistical Analysis

All statistical tests were 2 sided and evaluated with a significance level of $P < .05$. All analyses were conducted using Stata (version 15; StataCorp, College Station, TX). Continuous variables were presented as means (95% confidence intervals [CIs]) or medians (interquartile ranges), and categorical variables were presented as the number included and corresponding percentage. We compared differences between collision and contact and no- or limited-contact sport groups for demographic, training, and injury history variables using independent-samples *t* tests and χ^2 analyses.

Consistent with our first purpose, the primary grouping variable was sport type (collision and contact versus no- or limited-contact sport). To assess univariable differences between sport types, we compared PROMIS domain scores between groups using Mann-Whitney U tests. We then constructed a series of multivariable linear regression models. The outcome in each model was the PROMIS domain score, the predictor variable was sport type (collision and contact versus no- or limited-contact sport), and the covariates were factors that may have affected PROMIS outcomes (sex; age; history of time loss for bone, muscle, ligament, or tendon injury; history of bone fracture or dislocation; and history of concussion).

Consistent with the secondary purpose of our study and past findings,[9] we then constructed multivariable linear regression models for which our grouping variable was

Table 2. Univariable PROMIS Domain Comparisons Between Collision and Contact and No- or Limited-Contact Sport Groups

	Sport Group						
	Collision and Contact		No or Limited Contact		Mann-Whitney U		Effect Size (95%
PROMIS Domain	Median (IQR)	Mean ± SD	Median (IQR)	Mean ± SD	Z Test Statistic	*P* Value	Confidence Interval)
Mobility	0 (0, 0)	0.09 ± 0.36	0 (0, 0)	0.21 ± 0.56	2.49	.013	0.26 (0.02, 0.50)
Anxiety[a]	0 (0, 2)	1.55 ± 2.39	2 (0, 6)	3.26 ± 3.35	4.68	<.001	0.59 (0.35, 0.83)
Depressive symptoms[a]	0 (0, 1)	0.82 ± 1.76	0 (0, 2)	1.88 ± 3.14	2.79	.005	0.42 (0.18, 0.65)
Fatigue	1 (0, 3)	1.67 ± 2.38	1 (0, 3)	2.16 ± 2.72	1.71	.087	0.19 (−0.05, 0.43)
Peer relationships	15 (13, 16)	13.6 ± 4.07	15 (12, 16)	13.0 ± 4.18	−1.33	.183	0.14 (−0.10, 0.37)
Pain interference	0 (0, 2)	1.32 ± 2.44	0 (0, 2)	1.53 ± 2.90	−0.23	.81	0.08 (−0.16, 0.31)
Pain scale	0 (0, 2)	1.08 ± 1.52	0 (0, 1)	0.89 ± 1.67	−1.86	.06	0.12 (−0.13, 0.36)

Abbreviations: IQR, interquartile range; PROMIS, Patient-Reported Outcomes Measurement Information System.

[a] $P < .05$.

sport type (collision sport only versus no- or limited-contact sport) and our outcome variable was each PROMIS domain score among only male athletes who reported no history of concussion. The covariates in these models were age; history of time loss for bone, muscle, ligament, or tendon injury; and history of bone fracture or dislocation. Before performing regression analyses, we assessed collinearity within each model using variance inflation factors. If 2 variables were *collinear*, defined as a variance inflation factor of >2.5, only 1 variable was included in the model.

RESULTS

A total of 281 adolescents participated in the study, and 143 (51%) reported collision and contact sport participation in the year before the study. A greater proportion of females was in the no- or limited-contact sport group than in the collision and contact sport group (Table 1). In addition, a higher proportion of collision and contact sport athletes reported a history of time-loss bone, muscle, ligament, or tendon injury or a history of acute fracture or dislocation than did no- or limited-contact sport athletes. The groups did not differ in the proportion of athletes who had sustained a diagnosed concussion.

Univariable comparison revealed that those in the collision and contact sport group reported lower mobility, anxiety, and depressive symptom domain scores than the no- or limited-contact sport group (Table 2). After covariate adjustment, collision and contact sport participation was still significantly associated with lower anxiety and depressive symptom domain scores (Table 3). Some covariates were also significantly associated with PROMIS domain outcomes. Female sex was significantly associated with higher anxiety ($\beta = -1.77$, 95% CI = –2.52; –1.01, $P < .001$) and depressive ($\beta = -0.99$, 95% CI = –1.67, –0.07; $P = .004$) symptom scores, whereas a history of bone, muscle, ligament, or tendon injury was significantly associated with higher pain scale scores ($\beta = 0.69$, 95% CI = 0.25, 1.13; $P = .002$).

A total of 55 male athletes reported collision sport participation in the year preceding evaluation and no history of concussion (n = 45 football, n = 4 lacrosse, and n = 6 football and lacrosse). Thirty-eight male athletes reported no- or limited-contact sport participation in the year preceding evaluation and no history of concussion (Table 4). After covariate adjustment, collision sport participation was significantly associated with lower anxiety scores and higher pain scale scores (Table 5). No covariates were significantly associated with PROMIS outcomes. A greater proportion of those in the collision sport group also reported a history of time-loss bone, muscle, ligament, or tendon injury than among the no- or limited-contact sport athletes (46% versus 14%, $P = .001$).

DISCUSSION

Our cross-sectional investigation of uninjured high school athletes indicated that current collision and contact sport athletes reported lower levels of current anxiety and depressive symptoms than high school athletes who were currently participating in no-contact or limited-contact sports. Additionally, we found that collision and contact sport athletes had a more extensive orthopaedic injury history than no- or limited-contact sport athletes. The PROMIS outcomes, however, reflected attitudes in the week preceding assessment and, thus, represent a situational or state assessment rather than an enduring or trait assessment, so these findings should be interpreted

Table 3. Effect of Collision and Contact Sport Participation on Patient-Reported Quality of Life (PROMIS) Outcome Measures[a]

PROMIS Domain	β Coefficient	Standard Error	95% Confidence Interval	*P* Value
Mobility	−0.106	0.065	−0.235, 0.023	.11
Anxiety[b]	−1.031	0.388	−1.795, −0.268	.008
Depressive symptoms[b]	−0.753	0.346	−1.433, −0.072	.030
Fatigue	−0.190	0.347	−0.872, 0.492	.58
Peer relationships	0.430	0.574	−0.700, 1.560	.46
Pain interference	−0.120	0.371	−0.845, 0.610	.75
Pain scale	0.157	0.221	−0.278, 0.591	.48

Abbreviation: PROMIS, Patient-Reported Outcomes Measurement Information System.

[a] Covariates in each model were sex; age; past history of bone, muscle, ligament, or tendon injury; past history of bone fracture or dislocation; and concussion history.

[b] $P < .05$.

Table 4. Demographic and Injury History Characteristics Between Contact or Collision Sport Athletes and No- or Limited-Contact Sport Athletes with No History of Concussion[a]

Variable	Sport Group		*P* Value	Effect Size (95% Confidence Interval)
	Contact or Collision (n = 55)	No or Limited Contact (n = 38)		
Male athletes	55 (100)	38 (100)		
Age, y[b]	14.9 (14.2, 16.0)	14.9 (14.5, 16.1)	.88	0.13 (−0.28, 0.55)
Age group	13–15 y: 28 (51) 16–18 y: 27 (49)	13–15 y: 20 (53) 16–18 y: 18 (47)	.87	
Height, cm	174.2 ± 18.0	174.2 ± 7.4	.99	0.00 (−0.41, 0.42)
Weight, kg	71.1 ± 17.5	65.8 ± 19.1	.17	0.29 (−0.13, 0.71)
Average time training, h/wk	13.2 ± 5.5	10.9 ± 4.4	.06	0.46 (−0.03, 0.95)
Varsity athlete	19 (35)	8 (21)	.16	
History of bone, muscle, ligament, or tendon injury	25 (45)	5 (13)	.001	
History of acute fracture or dislocation	24 (44)	9 (24)	.05	
History of stress fracture	5 (9)	2 (5)	.48	
Organized sports played in the past year	Football: 45 (82) Football and lacrosse: 6 (11) Lacrosse: 4 (7)	Baseball: 2 (5) Cross-country: 8 (21) Golf: 5 (13) Marching band: 2 (5) Swimming: 4 (11) Tennis: 7 (18) Track & field: 10 (26)		
Lifetime years playing collision/contact sports	7.3 ± 3.0	4.1 ± 3.0	.001	1.09 (0.43, 1.74)

[a] Continuous variables are presented as mean ± SD; categorical variables are presented as the number included in each group (%).
[b] Data presented as median (interquartile range).

accordingly. Finally, we observed that among boys who reported no history of diagnosed concussion, collision sport participation (ie, football or lacrosse or both) was significantly associated with less reported anxiety but higher pain levels than those in no- or limited-contact sports. Given the cross-sectional design of our study, we cannot infer causation from these findings. Nevertheless, the results provide insights into the relative benefits and risks of participation in collision and contact sports during high school compared with no-contact or limited-contact sports.

The association between sport type and QoL outcomes in high school athletes is likely multifactorial and complex. Our finding that collision and contact sport participation in the year before the assessment was associated with reduced current anxiety and depressive symptom ratings relative to no- or limited-contact sport participation is therefore likely to be due to many factors, some of which we did not investigate (eg, socioeconomic status, sport availability at school, or race and ethnicity). Sports of all types allow youth to experience challenges and enjoyment and can help to increase self-esteem and decrease stress.[18] Team-oriented sports may offer added benefits, as researchers[19] documented fewer anxiety symptoms (eg, panic and agoraphobia) among these athletes than among athletes in individual sports. The collision and contact sport group primarily consisted of athletes who were involved in team sports, so the benefits of teamwork may be one source of possible differences in anxiety and depressive symptoms. Also, individuals who choose collision and contact sports may inherently have fewer anxiety and depressive symptoms than those who choose no- or limited-contact sports. For example, a higher proportion of female athletes were in the no- or limited-contact sport group than in the collision and contact sport group. Female athletes may have a higher prevalence of anxiety symptoms than male athletes,[20] and this group may have been more prone to reports of anxiety due to the proportional difference in sex. Although we adjusted for this variable in our multivariable analysis, the effects of sport participation and inherent traits such as biological sex should be considered when interpreting the association between QoL and sport participation.

In contrast to the potential benefit of fewer anxiety and depressive symptoms among the collision and contact sport

Table 5. Effect of Contact or Collision Sport Participation (Football and Lacrosse) on Patient-Reported Quality of Life (PROMIS) Outcome Measures Among Male Participants with No Reported History of Concussion[a]

PROMIS Domain	β Coefficient	Standard Error	95% Confidence Interval	*P* Value
Mobility	−0.125	0.083	−0.289, 0.039	.13
Anxiety[b]	−1.219	0.426	−2.066, −0.373	.005
Depressive symptoms	−0.522	0.265	−1.048, 0.004	.05
Fatigue	−0.281	0.521	−1.065, 1.009	.96
Peer relationships	0.739	1.039	−1.326, 2.805	.48
Pain interference	0.112	0.485	−0.853, 1.077	.82
Pain scale[b]	0.672	0.273	0.128, 1.215	.016

Abbreviation: PROMIS, Patient-Reported Outcomes Measurement Information System.
[a] Covariates in each model were age; history of bone, muscle, ligament, or tendon injury; and history of bone fracture or dislocation.
[b] $P < .05$.

athletes, the risks of participating in these sports were also evident in our findings. Specifically, a higher proportion of these athletes reported a history of time-loss orthopaedic injuries than did the no- or limited-contact athletes. This result aligns with the findings of other epidemiologic researchers[21,22] who described a greater prevalence of bone, joint, and muscle injuries among collision and contact athletes than among no- or limited-contact athletes. As such, it appears that collision and contact sport participation may be associated with both a greater risk of orthopaedic injury and fewer anxiety and depression symptoms. Interestingly, the proportion of participants within each group who reported a history of a diagnosed concussion did not differ. This was surprising given past studies[23,24] that demonstrated a higher risk of concussion among collision and contact sport high school athletes than among no- or limited-contact sport athletes. However, concussions can occur from a variety of sport- and recreation-related injuries in childhood and adolescence,[25] potentially leading to the similar proportions of those with a past concussion. The fact that we relied on self-reporting of a head injury or concussion rather than requiring a medically diagnosed concussion or brain injury could have influenced these results as well.

For our secondary purpose, we compared male athletes with no lifetime concussions who participated in collision sports and those who participated in no- or limited-contact sports. Unexpectedly, our results indicated that the collision sport athletes reported less anxiety than the no- or limited-contact sport athletes. It is evident that a concussion during adolescence can cause QoL alterations and that the effects can be long-lasting,[9,26,27] but the effects of repetitive head impacts are less clear. In our sample, collision sport participation was not associated with worse QoL during high school in the absence of a diagnosed concussion and, in fact, was associated with fewer anxiety symptoms. Causality cannot be inferred from this finding, however, as it suggests that either less anxious athletes are more likely to play collision sports or participating in collision sports is potentially beneficial, possibly due to the positive effects of team sports, or both.[19] Yet any potential benefit must be weighed against the potential relative risks. The higher pain scale scores among the collision sport male athletes may relate to their greater proportion of past orthopaedic injuries than among the no- or limited-contact male athletes. Given the small sample in this secondary analysis, future researchers need to confirm these results, but they do suggest that potential benefits and risks exist during high school collision sports.

Our study had a number of limitations, and our interpretation of the findings must be considered in light of them. The cross-sectional nature of the study eliminates the ability to establish any causality related to the effect of sport participation on QoL outcomes. Additionally, our participants were from a single school district. As such, our results cannot be generalized or extrapolated to other geographic locations. No female athletes reported collision sport participation in the year preceding assessment, so we could only examine boys in our secondary purpose. Future investigators should pursue valid comparisons by sex. Finally, our reliance on self-reported injury histories may have influenced the findings. Prospective injury-monitoring approaches may permit causation to be established between the factors we observed.

CONCLUSIONS

High school athletes who were currently participating in organized collision and contact sports reported fewer anxiety and depressive symptoms than those currently participating in no- or limited-contact sports. Furthermore, among boys with no reported history of diagnosed concussion, current collision sport participation was associated with lower state anxiety levels than among boys who participated in no- or limited-contact sports. However, our study design prohibited any interpretation related to long-term QoL in this cohort. Yet our results reinforce the need to reexamine assumptions that youth collision and contact sports are necessarily associated with negative QoL. The higher proportion of those with a history of orthopaedic injury among collision and contact sport athletes suggests that there may be physical risks to participation in these sports. Future researchers should better understand any causal relationship between contact sports and psychological well-being in young athletes, both in the short and long term, and determine the risks and benefits of playing different types of sports during high school.

REFERENCES

1. High school sports participation increases for 29th consecutive year. GlobeNewswire Web site. http://www.globenewswire.com/news-release/2018/08/24/1556351/0/en/High-School-Sports-Participation-Increases-for-29th-Consecutive-Year.html. Published 2018. Accessed August 27, 2019.
2. 2018-19 High school athletics participation survey. National Federation of State High School Associations Web site. https://www.nfhs.org/media/1020406/2018-19-participation-survey.pdf. Published 2019. Accessed June 17, 2020.
3. Bailey R. Physical education and sport in schools: a review of benefits and outcomes. *J Sch Health*. 2006;76(8):397–401. doi: 10.1111/j.1746-1561.2006.00132.x
4. Pate RR, Trost SG, Levin S, Dowda M. Sports participation and health-related behaviors among US youth. *Arch Pediatr Adolesc Med*. 2000;154(9):904–911. doi: 10.1001/archpedi.154.9.904
5. Merkel DL. Youth sport: positive and negative impact on young athletes. *Open Access J Sports Med*. 2013;4:151–160. doi: 10.2147/OAJSM.S33556
6. Pierpoint LA, LaBella CR, Collins CL, Fields SK, Comstock RD. Injuries in girls' soccer and basketball: a comparison of high schools with and without athletic trainers. *Inj Epidemiol*. 2018;5(1):29. doi: 10.1186/s40621-018-0159-6
7. Alosco ML, Kasimis AB, Stamm JM, et al. Age of first exposure to American football and long-term neuropsychiatric and cognitive outcomes. *Transl Psychiatry*. 2017;7(9):e1236. doi: 10.1038/tp.2017.197
8. Solomon GS, Kuhn AW, Zuckerman SL, et al. Participation in pre-high school football and neurological, neuroradiological, and neuropsychological findings in later life: a study of 45 retired National Football League players. *Am J Sports Med*. 2016;44(5):1106–1115. doi: 10.1177/0363546515626164
9. Meehan WP III, Taylor AM, Berkner P, et al. Division III collision sports are not associated with neurobehavioral quality of life. *J Neurotrauma*. 2016;33(2):254–259. doi: 10.1089/neu.2015.3930
10. Caccese JB, DeWolf RM, Kaminski TW, et al; CARE Consortium Investigators. Estimated age of first exposure to American football and neurocognitive performance amongst NCAA male student-

athletes: a cohort study. *Sports Med*. 2019;49(3):477–487. doi: 10.1007/s40279-019-01069-x
11. Brett BL, Huber DL, Wild A, Nelson LD, McCrea MA. Age of first exposure to American football and behavioral, cognitive, psychological, and physical outcomes in high school and collegiate football players. *Sports Health*. 2019;11(4):332–342. doi: 10.1177/1941738119849076
12. Deshpande SK, Hasegawa RB, Rabinowitz AR, et al. Association of playing high school football with cognition and mental health later in life. *JAMA Neurol*. 2017;74(8):909–918. doi: 10.1001/jamaneurol.2017.1317
13. Savica R, Parisi JE, Wold LE, Josephs KA, Ahlskog JE. High school football and risk of neurodegeneration: a community-based study. *Mayo Clin Proc*. 2012;87(4):335–340. doi: 10.1016/j.mayocp.2011.12.016
14. Hutchison MG, Di Battista AP, McCoskey J, Watling SE. Systematic review of mental health measures associated with concussive and subconcussive head trauma in former athletes. *Int J Psychophysiol*. 2018;132(pt A):55–61. doi: 10.1016/j.ijpsycho.2017.11.006
15. Kratz AL, Slavin MD, Mulcahey MJ, Jette AM, Tulsky DS, Haley SM. An examination of the PROMIS(®) pediatric instruments to assess mobility in children with cerebral palsy. *Qual Life Res*. 2013;22(10):2865–2876. doi: 10.1007/s11136-013-0397-6
16. Varni JW, Magnus B, Stucky BD, et al. Psychometric properties of the PROMIS® pediatric scales: precision, stability, and comparison of different scoring and administration options. *Qual Life Res*. 2014;23(4):1233–1243. doi: 10.1007/s11136-013-0544-0
17. Rice SG; American Academy of Pediatrics Council on Sports Medicine and Fitness. Medical conditions affecting sports participation. *Pediatrics*. 2008;121(4):841–848. doi: 10.1542/peds.2008-0080
18. Fraser-Thomas JL, Côté J, Deakin J. Youth sport programs: an avenue to foster positive youth development. *Phys Educ Sport Pedagogy*. 2005;10(1):19–40. doi: 10.1080/1740898042000334890
19. Ashdown-Franks G, Sabiston CM, Solomon-Krakus S, O'Loughlin JL. Sport participation in high school and anxiety symptoms in young adulthood. *Ment Health Phys Act*. 2017;12:19–24. doi: 10.1016/j.mhpa.2016.12.001
20. Patel DR, Omar H, Terry M. Sport-related performance anxiety in young female athletes. *J Pediatr Adolesc Gynecol*. 2010;23(6):325–335. doi: 10.1016/j.jpag.2010.04.004
21. Hootman JM, Dick R, Agel J. Epidemiology of collegiate injuries for 15 sports: summary and recommendations for injury prevention initiatives. *J Athl Train*. 2007;42(2):311–319.
22. Caine D, Maffulli N, Caine C. Epidemiology of injury in child and adolescent sports: injury rates, risk factors, and prevention. *Clin Sports Med*. 2008;27(1):19–50, vii. doi: 10.1016/j.csm.2007.10.008
23. Marar M, McIlvain NM, Fields SK, Comstock RD. Epidemiology of concussions among United States high school athletes in 20 sports. *Am J Sports Med*. 2012;40(4):747–755. doi: 10.1177/0363546511435626
24. Meehan WP III, d'Hemecourt P, Collins CL, Comstock RD. Assessment and management of sport-related concussions in United States high schools. *Am J Sports Med*. 2011;39(11):2304–2310. doi: 10.1177/0363546511423503
25. Bryan MA, Rowhani-Rahbar A, Comstock RD, Rivara F; Seattle Sports Concussion Research Collaborative. Sports- and recreation-related concussions in US youth. *Pediatrics*. 2016;138(1):e20154635. doi: 10.1542/peds.2015-4635
26. Fineblit S, Selci E, Loewen H, Ellis M, Russell K. Health-related quality of life after pediatric mild traumatic brain injury/concussion: a systematic review. *J Neurotrauma*. 2016;33(17):1561–1568. doi: 10.1089/neu.2015.4292
27. Kamins J, Bigler E, Covassin T, et al. What is the physiological time to recovery after concussion? A systematic review. *Br J Sports Med*. 2017;51(12):935–940. doi: 10.1136/bjsports-2016-097464

Address correspondence to David R. Howell, PhD, ATC, Sports Medicine Center, Children's Hospital Colorado, Department of Orthopedics, University of Colorado School of Medicine, 13123 East 16th Avenue, B060, Aurora, CO 80045. Address email to David.Howell@CUanschutz.edu.

American Journal of
Preventive Medicine

BRIEF REPORT

Concussion Rates in U.S. Middle School Athletes, 2015–2016 School Year

Zachary Y. Kerr, PhD, MPH,[1] Nelson Cortes, PhD,[2] Amanda M. Caswell, PhD, ATC, VATL,[2] Jatin P. Ambegaonkar, PhD, ATC, VATL,[2] Kaitlin Romm Hallsmith, MS, ATC, VATL,[2] A. Frederick Milbert, MEd,[3] Shane V. Caswell, PhD, ATC, VATL[2]

Introduction: Concussion incidence estimates in middle school sports settings are limited. This study examines concussion incidence in nine U.S. middle schools during the 2015–2016 school year.

Methods: Concussion data originated from nine public middle schools in Prince William County, Virginia, during the 2015–2016 school year. Certified athletic trainers collected concussion and athlete exposure (AE) data in school-sanctioned games and practices in boys' baseball, basketball, football, soccer, track, and wrestling; and girls' basketball, cheerleading, soccer, softball, track, and volleyball. Athletic trainers also acquired data on non–school sanctioned sport concussions. In 2017, concussion rates were calculated per 1,000 AEs. Injury rate ratios with 95% CIs compared rates between games and practices and by sex.

Results: Overall, 73 concussions were reported, of which 21.9% were from non–school sanctioned sport settings. The 57 remaining game and practice concussions were reported during 76,384 AEs, for a concussion rate of 0.75/1,000 AEs. Football had the highest concussion rate (2.61/1,000 AEs). Concussion rates were higher in games versus practices (injury rate ratio=1.83, 95% CI=1.06, 3.15), and in girls versus boys in sex-comparable sports, i.e., baseball/softball, basketball, soccer, and track (injury rate ratio=3.73, 95% CI=1.24, 11.23).

Conclusions: Current findings parallel those found in high school and college sports settings in that higher concussion rates were reported in girls and competitions. However, concussion rates exceeded those recently reported in high school and youth league settings, highlighting the need for continued research in the middle school sports setting. Given that one in five concussions were from non–school sanctioned sport settings, prevention efforts in middle school sports settings should consider sport and non-sport at-risk exposure.

Am J Prev Med 2017;53(6):914–918.

INTRODUCTION

The ability to compare concussion incidence among sports allows for better detection of where the greatest prevention needs are required. Multi-sport concussion-specific analyses at the collegiate and high school levels[1–3] suggest that football, wrestling, soccer, and lacrosse have the highest concussion rates. Estimates of concussion incidence in youth are limited, yet needed.[4] Most findings focus on the general population presenting at emergency departments[5–9] or within single-sport youth league settings.[9–17] Few studies have examined concussions in middle school sports.[18,19] One study examined 20 years of data (1988–2008) from one school, reporting 58 concussions and the highest rate in football.[18] The second study examined nine middle

From the [1]Department of Exercise and Sport Science, University of North Carolina, Chapel Hill, North Carolina; [2]Sports Medicine Assessment, Research and Testing (SMART) Laboratory, George Mason University, Manassas, Virginia; and [3]Prince William County Public Schools, Prince William County Virginia, Manassas, Virginia

Address correspondence to: Shane V. Caswell, PhD, ATC, VATL, George Mason University, 10900 University Blvd., MS 4E5, Manassas VA 20110. E-mail: scaswell@gmu.edu.

0749-3797/$36.00
https://doi.org/10.1016/j.amepre.2017.05.017

school football teams, reporting zero concussions.[19] Current estimates may differ because of increased concussion knowledge among clinicians, athletes, and coaches.[3] This study examines concussion incidence in nine U.S. middle schools during the 2015–2016 school year.

METHODS

Data originated from nine public middle schools in Prince William County, Virginia, during the 2015–2016 school year. Overall, 2,622 athlete seasons originated from 12 sports: boys' baseball, basketball, football, soccer, track, and wrestling; and girls' basketball, cheerleading, soccer, softball, track, and volleyball. All nine schools had each sport, with the exception of one not sponsoring softball.

Certified athletic trainers (ATs) collected injury and exposure data at all school-sanctioned games and practices. Injury data were collected via electronic medical record systems used for daily clinical practice. Exposure data were recorded daily via daily attendance using a phone application designed specifically for this study. Injuries were those that (1) occurred as a result of participation in an organized game or practice and (2) received medical attention by an AT, school nurse, or team physician. Team medical staff followed the Zurich Concussion guidelines.[20] An athlete exposure (AE) was defined as one athlete participating in one school-sanctioned game or practice.

Data were also collected on concussions that occurred during the sport season, but in non–school sanctioned sport settings, and were reported by athletes, their parents, or both to ATs, school nurses, or team physicians. School nurses and team physicians were instructed to inform the ATs serving as data collectors about such non–school sanctioned sport concussions. Specific injury mechanism information and AEs were not collected for non–school sanctioned sport concussions.

In 2017, concussion rates were calculated per 1,000 AEs overall, and by games and practices. Injury rate ratios (IRRs) compared concussion rates between games and practices and by sex. Sex comparisons were conducted for sports played by both boys and girls (i.e., baseball/softball, basketball, soccer, track). The IRRs with 95% CIs excluding 1.00 were deemed statistically significant. This study was approved by the George Mason University Human Subjects Review Board.

RESULTS

Seventy-three concussions were reported in the 2015–2016 school year. Most occurred in practice (50.7%), followed by games (27.4%) and non–school sanctioned sport settings (21.9%).

The 57 game and practice concussions were reported during 76,384 AEs (Table 1) for an overall concussion rate of 0.75/1,000 AEs (Table 2). Sports with the highest concussion rates were boys' football (2.61/1,000 AEs) and girls' soccer (1.30/1,000 AEs). Football also had the highest game and practice concussion counts and rates. However, in games, girls' basketball had a slightly higher concussion rate than girls' soccer (2.61/1,000 AEs and 2.47/1,000 AEs, respectively). In practices, concussion rates in girls' softball (0.90/1,000 AEs) and girls' cheerleading (0.84/1,000 AEs) were slightly higher than that of girls' soccer (0.80/1,000 AEs).

Table 1. Concussion and Athlete Exposure Counts for Middle School Sports, 2015–2016 School Year

	Concussions[a]		Athlete exposures[b]	
Sport	Game	Practice	Game	Practice
Boys				
Baseball	1	1	979	2,522
Basketball	0	1	1,762	3,944
Football	7	20	1,624	8,739
Soccer	1	0	1,664	4,846
Track	0	0	1,394	5,538
Wrestling	1	2	1,511	4,388
Sex-comparable total[c]	2	2	5,799	16,850
Total	10	24	8,934	29,977
Girls				
Basketball	5	1	1,917	4,896
Cheerleading	0	6	1,637	7,127
Soccer	4	3	1,617	3,769
Softball	0	2	722	2,232
Track	0	0	1,195	6,443
Volleyball	1	1	1,399	4,519
Sex-comparable total[c]	9	6	5,451	17,340
Total	10	13	8,487	28,986
Overall	20	37	17,421	58,963

[a]Does not include the 16 concussions that occurred to athletes in non–school sanctioned sport settings during their respective seasons (two boys' basketball, one boys' football, two boys' soccer, three girls' cheerleading, five girls' soccer, one girls' softball, one girls' track, and one girls' volleyball).
[b]An athlete exposure is defined as one athlete's participation in one game or practice; athlete exposures were not collected for non–school sanctioned sport concussions.
[c]Sex-comparable sports include baseball/softball, basketball, soccer, and track.

The overall concussion rate was higher in games than practices (1.15 vs 0.63/1,000 AEs, IRR=1.83, 95% CI=1.06, 3.15; Table 2). This finding was retained per sport, although only girls' basketball reported a significant finding.

Concussion rates did not vary between girls and boys (0.61 vs 0.87/1,000 AEs, IRR=0.70, 95% CI=0.41, 1.19; Table 2). However, when considering sex-comparable sports only, the concussion rate was higher in girls than boys (0.66 vs 0.18/1,000 AEs, IRR=3.73, 95% CI=1.24, 11.23). Findings were retained in sport-specific analyses, but only found significant in soccer.

Table 2. Concussion Rates and Injury Rate Ratios for Middle School Sports, 2015–2016 School Year

Sport	Concussion rates per 1,000 AEs (95% CI)			Injury rate ratios (95% CI)	
	Game	Practice	Total[a]	Game versus practice	Females versus males[b]
Boys					
Baseball	1.02 (0.00, 3.02)	0.40 (0.00, 1.17)	0.57 (0.00, 1.36)	2.58 (0.16, 41.19)	
Basketball	0.00	0.25 (0.00, 0.75)	0.18 (0.00, 0.52)	N/A	
Football	4.31 (1.12, 7.50)	2.29 (1.29, 3.29)	2.61 (1.62, 3.59)	1.88 (0.80, 4.45)	
Soccer	0.60 (0.00, 1.78)	0.00	0.15 (0.00, 0.45)	N/A	
Track	0.00	0.00	0.00	N/A	
Wrestling	0.66 (0.00, 1.96)	0.46 (0.00, 1.09)	0.51 (0.00, 1.08)	1.45 (0.13, 16.01)	
Sex-comparable total[b]	0.34 (0.00, 0.82)	0.12 (0.00, 0.28)	0.18 (0.00, 0.35)	2.91 (0.41, 20.63)	
Total	1.12 (0.43, 1.81)	0.80 (0.48, 1.12)	0.87 (0.58, 1.17)	1.40 (0.67, 2.92)	
Girls					
Basketball	2.61 (0.32, 4.89)	0.20 (0.00, 0.60)	0.88 (0.18, 1.59)	**12.77 (1.49, 109.31)**	5.03 (0.60, 41.74)
Cheer	0.00	0.84 (0.17, 1.52)	0.68 (0.14, 1.23)	N/A	
Soccer	2.47 (0.05, 4.90)	0.80 (0.00, 1.70)	1.30 (0.34, 2.26)	3.11 (0.70, 13.89)	**8.46 (1.04, 68.77)**
Softball	0.00	0.90 (0.00, 2.14)	0.68 (0.00, 1.62)	N/A	1.19 (0.17, 8.41)
Track	0.00	0.00	0.00	N/A	N/A
Volleyball	0.71 (0.00, 2.12)	0.22 (0.00, 0.66)	0.34 (0.00, 0.81)	3.23 (0.20, 51.64)	
Sex-comparable total[b]	1.65 (0.57, 2.73)	0.35 (0.07, 0.62)	0.66 (0.33, 0.99)	**4.77 (1.70, 13.41)**	**3.73 (1.24, 11.23)**
Total	1.18 (0.45, 1.91)	0.45 (0.20, 0.69)	0.61 (0.36, 0.86)	**2.63 (1.15, 5.99)**	0.70 (0.41, 1.19)
Overall	1.15 (0.64, 1.65)	0.63 (0.43, 0.83)	0.75 (0.55, 0.94)	**1.83 (1.06, 3.15)**	

Note: Boldface indicates statistical significance (95% CIs do not include 1.00).
[a]Total concussion rates do not include non–school sanctioned sport concussions.
[b]Sex-comparable sports include baseball/softball, basketball, soccer, and track.
AEs, athlete exposures; N/A, not available.

DISCUSSION

This study's data provide an updated examination of concussion incidence in middle school sports from the 2015–2016 school year. The study includes athletes from 12 sports in nine public middle schools in Prince William County, Virginia, and produced the largest data set of middle school concussions. As found in high school and collegiate sports,[1–3,21–23] middle school sports had higher concussion rates in games than practices, and in girls than boys. Football also had the highest concussion rate.[2,3,21–23] Sports such as boys' basketball, baseball, soccer, and track and girls' softball, track, and volleyball reported low counts. Future research should continue to examine middle school sports settings to validate and improve upon this study's findings.

The sport-specific concussion rates in this study exceed those reported in Beachy and Rauh[18] for all sports, including football (2.61 vs 0.35/1,000 AEs); boys' basketball (0.18 vs 0.09/1,000 AEs); and girls' soccer (1.30 vs 0.07/1,000 AEs). The findings may reflect increases in concussion education and detection among athletes, coaches, and medical staff, which may have led to better reporting.[3] Interestingly though, reported concussion rates were also generally larger than those reported in high school data from the 2015–2016 school year[24]; the football concussion rate exceeded recent estimates from youth football leagues (range of 0.62–1.76/1,000 AEs).[11,12,14,15] These findings may be due to variations in data collection efforts, although these studies also utilized onsite medical staff to collect data. Because of inherent differences related to factors such as resource allocation and scheduling, it may be difficult to compare concussion incidence across various sports settings. Nevertheless, findings stress that concussion prevention must occur within middle schools alongside high schools and youth league sports.

Interestingly, more than one fifth of concussions reported in the study occurred from non–school sanctioned sports settings. This proportion may be an underestimate given the reliance on reporting from athletes and parents. Also, there were no data on specifically where these concussions occurred. Examining concussion incidence across all youth activities in which athletes are at risk is warranted. It is also essential

to investigate how concussion prevention and management should extend beyond one particular setting. This is especially important as school systems are confronted with the responsibility of appropriately managing a child's concussion recovery regardless if it occurred within or outside school-sanctioned activities.

Limitations

The current study is not without limitations. This sample may not be generalizable to non-participating middle schools, particularly those that may sponsor other sports not included in this study. Also, because AEs were unit-based, injury rates could not be interpreted per minute or hour of participation. However, the use of AEs reduces burden of data collection on participating ATs. Although ATs are educated and trained to properly identify concussions, it is possible that concussions were missed because of athlete non-reporting associated with purposeful withholding,[25] lack of knowledge,[25] or receiving care outside of school. Last, analyses did not account for multi-sport athletes in middle school sports or at-risk exposure time from non–school sanctioned sport activities. Future research may help identify how such increased at-risk exposure affects risk of concussion and other injuries.

CONCLUSIONS

Patterns of concussion incidence present in other sports settings were observed in this study's middle school sports data,[1–3,21–23] which emphasizes the need for concussion prevention and management within middle schools to reduce incidence and severity. Given the call for estimates of concussion incidence in youth,[4] additional concussion surveillance data collection in middle school sports is needed. This setting may be more diverse and representative of the general population than some youth sports leagues, where fees and extra costs may inhibit participation from individuals with lower SES.

ACKNOWLEDGMENTS

We greatly appreciate the collaboration of the Prince William County Public Schools and support of the middle school principals, nurses, athletic coordinators, coaches, and athletic trainers. The Middle School ACHIEVES (Advancing Healthcare Initiatives for Underserved Students) Project aims to increase access to health care for underserved communities and is made possible with grant funding from the Potomac Health Foundation, of which S. Caswell is the Principal Investigator and N. Cortes, J. Ambegaonkar, and A. Caswell are co-investigators.

No financial disclosures were reported by the authors of this paper.

REFERENCES

1. Zuckerman SL, Kerr ZY, Yengo-Kahn A, Wasserman E, Covassin T, Solomon GS. Epidemiology of sports-related concussion in NCAA athletes from 2009-2010 to 2013-2014: incidence, recurrence, and mechanisms. *Am J Sports Med*. 2015;43(11):2654–2662. https://doi.org/10.1177/0363546515599634.
2. O'Connor KL, Baker MM, Dalton SL, Dompier TP, Broglio SP, Kerr ZY. Epidemiology of sport-related concussion in high school athletes: National Athletic Treatment, Injury and Outcomes Network, 2011/12-2013/14. *J Athl Train*. 2017;52(3):175–185. https://doi.org/10.4085/1062-6050-52.1.15.
3. Rosenthal JA, Foraker RE, Collins CL, Comstock RD. National high school athlete concussion rates from 2005-2006 to 2011-2012. *Am J Sports Med*. 2014;42(7):1710–1715. https://doi.org/10.1177/0363546514530091.
4. National Academy of Medicine. Sports-related concussions in youth: improving the science, changing the culture. www.nap.edu/catalog/18377/sports-related-concussions-in-youth-improving-the-science-changing-the. Published 2013. Accessed February 22, 2017.
5. Bakhos LL, Lockhart GR, Myers R, Linakis JG. Emergency department visits for concussion in young child athletes. *Pediatrics*. 2010;126(3):e550–e556. https://doi.org/10.1542/peds.2009-3101.
6. Jacobson NA, Buzas D, Morawa LG. Concussions from youth football results from NEISS hospitals over an 11-year time frame, 2002-2012. *Orthop J Sports Med*. 2013;1(7). https://doi.org/10.1177/2325967113517860.
7. Coronado VG, Haileyesus T, Cheng TA, et al. Trends in sports- and recreation-related traumatic brain injuries treated in U.S. emergency departments: the National Electronic Injury Surveillance System–All Injury Program (NEISS-AIP) 2001-2012. *J Head Trauma Rehabil*. 2015;30(3):185–197. https://doi.org/10.1097/HTR.0000000000000156.
8. Buzas D, Jacobson NA, Morawa LG. Concussions from 9 youth organized sports results from NEISS hospitals over an 11-year time frame, 2002-2012. *Orthop J Sports Med*. 2014;2(4). https://doi.org/10.1177/2325967114528460.
9. Bryan MA, Rowhani-Rahbar A, Comstock RD, Rivara F. Sports-and recreation-related concussions in U.S. youth. *Pediatrics*. 2016;138(1):e20154635. https://doi.org/10.1542/peds.2015-4635.
10. O'Kane JW, Spieker A, Levy MR, Neradilek M, Polissar NL, Schiff MA. Concussion among female middle-school soccer players. *JAMA Pediatr*. 2014;168(3):258–264. https://doi.org/10.1001/jamapediatrics.2013.4518.
11. Dompier TP, Kerr ZY, Marshall SW, et al. Incidence of concussion during practice and games in youth, high school, and collegiate American football players. *JAMA Pediatr*. 2015;169(7):659–665. https://doi.org/10.1001/jamapediatrics.2015.0210.
12. Kerr ZY, Zuckerman SL, Wasserman EB, Covassin T, Djoko A, Dompier TP. Concussion symptoms and return to play time in youth, high school, and college American football athletes. *JAMA Pediatr*. 2016;170(7):647–653. https://doi.org/10.1001/jamapediatrics.2016.0073.
13. Kerr ZY, Marshall SW, Simon JE, et al. Injury rates in age-only versus age-and-weight playing standard conditions in American youth football. *Orthop J Sports Med*. 2015;3(9). https://doi.org/10.1177/2325967115603979.
14. Kontos AP, Elbin R, Fazio-Sumrock VC, et al. Incidence of sports-related concussion among youth football players aged 8-12 years. *J Pediatr*. 2013;163(3):717–720. https://doi.org/10.1016/j.jpeds.2013.04.011.
15. Kerr ZY, Yeargin S, McLeod TCV, et al. Comprehensive coach education and practice contact restriction guidelines result in lower injury rates in youth American football. *Orthop J Sports Med*. 2015;3(7). https://doi.org/10.1177/2325967115594578.
16. Kontos AP, Elbin R, Sufrinko A, et al. Incidence of concussion in youth ice hockey players. *Pediatrics*. 2016;137(2):e20151633. https://doi.org/10.1542/peds.2015-1633.
17. Kerr ZY, Caswell SV, Lincoln AE, Djoko A, Dompier TP. The epidemiology of boys' youth lacrosse injuries in the 2015 season. *Inj Epidemiol*. 2016;3(1):3. https://doi.org/10.1186/s40621-016-0068-5.

18. Beachy G, Rauh M. Middle school injuries: a 20-year (1988-2008) multisport evaluation. *J Athl Train*. 2014;49(4):493–506. https://doi.org/10.4085/1062-6050-49.2.19.
19. Turbeville SD, Cowan LD, Asal NR, Owen WL, Anderson MA. Risk factors for injury in middle school football players. *Am J Sports Med*. 2003;31(2):276–281.
20. McCrory P, Meeuwisse WH, Aubry M, et al. Consensus statement on concussion in sport: the 4th International Conference on Concussion in Sport, Zurich, November 2012. *J Athl Train*. 2013;48(4):554–575. https://doi.org/10.4085/1062-6050-48.4.05.
21. Gessel LM, Fields SK, Collins CL, Dick RW, Comstock RD. Concussions among United States high school and collegiate athletes. *J Athl Train*. 2007;42(4):495–503.
22. Marar M, McIlvain NM, Fields SK, Comstock RD. Epidemiology of concussions among United States high school athletes in 20 sports. *Am J Sports Med*. 2012;40(4):747–755. https://doi.org/10.1177/0363546511435626.
23. Lincoln AE, Caswell SV, Almquist JL, Dunn RE, Norris JB, Hinton RY. Trends in concussion incidence in high school sports: a prospective 11-year study. *Am J Sports Med*. 2011;39(5):958–963. https://doi.org/10.1177/0363546510392326.
24. Comstock R, Currie D, Pierpoint L. Convenience sample summary report: National High School Sports-Related Injury Surveillance Study: 2015-2016 school year. www.ucdenver.edu/academics/colleges/PublicHealth/research/ResearchProjects/piper/projects/RIO/Documents/Conv%20Report%20Final%202015%2016%2009%203%2016.pdf. Published 2016. Accessed January 31, 2017.
25. Kerr ZY, Register-Mihalik JK, Marshall SW, Evenson KR, Mihalik JP, Guskiewicz KM. Disclosure and non-disclosure of concussion and concussion symptoms in athletes: review and application of the socio-ecological framework. *Brain Inj*. 2014;28(8):1009–1021. https://doi.org/10.3109/02699052.2014.904049.

THE JOURNAL OF
SEXUAL MEDICINE

ORIGINAL RESEARCH

TRANSGENDER HEALTH

Early Hormonal Treatment Affects Body Composition and Body Shape in Young Transgender Adolescents

Maartje Klaver, MD,[1] Renée de Mutsert, PhD,[2] Chantal M. Wiepjes, MD,[1] Jos W. R. Twisk, PhD,[3] Martin den Heijer, MD, PhD,[1] Joost Rotteveel, MD, PhD,[4] and Daniël T. Klink, MD, PhD[4,5]

ABSTRACT

Background: Transgender adolescents aspiring to have the body characteristics of the affirmed sex can receive hormonal treatment. However, it is unknown how body shape and composition develop during treatment and whether transgender persons obtain the desired body phenotype.

Aim: To examine the change in body shape and composition from the start of treatment with gonadotropin-releasing hormone agonists (GnRHa) until 22 years of age and to compare these measurements at 22 years with those of age-matched peers.

Methods: 71 transwomen (birth-assigned boys) and 121 transmen (birth-assigned girls) who started treatment from 1998 through 2014 were included in this retrospective study. GnRHa treatment was started and cross-sex hormonal treatment was added at 16 years of age. Anthropometric and whole-body dual-energy x-ray absorptiometry data were retrieved from medical records. Linear mixed model regression was performed to examine changes over time. SD scores (SDS) were calculated to compare body shape and composition with those of age-matched peers.

Outcomes: Change in waist-hip ratio (WHR), total body fat (TBF), and total lean body mass (LBM) during hormonal treatment. SDS of measures of body shape and composition compared with age-matched peers at 22 years of age.

Results: In transwomen, TBF increased (+10%, 95% CI = 7−11) while total LBM (−10%, 95% CI = −11 to −7) and WHR (−0.04, 95% CI = −0.05 to −0.02) decreased. Compared with ciswomen, SDS at 22 years of age were +0.3 (95% CI = 0.0−0.5) for WHR, and 0.0 (95% CI = −0.2 to 0.3) for TBF. Compared with cismen, SDS were −1.0 (95% CI = −1.3 to −0.7) for WHR, and +2.2 (95% CI = 2.2−2.4) for TBF. In transmen, TBF decreased (−3%, 95% CI = −4 to −1), while LBM (+3%, 95% CI = 1−4) and WHR (+0.03, 95% CI = 0.01−0.04) increased. Compared with ciswomen, SDS at 22 years of age were +0.6 (95% CI = 0.4−0.8) for WHR, and −1.1 (95% CI = −1.4 to −0.9) for TBF. Compared with cismen, SDS were −0.5 (95% CI = −0.8 to −0.3) for WHR, and +1.8 (95% CI = 1.6−1.9) for TBF.

Clinical Implications: Knowing body shape and composition outcomes at 22 years of age will help care providers in counseling transgender youth on expectations of attaining the desired body phenotype.

Strengths and Limitations: This study presents the largest group of transgender adults to date who started treatment in their teens. Despite missing data, selection bias was not found.

Conclusions: During treatment, WHR and body composition changed toward the affirmed sex. At 22 years of age, transwomen compared better to age-matched ciswomen than to cismen, whereas transmen were between reference values for ciswomen and cismen. **Klaver M, de Mutsert R, Wiepjes CM, et al. Early Hormonal Treatment Affects Body Composition and Body Shape in Young Transgender Adolescents. J Sex Med 2018;15:251−260.**

Received September 20, 2017. Accepted December 14, 2017.

[1]Department of Endocrinology and Center of Expertise on Gender Dysphoria, VU University Medical Center, Amsterdam, The Netherlands;

[2]Department of Clinical Epidemiology, Leiden University and Medical Center, Leiden, the Netherlands;

[3]Department of Clinical Epidemiology, VU University, Amsterdam, The Netherlands;

[4]Department of Pediatrics, Division of Endocrinology, VU University Medical Center, Amsterdam, The Netherlands;

[5]Department of Pediatrics, Division of Pediatric Endocrinology, Ghent University Hospital, Ghent, Belgium

https://doi.org/10.1016/j.jsxm.2017.12.009

Key Words: Transgender Persons; Adolescent; Gonadotropin-Releasing Hormone Analogues; Cross-Sex Hormonal Treatment; Body Composition; Body Shape

INTRODUCTION

Adolescents with gender dysphoria aspire to have body characteristics that are similar to those of the affirmed sex. From 12 years of age, adolescents with male-to-female gender dysphoria, referred to as transwomen, and adolescents with female-to-male gender dysphoria, referred to as transmen, can be treated with gonadotropin-releasing hormone analogues (GnRHa) to suppress puberty. Subsequently, at 16 years of age and if the person still pursues gender-affirming treatment, cross-sex hormonal treatment (CHT) is added to induce the secondary sexual characteristics of the affirmed sex.[1]

During puberty, with increasing sex steroid levels, girls develop more body fat that is deposited mainly in the gluteal and femoral region (so-called gynoid region).[2,3] This leads to a female body shape with a low waist-to-hip ratio (WHR).[3,4] Pubertal boys obtain more lean body mass (LBM) and store body fat mainly in the abdominal region (also referred to as the android region),[3] resulting in a male body shape with a higher WHR than seen in girls.[3,4] It is unknown how total and regional body fat, LBM, and body shape develop in transgender adolescents treated with GnRHa and CHT, and whether this results in a similar body composition and body shape as the affirmed sex in young adulthood.

Therefore, the 1st aim of this study was to examine the effects of treatment with GnRHa and CHT on total body and regional body fat, LBM, and body shape in adolescents with gender dysphoria. A 2nd aim of this study was to compare the achieved amount of total and regional body fat, LBM, and WHR of these transwomen and transmen at 22 years of age with reference values of the affirmed sex and to examine whether they obtained the desired body composition and body shape in young adulthood. A 3rd aim was to examine the influence of pubertal stage at start of treatment on the achieved body composition and body shape at 22 years.

METHODS

Study Design and Study Population

We retrospectively reviewed the medical records of all adolescents diagnosed with gender dysphoria (*Diagnostic and Statistical Manual of Mental Disorders, 4th Edition, Text Revision*[5]) at the VU University Medical Center (Amsterdam, the Netherlands) until December 2015. All persons who started hormonal treatment before 18 years of age, started the treatment protocol as described below,[1] had undergone whole-body dual-energy x-ray absorptiometry (DXA) during treatment, and according to their age had their medical checkups in young adulthood (>20.5 years) were eligible for this study. Data obtained during routine medical checkups on anthropometry, laboratory measurements, and whole-body DXA were collected at 3 time points: start of GnRHa, addition of CHT, and result at 22 years of age (range = 20.5−23.5 years). The local ethics committee approved the study and the necessity for informed consent was waived.

Treatment Protocol

The treatment protocol, also referred to as the Dutch protocol, has been published in detail.[1] At a minimum age of 12 years and stage B2 (breast) for girls and Tanner stage G3 (genital) for boys, subcutaneous GnRHa 3.75 mg for 4 weeks was started. From 16 years of age, CHT was added with increasing doses to initiate pubertal development. Transwomen were prescribed oral 17β-estradiol starting at 5 μg per kilogram of body weight per day, which was increased by 5 μg/kg per day every 6 months until the maintenance dose of 2 mg/day was reached. Transmen used initially mixed testosterone esters (Sustanon; Organon Pharmaceuticals, Oss, The Netherlands) intramuscularly starting at 25 mg per square meter of body surface area every 2 weeks, which was increased by 25 mg/m^2 every 6 months until the maintenance dose of 250 mg every 3 to 4 weeks was achieved. When GnRHa were started after 16 years of age, CHT was added after 3 to 6 months with a start dosage of 17β-estradiol 1 mg/day or intramuscular Sustanon 75 mg/week. After 6 months, this was increased to 17β-estradiol 2 mg/day in transwomen and Sustanon 250 mg every 3 to 4 weeks in transmen. From 18 years, patients were eligible for gonadectomy, after which treatment with GnRHa ceased. From the start of treatment, patients were advised to maintain a healthy lifestyle with sportive activities and an adequate calcium intake to prevent bone loss.

Anthropometry and Whole-Body DXA

At each visit, body height, body weight, waist circumference, and hip circumference were measured. Body height was measured to the nearest 0.1 cm using a Harpenden stadiometer. Body weight was measured while the subject wore only underwear without shoes to the nearest 0.1 kg. Body mass index (BMI) was calculated as weight in kilograms divided by height in meters squared. Waist circumference, defined as the smallest abdominal circumference, and hip circumference, measured at the level of the trochanter major, were determined with a tape measure to the nearest 0.1 cm. From these 2 measurements, the

WHR was calculated, which was used as a measure for body shape.

Whole-body and regional body fat, LBM, and total mass were measured using DXA. Until 2002, the Hologic QDR 2000 (Hologic Inc, Bedford, MA, USA) was used. From 2002, a Hologic Delphi apparatus (Hologic Inc) was used with software version 8.26, which was updated in 2005 to version 12.3.3. In 2011, the Hologic Delphi was replaced by a Hologic Discovery 13.1 (Hologic Inc), which was updated to version 3.3 in 2012 and to version 4.5.3 in 2015. During the review of medical records, DXA scans were included when they were obtained within 4 months before or after the start of GnRHa or CHT or within 1.5 years before or after the 22nd birthday. All available DXA scans from participants were obtained and reanalyzed with the most recent software version (version 13.5.3). The android region and gynoid region were defined using the software provided by Hologic. The lower boundary of the android region coincides with the upper edge of the pelvis and the height equals 20% of the distance from the upper edge of the pelvis to the bottom of the chin. The upper boundary of the gynoid region is below the upper edge of the pelvis by 1.5 times the height of the android region. The gynoid region equals twice the height of the android region.

Statistical Analyses

STATA 13.1 (StataCorp, College Station, TX, USA) was used for statistical analyses. Linear mixed model regression analyses with observations clustered within participants were performed to examine mean changes in measurements of body composition and body shape from the start of GnRHa to 22 years of age. Linear mixed models also properly deal with missing data.[6] When outcome variables were not normally distributed, the natural logarithm was obtained for analyses. Analyses were adjusted for Tanner stage at start of treatment, and BMI at start of treatment using centered variables. Student t-tests were used to examine whether changes in body shape and body composition during GnRHa monotherapy differed from changes after the addition of CHT.

Tanner stage at start of treatment, time, and the interaction between Tanner stage and time were added to the linear mixed model regression analysis to examine the influence on the achieved body composition and body shape at 22 years of age. In transwomen, Tanner stage at start was defined by testes volume and this resulted in the following categories: less than or equal to 8 mL (early puberty), 10 to 15 mL (mid-puberty), and at least 20 mL (late puberty).[7,8] In transmen, Tanner stage at start of treatment was defined by breast development.[7] Because a small number of transmen had Tanner stage II (n = 3) or III (n = 8) at the start of CHT, the 2 groups were classified as starters in early and mid-puberty. Transmen with Tanner stage IV or V were classified as starters in late puberty. Analyses were adjusted for BMI at 22 years of age.

We calculated standard deviation scores (SDS) to compare measures of body composition and body shape in participants at 22 years of age with reference values from age-matched peers with no (treatment for) gender dysphoria, also referred to as ciswomen and cismen. In transwomen and transmen, SDS were calculated for both ciswomen and cismen in order to compare measures with the at birth assigned sex and the affirmed sex. For the mean age of start of GnRHa (15 years in both transwomen and transmen) and for the age of 22 years, age-specific reference values [4,9-12] were retrieved from literature.

RESULTS

71 transwomen and 121 transmen who started with GnRHa and CHT from 1998 through 2014 were included in the present analyses (Figure 1). The general characteristics of the participants are presented in Table 1.

Change in Body Shape and Body Composition During Treatment in Transwomen

The results of the mixed model analyses showed that in transwomen waist circumference (+8 cm, 95% CI = 5−10, $P < .001$) and hip circumference (+17 cm, 95% CI = 13−21, $P < .001$) increased and WHR decreased (−0.04, 95% CI = −0.05 to −0.02, $P < .001$). The percentage of total body fat increased (+9%, 95% CI = 8−11, $P < .001$) and thus the percentage of LBM decreased (−9%, 95% CI = −11 to −8, $P < .001$; Table 2, Figure 2). Percentage of body fat increased in the android region (+9%, 95% CI = 6−12, $P < .001$) and the gynoid region (+11%, 95% CI = 9−12, $P < .001$; Table 2). Changes in body composition and body shape measurements were not different after adjustment for Tanner stage at start of treatment, or BMI at start of treatment.

Change in Body Shape and Body Composition During Treatment in Transmen

Transmen showed increases in waist circumference (+6 cm, 95% CI = 4−8, $P < .001$), hip circumference (+5 cm, 95% CI = 2−7, $P < .001$), and WHR (+0.03, 95% CI =

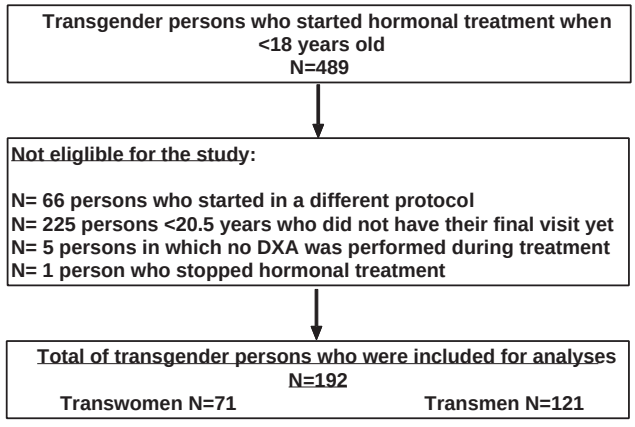


Figure 1. Flowchart of participant inclusion process. DXA = whole-body dual-energy x-ray absorptiometry.

Table 1. Characteristics of adolescents who started treatment with GnRHa and CHT at the VU University Medical Center from 1998 through 2014*

	Transwomen	Transmen
Adolescents	71	121
Age at start of GnRHa (y)	14.5 ± 1.8	15.3 ± 2.0
Age at start of CHT (y)	16.4 ± 1.1	16.9 ± 0.9
Ethnicity, %		
Caucasian	98	94
Asian	1	2
Black American	1	2
Occurrence of menarche, %†	–	84
BMI at start (kg/m^2)	19.8 (18.0–22.0)	20.6 (19.1–23.1)
Duration of GnRHa monotherapy (y)	2.1 (1.0–2.8)	1.0 (0.5–2.9)
Duration of GnRHa + CHT (y)	3.1 (2.5–3.6)	2.4 (2.0–3.1)
Duration of CHT monotherapy (y)	2.8 (1.6–3.4)	3.0 (1.9–3.4)
E2 level at start of GnRHa (pmol/L)	50 (20–79)	112 (70–202)
E2 level at start of CHT (pmol/L)	25 (20–31)	28 (23–36)
E2 level at 22 y of age (pmol/L)	121 (81–154)	70 (43–135)
T level at start of GnRHa (nmol/L)	10.0 (4.3–14.0)	1.0 (1.0–1.3)
T level at start of CHT (nmol/L)	1.0 (1.0–1.0)	1.0 (1.0–1.0)
T level at 22 y of age (nmol/L)	1.0 (0.8–1.0)	16.0 (8.8–37.0)
SDS at start vs ciswomen		
BMI	0.1 (−0.1 to 0.4)	0.4 (0.4–0.8)
Waist	0.6 (0.3–0.9)	0.6 (0.3–0.8)
Hip	0.0 (−0.2 to 0.2)	1.1 (0.1–0.5)
WHR	0.9 (0.6–0.9)	0.2 (0.0–0.3)
Total body fat	−0.9 (−1.0 to −0.7)	−0.1 (−0.2 to 0.1)
Lean body mass	2.5 (2.0–2.8)	1.3 (1.0–1.5)
SDS at start vs cismen		
BMI	0.4 (0.1–0.7)	0.7 (0.7–1.1)
Waist	0.1 (−0.2 to 0.4)	0.1 (−0.2 to 0.3)
Hip	0.3 (0.0–0.6)	0.7 (0.4–0.8)
WHR	−0.2 (−0.3 to 0.0)	−1.0 (−1.0 to −0.7)
Total body fat	1.6 (1.5–1.8)	2.0 (1.9–2.0)
Lean body mass	−0.7 (−0.9 to −0.6)	−1.1 (−1.1 to −1.0)

BMI = body mass index; CHT = cross-sex hormonal treatment; E2 = estradiol; GnRHa = gonadotropin-releasing hormone analogues; SDS = SD score; T = testosterone; WHR = waist-to-hip ratio.
*Data are presented as number, mean ± SD, median (interquartile range), or SDS (95% CI).
†Data were missing for 8% of transmen.

0.01−0.04, $P < .002$). Percentage of total body fat decreased (−3%, 95% CI = −4 to −2, $P < .001$), so percentage of LBM increased (+3%, 95% CI = 2−4, $P < .001$; Table 2, Figure 2). Percentage of body fat decreased in the gynoid region (−5%, 95% CI = −6 to −3, $P < .001$) with no change in the android region (+1%, 95% CI = 0−3, $P = .18$; Table 2). Changes in body composition and body shape measurements were not different after adjustment for Tanner stage at start of treatment, or BMI at start of treatment.

Body Shape and Body Composition at 22 Years Compared With Peers

SDS of body shape and body composition in transgender persons for both ciswomen and cismen are presented in Table 3.

Influence of Tanner Stage at Start on Body Shape and Body Composition at 22 Years

Transmen who started CHT in early or mid-puberty had a higher WHR than transmen who started treatment in late puberty. Transwomen tended to have a lower WHR starting in early or mid-puberty than those who started in late puberty. No differences in percentage of total body fat or percentage of total LBM were found across Tanner stages at start of treatment in transwomen and transmen (Table 4).

DISCUSSION

This study of 71 transwomen and 121 transmen shows that measurements of body shape and body composition change

Table 2. Measurements of body shape and body composition at the start of GnRHa, the start of CHT, and at 22 years of age in transwomen (n = 71) and transmen (n = 121)

	Start of GnRH	Start of CHT	22 y of age
Transwomen			
Body weight (kg)	58 (56–61)	66 (63–69)	76 (71–82)
BMI (kg/m^2)	20.2 (19.4–20.9)	21.3 (20.5–22.0)	23.2 (21.6–24.8)
Waist circumference (cm)*	71 (69–73)	75 (73–77)	79 (77–82)
Hip circumference (cm)*	89 (87–91)	95 (93–97)	106 (102–110)
WHR*	0.81 (0.79–0.82)	0.79 (0.78–0.80)	0.77 (0.75–0.79)
Body fat			
Total body (%)*	25 (23–26)	31 (29–32)	34 (32–36)
Android (%)*	23 (21–25)	28 (26–31)	32 (28–36)
Gynoid (%)*	29 (27–30)	36 (34–38)	40 (38–42)
Lean body mass			
Total body (%)*	75 (74–77)	69 (68–71)	66 (64–68)
Transmen			
Body weight (kg)	58 (56–61)	63 (60–65)	69 (66–71)
BMI (kg/m^2)	21.6 (20.9–22.3)	22.5 (21.7–23.2)	23.9 (23.0–24.7)
Waist circumference (cm)*	71 (69–72)	73 (71–74)	77 (75–79)
Hip circumference (cm)*	92 (90–93)	95 (93–97)	96 (94–99)
WHR*	0.77 (0.76–0.78)	0.76 (0.75–0.77)	0.80 (0.78–0.82)
Body fat			
Total body (%)*	30 (29–31)	33 (32–35)	27 (26–28)
Android (%)*	29 (27–30)	33 (32–35)	30 (28–32)
Gynoid (%)*	36 (35–37)	39 (38–40)	31 (30–33)
Lean body mass			
Total body (%)*	70 (69–71)	67 (66–68)	73 (72–74)

BMI = body mass index; CHT = cross-sex hormonal treatment; GnRHa = gonadotropin-releasing hormone analogues; WHR = waist-to-hip ratio.
*Percentages of missing data for anthropometrics were 11% in transwomen and 18% in transmen for start of GnRHa, 10% in transwomen and 13% in transmen for start of CHT, and 71% in transwomen and 76% for visit at 22 years of age. For measurements of body composition examined by whole-body dual-energy x-ray absorptiometry, percentages of missing data were 12% in transwomen and 11% in transmen for start of GnRHa, 36% in transwomen and 45% in transmen for start of CHT, and 64% in transwomen and 65% in transmen at 22 years of age.

toward the values of the affirmed sex during treatment with GnRHa and CHT. As a result of these changes, in young adult transwomen at 22 years of age, SDS for WHR, body fat, and LBM showed greater similarity to ciswomen than to cismen. In transmen at the same age, SDS for WHR, body fat, and LBM were between reference values for ciswomen and cismen. The achieved body fat and LBM at 22 years were similar across different Tanner stages at start of treatment in transwomen and transmen. However, in transmen, an earlier Tanner stage at start of treatment appeared to be associated with a closer resemblance of body shape to their affirmed sex at 22 years, and this tended to be the same in transwomen.

Compared with ciswomen and cismen in adolescence, transgender adolescents who are treated with GnRHa and CHT exhibit greater changes in body composition. A larger increase in percentage of body fat has been seen in transwomen compared with ciswomen within the same lifespan.[10] Transmen exhibit decreased percentage of body fat, whereas percentage of body fat in cismen remains stable during puberty.[10] This observation is likely explained by the fact that prepubertal girls already have more body fat than prepubertal boys, and therefore transgender persons have a different body composition at the start of hormonal treatment than age-matched peers of the affirmed sex.[13] Also, compared with adult transgender persons treated with CHT,[14,15] larger changes in body shape and body composition are seen in transgender persons who start in adolescence. Moreover, transgender persons who start treatment in adolescence have at 22 years of age a body composition that approaches the affirmed sex more closely than transgender persons who start CHT in adulthood.[15]

The appropriate moment of starting gender reassignment continues to be a topic of debate.[16] The findings of this study favor starting treatment in an early stage of puberty, because this appears to be associated with a closer resemblance of body shape to the affirmed sex at 22 years. This observation can be explained in part by the fact that in early puberty there is no distinct sex-specific body fat distribution, but also other factors can contribute to this difference. For instance, the period of gonadal suppression is generally considered a period of status quo. However, it can be postulated that a body of a transwomen

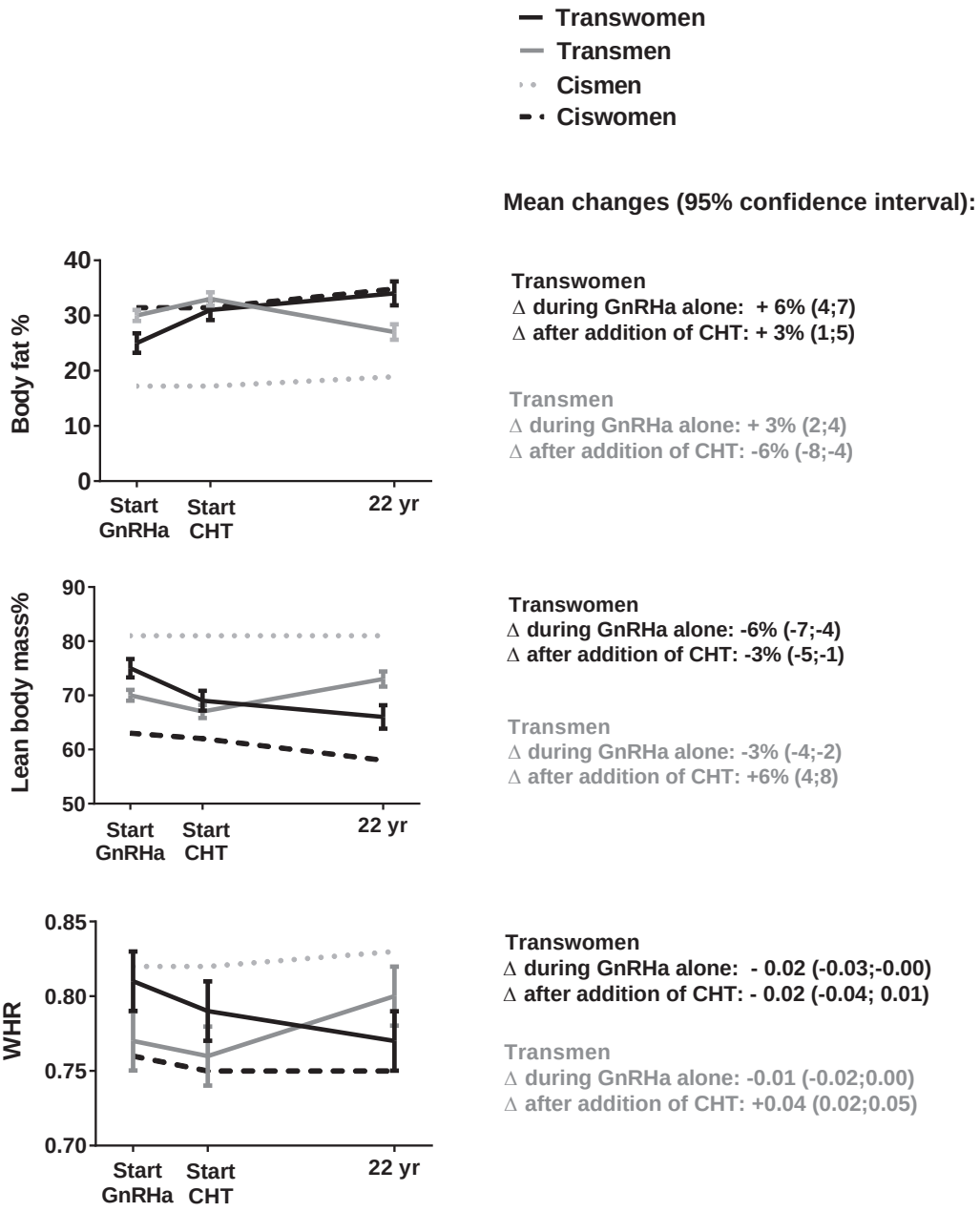


Figure 2. Changes (+95% CI) in percentage of total body fat, percentage of total body lean mass, and WHR in transwomen and transmen from the start of GnRHa to 22 years of age examined with mixed-model analyses using reference values for ciswomen and cismen. Reference data for ciswomen and cismen are shown for percentage of body fat,[10] percentage of lean body mass,[10] and WHR.[4] Mean differences between changes during GnRHa alone and changes after the addition of CHT in transwomen were +3 (95% CI −0 to 5) for percentage of total body fat, −3 (95% CI −5 to 0) for percentage of lean body mass, and 0.00 (95% CI −0.02 to 0.03) for WHR and in transmen were +9 (95% CI 7–11) for percentage of total body fat, −9 (95% CI −11 to −7) for percentage of lean body mass, and −0.04 (95% CI −0.06 to −0.02) for WHR. CHT = cross-sex hormonal treatment; GnRHa = gonadotrophin-releasing hormone analogues; WHR = waist-to-hip ratio.

in a prolonged androgen-deprived state might respond differently to estrogens than a body that has had greater androgen exposure. Although programming properties of androgens are most pronounced prenatally, during puberty programming effects have been described in the brain.[17] Therefore, other sex dimorphic organs such as body fat might also undergo androgen-induced programming that results in a lesser response to estrogens. To achieve a better outcome (ie, greater resemblance to the affirmed sex), both this phase of gonadal suppression can be modulated as the phase of gonadal suppression combined with CHT. Although in this study the period of CHT did not differ much between pubertal groups, previous

Table 3. SDS of body shape and body composition in trans persons at 22 years of age for ciswomen and cismen

	Transwomen (n = 71)		Transmen (n = 121)	
	SDS for ciswomen	SDS for cismen	SDS for ciswomen	SDS for cismen
BMI	+0.4 (−0.2 to 1.0)	+0.4 (−0.2 to 1.0)	+0.5 (0.3–0.8)	+0.6 (0.3–0.8)
Waist circumference	+1.1 (0.8–1.6)	−0.1 (−0.4 to 0.3)	+0.8 (0.5–1.1)	−0.1 (−0.4 to 0.3)
Hip circumference	+1.4 (1.0–1.9)	+1.8 (1.3–2.4)	+0.2 (−0.2 to 0.5)	+0.3 (−0.2 to 0.7)
WHR	+0.3 (0.0–0.5)	−1.0 (−1.3 to −0.7)	+0.6 (0.4–0.8)	−0.5 (−0.8 to −0.3)
Total body fat	−0.4 (−0.8 to 0.0)	+2.1 (2.1–2.3)	−1.1 (−1.4 to −0.9)	+1.8 (1.6–1.9)
Android fat	−0.1 (−0.3 to 0.3)	+0.8 (0.6–1.1)	−0.3 (−0.4 to 0.0)	+0.7 (0.5–0.9)
Gynoid fat	−0.1 (−0.4 to 0.3)	+2.1 (1.8–2.3)	−1.7 (−1.8 to −1.5)	+0.9 (0.8–1.1)
Lean body mass	+1.3 (0.8–1.8)	−1.4 (−1.5 to −1.2)	+3.0 (2.8–3.3)	−0.9 (−0.9 to −0.8)

BMI = body mass index; SDSs = SD scores; WHR = waist-to-hip ratio.

Table 4. Measurements of body shape and body composition at 22 years of age shown per category of Tanner stage at start of treatment and differences between those categories

	Tanner stages at start of treatment			Difference between Tanner stages	
Transwomen	Early puberty (n = 16)	Mid-puberty (n = 21)	Late puberty (n = 34)	ΔEarly vs mid-puberty	ΔEarly vs late puberty
Age at start (y)	13.2 (12.8–13.5)	13.3 (13.1–14.0)	15.6 (14.3–17.2)		
Duration of GnRHa monotherapy (y)	2.7 (2.4–3.1)	2.7 (2.1–3.1)	1.0 (0.6–1.6)		
Duration of GnRHa + CHT (y)	3.0 (2.6–3.6)	3.1 (2.6–3.7)	3.1 (2.3–3.5)		
BMI (kg/m^2)	23.8 (21.5–26.2)	20.9 (18.7–23.1)	24.2 (22.8–25.6)	2.9 (−0.3 to 6)	−0.4 (−3.1 to 2.3)
SDS for ciswomen	0.6 (−0.2 to 1.5)	−0.4 (−1.2 to 0.4)	2.1 (0.3–1.3)		
WHR	0.75 (0.71–0.79)	0.75 (0.69–0.80)	0.78 (0.76–0.80)	0.00 (−0.07 to 0.07)	−0.03 (−0.08 to 0.02)
SDS for ciswomen	0.0 (−0.5 to 0.5)	0.0 (−0.8 to 0.6)	0.4 (0.1–0.6)		
Total body fat (%)	33 (29–38)	30 (25–35)	34 (32–36)	3 (−4 to 10)	−1 (−6 to 4)
SDS for ciswomen	−0.2 (−0.8 to 0.4)	−0.6 (−1.4 to 0.0)	−0.1 (−0.4 to 0.1)		
Lean body mass (%)	67 (62–72)	70 (65–75)	66 (64–68)	−3 (−10 to 4)	+1 (−4 to 6)
SDS for ciswomen	1.5 (0.3–2.8)	2.3 (1.0–3.5)	1.3 (0.8–1.8)		

Transmen	Early to mid-puberty (n = 11)	Late puberty (n = 110)	ΔEarly to mid vs late puberty
Age at start (y)	12.3 ± 0.5	15.6 ± 1.7	
Duration of GnRHa monotherapy (y)	3.5 (3.4–3.9)	0.9 (0.5–2.0)	
Duration of GnRHa + CHT (y)	2.2 (2.0–2.5)	2.5 (1.9–3.2)	
BMI (kg/m^2)	22.5 (19.8–25.1)	23.9 (23.0–24.7)	−1.4 (−4.2 to 1.4)
SDS for cismen	0.1 (−0.9 to 1.1)	0.7 (0.3–1.0)	
WHR	0.85 (0.80–0.90)	0.79 (0.77–0.80)	0.06 (0.02 to 0.12)
SDS for ciswomen	0.3 (−0.5 to 1.2)	−0.7 (−1.0 to −0.5)	
Total body fat (%)	26 (21–31)	27 (26–29)	−1 (−6 to 4)
SDS for cismen	1.7 (1.3–2.0)	1.8 (1.7–1.9)	
Lean body mass (%)	74 (69–79)	73 (72–74)	1 (−4 to 6)
SDS for cismen	−0.8 (−1.1 to −0.4)	−0.9 (−0.9 to −0.8)	

BMI = body mass index; CHT = cross-sex hormonal treatment; GnRHa = gonadotropin-releasing hormone analogues; SDS = SD score; WHR = waist-to-hip ratio.

studies have shown that in adult transwomen[18]—but not in transmen[19]—a longer period of CHT establishes ongoing changes in body composition after the 1st year of treatment, which also could be applicable to body fat distribution and body shape. Therefore, one can postulate that starting CHT earlier than the current recommended age of 16[20] would improve outcomes in young adulthood.

We observed that transwomen at 22 years more closely resembled the affirmed sex regarding total body fat than transmen did. Body composition is the result of an interplay among multiple factors, including hormonal status, genetic predisposition, and lifestyle features such as diet and physical activity levels. A variance in these factors could explain this difference in treatment results between transwomen and transmen. Before the start of GnRHa, this difference was already present. Transmen had a percentage of total body fat very similar to that of ciswomen (SDS = −0.1), but transwomen had a percentage of body fat closer to that of ciswomen (SDS = −0.9) than to that of cismen (SDS = 1.6). The cause of the increased percentage of body fat and BMI in transwomen is unknown, but it can be postulated that psychological stress from gender dysphoria and an inactive lifestyle[21] could have contributed. Alternatively, the genetic predisposition, independent of hormonal treatment, of transmen could have resulted in less comparable values to their affirmed sex at 22 years than that of transwomen.[22] It can be debated whether the 46,XX karyotype[23] or other autosomal genes[24] are responsible.

As shown in Figure 2, the largest changes in transwomen were seen during GnRHa monotherapy (ie, during the suppression of testosterone). In transmen, the largest change was observed after the start of testosterone treatment. Therefore, it can be postulated that the suppression or addition of testosterone has a greater impact than the suppression or addition of estradiol. However, it is unclear whether the larger increase in body fat during GnRHa monotherapy in transwomen is due solely to the direct absence of testosterone action or the hypogonadal state itself. Indeed, also in transmen an increase in body fat was seen during GnRHa monotherapy. The larger change in body fat in transmen after the addition of testosterone could be due to the known lipolytic and anabolic effects of testosterone,[25] which in turn could be enhanced by practicing more sports or strength training in later adolescence.

This study is the 1st examining the effects of GnRHa and CHT on body shape and body composition in a large cohort of transgender adults who started treatment in their teens. A limitation of our study is the presence of missing data, which is seen more often in studies with a retrospective design when data are collected during regular patient care. However, our missing data were missing ad random and analyses between persons with missing data and persons without missing data did not indicate selection bias. Another limitation of our study is that (change in) body fat and LBM are dependent on other factors such as diet and physical activity and those were not systematically recorded.

This study provides insight into the effects of GnRHa and CHT on body shape and body composition during treatment in adolescence. Transgender adolescents starting treatment can be better informed on the extent to which their bodies will change and on what to expect of the results of such changes in young adulthood. Also, the finding that an early start of treatment in transmen results in a body shape that more closely resemble those of the affirmed sex than a late start supports early medical intervention.

The transgender persons included in this study were the first adolescents to be treated with the Dutch protocol.[1] Despite the favorable results on body shape and body composition, future research is warranted to improve current treatment protocols or devise alternative treatment protocols when GnRHa are not available.[26] Further, this study focused on body shape and body composition as important features of physical appearance, but it would also be of interest to determine the effects on body composition in the context of cardiovascular risk. For example, visceral fat seems to play an important role in the onset of insulin resistance and the metabolic syndrome,[27,28] and thus the study of changes in subcutaneous and visceral fat depots and their relation to cardiometabolic outcome could be relevant in the future.

CONCLUSION

Transwomen who started treatment with GnRHa and CHT in adolescence showed an increase in body fat with a decrease in LBM and WHR. At 22 years of age, their body shape and body composition were more similar to those of ciswomen than to those of cismen. In transmen, an increase in LBM and WHR was seen with a decrease in body fat during therapy, and at 22 years they had values between reference values for ciswomen and cismen. The achieved body fat and LBM at 22 were similar across different Tanner stages at start of treatment in transwomen and transmen. However, in transmen, start of treatment in early or mid-puberty resulted in a body shape more similar to that of the affirmed sex, and this tended to be the same in transwomen.

Corresponding Author: D. T. Klink, MD, PhD, Department of Pediatrics, Division of Endocrinology, VU Medical Center, De Boelelaan 1117, 1081 HZ Amsterdam, The Netherlands. Tel: 0031-(0)-20442550; Fax: 0031-(0)-204445254; E-mail: d.klink@vumc.nl

Conflicts of Interest: The authors report no conflicts of interest.

Funding: None.

STATEMENT OF AUTHORSHIP

Category 1

(a) Conception and Design
M. Klaver; R. de Mutsert; D. T. Klink

(b) Acquisition of Data
M. Klaver; C. M. Wiepjes; J. Rotteveel; D. T. Klink

(c) Analysis and Interpretation of Data
M. Klaver; R. de Mutsert; J. W. R. Twisk; M. den Heijer; D. T. Klink

Category 2

(a) Drafting the Article
M. Klaver; R. de Mutsert; D. T. Klink

(b) Revising It for Intellectual Content
M. Klaver; R. de Mutsert; C. M. Wiepjes; J. W. R. Twisk; M. den Heijer; J. Rotteveel; D. T. Klink

Category 3

(a) Final Approval of the Completed Article
M. Klaver; R. de Mutsert; C. M. Wiepjes; J. W. R. Twisk; M. den Heijer; J. Rotteveel; D. T. Klink

REFERENCES

1. Delemarre-van de Waal HA, Cohen-Kettenis PT. Clinical management of gender identity disorder in adolescents: a protocol on psychological and paediatric endocrinology aspects. **Eur J Endocrinol 2006;155:131-137.**
2. Wells JCK. Sexual dimorphism of body composition. **Best Pract Res Clin Endocrinol Metab 2007;21:415-430.**
3. Taylor RW, Grant AM, Williams SM, et al. Sex differences in regional body fat distribution from pre- to postpuberty. **Obesity 2010;18:1410-1416.**
4. Fredriks AM, van Buuren S, Fekkes M, et al. Are age references for waist circumference, hip circumference and waist-hip ratio in Dutch children useful in clinical practice? **Eur J Pediatr 2005;164:216-222.**
5. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4th ed, text rev. Washington, DC: American Psychiatric Association; 2000.
6. Twisk JWR, de Boer M, de Vente W, et al. Multiple imputation of missing values was not necessary before performing a longitudinal mixed-model analysis. **J Clin Epidemiol 2013; 66:1022-1028.**
7. Tanner JM, Whitehouse RH. Clinical longitudinal standards for height, weight, height velocity, weight velocity, and stages of puberty. **Arch Dis Child 1976;51:170-179.**
8. Prader A. Testicular size: assessment and clinical importance. **Triangle 1966;7:240-243.**
9. Fredriks AM, van Buuren S, Wit JM, et al. Body index measurements in 1996–7 compared with 1980. **Arch Dis Child 2000;82:107-112.**
10. van der Sluis IM, de Ridder MAJ, Boot AM, et al. Reference data for bone density and body composition measured with dual energy x ray absorptiometry in white children and young adults. **Arch Dis Child 2002;87:341-347.**
11. Stults-Kolehmainen MA, Stanforth PR, Bartholomew JB. Fat in android, trunk, and peripheral regions varies by ethnicity and race in college aged women. **Obesity 2012; 20:660-665.**
12. Stults-Kolehmainen MA, Stanforth PR, Bartholomew JB, et al. DXA estimates of fat in abdominal, trunk and hip regions varies by ethnicity in men. **Nutr Diabetes 2013;3:1-6.**
13. Taylor RW, Gold E, Manning P, et al. Gender differences in body fat content are present well before puberty. **Int J Obes 1997; 21:1082-1084.**
14. Klaver M, Dekker MJHJ, de Mutsert R, et al. Cross-sex hormone therapy in transgender persons affects total body weight, body fat and lean body mass: a meta-analysis. **Andrologia** https://doi.org/10.1111/and.12660. E-pub ahead of print.
15. Klaver M, Dekker MJHJ, Schreiner T, et al. Cross-sex hormone therapy and the effects on fat distribution in transgender persons. Presented at: 24th WPATH Symposium; Amsterdam, The Netherlands; 2016.
16. de Vries ALC, Klink D, Cohen-Kettenis PT. What the primary care pediatrician needs to know about gender incongruence and gender dysphoria in children and adolescents. **Pediatr Clin North Am 2016;63:1121-1135.**
17. Schulz KM, Sisk CL. The organizing actions of adolescent gonadal steroid hormones on brain and behavioral development. **Neurosci Biobehav Rev 2016;70:148-158.**
18. Mueller A, Zollver H, Kronawitter D, et al. Body composition and bone mineral density in male-to-female transsexuals during cross-sex hormone therapy using gonadotrophin-releasing hormone agonist. **Clin Endocrinol Diabetes 2011; 119:95-100.**
19. Mueller A, Haeberle L, Zollver H, et al. Effects of intramuscular testosterone undecanoate on body composition and bone mineral density in female-to-male transsexuals. **J Sex Med 2010;7:3190-3198.**
20. Hembree WC, Cohen-Kettenis P, Delemarre-van de Waal HA, et al. Endocrine treatment of transsexual persons: an Endocrine Society clinical practice guideline. **J Clin Endocrinol Metab 2009;94:3132-3154.**
21. Jones BA, Arcelus J, Bouman WP, et al. Barriers and facilitators of physical activity and sport participation among young transgender adults who are medically transitioning. **Int J Transgend 2017;18:227-238.**
22. Samaras K, Spector TD, Nguyen TV, et al. Independent genetic factors determine the amount and distribution of fat in women after the menopause. **J Clin Endocrinol Metab 1997;82:781-785.**
23. Arnold AP. The organizational-activational hypothesis as the foundation for a unified theory of sexual differentiation of all mammalian tissues. **Horm Behav 2009; 55:570-578.**

24. Chagnon YC, Rice T, Pérusse L, et al. Genomic scan for genes affecting body composition before and after training in Caucasians from HERITAGE. **J Appl Physiol 2001; 90:1777-1787.**

25. Geer EB, Shen W. Gender differences in insulin resistance, body composition, and energy balance. **Gend Med 2009; 6:60-75.**

26. Tack LJW, Heyse R, Craen M, et al. Consecutive cyproterone acetate and estradiol treatment in late-pubertal transgender female adolescents. **J Sex Med 2017;14:747-757.**

27. Després JP, Lemieux I. Abdominal obesity and metabolic syndrome. **Nature 2006;444:881-887.**

28. Tchernof A, Després JP. Pathophysiology of human visceral obesity: an update. **Physiol Rev 2013;93:359-404.**

Intended for healthcare professionals

Papers

Acute injuries in soccer, ice hockey, volleyball, basketball, judo, and karate: analysis of national registry data

BMJ 1995; 311 doi: https://doi.org/10.1136/bmj.311.7018.1465 (Published 02 December 1995) Cite this as: BMJ 1995;311:1465

Urho M Kujala, chief physician[a], Simo Taimela, research assistant[a], Ilkka Antti-Poika, consultant orthopaedic surgeon[a], Sakari Orava, consultant orthopaedic surgeon[a], Risto Tuominen, senior researcher[b], Pertti Myllynen, senior lecturer in orthopaedics and traumatology[c]

[a] *Unit for Sports and Exercise Medicine, Institute of Biomedicine, University of Helsinki, Toolo Sports Hall, Mannerheimintie 17, FIN-00250 Helsinki, Finland*

[b] *Department of Public Health, University of Helsinki, FIN-00290, Helsinki,*

[c] *Department of Orthopaedics and Traumatology, Helsinki University Central Hospital, FIN-00260 Helsinki*

Correspondence to: Dr Kujala.

Abstract

Objective: To determine the acute injury profile in each of six sports and compare the injury rates between the sports.

Design: Analysis of national sports injury insurance registry data.

Setting: Finland during 1987-91.

Subjects: 621691 person years of exposure among participants in soccer, ice hockey, volleyball, basketball, judo, or karate.

Main outcome measures: Acute sports injuries requiring medical treatment and reported to the insurance company on structured forms by the patients and their doctors.

Results: 54186 sports injuries were recorded. Injury rates were low in athletes aged under 15, while 20-24 year olds had the highest rates. Differences in injury rates between the sports were minor in this adult age group. Overall injury rates were higher in sports entailing more frequent and powerful body contact. Each sport had a specific injury profile. Fractures and dental injuries were most common in ice hockey and karate and least frequent in volleyball. Knee injuries were the most common cause of permanent disability.

Conclusions: Based on the defined injury profiles in the different sports it is recommended that sports specific preventive measures should be employed to decrease the number of violent contacts between athletes, including improved game rules supported by careful refereeing. To prevent dental injuries the wearing of mouth guards should be encouraged, especially in ice hockey, karate, and basketball.

Key messages

Key messages

Many sports injuries result from true accidents but others are preventable

Injury rates are low in child athletes and highest in young adults

Every sport has a specific injury profile

Preventive measures should be specific to the sport concerned and include those aimed at decreasing the number of violent contacts between athletes

Introduction

The growing popularity of sports and exercise is focusing attention on the injuries that may occur in addition to the health benefits.[1] [2] [3] [4] [5] [6] Treating sports injuries may be expensive, so preventive strategies and measures are required on economic as well as medical grounds.[7] [8] [9] Several epidemiological surveys have outlined the frequency and types of injuries in various sports, but study comparisons are complicated by the different injury criteria used as well as by inconsistency in data collection and recording.[10] The risk of acute injury varies enormously. Most endurance sports are extremely safe, whereas formula 1 car racing killed 69 of a small group of drivers between 1950 and 1994. Injury rates in popular team games such as soccer, volleyball, basketball, and ice hockey lie between these extremes.[11] Martial arts such as judo and karate are also becoming popular, and the associated risks may be greater than in most team games.[11] [12] Though endurance sports may cause the highest rates of stress injury, these rarely result in permanent disability.

Before embarking on a programme to prevent sports injuries we must first define the extent of the problem and identify the mechanisms and factors involved. Then we must introduce measures likely to reduce the risks and monitor their effects. Research shows that strategies to prevent sports injuries may be useful and that most interventions effective enough to measurably alter injury profiles in various sports entail changing rules or improving equipment.[5] [13] [14] In soccer, safety interventions and improved treatment of injuries and rehabilitation may prevent future injury.[15] [16]

We analysed the types and severity of acute injuries in some common team games (soccer, ice hockey, volleyball, basketball) as well as in judo and karate and compared the apparent injury risks between these sports. This information is crucial for prioritising measures in sports injury prevention.

Subjects and methods

From 1987 to 1991 anyone in Finland intending to compete in soccer, ice hockey, volleyball, basketball, judo, or karate was obliged to obtain a licence from the appropriate sports association. During the study period all licences issued to soccer and ice hockey players as well as those issued to judo and karate competitors were linked to an insurance policy from a single company (Pohjola Insurance Company Ltd) covering acute onset sports injuries. Among basketball players the insurance was not compulsory. For volleyball players the insurance was compulsory

from 1987 to 1990 but not during 1991. However, about two thirds of basketball and volleyball players had the insurance linked to their sports licence even when it was not compulsory. This study is therefore based on 621691 person years of exposure among athletes with a sports licence linked to insurance (see table 1). Exact data on age and sex of the insured athletes at the beginning of each person year of exposure were available for 1990 and 1991 in all the sports except basketball. Thus the analysis of injury rates by age and sex was limited to five sports and two years (23363 injuries during 250291 person years of exposure; see table 2).

Sport	Person years of exposure	No of Injury rate	Injury rate 95% confidence interval)+
Soccer	296646	26330	89 (88 to 90)
Ice hockey	179798	16836	94 (92 to 95)
Volleyball	87668	5235	60 (58 to 61)
Basketball	39541	3472	88 (85 to 91)
Judo	9936	1163	117 (111 to 123)
Karate	8102	1150	142 (134 to 150)

+Injuries per 1000 person years of exposure

Table 1

Person years of exposure, numbers of injuries,and injury rate in six sports in Finland (sports insurance data 1987-91)

	No of injuries/years of exposure [injury rate+ (95% confidence interval)]				
Sport	**Age <15**	**Age 15-19**	**Age 20-24**	**Age 25-34**	**Age >34**
Soccer	1591/63591 [25 (24 to 26)]	2474/26698 [93 (89 to 96)]	2837/11162 [254 (246 to 262)]	4344/22553 [193 (187 to 198)]	766/11946 [64 (60 to 69)]
Males	1273/56623 [23 (21 to 24)]	2065/21788 [95 (91 to 99)]	2602/9830 [265 (256 to 273)]	4183/21464 [195 (190 to 200)]	761/11768 [65 (60 to 69)]
Females	318/6968 [46 (41 to 51)]	409/4910 [83 (76 to 91)]	235/1332 [176 (156 to 197)]	161/1089 [148 (127 to 169)]	5/178 [28 (9 to 64)]

No of injuries/years of exposure [injury rate + (95% confidence interval)]					
Sport	**Age <15**	**Age 15-19**	**Age 20-24**	**Age 25-34**	**Age >34**
Ice hockey	1202/33393 [36 (34 to 38)]	2875/19688 [146 (141 to 151)]	1829/7048 [260 (249 to 270)]	1705/8576 [199 (190 to 207)]	235/6727 [35 (31 to 39)]
Males	1186/32958 [36 (34 to 38)]	2823/19027 [148 (143 to 153)]	1796/6759 [266 (255 to 276)]	1685/8262 [204 (195 to 213)]	231/6576 [35 (31 to 40)]
Females	16/435 [37 (21 to 59)]	52/661 [79 (59 to 102)]	33/289 [114 (78 to 151)]	20/314 [64 (39 to 97)]	4/151 [27 (7 to 67)]
Volleyball	94/8202 [12 (9 to 14)]	391/7688 [51 (46 to 56)]	593/2753 [215 (200 to 231)]	1154/7962 [145 (137 to 153)]	364/5171 [70 (63 to 77)]
Males	25/3838 [6 (4 to 10)]	185/3571 [52 (45 to 60)]	345/1461 [236 (214 to 258)]	815/5255 [155 (145 to 165)]	254/3816 [67 (59 to 75)]
Females	69/4364 [16 (12 to 20)]	206/4117 [50 (43 to 57)]	248/1292 [192 (170 to 213)]	339/2707 [125 (113 to 138)]	110/1355 [81 (67 to 97)]
Judo	34/1142 [30 (21 to 41)]	154/1524 [101 (86 to 116)]	128/601 [213 (180 to 246)]	108/684 [158 (131 to 185)]	25/289 [88 (58 to 128)]
Males	21/969 [22 (14 to 33)]	111/1228 [90 (75 to 108)]	95/436 [218 (179 to 257)]	81/496 [163 (131 to 196)]	22/260 [85 (54 to 125)]
Females	13/173 [75 (41 to 125)]	43/296 [145 (105 to 185)]	33/165 [200 (139 to 261)]	27/188 [144 (94 to 194)]	3/29 [103 (22 to 273)]
Karate	11/352 [31 (16 to 55)]	91/896 [102 (82 to 121)]	141/572 [247 (211 to 282)]	185/885 [209 (182 to 236)]	32/188 [170 (116 to 224)]
Males	9/304 [30 (14 to 56)]	63/664 [95 (74 to 120)]	123/417 [295 (251 to 339)]	148/710 [208 (179 to 238)]	27/158 [171 (112 to 230)]
Females	2/48 [42 (5 to 143)]	28/232 [121 (79 to 163)]	18/155 [116 (66 to 167)]	37/175 [211 (151 to 272)]	5/30 [167 (57 to 347)]

+Injuries per 1000 person years of exposure

Table 2

Injuries per person years of exposure and injury rates (plus 95% confidence intervals) by age and sex in soccer, ice hockey, volleyball, judo, and karate among competitive athletes in Finland in 1990 and 1991

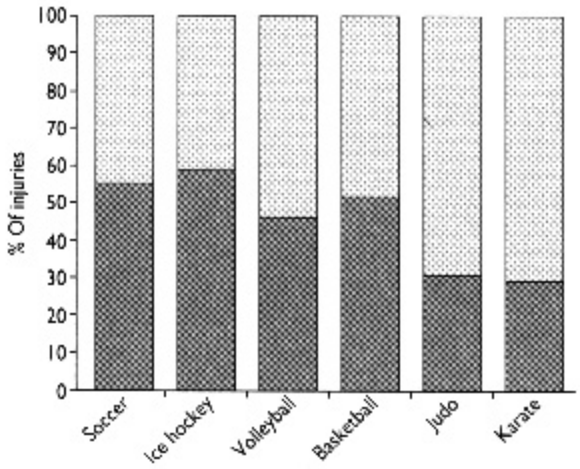
The injury criteria remained similar throughout. The sports insurance covered all traumatic acute injuries during competitions and training. The injury criteria also included all injuries of sudden onset, such as those that usually have no clear external accidental cause—for example, muscular strains.

The insurance company paid the medical costs of treatment after the injured athlete completed the injury report and the treating physician the medical accident report. Data on each injury, based on the two reports, were entered into a computer database by means of a structured format. Before paying the medical costs the insurance company checked the two reports for agreement. In cases of disagreement or incomplete information the insurance company sought clarification. This increased the validity of the data. The structured format of each injury report included age at the time of injury, type of sports event, circumstance of injury (training or competition), type of injury and mechanism, and injured body part. Data on payments made as death benefits or permanent disability benefits after sports injuries were also recorded. The insurance company and the sports associations consented to our using the data (without personal identification codes).

Statistical analyses—For each sport we calculated injury rates per 1000 person years of exposure (plus 95% confidence intervals) by age and sex as well as by types of injuries, anatomical locations of the injuries, and circumstances of the injuries.

Results

A total of 54186 acute sports injuries (48256 in males, 5930 in females) were recorded during the five years of the study. Karate and judo had the highest injury rates, followed by ice hockey, soccer, and basketball. Volleyball had the lowest injury rate (table 1). In the team games 46-59% of the injuries occurred during competitions, whereas in judo and karate around 70% occurred in training (figure). From the data for 1990 and 1991 the injury rates were clearly highest among 20-24 year old athletes (table 2). Sex differences in injury rates were less obvious, though among 20-24 year olds men had a higher injury rate than women in each sport.



Proportions of injuries occurring in competitions (dark areas) and training sessions (light areas) in different sports

Most injuries were to the lower limbs in soccer (66.8%), volleyball (57.4%), and basketball (56.0%), whereas upper limb injuries were most common in judo (37.6%). Sites other than limbs, including the teeth, were injured most often in karate and ice hockey (table 3). Sprains, strains, and bruises were the most common types of injury (table 4). Non-dental fractures accounted for 4.0-10.8% of injuries overall, occurring most often in karate, judo, and ice hockey and least often in volleyball (table 4). Dislocations were proportionally more frequent in judo and karate (table 4).

			%+ (No) [injury rate++ (95% confidence interval)]			
Injury location	Soccer	Ice hockey	Volleyball	Basketball	Judo	Karate
Lower limb	66.8 (17577) [59 (58 to 60)]	35.8 (6031) [34 (33 to 34)]	57.4 (3005) [34 (33 to 36)]	56.0 (1945) [49 (47 to 51)]	38.6 (449) [45 (41 to 49)]	37.3 (429) [53 (48 to 58)]
Thigh	10.6 (2803) [9 (9 to 10)]	6.1 (1031) [6 (5 to 6)]	2.3 (118) [1 (1 to 2)]	2.5 (87) [2 (2 to 3)]	1.9 (22) [2 (1 to 3)]	4.3 (49) [6 (4 to 8)]
Knee	21.5 (5657) [19 (19 to 20)]	17.2 (2888) [16 (16 to 17)]	19.0 (993) [11 (11 to 12)]	15.8 (547) [14 (13 to 15)]	20.2 (235) [24 (21 to 27)]	11.0 (127) [16 (13 to 18)]
Legsection	5.9 (1557) [5 (5 to 6)]	1.9 (325) [2 (2 to 2)]	1.8 (94) [1 (1 to 1)]	2.0 (68) [2 (1 to 2)]	2.0 (23) [2 (1 to 3)]	3.3 (38) [5 (3 to 6)]
Ankle	20.5 (5397) [18 (18 to 19)]	7.9 (1325) [7 (7 to 8)]	31.1 (1626) [19 (18 to 19)]	31.4 (1090) [28 (26 to 29)]	8.3 (97) [10 (8 to 12)]	7.7 (88) [11 (9 to 13)]
Foot	7.7 (2015) [7 (7 to 7)]	2.4 (404) [2 (2 to 2)]	2.9 (153) [2 (1 to 2)]	4.0 (138) [3 (3 to 4)]	6.0 (70) [7 (5 to 9)]	10.7 (123) [15 (13 to 18)]
Lower limb, other@	0.6 (148) [0 (0 to 1)]	0.3 (58) [0 (0 to 0)]	0.4 (21) [0 (0 to 0)]	0.4 (15) [0 (0 to 1)]	0.2 (2) [0 (0 to 0)]	0.3 (4) [0 (0 to 1)]
Upper limb	12.1 (3189) [11 (10 to 11)]	31.5 (5306) [30 (29 to 30)]	22.4 (1174) [13 (13 to 14)]	19.3 (670) [17 (16 to 18)]	37.6 (437) [44 (40 to 48)]	26.3 (303) [37 (33 to 42)]

			%+ (No) [injury rate++ (95% confidence interval)]			
Upper arm and shoulder	3.0 (800) [3 (3 to 3)]	10.2 (1724) [10 (9 to 10)]	9.3 (489) [6 (5 to 6)]	2.6 (92) [2 (2 to 3)]	20.0 (233) [24 (21 to 26)]	6.1 (70) [9 (7 to 11)]
Forearm and elbow	1.3 (334) [1 (1 to 1)]	4.6 (771) [4 (4 to 5)]	1.6 (84) [1 (1 to 1)]	1.3 (46) [1 (1 to 2)]	7.7 (90) [9 (7 to 11)]	3.8 (44) [5 (4 to 7)]
Palm and wrist	3.6 (948) [3 (3 to 3)]	8.2 (1382) [8 (7 to 8)]	2.1 (111) [1 (1 to 2)]	3.8 (131) [3 (3 to 4)]	3.4 (40) [4 (3 to 5)]	6.0 (69) [9 (7 to 11)]
Fingers	3.7 (978) [3 (3 to 4)]	6.4 (1074) [6 (6 to 6)]	9.1 (477) [5 (5 to 6)]	11.1 (386) [10 (9 to 11)]	4.5 (52) [5 (4 to 7)]	9.3 (107) [13 (11 to 16)]
Upper limb, other@	0.5 (129) [0 (0 to 1)]	2.1 (355) [2 (2 to 2)]	0.2 (13) [0 (0 to 0)]	0.4 (15) [0 (0 to 1)]	1.9 (22) [2 (1 to 3)]	1.1 (13) [2 (1 to 2)]
Other sites	21.1 (5564) [19 (18 to 19)]	32.7 (5499) [31 (30 to 31)]	20.2 (1056) [12 (11 to 13)]	24.7 (857) [22 (20 to 23)]	23.8 (277) [28 (25 to 31)]	36.3 (418) [52 (47 to 56)]
Teeth	2.8 (737) [2 (2 to 3)]	7.1 (1196) [7 (6 to 7)]	2.0 (106) [1 (1 to 1)]	5.2 (182) [5 (4 to 5)]	2.7 (31) [3 (2 to 4)]	6.4 (74) [9 (7 to 11)]
Eye	0.9 (245) [0 (0 to 1)]	1.2 (202) [1 (1 to 2)]	1.9 (101) [1 (1 to 1)]	3.0 (103) [3 (2 to 3)]	0.9 (11) [1 (0 to 2)]	2.1 (24) [3 (2 to 4)]
Head and neck, other sites	4.9 (1282) [4 (4 to 5)]	8.6 (1448) [8 (8 to 8)]	5.6 (291) [3 (3 to 4)]	7.4 (256) [6 (6 to 7)]	6.3 (73) [7 (6 to 9)]	10.9 (125) [15 (13 to 18)]
Thorax and abdomen	2.6 (693) [2 (2 to 3)]	3.8 (632) [4 (3 to 4)]	0.9 (47) [1 (0 to 1)]	1.5 (52) [1 (1 to 2)]	4.2 (49) [5 (4 to 6)]	4.7 (54) [7 (5 to 8)]
Back	5.8 (1531) [5 (5 to 5)]	7.2 (1220) [7 (6 to 7)]	8.7 (458) [5 (5 to 6)]	5.4 (187) [5 (4 to 5)]	7.9 (92) [9 (7 to 11)]	9.7 (111) [14 (11 to 16)]

			%+ (No) [injury rate++ (95% confidence interval)]			
Pelvis and hip	2.5 (665) [2 (2 to 2)]	2.7 (447) [2 (2 to 3)]	0.5 (28) [0 (0 to 0)]	0.9 (3) [1 (0 to 1)]	0.7 (8) [1 (0 to 1)]	1.5 (17) [2 (1 to 3)]
Multiple sites or unknown	1.6 (411) [1 (1 to 2)]	2.1 (354) [2 (2 to 2)]	0.5 (25) [0 (0 to 0)]	1.4 (47) [1 (1 to 2)]	1.1 (13) [1 (1 to 2)]	1.1 (13) [2 (1 to 2)]

- +Percentage of injuries in sports studied
- (sectionLeg refers to anatomical area between knee and ankle.
- ++Injuries per 1000 person years of exposure.
- @Other or multiple sites: this category of upper limb includes clavicle injuries.

TABLE III

Anatomical locations of injuries by sports and injury rates (plus 95% confidence intervals) (sports insurance data 1987-91)

			%+ (No) [injury rate++ (95% confidence interval)]			
Injury type	Soccer	Ice hockey	Volleyball	Basketball	Judo	Karate
Sprains and strainssection	56.1 (14769) [50 (49 to 51)] 36.9 (6213) [35 (34 to 35)]	74.6 (3905) [45 (43 to 46)]	61.3 (2128) [54 (52 to 56)]	59.8 (696) [70 (65 to 75)]	44.7 (514) [63 (58 to 69)]	
Knee	16.4 (4321) [15 (14 to 15)] 11.2 (1886) [11 (10 to 11)]	15.4 (807) [9 (9 to 10)]	12.4 (432) [11 (10 to 12)]	15.3 (178) [18 (15 to 21)]	9.0 (104) [13 (10 to 15)]	

			%+ (No) [injury rate++ (95% confidence interval)]			
Ankle	16.7 (4401) [15 (14 to 15)]	4.8 (812) [5 (4 to 5)]	29.2 (1527) [17 (17 to 18)]	29.5 (1023) [26 (24 to 27)]	6.5 (76) [8 (6 to 9)]	6.3 (72) [9 (7 to 11)]
Bruises and wounds	31.1 (8184) [28 (27 to 28)] 42.0 (7069) [39 (38 to 40)]	16.3 (853) [10 (9 to 10)]	22.2 (773) [20 (18 to 21)]	23.1 (269) [27 (24 to 30)]	35.1 (404) [50 (45 to 55)]	
Fractures	9.6 (2515) [8 (8 to 9)]	17.1 (2885) [16 (16 to 17)]	5.9 (311) [4 (3 to 4)]	12.6 (436) [11 (10 to 12)]	11.3 (131) [13 (11 to 15)]	16.9 (194) [24 (21 to 27)]
Fractures other than dental	7.0 (1844) [6 (6 to 7)]	10.6 (1785) [10 (9 to 10)]	4.0 (212) [2 (2 to 3)]	7.6 (265) [7 (6 to 8)]	8.8 (102) [10 (8 to 12)]	10.8 (124) [15 (13 to 18)]
Foot and ankle@	25.0 (461) [2 (1 to 2)]	11.3 (201) [1 (1 to 1)]	29.2 (62) [1 (1 to 1)]	18.5 (49) [1 (1 to 2)]	34.3 (35) [4 (2 to 5)]	21.0 (26) [3 (2 to 4)]
Lower limb, proximal from ankle@	16.1 (297) [1 (1 to 1)]	6.8 (122) [1 (1 to 1)]	7.5 (16) [0 (0 to 0)]	3.8 (10) [0 (0 to 0)]	3.9 (4) [0 (0 to 1)]	4.0 (5) [1 (0 to 1)]
Fingers, palm, and wrist@	35.5 (654) [2 (2 to 2)]	48.1 (858) [5 (4 to 5)]	53.8 (114) [1 (1 to 2)]	57.0 (151) [4 (3 to 4)]	18.6 (19) [2 (1 to 3)]	36.3 (45) [6 (4 to 7)]
Upper limb proximal from wrist@	8.4 (154) [1 (0 to 1)]	21.6 (386) [2 (2 to 2)]	0.9 (2) [0 (0 to 0)]	4.2 (11) [0 (0 to 0)]	31.4 (32) [3 (2 to 3)]	5.6 (7) [1 (0 to 2)]

			%+ (No) [injury rate++ (95% confidence interval)]			
Other, non-dental@	15.1 (278) [1 (1 to 1)]	12.2 (218) [1 (1 to 1)]	8.5 (18) [0 (0 to 0)]	16.6 (44) [1 (1 to 1)]	11.8 (12) [1 (1 to 2)]	33.1 (41) [5 (4 to 7)]
Dental fractures	2.5 (671) [2 (2 to 2)]	6.5 (1100) [6 (6 to 6)]	1.9 (99) [1 (1 to 1)]	4.9 (17) [4 (4 to 5)]	2.5 (29) [3 (2 to 4)]	6.1 (70) [9 (6 to 11)]
Dislocations	0.8 (215) [1 (1 to 1)]	1.1 (191) [1 (1 to 1)]	1.4 (73) [1 (1 to 1)]	1.7 (59) [1 (1 to 2)]	3.9 (45) [5 (3 to 6)]	1.1 (13) [2 (1 to 2)]
Knee||	0.3 (76) [0 (0 to 0)]	0.2 (39) [0 (0 to 0)]	0.3 (17) [0 (0 to 0)]	0.5 (16) [0 (0 to 1)]	1.0 (12) [1 (1 to 2)]	0.1 (1) [0 (0 to 0)]
Shoulder and elbow	0.2 (52) [0 (0 to 0)]	0.7 (113) [1 (1 to 1)]	0.2 (11) [0 (0 to 0)]	0.3 (10) [0 (0 to 0)]	1.9 (22) [2 (1 to 3)]	0.3 (3) [0 (0 to 1)]
Fingers	0.1 (31) [0 (0 to 0)]	0.1 (13) [0 (0 to 0)]	0.5 (25) [0 (0 to 0)]	0.3 (12) [0 (0 to 0)]	0.3 (4) [0 (0 to 1)]	0.3 (3) [0 (0 to 1)]
Others or unknown	2.5 (647) [2 (2 to 2)]	2.8 (478) [3 (2 to 3)]	1.8 (93) [1 (1 to 1)]	2.2 (76) [2 (1 to 2)]	1.9 (22) [2 (1 to 3)]	2.2 (25) [3 (2 to 4)]

- +Percentage of injuries in sports studied
- ++Injuries per 1000 person years of exposure.
- sectionIncludes ligament, tendon, and muscle ruptures.
- @Percentages of non-dental fractures.

TABLE IV

Injury types by sports and injury rates (plus 95% confidence intervals) (sports insurance data 1987-91)

No death benefit for an accidental sports injury was awarded during the study. There was one neck fracture in an ice hockey player leading to tetraplegia. Benefit in respect of various degrees of permanent disability (that is, at least 5% disability) was awarded in 102 cases. Fifty nine of these occurred in soccer (0.22% of all soccer injuries), 24 in ice hockey (0.14%), 11 in volleyball (0.21%), four in basketball (0.11%), two in judo (0.17%), and two in karate (0.17%); 92 occurred in males and 68 during competitions. The most common injury was a sprain or strain (66 cases), while 16 injuries were fractures. The knee was the most common location for injuries resulting in permanent disability (64 cases).

Discussion

We have defined the acute injury profiles in six sports on the basis of 54186 injuries examined by physicians and reported to a national sports insurance company. However, not all treated injuries are reported to the insurance company and many minor injuries that are self treated also go unreported. Thus our data underestimate the true injury rates in each sport.

The overall sex difference in injury risk was small but the age difference was clear. Athletes aged 20-24 years had the highest risk, probably because training and competition are most intense at this age. We did not have records on exact hours of exposure and so could not calculate the exact injury risk per hour of training or competition. Our findings agree with earlier reports that injuries in young team players are less frequent than in adults.[17] [18] [19] In judo the reason for the unexpectedly high injury rate among young girls was probably that as a minority group in many clubs they often have to train with boys and men. In adult team games entailing various types of bodily contact between athletes men probably train more but tend to have a rougher style than women. This also may partly explain the sex difference in injury risk.

Athletes usually spend far more time training than competing. About half of the injuries to team game athletes occurred in competitions. Hence competitions plainly entail a higher risk of injury per hour than training.

When the analysis of injury rates was restricted to 20-24 year old athletes only small differences were found between sports. The overall injury risk was lowest in volleyball and highest in ice hockey, judo, and karate. Our findings agree with other reports that violent bodily contact between athletes increases the risk of injury[20] [21] but that use of protective equipment may reduce the difference in injury outcomes between sports. Comparison of injury rates with those in other studies is complicated by methodological differences. The ranking of injury risks in different sports may also vary with local circumstances in the study area, such as the age distribution and level of the teams playing. De Loes and Goldie reported that soccer players had a clearly higher injury risk than ice hockey players in one municipality in south west Sweden.[11] But our finding that the overall injury rate was higher in ice hockey players than footballers in Finland agrees with national data from Sweden.[22] In our study the differences in overall injury rates between the sports were partly explained by the differing age distributions of the athletes.

All the sports studied had higher acute injury rates than reported among endurance athletes.[23] During the study period participants in motor sports had similar compulsory national sports licence insurance. In 1990 and 1991 among male participants in various competitive motor sports the injury rate per 1000 person years of exposure was 182 (95% confidence interval 171 to 194) and was highest in the youngest age groups—278 (223 to 333) among participants aged under 15 and 245 (220 to 269) among 15-19 year olds. From 1987 to 1991 fractures other than dental accounted for 29% of all injuries to motor sports participants. These data confirm that the relative safety of junior sports is not extended to all types of motor sports.

The types and anatomical distribution of injuries, as well as the rarity of severe injuries, corresponded with earlier findings.[24] As expected, the most common injuries were sprains, strains, and bruises. As found in earlier studies,

knee injury was the most common cause of permanent disability,**9** defined simply as an impairment of optimal function. Fractures seldom resulted in permanent disability, though the number of fractures (highest in ice hockey, judo, and karate) highlight the risk for high energy injuries.

High puck velocities, aggressive stick use, and body checking (collisions) account for most ice hockey injuries.**25** Catastrophic ice hockey injuries seem to be less frequent in Finland than North America,**26** possibly because of the larger rinks and less aggressive style in Europe. To avoid these injuries as far as possible, aggressive checking—particularly from behind the player and near the rink boards—should be minimised by game rules and strict refereeing.**25** Aggressive stick use may partly account for the high number of hand and wrist fractures in our study. Though facial injuries are common, they have declined with the more routine use of helmets and facemasks.**25** In ice hockey and many other sports mouth guards would substantially reduce dental injuries and should be designed according to the characteristics of each sport.

The injury profiles of the sports differed widely. To avoid injuries preventive measures should be specific to each sport. In general there should be greater focus on diminishing rough and violent contact between athletes.

Footnotes

- Funding Finnish Ministry of Education.
- Conflict of interest None.

References

1. Powell KE, Thompson PD, Caspersen CJ, Kendrick JS.Physical activity and the incidence of coronary heart disease.*Annu Rev Public Health*1987;**8**:253–87.
2. Helmrich SP, Ragland DR, Leung RW, Paffenbarger RS.Physical activity and reduced occurrence of non-insulin-dependent diabetes mellitus.*N Engl J Med*1991;**325**:147–52.
3. Sarna S, Sahi T, Koskenvuo M, Kaprio J.Increased life expectancy of world class male athletes.*Med Sci Sports Exerc*1993;**25**:237–44.
4. Fentem PH.Benefits of exercise in health and disease.*BMJ*1994;**308**:1291–5.
5. Torg JS, Vegso JJ, Sennelt B, Das M.The national football head and neck injury registry.*JAMA*1985;**254**:3439–43.
6. Kujala UM, Kaprio J, Sarna S.Osteoarthritis of the weightbearing joints of the lower limbs in former elite male athletes.*BMJ*1994;**308**:231–4.
7. De Loes M.Medical treatment and costs of sports-related injuries in total population.*Int J Sports Med*1990;**11**:66–72.
8. Sandelin J, Santavirta S, Lattila R, Vuolle P, Sarna S.Sport injuries in a large urban population: occurrence and epidemiologic aspects.*Int J Sports Med*1987;**8**:61–6.
9. Inklaar H.Soccer injuries. I: incidence and severity.*Sports Med*1994;**18**:55–73.
10. Walter SD, Sutton JR, McIntosh JM, Connolly C.The aetiology of sport injuries. A review of methodologies.*Sports Med*1985;**2**:47–58.
11. De Loes M, Goldie I.Incidence rate of injuries during sport activity and physical exercise in a rural Swedish municipality: incidence rates in 17 sports.*Int J Sports Med*1988;**9**:461–7.
12. McLatchie GR, Davies JE, Caulley JH.Injuries in karate–a case for medical control.*J Trauma*1980;**20**:956–8.
13. Johnson RJ, Ettlinger CF.Alpine ski injuries: changes through the years.*Clin Sports Med*1982;**1**:181–97.
14. Sim FH, Simonet WT, Melton LJ, Lehn TA.Ice hockey injuries.*Am J Sports Med*1987;**15**:86–96.

15. Ekstrand J, Gillquist J.The avoidability of soccer injuries.*Int J Sports Med*1983;**4**:124–8.

16. Ekstrand J, Gillquist J, Liljedahl SO.Prevention of soccer injuries. Supervision by doctor and physiotherapist.*Am J Sports Med*1983;**11**:116–20.

17. Hayes D.An injury profile for hockey.*Canadian Journal of Applied Sports Science*1978;**3**:61–4.

18. Nilsson S, Rooas A.Soccer injuries in adolescents.*Am J Sports Med*1978;**6**:358–61.

19. Baxter-Jones A, Maffulli N, Helms P.Low injury rates in elite athletes.*Arch Dis Child*1993;**68**:130–2.

20. Backx FJG, Beijer HJM, Bol E, Erich WBM.Injuries in high-risk persons and high-risk sports.*Am J Sports Med*1991;**19**:124–30.

21. Watson AWS.Incidence and nature of sports injuries in Ireland. Analysis of four types of sport.*Am J Sports Med*1993;**21**:137–43.

22. Folksam.*Sports injuries 1976–1983*. Uddevalla, Sweden: Bohuslaningens Boktryckeri AB,1985. (144 pages.)

23. Kujala UM, Nylund T, Taimela S.Acute injuries in orienteers.*Int J Sports Med*1995;**16**:122–5.

24. Kujala UM, Heinonen OJ, Lehto M, Jarvinen M, Bergfeld JA.Equipment, drugs and problems of the competition and team physician.*Sports Med*1988;**6**:197–209.

25. Daly PJ, Sim FH, Simonet WT.Ice hockey injuries. A review.*Sports Med*1990;**10**:122–31.

26. Tator CH, Edmonds VE, Lapczak L, Tator IB.Spinal injuries in ice hockey players, 1966–1987.*Can J Surg*1991;**34**:63–9.

Reaction Times of Preschool Children on the Ruler Drop Test: A Cross-Sectional Study With Reference Values

Article *in* Perceptual and Motor Skills · July 2018

DOI: 10.1177/0031512518789563

CITATIONS
3

READS
2,886

4 authors, including:

Pedro Ángel Latorre Román
Universidad de Jaén
267 PUBLICATIONS **2,296** CITATIONS
SEE PROFILE

Felipe García Pinillos
University of Granada
180 PUBLICATIONS **1,771** CITATIONS
SEE PROFILE

Jesús Salas Sánchez
Universidad Internacional de La Rioja
39 PUBLICATIONS **150** CITATIONS
SEE PROFILE

Article

Perceptual and Motor Skills
0(0) 1–13

DOI: 10.1177/0031512518789563
journals.sagepub.com/home/pms

Reaction Times of Preschool Children on the Ruler Drop Test: A Cross-Sectional Study With Reference Values

Pedro Ángel Latorre-Roman[1], Alejandro Robles-Fuentes[1], Felipe García-Pinillos[2], and Jesús Salas-Sánchez[3]

Abstract

Reaction time (RT) tasks assess several brain functions, and a slow RT can be due to various brain diseases, disorders, and acquired conditions. This study examined age and gender differences in RTs of Spanish preschool children on the ruler drop test (RDT) and presents norm-referenced results. Participants were 3,741 children (1,845 girls and 1,896 boys; mean [M] age = 55.93, standard deviation [*SD*] = 11.14 months; M body mass index = 15.94, *SD* = 1.91 kg/m^2), selected from 51 schools in southern Spain. We measured RT with the RDT, and we collected both right- and left-hand data. We expressed normative mean RDT values of both hands according to gender and age in percentiles. Based on mean RDT scores, girls exhibited a poorer performance than boys aged 4 years ($p = .032$, Cohen's $d = -0.122$) and 5 years ($p = .001$, Cohen's $d = -0.194$). For the whole group, RDT performance was faster with increased age, from the age of five years.

Keywords

early childhood, fitness, health, reaction time

[1]Department of Didactics of Corporal Expression, University of Jaén, Spain
[2]Department of Physical Education, Sport and Recreation, Universidad de La Fontera, Temuco, Chile
[3]Universidad Autónoma de Chile, Chile

Corresponding Author:
Jesús Salas-Sánchez, Universidad Autónoma de Chile, Providencia, Chile.
Email: jesussalas644@gmail.com

Introduction

The preschool age is characterized by significant changes in the acquisition of fundamental motor skills and nervous system maturation (Tanaka, Hikihara, Ohkawara, & Tanaka, 2012). The development of motor competence during infancy and childhood is influenced by the individual child's growth and morphological, physiological, and neuromuscular characteristics (Venetsanou & Kambas, 2009). The mastery of fundamental motor skills contributes to children's physical, cognitive, and social development and is essential for an active lifestyle (Lubans, Morgan, Cliff, Barnett, & Okely, 2010). Reaction time (RT), the speed of movement, and agility are some components of motor skills related to fitness (Moradi & Esmaeilzadeh, 2015). RT is the measure of time taken to respond, from the onset of a "Go" signal until a response is made.

Motor-cognitive response time is expressed as RT, which represents both the speed of information processing and the motor response of coordinated peripheral movements (Mishra, Dasgupta, Mohan, Aranha, & Samuel, 2018). RT can be assessed in young children to evaluate developmental changes and individual differences in sustained attention and organization of behavior (Weissberg, Ruff, & Lawson, 1990). Moreover, RT tasks assess several brain functions, such as different attentional processes (Stuss et al., 2005), interhemispheric transfer, age-related changes in cognition (Anstey et al., 2007), cognitive flexibility (Hillman et al., 2014), and motor and cognitive processing speed (Aranha, Moitra, et al., 2017). Simple RT measures correlate significantly with the measures of general intelligence (*g* factor) and are considered elementary measures of cognition (Woodley, Te Nijenhuis, & Murphy, 2014).

A slow RT can result from various brain diseases, disorders and acquired conditions that affect white matter conduction, gray matter neurotransmission, and efficiency of cognitive neural networks (Klotz, Johnson, Wu, Isaacs, & Gilbert, 2012). Therefore, RT deficits have been well demonstrated in children with attention deficit hyperactivity disorder (ADHD; Sjöwall, Roth, Lindqvist, & Thorell, 2013), developmental coordination disorder (DCD; Gama, Ferracioli, Hiraga, & Pellegrini, 2016; Johnston, Burns, Brauer, & Richardson, 2002), dyslexia (Kaltner & Jansen, 2014), and autism spectrum disorder (ASD; Herrero & Crocetta, 2015). Poor RT has also been associated with childhood obesity and high body mass index (BMI) values (Gentier et al., 2013; Skurvydas et al., 2009). Conversely, a high level of moderate and vigorous physical activity has been associated with improved RT performance in children (Syväoja, Tammelin, Ahonen, Kankaanpää, & Kantomaa, 2014).

There are only a limited number of studies that have correlated physical and motor fitness with growth and maturity (Malina & Katzmarzyk, 2006), particularly on RT in preschool children, as most studies of RT have been conducted in children over six years of age (Aranha, Saxena, et al., 2017; Madsen et al., 2011; Manna, Pan, & Chowdhury, 2014; Tamnes, Fjell,

Westlye, Ostby, & Walhovd, 2012). Moreover, there is limited information about the reliability and validity of fitness and motor tests in preschool children (Ortega et al., 2015), though reliable measures of fitness and motor tests are necessary in order to investigate the relationship between physical fitness and health in this population (Latorre Román et al., 2015; Ortega et al., 2015).

RT is typically assessed using computerized neuropsychological testing software (Baisch, Cai, Li, & Pinheiro, 2017; Moradi & Esmaeilzadeh, 2017). However, high cost and requirements for professional management in estimating RT make this method inapplicable in a school setting. A simple, less expensive measure of RT that can be used to replace the computer assessment is the traditional ruler drop test (RDT) and, although the RDT measures RT plus movement time, it continues to be an acceptable means of measuring simple RT (Del Rossi, Malaguti, & Del Rossi, 2014). The RDT has acceptable reliability and criterion validity (Aranha, Sharma, Joshi, & Samuel, 2015; Eckner, Whitacre, Kirsch, & Richardson, 2009), and RDT reliability in preschool children has previously been reported. In prior analyses of reliability using test–retest, descriptive results (i.e., mean [M] and ± standard deviation [*SD*]) for pretest and retest were 38.43 ± 7.86 and 37.56 ± 9.75 cm ($p = .264$), respectively (Latorre Román et al., 2015). Moreover, prior researchers found an intraclass correlation coefficient (ICC) equal to 0.744 (95% confidence interval [0.836, 0.602]), the Bland–Altman graphic showed limits of agreement (2 *SD*) of 13.8 and −13.6 cm, and the mean of the differences was equal to 0.10 ± 6.87 cm (Latorre Román et al., 2015). In addition, the RDT has also been used in previous studies among school-aged children (Aranha, Saxena, et al., 2017; Fong, Ng, & Chung, 2013; Manna et al., 2014). In this regard, a previous study showed no significant differences between boys and girls in RDT for age groups 6-12 years; in addition, the RDT performance increased with age, and a significant change occurred between six and eight years of age (Aranha, Saxena, et al., 2017).

Despite being a significant indicator of function, behavior, and performance, RT has been infrequently employed in school settings to identify children with slowed motor cognitive processing. We assert that an early identification of delayed RT would help teachers and parents address children's functional deficits and quality of life (Aranha, Saxena, et al., 2017). Yet, to the best of our knowledge, there is no information available about reference values of RDT in preschool children. This study hypothesized no significant gender differences in RDT, but we expected variables, such as age and BMI, to influence a participant's RDT performance, with improved RDT performance with increasing age and worsening of RDT performance with increased BMI. Therefore, the main purpose of this study was to examine age and gender differences in RDT and determine and present reference RDT values for Spanish preschool children.

Method

Participants

A total of 3,741 children participated in this study (1,845 girls and 1,896 boys; M age = 55.93, $SD = 11.14$ months; M BMI = 15.94, $SD = 1.91$ kg/m^2); they were selected from 51 schools in southern Spain. This was a convenience sample selected from a large geographic area (both urban and rural) of Andalusia (Spain). Inclusion criteria were schooling in early childhood and being free from physical or intellectual disabilities. Parents of all child participants voluntarily signed an informed consent form prior to their child's participation in this study. The study was completed in accordance with the norms of the Declaration of Helsinki (2013 version) and was approved by the ethics committee of the University of Jaén (Spain).

Instrumentation. Body height (in cm) was measured with a stadiometer (Seca 222, Hamburg, Germany) and body mass (in kg) with a weighing scale (Seca 899, Hamburg, Germany). BMI was calculated by dividing body mass (kg) by height^2 (m). Waist circumference (WC) was established using a Seca Ergonomic Circumference Measuring Tape SE201 (Seca, Germany). To measure the RT, we used the RDT, which aims to measure the RT and eye-hand coordination. We used a 50- to 60-cm-long ruler and repeated the RDT three times with each hand, determining the average score of each hand for subsequent statistical analysis. The RT conversion is performed using the formula for a body in free fall under the influence of gravity ($d = ½\ gt^2$). The test score was the distance reached, with a lower distance indicating better performance. Regarding time conversion, the test score was the running time, with a longer time corresponding to a poorer performance.

Procedure. We used the following standardized testing procedure for RDT: The child was invited to sit on a chair with their hand kept in the mid-prone position, elbow flexed to 90°, and forearm supported on a table, with the open hand at the edge of the surface. The ruler was suspended vertically by the examiner, such that the 0-cm mark on the ruler coincided with the borders of the fingers (Figure 1). The ruler was then dropped between two fingers without prior intimation, and the subjects were asked to grasp it as quickly as possible. The order of testing of each hand was randomized. The research team conducted a demonstration. The children also performed some familiarization trials for the RDT. The children were encouraged to reach the best score possible. A week later, 84 children (included in the previous data collection) performed the same test (retest).

Data analysis. Data were analyzed using SPSS, v.19.0 for Windows (SPSS Inc., Chicago, IL) and the R statistical program (R Development Core Team, 2016)

Figure 1. Ruler drop test.

with the Generalized Additive Model for Location, Scale and Shape (GAMLSS) package (Rigby & Stasinopoulos, 2006). The statistical significance level was set at $p < .05$. Descriptive data were reported as Ms and *SD*s. Tests of normal distribution and homogeneity (Kolmogorov–Smirnov and Levene's tests) were conducted on all data before analysis. Differences between gender and age groups were analyzed using analysis of variance adjusted by the Bonferroni test. The magnitudes of the differences between values were also interpreted using Cohen's *d* effect size (Cohen, 1988). Effect sizes are reported as: trivial (<0.2), small (0.2–0.49), medium (0.5–0.79), and large (≥0.8) (Cohen, 1988). Pearson's correlation analysis was performed between the RDT and the anthropometric variables, adjusting for age and gender. A reliability pretest–posttest analysis was performed using intraclass correlation coefficients (ICCs). According to the classifications proposed by Shrout and Fleiss (1979), a very good correlation is represented by ICC > 90; a good correlation lies between $0.71 < \text{ICC} < 0.90$, a moderate ICC between $0.51 < \text{ICC} < 0.70$, and a poor ICC between $0.31 < \text{ICC} < 0.50$. The percentile curves were calculated as a function of age stratified by gender, using the mbda, mu, sigma, power exponential (LMSP) method, assuming a Box–Cox power exponential distribution, a generalized model of the lambda, mu, sigma (LMS) method. This approach has been implemented in the GAMLSS package in R software (Stasinopoulos & Rigby, 2007).

Results

Table 1 shows these participants' anthropometric characteristics and RDT performances, subgrouped by gender. Boys presented a higher mean BMI than girls, and girls exhibited significantly poorer RT scores in the RDT than did boys when comparing the mean score of both hands and the individual scores of each hand (right and left). Table 2 shows the mean RT performance by both age-group and gender. Girls displayed a significantly poorer performance from the age of 4 years ($p = .032$) to 5 years ($p = .001$) than did same aged boys. In relation to age, for the whole group, the RT performance increased with increasing age from 5 years. Pearson's correlation analysis indicated a weak negative correlation between the RT and age ($r = -.052$, $p = .002$), height ($r = -.111$, $p < .001$), and weight ($r = -.100$, $p < .001$). The 0.4th, 2nd, 10th,

Table 1. Anthropometric Characteristics and RDT Performance According to Sex.

	All ($n = 3,741$) Mean (*SD*)	Boys ($n = 1,896$) Mean (*SD*)	Girls ($n = 1,845$) Mean (*SD*)	*p*	Cohen's *d*
Age (years)	55.93 (11.14)	55.71 (11.11)	56.16 (11.16)	.211	−0.041
Body mass (kg)	19.39 (4.27)	19.63 (4.37)	19.14 (4.16)	<.001	0.114
Body height (cm)	109.18 (8.35)	109.48 (8.41)	108.87 (8.29)	.023	0.073
Body mass index (kg/m^2)	15.94 (1.91)	16.03 (1.93)	15.85 (1.89)	.003	0.099
WC (cm)	56.72 (7.90)	56.86 (7.68)	56.57 (8.11)	.243	0.037
RDT right hand (cm)	33.27 (11.89)	32.56 (11.77)	33.99 (11.97)	<.001	−0.083
RDT left hand (cm)	31.95 (11.19)	31.43 (11.01)	32.49 (11.36)	.004	−0.012
RDT average (cm)	32.43 (10.56)	31.86 (10.40)	33.02 (10.70)	.001	−0.053

Note. *SD* = standard deviation; WC = waist circumference; RDT = ruler drop test.

Table 2. RDT Average by Age Groups and Sex.

Age (years)	All		Boys		Girls		*p*	Cohen's *d*
	n	Mean (*SD*)	*n*	Mean (*SD*)	*n*	Mean (*SD*)		
3	931	32.56 (11.56)$_a$	483	32.43 (11.08)$_a$	448	32.70 (11.26)$_a$	.697	−0.024
4	1,175	33.30 (10.81)$_a$	604	32.66 (10.53)$_a$	571	33.98 (11.07)$_a$	.032	−0.122
5	1,302	32.25 (10.07)$_a$	659	31.29 (9.86)$_{a,b}$	643	33.24 (10.19)$_a$	.001	−0.194
6	329	29.73 (9.35)$_b$	148	29.26 (9.47)$_b$	181	30.12 (9.26)$_b$	.461	−0.092
p (intra group)		<.001		<.001		<.001		

Note. Values with different subscript letters indicate significant differences ($p < .05$) in post hoc analysis Bonferroni. The data are displayed in cm. *SD* = standard deviation; RDT = ruler drop test.

25th, 50th, 75th, 90th, 98th, and 99.6th percentile curves were computed for RDT, averaging the results of both hands, according to gender and age (Table 3, Figures 2 and 3).

Discussion

This study hypothesized no significant gender differences in RDT, but we expected variables, such as age and BMI, to influence a participant's RDT performance such that RDT performance would improve with increasing age and BMI. However, our results unexpectedly revealed that RDT performance differed between 4- and 5-year-old boys and girls, with girls exhibiting a poorer performance than boys. There were no significant gender differences between the genders for 3-year-olds. Few studies have focused on preschool children over this age range. Nonetheless, a previous study showed that males had a significantly faster mean RT than females across the life span (Dykiert, Der, Starr, & Deary, 2012). Surnina and Levedeva (2001) found that even in preschool children, gender dimorphism manifested itself in the reaction rate, with quick reactions occurring more often in boys than girls. Conversely,

Table 3. Percentiles Values of RDT Average in Boys and Girls.

Age (months)	Percentile	0.4	2	10	25	50	75	90	98	99.6
36	Boys	7.59	11.02	16.61	22.39	30.67	39.34	45.91	52.80	57.41
	Girls	8.17	11.56	17.05	22.51	30.04	38.41	45.76	54.81	61.66
42	Boys	9.59	12.86	18.24	23.82	31.89	40.68	47.75	55.64	61.17
	Girls	8.71	12.53	18.54	24.34	32.12	40.62	48.09	57.30	64.29
48	Boys	8.49	12.47	18.60	24.44	32.25	40.45	47.20	55.00	60.62
	Girls	8.94	13.06	19.39	25.32	33.03	41.34	48.64	57.69	64.57
54	Boys	8.06	12.63	19.31	25.21	32.47	40.06	46.79	55.20	61.65
	Girls	9.09	13.50	20.10	26.10	33.68	41.73	48.82	57.65	64.39
60	Boys	9.60	13.73	19.86	25.25	31.80	38.91	45.77	55.12	62.82
	Girls	9.11	13.75	20.51	26.46	33.79	41.46	48.26	56.75	63.26
66	Boys	11.29	14.47	19.42	24.11	30.31	37.36	44.17	53.45	61.15
	Girls	8.86	13.60	20.31	26.04	32.91	40.02	46.34	54.28	60.40
72	Boys	10.51	13.44	18.07	22.63	28.92	35.89	42.01	49.51	55.20
	Girls	8.36	13.03	19.47	24.82	31.05	37.42	43.12	50.34	55.91
78	Boys	7.95	11.15	16.07	20.86	27.38	34.06	39.19	44.70	48.45
	Girls	7.55	11.95	17.84	22.60	28.00	33.46	38.38	44.65	49.52

Note. The data are displayed in cm. RDT = ruler drop test.

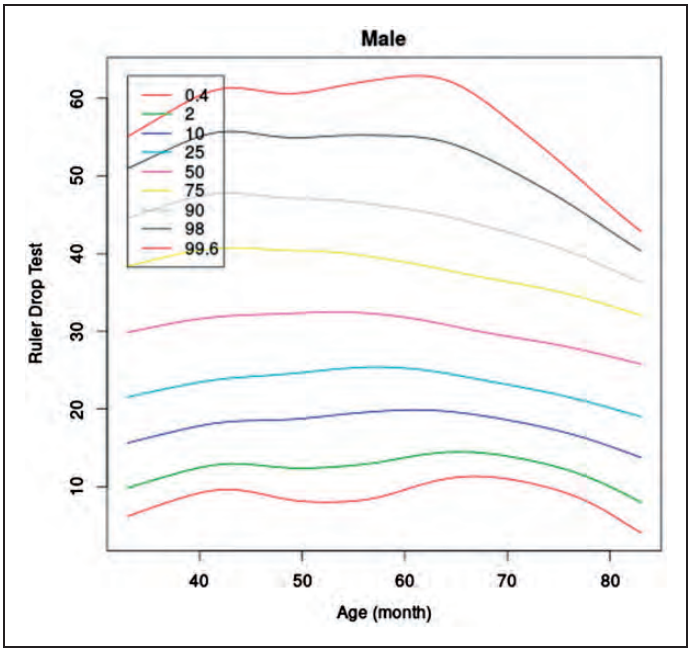


Figure 2. Percentile curves for RDT (cm) for boys.

Aranha, Saxena, et al (2017) found no significant gender differences in RDT in children aged 6-12 years, which can be explained by the fact that, in this study, there were no significant differences in BMI between genders.

Regarding age, the RT values decreased with increasing age, beginning only from five years of age. Our findings add to past research, showing that at age 5-6 years, RT decreased (Surnina & Lebedeva, 2001). Likewise, Bucsuházy and Semela (2017) noted significant differences between children 3-5 and 6-7 years old. In addition, Aranha, Saxena, et al., 2017 found that RDT values were similar in children 6-8 and 10-12 years old, while significant differences emerged from 8-10 years old. Therefore, these systems mature in childhood and, consequently, a shorter and less variable RT is part of typical development (Klotz et al., 2012). Consistent with Kiselev (2015), we assume that the age-related differences in processing speed can be understood in relation to the heterochronicity of child brain development and the specific mechanisms related to brain maturation.

In contrast, the influences of anthropometric characteristics, such as body mass, body height, BMI, and WC, on the RT in preschool children are less well understood. Several studies demonstrated that obesity and being overweight reduced RT performance (Gentier et al., 2013; Skurvydas

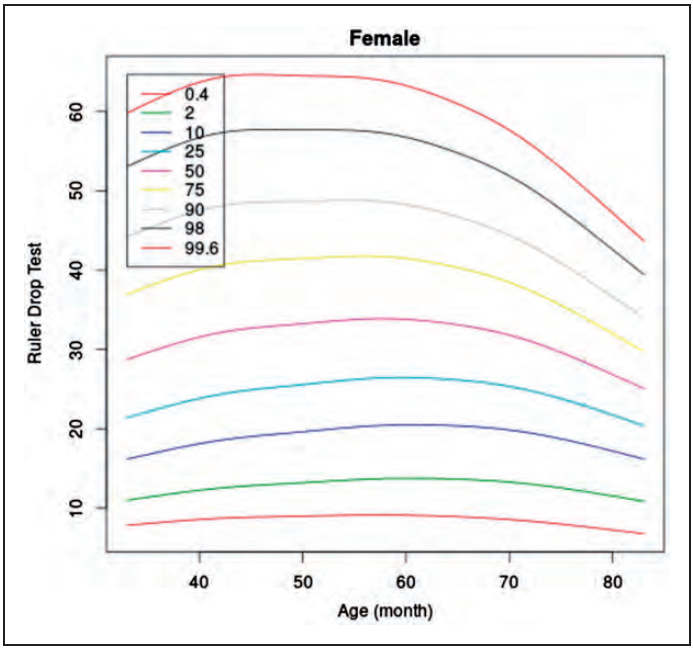


Figure 3. Percentile curves for RDT (cm) for girls.

et al., 2009). This study suggests that RDT in healthy preschool children aged 3-6 years are correlated with parameters of physical growth, such as body mass and body height, although Pearson's values were very low. Similar results were reported by Aranha, Saxena, et al. (2017) who found moderate correlations between RDT and body height ($r = -.33$) and body mass ($r = -.28$) in school children. Moreover, in this study, no correlations between BMI, WC, and RDT were noted, concurring with recent studies that investigated the relationship between RT and weight status in children (Aranha, Saxena, et al., 2017; Esmaeilzadeh, 2014; Moradi & Esmaeilzadeh, 2017). Finally, this study provides age- and gender-adjusted reference values for RDT, in Spanish preschool children.

The main limitation of this study was its cross-sectional design. Following children over time in longitudinal research would provide further needed data regarding time-related developmental changes as affected by other variables in individual children. However, a strength of this study was the large population sample we gathered. Of course, this study might also have been improved by the collection of additional relevant data that might affect RT development and manifestation in preschool children.

Conclusion

This study revealed unexpected gender differences (with boys superior at ages 4–5 years) and expected age differences in RT performance, and we provide reference values for the RDT that now allow future individual child comparisons on this simple RT task. Advantages of the RDT include that it requires no training and has a relatively short duration and that instrument requirements are simple and accessible (Aranha, Saxena, et al., 2017). The RDT may now be more frequently used for monitoring RT development in preschoolers, with percentile values of this study providing reference values for teachers and coaches working with children aged 3-6 years.

Acknowledgments

The authors gratefully thank all the schools of Andalusia who have collaborated in this study.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Human Subjects Approval Statement

The study was completed in accordance with the norms of the Declaration of Helsinki (2013 version) and was approved by the Ethics Committee of the University of Jaén (Spain).

References

Anstey, K. J., Mack, H. A., Christensen, H., Li, S. C., Reglade-Meslin, C., Maller, J., . . . Sachdev, P. (2007). Corpus callosum size, reaction time speed and variability in mild cognitive disorders and in a normative sample. *Neuropsychologia*, *45*(8), 1911–1920. doi:10.1016/j.neuropsychologia.2006.11.020

Aranha, V. P., Moitra, M., Saxena, S., Narkeesh, K., Arumugam, N., & Samuel, A. (2017). Motor cognitive processing speed estimation among the primary schoolchildren by deriving prediction formula: A cross-sectional study. *Journal of Neurosciences in Rural Practice*, *8*(1), 79–83.

Aranha, V. P., Saxena, S., Moitra, M., Narkeesh, K., Arumugam, N., & Samuel, A. J. (2017). Reaction time norms as measured by ruler drop method in school-going South Asian children: A cross-sectional study. *HOMO-Journal of Comparative Human Biology*, *68*(1), 63–68. doi:10.1016/j.jchb.2016.12.001

Aranha, V. P., Sharma, K., Joshi, R., & Samuel, A. J. (2015). Catch the moving ruler and estimate reaction time in children. *Indian Journal of Medical & Health Sciences*, *2*(1), 23–26. doi:10.21088/ijmhs.2347.9981.2115.4

Baisch, B., Cai, S., Li, Z., & Pinheiro, V. (2017). Reaction time of children with and without autistic spectrum disorders. *Open Journal of Medical Psychology*, *6*, 166–178. doi:10.4236/ojmp.2017.62014

Bucsuházy, K., & Semela, M. (2017). Case study: Reaction time of children according to age. *Procedia Engineering*, *187*, 408–413.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences. *Statistical Power Analysis for the Behavioral Sciences*. doi:10.1234/12345678

Del Rossi, G., Malaguti, A., & Del Rossi, S. (2014). Practice effects associated with repeated assessment of a clinical test of reaction time. *Journal of Athletic Training*, *49*(3), 356–359. doi:10.4085/1062-6059-49.2.04

Dykiert, D., Der, G., Starr, J. M., & Deary, I. J. (2012). Sex differences in reaction time mean and intraindividual variability across the life span. *Developmental Psychology*, *48*(5), 1262–1276. doi:10.1037/a0027550

Eckner, J. T., Whitacre, R. D., Kirsch, N. L., & Richardson, J. K. (2009). Evaluating a clinical measure of reaction time: An observational study. *Perceptual and Motor Skills*, *108*, 717–720. doi:10.2466/PMS.108.3.717-720

Esmaeilzadeh, S. (2014). Reaction time: Does it relate to weight status in children? *HOMO-Journal of Comparative Human Biology*, *65*(2), 171–178. doi:10.1016/j.jchb.2013.09.007

Fong, S. S. M., Ng, S. S. M., & Chung, L. M. Y. (2013). Health through martial arts training: Physical fitness and reaction time in adolescent Taekwondo practitioners. *Health*, *5*(6), 1–5. doi:10.4236/health.2013.56A3001

Gama, D. T., Ferracioli, M. D. C., Hiraga, C. Y., & Pellegrini, A. M. (2016). Value of pre-cue information for motor tasks performed by children with developmental coordination disorder (DCD). *Motriz: Revista de Educacao Fisica*, *22*(3), 138–143. doi:10.1590/S1980-6574201600030004

Gentier, I., Augustijn, M., Deforche, B., Tanghe, A., De Bourdeaudhuij, I., Lenoir, M., & D'Hondt, E. (2013). A comparative study of performance in simple and choice reaction time tasks between obese and healthy-weight children. *Research in Developmental Disabilities*, *34*(9), 2635–2641. doi:10.1016/j.ridd.2013.04.016

Herrero, D., & Crocetta, T. B. (2015). Total reaction time performance of individuals with autism after a virtual reality task. *International Journal of Neurorehabilitation*, *2*(5). doi:10.4172/2376-0281.1000189

Hillman, C. H., Pontifex, M. B., Castelli, D. M., Khan, N. A., Raine, L. B., Scudder, M. R., . . . Kamijo, K. (2014). Effects of the FITKids randomized controlled trial on executive control and brain function. *Pediatrics*, *134*(4), e1063–e1071. doi:10.1542/peds.2013-3219

Johnston, L. M., Burns, Y. R., Brauer, S. G., & Richardson, C. A. (2002). Differences in postural control and movement performance during goal directed reaching in children with developmental coordination disorder. *Human Movement Science*, *21*(5–6), 583–601. doi:10.1016/S0167-9457(02)00153-7

Kaltner, S., & Jansen, P. (2014). Mental rotation and motor performance in children with developmental dyslexia. *Research in Developmental Disabilities*, *35*(3), 741–754. doi:10.1016/j.ridd.2013.10.003

Kiselev, S. (2015). Age-related differences in processing speed in preschool children. *The Open Behavioral Science Journal*, *9*(1), 23–31.

Klotz, J. M., Johnson, M. D., Wu, S. W., Isaacs, K. M., & Gilbert, D. L. (2012). Relationship between reaction time variability and motor skill development in ADHD. *Child Neuropsychology*, *18*(6), 576–585. doi:10.1080/09297049.2011.625356

Latorre Román, P. Á., López, D. M., Sánchez, M. F., Sánchez, J. S., Coronas, F. M., & García-Pinillos, F. (2015). Test-retest reliability of a field-based physical fitness assessment for children aged 3-6 years. *Nutricion Hospitalaria*, *32*(4), 1683–1688. doi:10.3305/nh.2015.32.4.9486

Lubans, D. R., Morgan, P. J., Cliff, D. P., Barnett, L. M., & Okely, A. D. (2010). Fundamental movement skills in children and adolescents: Review of associated health benefits. *Sports Medicine (Auckland, N.Z.)*, *40*(12), 1019–35. doi:10.2165/11536850-000000000-00000

Madsen, K. S., Baaré, W. F. C., Skimminge, A., Vestergaard, M., Siebner, H. R., & Jernigan, T. L. (2011). Brain microstructural correlates of visuospatial choice reaction time in children. *NeuroImage*, *58*(4), 1090–1100. doi:10.1016/j.neuroimage.2011.07.032

Malina, R. M., & Katzmarzyk, P. T. (2006). Physical activity and fitness in an international growth standard for preadolescent and adolescent children. *Food and Nutrition Bulletin*, *27*(4 Suppl): S295–S313.

Manna, I., Pan, S. R., & Chowdhury, M. (2014). Anthropometric, physical, cardiorespiratory fitness and lipids and lipoproteins profile of young Indian children of 10-16 years age group. *American Journal of Sports Science and Medicine*, *2*(4), 154–160. doi:10.12691/ajssm-2-4-7

Mishra, R., Dasgupta, A., Mohan, V., Aranha, V. P., & Samuel, A. J. (2018). Increasing cardiopulmonary aerobic activity improves motor cognitive response time: An inference from preliminary one-group pretest-posttest quasi-experimental study. *Indian Heart Journal*, *70*(1), 128–129. doi:10.1016/J.IHJ.2017.11.010

Moradi, A., & Esmaeilzadeh, S. (2015). Association between reaction time, speed and agility in schoolboys. *Sport Sciences for Health*, *11*(3), 251–256. doi:10.1007/s11332-015-0230-4

Moradi, A., & Esmaeilzadeh, S. (2017). Simple reaction time and obesity in children: Whether there is a relationship? *Environmental Health and Preventive Medicine*, *22*(1), 2. doi:10.1186/s12199-017-0612-0

Ortega, F. B., Cadenas-Sánchez, C., Sánchez-Delgado, G., Mora-González, J., Martínez-Téllez, B., Artero, E. G., . . . Ruiz, J. R. (2015). Systematic review and proposal of a field-based physical fitness-test battery in preschool children: The PREFIT battery. *Sports Medicine*, *45*(4), 533–555.

R Development Core Team. (2016). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. doi:10.1038/sj.hdy.6800737

Rigby, R. A., & Stasinopoulos, D. M. (2006). Using the Box-Cox t distribution in GAMLSS to model skewness and kurtosis. *Statistical Modelling*, *6*(3), 209–229. doi:10.1191/1471082X06st122oa

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, *86*(2), 420–428. doi:10.1037/0033-2909.86.2.420

Sjöwall, D., Roth, L., Lindqvist, S., & Thorell, L. B. (2013). Multiple deficits in ADHD: Executive dysfunction, delay aversion, reaction time variability, and emotional deficits. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *54*(6), 619–627. doi:10.1111/jcpp.12006

Skurvydas, A., Gutnik, B., Zuoza, A. K., Nash, D., Zuoziene, I. J., & Mickeviciene, D. (2009). Relationship between simple reaction time and body mass index. *HOMO-Journal of Comparative Human Biology*, *60*(1), 77–85. doi:10.1016/j.jchb.2008.06.006

Stasinopoulos, D. M., & Rigby, R. A. (2007). Generalized additive models for location scale and shape (GAMLSS) in R. *Journal of Statistical Software*, *23*(7), 1–46.

Stuss, D. T., Alexander, M. P., Shallice, T., Picton, T. W., Binns, M. A., Macdonald, R., . . . Katz, D. I. (2005). Multiple frontal systems controlling response speed. *Neuropsychologia*, *43*(3), 396–417. doi:10.1016/j.neuropsychologia.2004.06.010

Surnina, O. E., & Lebedeva, E. V. (2001). Sex- and age-related differences in the time of reaction to moving object in children and adults. *Human Physiology*, *27*(4), 436–440. doi:10.1023/A:1010958602326

Syväoja, H. J., Tammelin, T. H., Ahonen, T., Kankaanpää, A., & Kantomaa, M. T. (2014). The associations of objectively measured physical activity and sedentary time with cognitive functions in school-aged children. *PLoS One*, *9*(7), e103559. doi:10.1371/journal.pone.0103559

Tamnes, C. K., Fjell, A. M., Westlye, L. T., Ostby, Y., & Walhovd, K. B. (2012). Becoming consistent: Developmental reductions in intraindividual variability in reaction time are related to white matter integrity. *Journal of Neuroscience*, *32*(3), 972–982. doi:10.1523/JNEUROSCI.4779-11.2012

Tanaka, C., Hikihara, Y., Ohkawara, K., & Tanaka, S. (2012). Locomotive and non-locomotive activity as determined by triaxial accelerometry and physical fitness in Japanese preschool children. *Pediatric Exercise Science*, *24*, 420–434.

Thomas, J., Silverman, S., & Nelson, J. (2015). *Research methods in physical activity, 7E*. Champaign, IL: Human Kinetics.

Venetsanou, F., & Kambas, A. (2009). Environmental factors affecting preschoolers' motor development. *Early Childhood Education Journal*, *37*(4), 319–327. doi:10.1007/s10643-009-0350-z

Weissberg, R., Ruff, H. A., & Lawson, K. R. (1990). The usefulness of reaction time tasks in studying attention and organization of behavior in young children. *Journal of Developmental and Behavioral Pediatrics: JDBP*, *11*(2), 59–64. doi:10.1097/00004703-199004000-00004

Woodley, M. A., Te Nijenhuis, J., & Murphy, R. (2014). Were the Victorians cleverer than us? The decline in general intelligence estimated from a meta-analysis of the slowing of simple reaction time. *Intelligence*, *41*(6), 843–850. doi:10.1016/j.intell.2013.04.006.

Author Biographies

Pedro Ángel Latorre-Roman is a professor in the Department of Didactics of Corporal Expression, University of Jaén, Spain.

Alejandro Robles-Fuentes is a professor and director of municipal sports schools in Santiago de la Espada, Jaén, Spain.

Felipe García-Pinillos is a professor in the Department of Physical Education, Sport and Recreation, Universidad de La Fontera, Temuco, Chile.

Jesús Salas-Sánchez is a professor and an associate researcher in Universidad Autó noma de Chile.

Trends in Triathlon Performance: Effects of Sex and Age

Article *in* Sports Medicine · June 2013

DOI: 10.1007/s40279-013-0067-4 · Source: PubMed

CITATIONS
95

READS
2,239

3 authors:

Romuald Lepers
University of Burgundy
317 PUBLICATIONS **9,185** CITATIONS
SEE PROFILE

Paul J Stapley
University of Wollongong
140 PUBLICATIONS **2,105** CITATIONS
SEE PROFILE

Beat Knechtle
874 PUBLICATIONS **12,124** CITATIONS
SEE PROFILE

Sports Med (2013) 43:851–863
DOI 10.1007/s40279-013-0067-4

REVIEW ARTICLE

Trends in Triathlon Performance: Effects of Sex and Age

Romuald Lepers · Beat Knechtle · Paul J. Stapley

Published online: 25 June 2013

Abstract The influences of sex and age upon endurance performance have previously been documented for both running and swimming. A number of recent studies have investigated how sex and age influence triathlon performance, a sport that combines three disciplines (swimming, cycling and running), with competitions commonly lasting between 2 (short distance: 1.5-km swim, 40-km cycle and 10-km run) and 8 h (Ironman distance: 3.8-km swim, 180-km cycle and 42-km run) for elite triathletes. Age and sex influences upon performance have also been investigated for ultra-triathlons, with distances corresponding to several Ironman distances and lasting several days, and for off-road triathlons combining swimming, mountain biking and trail running. Triathlon represents an intriguing alternative model for analysing the effects of age and sex upon endurance and ultra-endurance (>6 h) performance because sex differences and age-related declines in performance can be analysed in the same individuals across the three separate disciplines. The relative participation of both females and masters athletes (age >40 years) in triathlon has increased consistently over the past 25 years. Sex differences in triathlon performance are also known to differ between the modes of locomotion adopted (swimming, cycling or running) for both elite and non-elite triathletes. Generally, time differences between sexes in swimming have been shown to be smaller on average than during cycling and running. Both physiological and morphological factors contribute to explaining these findings. Performance density (i.e. the time difference between the winner and tenth-placed competitor) has progressively improved (time differences have decreased) for international races over the past two decades for both males and females, with performance density now very similar for both sexes. For age-group triathletes, sex differences in total triathlon performance time increases with age. However, the possible difference in age-related changes in the physiological determinants of endurance and ultra-endurance performances between males and females needs further investigation. Non-physiological factors such as low rates of participation of older female triathletes may also contribute to the greater age-related decline in triathlon performance shown by females. Total triathlon performance has been shown to decrease in a curvilinear manner with advancing age. However, when triathlon performance is broken down into its three disciplines, there is a smaller age-related decline in cycling performance than in running and swimming performances. Age-associated changes in triathlon performance are also related to the total duration of triathlon races. The magnitude of the declines in cycling and running performances with advancing age for short triathlons are less pronounced than for longer Ironman-distance races. Triathlon distance is also important when considering how age affects the rate of the decline in performance. Off-road triathlon performances display greater decrements with age than road-based triathlons, suggesting that the type of discipline (road vs. mountain bike cycling and road vs. trail running) is an important factor in age-associated changes in triathlon performance.

R. Lepers (✉)
INSERM U1093, Université de Bourgogne, Faculty of Sport Science, BP 27877, 21078 Dijon cedex, France
e-mail: romuald.lepers@u-bourgogne.fr

B. Knechtle
Institute of General Practice and Health Services Research, University of Zurich, Zurich, Switzerland

P. J. Stapley
School of Health Sciences, Neural Control of Movement Laboratory, Faculty of Science, Medicine and Health, University of Wollongong, Wollongong, Australia

Finally, masters triathletes have shown relative improvements in their performances across the three triathlon disciplines and total triathlon event times during Ironman races over the past three decades. This raises an important issue as to whether older male and female triathletes have yet reached their performance limits during Ironman triathlons.

1 Introduction

Triathlon is a unique endurance sport that combines three disciplines (swimming, cycling and running) over a variety of distances [1]. Triathlon first appeared at the end of the 1970s and, in only one quarter of a century, has developed into a well organized sport with global participation. The first Hawaii Ironman triathlon, consisting of a 3.8-km swim, 180-km cycle and 42-km run had 12 male participants in 1978. Currently, more than 1,700 triathletes (<30 % females), most of whom have to qualify at one of 25 Ironman triathlons worldwide, participate in this event, which has become the Ironman World Championship [2]. Short- or Olympic-distance triathlons (1.5-km swim, 40-km cycle and 10-km run) have also rapidly grown in popularity during approximately the same period, with the first World Championship being held in Avignon (France) in 1989. Triathlon was officially accepted into the Olympic games as a full sport in Sydney in 2000 [3]. Ultra-triathlons, consisting of distances greater than Ironman distance, and including double Ironman triathlons (7.6-km swim, 360-km cycle and 84-km run) and deca-Ironman triathlons (38-km swim, 1,800-km cycle and 420-km run), also appeared during the 1980s, with the first double Ironman triathlon held in 1985 in Huntsville (USA) [4–7]. More recently, in addition to conventional road-based triathlon, off-road triathlon, combining swimming, mountain biking and trail running, has also established itself in the field of endurance sports, and has grown very rapidly in popularity [8, 9].

Triathlon performance has been studied from numerous perspectives by a number of research groups over approximately the same period that the sport has grown. Research has investigated physiological [1, 10–14], biomechanical [15–17], training [16, 18–20], nutritional [21] or medical [22] aspects of triathlon performance. Although Ironman triathlon is a relatively new ultra-endurance event, recently elite Ironman triathletes appear to have reached their performance limits [2, 23] as has been observed in more traditional sports such as marathon running, where performance times have plateaued [24, 25]. With the increase in popularity of endurance sports worldwide, endurance sports have attracted a greater participation of female and masters athletes (age >40 years) during recent times [24–27].

This review addresses the specific aspects of (i) sex differences and (ii) age-related declines in triathlon performance. Sex differences in performance will be presented and discussed for each of the three disciplines and total event times for different distances and for both elite and age-group triathletes. Age-related declines will be outlined by focusing on the performances of the best master triathletes because these athletes represent a unique model for studying the effects of high levels of physical training in older individuals.

2 Sex Differences in Performance

2.1 Female Rates of Participation

Independent of age, the number of females competing in triathlon has increased progressively since the 1980s. For example, between 2000 and 2011, female USA Triathlon membership has grown from 27 % of the total number of annual members to more than 38 % [28]. Factors leading to this growth include society's increasing acceptance of 'active' females; females feeling more comfortable living an active lifestyle [29] and the growth of females-only events, good examples of which are the 'Danskin' and 'Trek Triathlon' Series in the USA.

The rates of female participation appear to decrease with an increase in distance of triathlon races. For example, in 2010, females accounted for 26, 19 and 13 % of athletes competing at short-, half-Ironman and full-Ironman distance triathlons held in the region of Zurich (Switzerland) [30–32]. However, for ultra-triathlons, female participation remains relatively low, representing generally less than 10 % of overall participant numbers [4, 5, 33].

The number of females finishing the Hawaii Ironman World Championship increased from 20 in 1981 (6 % of participants) to more than 470 in 2010 (27 % of participants) [34]. In Europe, there was also a progressive rise in the number of female Ironman finishers, but overall rates were lower. In 2011, females accounted for 13 % of the field at Ironman Switzerland [32]. Lower rates of female participation in Europe may be because the European Ironman triathlon is relatively younger than the Hawaii Ironman triathlon. Indeed, the first Hawaii Ironman was held in 1978, while the first European Ironman, Ironman Switzerland, was held in 1995. Female participation in Ironman triathlon remains slightly lower than in traditional endurance events such as marathon running, but greater than ultra-endurance events such as a 161-km ultra-marathon. For example, Lepers and Cattagni [24] showed that the relative participation of females at the New York City Marathon increased over the last 30 years from 17 to 33 % of the total field. In contrast, females accounted for 20 % of

the finishers in the same time period at 161-km ultramarathons in North America [35].

2.2 Physiological and Morphological Considerations

Physiological and morphological characteristics may account for the sex differences seen in triathlon performance. Until the 1970s, data were scarce with regard to the physiological determinants of females in endurance sports. Even though there are fewer data for females than for males, studies have shown that maximal oxygen uptake ($\dot{V}O_{2max}$), lactate threshold and running economy interact similarly in females as determinants of endurance performance as they do in males [36].

Current explanations for sex differences in $\dot{V}O_{2max}$ among elite athletes, when expressed relative to body mass, provide two major findings [37, 38]. First, elite females have more (<13 vs. <5 %) body fat than males [39]. Indeed, much of the difference in $\dot{V}O_{2max}$ between males and females disappears when it is expressed relative to lean body mass [37, 38]. Second, the hemoglobin concentration of elite athletes is 5–10 % lower in females than in males [40]. Concerning lactate threshold, there is no reason to believe that values should be lower in females than in males because mitochondrial adaptations in the skeletal muscles of highly trained male and female athletes appear to be similar [41]. Finally, the average oxygen cost to run a given speed, (i.e. running economy) by groups of elite male and female athletes is similar and appears to play the same role in determining success in endurance performance [42]. Therefore, the major physiological reason that explains the slower record performances by females compared with males is probably the lower $\dot{V}O_{2max}$ values observed in females.

Physiological differences between male and female triathletes have been infrequently investigated [14, 39, 43]. The differences in the physiological responses in cycling and the energy cost of running after cycling have been compared in males and females for both elite junior and senior triathletes [14]. These authors showed that senior females had a significantly higher cycling peak power output than their junior counterparts. Additionally, senior males had a higher ventilatory threshold than junior males, whereas the ventilatory threshold was similar in junior and senior females.

A recent study conducted on junior triathletes has suggested that morphological characteristics of triathletes have evolved since the late 1990s [44]. In 2011, both male and female junior elite triathletes appeared more ectomorphic than their 1997 counterparts. Nowadays, junior triathletes are proportionally lighter, with significantly smaller flexed arm and thigh girths, and femur breadths. The junior males in 2011 also had significantly longer segmental lengths and lower endomorphic values than their 1997 counterparts. However, more research is required to validate these findings for senior triathletes.

Male triathletes possess a larger muscle mass, greater muscular strength and lower relative body fat than female triathletes [45]. Low body fat is an important predictor variable for total time performance in triathlon. For example, Knechtle et al. [46] showed that low body fat was associated with faster race times in male Ironman triathletes but not in females. Males possess on average 7–9 % less percent body fat than females, which is likely an advantage for males. Therefore, it appears that sex differences in percentage body fat, oxygen-carrying capacity and muscle mass may be major factors for sex differences in overall triathlon performance. Menstrual cycle, and possibly pregnancy, may also impact training and racing in female athletes, factors that do not affect males [47].

2.3 Sex Differences in Triathlon Performance

Sex differences in endurance and ultra-endurance (>6 h) [48] performance have received considerable attention over the past few decades, but the majority of studies have focused on running performances [49, 50]. Interestingly, Speechly et al. [51] reported that females, who were matched with males for a 42-km run, were faster than males in a 90-km race. Similarly, it has been shown that females and males who were matched for 50-km trail running performance also performed similarly in trail runs of 80- and 161-km distances [52]. Despite the suggestion in 1992 that females may one day outrun males in competitive ultra-endurance events [53], elite males appear to run approximately 10–12 % faster than elite females across all endurance running race distances up to marathon, with the sex difference narrowing as the race distance increases [49, 50]. However, at distances greater than 100 km, such as the 161-km ultramarathon, the difference seems even larger, with females 20–30 % slower than males [49, 54, 55].

Sex differences in triathlon performance have previously been described for elite triathletes [2, 9, 31, 56, 57] and for non-elite triathletes of different age groups [30, 34, 58] for long distances (distances greater or equal to half-Ironman) [2, 31, 34, 56, 57] and for off-road triathlon [9]. Surprisingly, very few data are available for short-distance triathlons, especially for high-level international races. One reason may be that the top international short-distance triathlons (i.e. World Championship Series events or Olympics) have all been draft legal for several years and therefore it is difficult to find a reference of high-level Olympic distance triathlon without drafting for comparison. The increase in non-drafting international distance triathlons such the *Hy-Vee* triathlon, which offers

Table 1 Mean sex differences in time performance for swimming, cycling, running and total time at different national and international triathlons

Event	Sex difference in time performance (%)			
	Swim	Cycle	Run	Total
Short distance (1.5–40–10 km): [30, 79]				
Zurich (Switzerland) from 2000 to 2010				
Top five elite overall	15.2	13.4	17.1	14.8
Top five AG, from 18 to 54 years	18.5	15.5	18.5	17.1
World Championship from 2009 to 2011				
Top ten AG, from 18 to 64 years	13.3	10.7	7.5	12.0
Half Ironman (1.9–90–21 km): [31, 79]				
Rapperswil (Switzerland) from 2007 to 2010				
Top five elite overall	14.1	12.3	12.5	12.6
Top five AG, from 18 to 54 years	22.3	16.4	19.2	17.6
World Championship from 2009 to 2011				
Top ten AG, from 18 to 64 years	12.4	11.2	14.5	12.6
Off-road triathlon (1.5–30–10 km): [9]				
World championship (Maui, USA) from 2007 to 2009				
Top ten elite overall	12.4	19.6	18.4	18.2
Ironman (3.8–180–42 km): [2, 32, 34]				
World championship (Kona, Hawaii, USA) from 1988 to 2007				
Top ten elite overall	9.8	12.7	13.3	12.6
Top ten AG, from 18 to 64 years	12.1	15.4	18.2	15.8
Zurich (Switzerland) from 1995 to 2010				
Top ten elite overall	14.0	13.2	18.2	14.9

AG age groups

significant prize money of $US151,000 for both male and female winners and represents a highly competitive platform for professional triathletes, should help to better determine sex differences in short-distance triathlon performance when athletes compete on a level (non-drafting) playing field in the future.

Table 1 synthesizes the data from the literature that have focused on sex differences in triathlon performances, including international and national triathlon races, from age groups and elite triathletes. Over the past 25 years, sex differences in total triathlon performance (across triathlons of different distances and format) have varied from 12 to 18.2 % (far right column, Table 1), depending on the level of the triathletes (elite vs. age group) or the distance/format of the race. However, overall, data have shown that the sex difference in triathlon performance has narrowed across the years [2]. Table 2 presents the most recent data for sex differences in triathlon performance for international elite triathletes participating in the most competitive races at each triathlon distance. By 2012, values narrowed so much that, for the top ten triathletes overall in three top-level international road-based triathlon races, the difference between males and females was 11.3 % for the Hawaii Ironman triathlon, 14.1 % for the Olympics and 9.3 % for the *Hy-Vee* short distance triathlon (see Table 2). Non road-based triathlon produced a greater sex difference (discussed in more detail below). Interestingly, for ultra-triathlons, it has been shown that with increasing length of the event, the best females became relatively slower compared with the best males [4]. Indeed, if the world's best performances are considered, males were 19 % faster than the females in both Double and Triple Ironman distance, and 30 % faster in the Deca-Ironman distance [4]. However, non-physiological factors may have contributed to these observations, in particular a fewer overall number of female finishers in ultra-triathlons, compared with formats of more standard length [4].

2.4 Sex Differences in Triathlon Swimming Performance

The average sex difference in swimming performance during triathlon for race distances between 1.5 and 3.8 km ranged between approximately 10 and 15 % for elite triathletes (see Table 1). The sex difference in triathlon swimming performance is consistent with values found for different pool swimming events. It has been shown that sex differences in pool swimming performances become progressively smaller with increasing distance from 50 m (19 %) to 1,500 m (11 %) [59]. This seems to also be the

Table 2 Mean percentage differences in times for swimming, cycling, running and total event between the top ten females and males and time difference between the winner and tenth-placed athlete in 2012 at four international triathlons: Hawaii Ironman triathlon World championship (Kona, Hawaii, USA); Olympics Triathlon (London, UK); *Hy-Vee* short distance triathlon (Des Moines, Iowa, USA); World championship off road (Xterra) triathlon (Maui, Hawaii, USA). (unpublished personal data)

Event	Sex difference in performance in top ten athletes in 2012 (mean ± SD)				Difference between tenth and first (%)	
	Swim	Cycle	Run	Total	Male	Female
Hawaii Ironman Triathlon (3.8–180–42 km)	14.1 ± 7.9	13.1 ± 2.3	7.3 ± 2.9	11.3 ± 0.5	3.6	4.0
Olympics Triathlon (1.5–40–10 km) with drafting	11.8 ± 2.0	11.3 ± 0.6	14.7 ± 0.8	14.1 ± 7.9	1.5	1.4
Hy-Vee Triathlon (1.5–40–10 km) without drafting	8.6 ± 4.8	10.2 ± 3.5	8.6 ± 4.4	9.3 ± 0.5	4.1	3.3
World Championship Off-Road Triathlon (1.5–30–10 km)	15.2 ± 15.5	22.6 ± 4.4	15.1 ± 6.7	17.3 ± 2.9	4.6	12.2

case for triathlon, as Table 1 shows that the longer the distance of the swim component of a triathlon race the smaller the sex difference, particularly when the top ten age-group athlete performances are considered. In ultra-endurance outdoor swimming events, female swimmers have tended to reduce the gap with their male counterparts. For example, time differences between male and female swimming records are 6.7 % for the 32-km 'English Channel Swim' and 2.3 % for the 26-km 'Marathon Swim in Lake of Zurich' in Switzerland, respectively [60, 61]. However, the sex difference in performance between the best male and female ultra-swimmers is more generally close to 11–12 %, which corresponds to values observed for swimming in triathlon [60–62]. It has been shown that the elite male and female triathletes completed the 3.8-km swim stage of the Hawaii Ironman triathlon ~10 % slower than the elite swimmer specialists for the same distance at the Waikiki Rough-water Swim race [2]. This difference could be explained by better propelling efficiency in elite swimmers than in elite triathletes [63].

Analysis from 2005 showed that the sex difference in triathlon swimming appears lower than in cycling and running at the Hawaii Ironman triathlon and at the off-road triathlon Xterra World Championship [2, 9, 34]. The difference between swimming and the two other disciplines could be explained in part by the biological difference in relative body fat (7–9 % higher in females) [64, 65]. Performance in activities that involve supporting one's weight (e.g. running) may be more susceptible to a greater body fat percentage than water-based events such as swimming, during which body fat can increase buoyancy and therefore improve performance [65]. Buoyancy is also improved in females through a lower 'underwater torque', which can be defined loosely as the tendency for the feet to sink [66]. In addition, in contrast with running, where the energy cost appears to be similar between females and males, the energy cost of freestyle swimming has been shown to be significantly higher (i.e. lower economy) in males compared with females [66, 67]. At a velocity of 1 m/s, there are differences in drag force and coefficient of drag between males and females [68]. The energy cost of swimming depends essentially on the propelling efficiency of the arm stroke and hydrodynamic resistance, but it has been suggested that differences in energy costs of swimming between sexes are mainly to be attributed to differences in hydrodynamic resistance [69]. In contrast, differences in energy cost of swimming across ages may be attributed also to changes in the propelling efficiency of the arm stroke [69]. Females have smaller body size (resulting in smaller body drag), smaller body density (greater fat percent) and shorter lower limbs, resulting in a more horizontal and streamlined position and therefore a smaller underwater torque [64, 66].

2.5 Sex Difference in Triathlon Cycling Performance

Sex differences in triathlon cycling vary from 12 to 16% according to the level of expertise of participating triathletes for road-based triathlons (Table 1). Data gained from analysing the performances of male and female participants in the 180-km cycling leg of an Ironman triathlon are unique datasets, as official time-trial road-cycling championships generally take place on distances much shorter than 180 km, with distances also being greater for males (<40 km) than for females (<25 km). Moreover, there is a paucity of data concerning sex differences in road or track cycling performance [70–72]. In track cycling, where females are generally weaker than males in terms of power/weight ratios, the performance gap between males and females appears to be constant (<11 %) and independent of the race distance from 200 to 1,000 m [70]. In ultra-cycling events, such as the 'Race Across America', sex difference in performance was around 15 % among top competitors [71]. Greater muscle mass and aerobic capacity in males, even expressed relative to the lean body mass [73], may

represent an advantage during long-distance cycling, especially on a relative flat course such as Ironman cycling, where cycling approximates to a non-weight-bearing sport. Indeed, it has been shown that absolute power output (which is greater for males than for females) is associated with successful cycling endurance performance because the primary force inhibiting forward motion on a flat course is air resistance [74].

Interestingly, for elite triathletes, the sex difference in mountain bike cycling during off-road triathlon (<20 %) is greater than cycling sex differences in conventional road-based events [9]. Mountain biking differs in many ways from road cycling. Factors other than aerobic power and capacity, such as off-road cycling economy, anaerobic power and capacity, and technical ability might influence off-road cycling performance [75]. Bouts of high-intensity exercise frequently encountered during the mountain biking leg of off-road triathlon (lasting <1 h 30 min for elite males and <2 h for elite females) can result from (1) having to overcome the constraints of gravity associated with steep climbs, (2) variable terrain necessitating wider tires and thus greater rolling resistance, and (3) isometric muscle contractions associated with the needs of more skilled bike-handling skills, not so often encountered in road cycling. However, in particular, lower power-to-weight ratios for female than for male triathletes inevitably leave them at a disadvantage during steep climbs [76, 77]. Moreover, the increased bike-handling skills, required especially during downhill mountain biking, may be impaired by relatively lower female arm or leg isometric muscle strength, meaning that greater sex differences in off-road triathlon could be attributed to decreased ability to control the bicycle in females. However, this assumption needs to be confirmed with specific investigations of the effect of sex on technical ability in mountain biking [9].

As percentage differences in time do not equate to percent differences in power output, due to non-linear relationships between speed and power output from air or water resistance [78], the magnitude of the sex difference has also been examined by calculating the percentage difference between males and females in estimated power outputs for each discipline [2, 9, 34, 79]. In this case, differences in estimated power output between the sexes are greater for triathlon cycling than for the swimming and running disciplines. For example, sex differences in power output in swimming, cycling and running were estimated to be 28, 39 and 33 %, respectively, for Ironman triathlon, and 30, 45 and 33 % for off-road triathlon [9]. Because power output is proportional to oxygen uptake, the magnitude of sex differences in power output provides a more realistic representation of underlying sex differences in physiological capacity [80].

2.6 Sex Differences in Triathlon Running Performance

During the 1988–2007 period, the top ten elite males have run the Hawaii Ironman marathon on average 13.3 % faster than the top ten females [2]. In contrast, during the same period, elite female triathletes have improved their marathon running times by 0.8 min/year while times have remained stable for the males. While these improvements in female performances compared with males are impressive, they remain as yet unexplained, which is strange when it is considered that females have benefited from similar training and nutritional advances [21]. If females continue to improve their running performance at Ironman, they could reduce the sex difference in the marathon leg and therefore their overall performance times. For example, at the 2012 Hawaii Ironman, the difference in marathon times between the top ten elite males and the top ten elite females reduced to 7.3 % (males: <3 h 05 min; females: <3 h 20 min) (Table 2). Another notable example is that the female winner of the famous Roth Ironman (Germany) in 2011 ran only 2.6 % slower than the winning male (2 h 44 min vs. 2 h 40 min). At the same race, both female and male winners bettered the Ironman distance performance world record (8 h 18 min vs. 7 h 41 min) [81]. Thus, it appears that the marathon running leg of Ironman triathlon has become the discipline in which female elite triathletes have most reduced the gap with their male counterparts. Interestingly, Lepers [2] showed that the sex difference in running at the Hawaii Ironman marathon for the top ten elite finishers was similar to that recorded during the New York marathon, suggesting above all, that the swim and cycle legs of the triathlon do not exacerbate the sex difference in running a marathon. Thus, the physiological differences between males and females in running performance that have been identified to occur in a regular marathon still persist in the marathon of an Ironman.

Contrary to the knowledge of running performance differences between sexes for the marathon distance, few analyses have been conducted between males and females who have run 10 km during an Olympic distance triathlon. A mean sex difference of 17 % was found for the Zurich Switzerland Olympic distance triathlon during the 2000–2010 period [30]. However, as Table 2 shows, when data from 2012 only are considered, there is a considerable difference even within elite-level Olympic distance run performances, between males and females. A difference of 12.3 % existed during the draft-legal Olympic short-distance race in London, whereas the difference was lower (8.6 %) at the *Hy-Vee* international triathlon in Des Moines (draft-free). It would appear that the format of the triathlon (drafting or not) affects males differently to females, with the possibility that males benefit in terms of their

subsequent running performance off the bike, to a greater extent when riding in a peloton. However, running performances in males and females need to be more closely analysed in relation to the effort produced during the cycle leg of an Olympic distance triathlon.

2.7 Performance Density in Triathlon Results

Performance density (i.e. the time difference between the winner and tenth-placed competitor) has been quantified by considering overall performance times in triathlon [2]. For example, at the Hawaii Ironman triathlon between 1981 and 2008, the average time difference between first and tenth place was smaller for male (5.8 %) than for female athletes (7.5 %) [2]. During the past 5 years, the performance density has decreased for both male (<3.1 %) and female athletes (<5.7 %), a trend that would suggest that in the future, high-performing female athletes may be as performance dense as the males. In fact, Table 2 indicates that in 2012 performance density was similar across the two sexes for top international road-based triathlon events, regardless of format (Hawaii, Olympics and *Hy-Vee*). A notable exception to this was the Xterra World Championships (off-road triathlon), in which female performance density (12.2 %) was considerably lower than the male performance density (4.6 %). As explained earlier (Sect. 2.5), this can likely be explained by the specific physiological and technical abilities needed for the mountain bike leg of Xterra, particularly explosive strength, which favours males over females [82]. The small performance density observed for the London Olympics compared with other events, for both males and females, may be explained by the drafting component of the cycling leg that led to a grouped start of a number of athletes (approximately half of the field) at the beginning of the run leg. It must also be considered that a number of nations fielded more than one athlete, and identified a designated leader who was 'protected' during the swim/cycle legs. This would also have led to a higher overall athlete density in terms of overall performance times in both males and females.

2.8 Effects of Age on Sex Differences

Physiological (e.g. lower muscle strength and oxygen-carrying capacity), morphological, (e.g. greater percentage of body fat, lower muscle mass) and functional capacities are well known to change with advancing age in both males and females [83]. In addition, is has been shown that, after 55 years of age, the decline in endurance performance is more pronounced in females than in males [84]. It is thus likely that sex differences in triathlon performance become more pronounced with advancing age. A greater age-related decline in performance in females than in males has been previously observed in swimming and running [59, 85]. Results from the Hawaii Ironman triathlon showed that the sex difference in total event performance time increased significantly with advancing age from 55 years during the 2006–2008 period [34]. Male triathletes aged 60 years were, on average, 27 % slower than those who were between 30 and 40 years of age, while the difference was 38 % for females. Sex differences in performance during a Swiss regional Olympic distance triathlon became greater over the age of 35 years [30]. The occurrence of this large sex difference was somewhat earlier with respect to athletes' age than recorded during the Hawaii Ironman triathlon [34]. Possible explanations for this include the greater competitive level of older participants at the Hawaii Ironman triathlon because athletes are required to qualify as it is a World Championship event. There is a need for further research to understand if exercise duration exerts an influence upon sex differences across older age groups by analysing performances of athletes in a wider range of triathlon events, from Sprint distance, through Olympic distance and Half-ironman, to Ironman distances [79].

The exact reasons for these sex-related differences with advancing age are currently not clear but may result from physiological, sociological and psychological changes [38, 86, 87]. For example, a greater decline of one or more physiological determinants of endurance performance for females compared with males, (e.g. $\dot{V}O_{2max}$, lactate threshold or exercise economy), or a difference in age-related changes in body composition (increase in percentage body fat and loss of muscular mass), hormonal changes and fluid balance changes (e.g. decline in the thirst mechanism), could affect triathlon performance [38, 83, 84, 88]. In addition, differences in terms of years of training, training volume and intensity between elderly male and female triathletes performing Ironman triathlon may exist, but further research is needed to clarify this. However, interpretation of cross-sectional comparisons of triathlon performance times across ages and sexes must be made carefully. It is likely that, compared with males, there are fewer females competing in triathlon events, especially in the older age categories. For example, the percentage of females participating at the Hawaii Ironman triathlon during 2006–2008 corresponded on average to 27 %, but 'finisher' females in the age group 60–64 years represented only 3 % of the females field [34]. This participation difference may diminish over the next couple of decades as has been observed in marathon running [24, 26], such that because the well trained females move up to the older age groups, the improvement of the oldest females may actually surpass the oldest males.

3 Age-Related Declines in Triathlon Performance

Age-related declines in endurance and ultra-endurance performance have been well described in the literature for

Table 3 Total time records (actualized with 2012 data) and corresponding split times for male and female age groups at the Hawaii Ironman Triathlon between 1986 and 2012

	Age groups (years)									
	18–39	40–44	45–49	50–54	55–59	60–64	65–69	70–74	75–79	>80
Male										
3.8-km swim (h:min:s)	51:56	51:48	56:55	1:03:32	1:07:09	1:16:20	1:14:11	1:47:46	1:37:47	1:49:34
180-km cycle (h:min:s)	4:24:05	4:39:16	5:04:47	4:51:44	5:00:17	5:19:17	5:42:08	5:47:33	6:39:35	7:42:08
42-km run (h:min:s)	2:44:02	2:53:28	3:04:21	3:24:51	3:34:03	3:25:28	4:14:52	3:52:47	4:55:43	5:41:51
Total (h:min:s)	8:03:56	8:24:32	9:11:24	9:26:23	9:47:29	10:08:15	11:19:07	11:45:05	13:27:50	15:38:25
Total change (%)	100	104	114	117	121	126	140	146	167	194
(year)	(2011)	(1994)	(2009)	(2006)	(2005)	(2010)	(2011)	(2011)	(2005)	(2012)
Female										
3.8-km swim (h:min:s)	54:31	1:13:52	1:06:21	1:08:08	1:06:18	1:32:16	1:21:02	1:37:54	1:45:05	–
180-km cycle (h:min:s)	4:52:06	5:25:00	5:06:07	5:31:56	5:35:36	6:27:46	6:47:28	7:24:33	7:25:17	–
42-km run (h:min:s)	3:03:05	3:17:48	3:09:18	3:47:23	4:00:08	4:05:22	4:59:01	6:07:02	6:19:43	–
Total (h:min:s)	8:54:02	10:02:35	9:26:25	10:35:59	10:51:43	12:17:24	13:16:32	15:19:19	15:54:16	–
Total change (%)	100	113	106	119	122	138	149	172	179	
(year)	(2009)	(2010)	(2012)	(2005)	(2010)	(2010)	(2010)	(2000)	(2005)	

The change in percentage compared with the fastest total time (age group 18–39 years) is given for each different age group

running [24, 35, 38, 54, 83, 89–91], cycling [92–95], swimming [83, 85] and more recently, for triathlon [34, 79, 92, 96–98]. Endurance and ultra-endurance performance appears to be maintained until approximately 35–40 years of age, followed by modest decreases until 50 years of age and a progressive decrease in performance thereafter [83, 84, 88]. The greatest declines in endurance and triathlon performance occur after the age of 70 years (Table 3). Physiological factors contribute to age-related declines in endurance performance in older athletes and there is obviously an interaction between training behaviour and performance in older athletes [38, 83, 86, 88].

It has been found that the age of peak performance in Ironman triathlon is around 33–34 years for both males and females [32, 57], which seems to be older than the age of peak overall performance of marathoners [99]. However, this finding observed for the Ironman Switzerland triathlon needs be confirmed for other Ironman events. For non-elite triathletes, the fastest race times are usually achieved between 25 and 44 years for both Ironman and Ultra-triathlons [32, 57, 97].

3.1 Effects of the Mode of Locomotion

Total triathlon performance decreases progressively in a curvilinear manner with advancing age. However, there is a smaller age-related decline in cycling performance than in running and swimming performances for both short- and long-distance triathlons [8, 96, 97]. These findings suggest that age-related declines in endurance performance are specific to the mode of locomotion, although the cause for such mode-specificity is not clear. Where does the focus of this locomotor mode-specificity with age lie? Our group has previously proposed that mechanical power could explain these age-related differences in cycling and running [98]. According to the formula $P = k.V$ (k: constant), mechanical power output (P) is dependent upon velocity (V) during running, whereas it is dependent upon the third power of velocity during cycling ($P = k.V$ [3]). Following this, we previously suggested that, as changes in aerobic capacity with age are tied to reductions in P, those reductions during running and cycling with age would give rise to lower cycling velocities than running velocities [98]. A number of explanations can be given to explain the smaller decline in cycling performance during triathlon compared with the declines in the other two disciplines. These include a lesser reduction in lactate threshold or economy during cycling, or a greater muscular fatigue during running with age, although these propositions remain to be validated. Some authors have though, attempted to explain this age-related phenomenon in cycling by proposing that the 'training stimulus', the ability

or will to train, is reduced in running compared with cycling [85]. In particular, this may be because running is associated with a greater amount of orthopaedic injuries, which limits the ability to train in running [100]. The training stimulus would therefore be maintained in cycling due to the less traumatic nature of the endurance activity and therefore with age, triathletes would tend to cycle more than run. However, this proposition remains speculative and requires further investigation.

3.2 Effects of Discipline Duration

The duration of a triathlon race exerts an important influence on the age-related changes in triathlon performance [98]. Age-related declines in swimming performance are not influenced by triathlon duration, with the magnitudes of decreases in swimming performance similar for Olympic distance versus Ironman triathlon. However, in contrast to swimming, the magnitude of the declines in cycling and running performances with advancing age during Olympic distance triathlon were less pronounced than during Ironman [98]. For age-group competitors of 70–74 years, total finishing time is approximately 3 hours for the Olympic distance triathlon, whereas it is around 15 h for an Ironman (Table 3). Certainly, the Ironman triathlon induces greater neuromuscular fatigue in cycling and running than the Olympic-distance event [101]. Furthermore, muscle damage during a 10-km run of the Olympic-distance triathlon is limited compared with that which occurs during an Ironman marathon. Greater muscle fatigability and greater sensibility to muscle damage of older triathletes needs further investigation to determine if these factors underlie the greater declines in cycling and running performance of older triathletes.

3.3 Road-Based Versus Off-Road Triathlon

It has recently been shown that the rate of the decline in performance for off-road triathlon is greater than for road-based triathlon [8]. This suggests that the type of discipline (road vs. mountain bike cycling and road vs. trail running) exerts an important influence on the magnitude of the age-associated changes in triathlon performance [8]. The specific aspects of mountain biking and trail running may explain why age-related declines in off-road triathlon are more pronounced for off-road than for conventional on-road triathlon. In particular, a decrease in power-to-weight ratio that has been shown to occur with age [102, 103] leaves older athletes with a distinct advantage during steep climbs often encountered in off-road triathlons. Moreover, despite a preservation of muscle strength in comparison with sedentary persons [104], older athletes may be disadvantaged in their technical bike-handling skills due to lower arm/leg muscle strength. Poorer running performances in older athletes, especially when variations in intensity are required, may be explained by a reduced ability to modify the biomechanical components of running (e.g. support phases, stride frequency, changes in stride length, etc.) [8].

3.4 Improvements in Triathlon Performance of Elderly Triathletes

An increase in participation of both male and female athletes older than 40 years over the past few decades has been reported for marathon and ultra-marathon running, such as 161-km ultra-marathons and for 100-km running in Switzerland [24, 25, 35, 90]. A relative increase in participation of masters triathletes has also been observed during the past decade for short- and long-distance triathlon, while the participation of triathletes younger than 40 years of age has decreased [27, 28, 30, 58]. The relative increase in participation of master triathletes at the Ironman distance triathlons has been accompanied by an improvement in their performance [27, 105]. We have previously presented data for the 1986–2010 period showing that swimming, cycling, running and total time performances at the Hawaii Ironman Triathlon improved for male triathletes older than 44 years and female triathletes older than 40 years [27]. During the 25-year study period, the total time decreased by <21 % (from 14 h 18 min to 11 h 16 min) for the best male finishers in the age group 60–64 years, and by the same percentage (from 14 h 38 min to 11 h 30 min) for the best finishers in the age group 40–44 years [27]. Similar findings have been observed for a Hawaii Ironman qualifying race such as the Switzerland Ironman [105]. In Table 3, we show that, when the most recent data (Ironman Hawaii 2011 and 2012) are added, improvements of older age groups, especially for the males has continued, and for some age groups, is quite startling. For example, the best male finisher in the 70–74 age group has bettered the best time set in 2010 (12 h 41 min) by almost 1 hour (11 h 45 min). Other, more modest improvements have been recorded in the male 65–69 and >80 age groups, and the female 45–49 age group with respect to the 2010 data [27]. However, what is clear is that, if the most recent (2012) data are compared with those of 1986, age groups show improvements of between 114 and 194 % for the males aged over 45 years, and 106 and 179 % for females aged over 45 years, and that these improvements seem to be continuing, most significantly for the oldest athletes.

Several reasons may explain the improved performance of master triathletes. These include an improvement of training facilities, coaching, training techniques, nutritional strategies and equipment [84, 86], as well as the possibility

that they have had better and prolonged access to the facilities required to train effectively. It is also obvious that the higher participation rates of master athletes increases the possibility of them obtaining better results due to the competitive nature of the sport. Additionally, athletes that have attained a number of good results and positive outcomes have increased levels of motivation to train and compete in endurance events [106]. Accordingly, the better physical condition of older athletes is likely to increase competitive spirit, participation and performance [84, 86]. The advancement of training quality is also a possible reason for these improved performance trends. Earlier studies suggested the performance decline with advancing age was due to decreased training volumes and intensity [86]. Other, more sociological, factors may also play a role in reducing the ability and motivation to train. These include greater work demands, demands of the family and a diminished inherent drive to push oneself or even train [83, 84, 86]. An important factor is also the greater amount of time needed by the body to recover from hard, physical effort resulting from training as humans age. Indeed, training for an Ironman triathlon is very demanding and, in order to place in the top ten of an age group category, extremely high training volumes and intensities are required [18, 107–109]. However, it has been suggested that masters athletes could optimize their quality of training, so they could reduce their training volume to save time for adequate recovery and remain injury free [84]. These studies suggest that masters triathletes have probably not reached their limits in ultra-endurance performance. Unfortunately, no information is available concerning the sporting background of these successful masters triathletes. Nevertheless, even if recreational masters athletes had relatively short training histories and little experience in triathlon, these successful athletes probably have life-long histories of physical activity [110–113].

4 Conclusions

Sex differences in triathlon performance differ generally between the locomotion modes, with lower differences seen in swimming than in cycling and running for both elite and non-elite triathletes. In elite triathletes, sex differences in performance ($\sim$10–14 %) are in agreement with values generally observed in endurance sports. At present, the performance density in highly competitive international triathlon races appears to have become similar between elite male and female triathletes. Marathon running during an Ironman triathlon appears to be the discipline where female elite triathletes have reduced the gap with the males to the greatest extent over the recent years. Sex-related differences increase with advancing age most likely due to physiological, sociological and psychological changes. However, these differences should decrease in the future, with the increase in participation of elderly female triathletes.

Age-related declines in triathlon performance depend on the locomotion mode, the exercise duration (short- vs. long-distance triathlon) and the triathlon format (off-road vs. road-based triathlon). The participation of older triathletes has increased over the past 25 years and will probably continue to grow in the future. This increase in participation has been accompanied by an improvement in masters triathlon performance. The question of whether older triathletes have yet reached limits in their performance should therefore be examined in more detail. Further studies investigating training regimes, competition experience or socio-demographic factors are needed to gain better insights into the phenomenon of the relative improvements in endurance and ultra-endurance performance with advancing age.

Acknowledgments This research project was not funded. All authors declare that there were no conflicts of interests.

References

1. Bentley DJ, Millet GP, Vleck VE, MacNaughton LR. Specific aspects of contemporary triathlon. Sports Med. 2002;32:1–15.
2. Lepers R. Analysis of Hawaii ironman performances in elite triathletes from 1981 to 2007. Med Sci Sports Exerc. 2008;40:1828–34.
3. International Triathlon Union. Olympic history. http://www.triathlon.org/olympics/history.
4. Knechtle B, Knechtle P, Lepers R. Participation and performance trends in ultra-triathlons from 1985 to 2009. Scand J Med Sci Sports. 2011;e82–90.
5. Lenherr R, Knechtle B, Rüst CA, Rosemann T. Lepers R. From double iron to double deca iron ultra-triathlon—a retrospective data analysis from 1985 to 2011. Physical Sport Culture and Sport Studies Res. 2012;54:55–67.
6. Lepers R, Knetchle B, Knechtle P, Rosemann T. Analysis of ultra-triathlon performances. Open Access J Sports Med. 2011;2:131–6.
7. Rust CA, Knechtle B, Knechtle P, Lepers R, Rosemann T, Onywera V. European athletes dominate performances in double iron ultra-triathlons—a retrospective data analysis from 1985 to 2010. Eur J Sport Sci. http://dx.doi.org/10.1080/17461391.2011.641033.
8. Lepers R, Stapley P. Age-related changes in conventional road versus off-road triathlon performance. Eur J Appl Physiol. 2011;111(8):1687–94.
9. Lepers R, Stapley P. Differences in gender and performance in off-road triathlon. J Sports Sci. 2010;28(14):1555–62.
10. O'Toole ML, Douglas PS. Applied physiology of triathlon. Sports Med. 1995;19(4):251–67.
11. Hausswirth C, Lehénaff D. Physiological demands of running during long distance runs and triathlons. Sports Med. 2001;31(9):679–89.
12. Laursen PB, Rhodes EC. Factors affecting performance in an ultraendurance triathlon. Sports Med. 2001;31(3):195–209.

13. Millet GP, Dréano P, Bentley DJ. Physiological characteristics of elite short- and long-distance triathletes. Eur J Appl Physiol. 2003;88(4–5):427–30.
14. Millet GP, Bentley DJ. The physiological responses to running after cycling in elite junior and senior triathletes. Int J Sports Med. 2004;25(3):191–7.
15. Le Meur Y, Thierry B, Rabita G, Dorel S, Honnorat G, Brisswalter J, Hausswirth C. Spring-mass behaviour during the run of an international triathlon competition. Int J Sports Med. 2013 (Epub ahead of print).
16. Millet GP, Vleck VE. Physiological and biomechanical adaptations to the cycle to run transition in Olympic triathlon: review and practical recommendations for training. Br J Sports Med. 2000;34(5):384–90 (Review).
17. Millet GP, Millet GY, Hofmann MD, Candau RB. Alterations in running economy and mechanics after maximal cycling in triathletes: influence of performance level. Int J Sports Med. 2000;21(2):127–32.
18. O'Toole ML. Training for ultraendurance triathlons. Med Sci Sports Exerc. 1989;21:S209–13.
19. Laursen PB. Long distance triathlon: demands, preparation and performance. J Hum Sport Exerc. 2011;6(2):247–63.
20. Hausswirth C, Brisswalter J. Strategies for improving performance in long duration events: Olympic distance triathlon. Sports Med. 2008;38(11):881–91.
21. Jeukendrup AE, Jentjens RL, Moseley L. Nutritional considerations in triathlon. Sports Med. 2005;35(2):163–81.
22. Dallam GM, Jonas S, Miller TK. Medical considerations in triathlon competition: recommendations for triathlon organizers, competitors and coaches. Sports Med. 2005;35(2):143–61.
23. Desgorces FD, Berthelot G, El Helou N, Thibault V, Guillaume M, Tafflet M, Hermine O, Toussaint JF. From Oxford to Hawaii ecophysiological barriers limit human progression in ten sport monuments. PLoS One. 2008;3(11):e3653.
24. Lepers R, Cattagni T. Do older athletes reach limits in their performance during marathon running? Age (Dordr). 2012;34:773–81.
25. Jokl P, Sethi PM, Cooper AJ. Master's performance in the New York City Marathon 1983–1999. Br J Sports Med. 2004;38(4):408–12.
26. Hunter SK, Stevens AA. Sex differences in marathon running with advanced age: physiology or participation? Med Sci Sports Exerc. 2013;45(1):148–56.
27. Lepers R, Rüst C, Stapley P, Knechtle B. Relative improvements in endurance performance with age: evidence from 25 years of Hawaii ironman racing. Age (Dordr). 2013;35(3):953–62.
28. USA Triathlon. Triathlon participation, growth trends and demographics. Retrieved 13August, 2012, from http://www.usatriathlon.org/about-multisport/demographics.aspx.
29. Deaner RO. Distance running as an ideal domain for showing a sex difference in competitiveness. Arch Sex Behav. 2013;42(3):413–28.
30. Etter F, Knechtle B, Bukowski A, Rüst CA, Rosemann T, Lepers R. Age and gender interactions in short distance triathlon performance. J Sport Sci. (in press).
31. Knechtle B, Rüst CA, Roseman T, Lepers R. Age- and gender-related differences in half-Ironman triathlon performances—the Ironman 70.3 Switzerland from 2007 to 2010. Open Access J Sports Med. 2012;3:59–66.
32. Rüst CA, Knechtle B, Knechtle P, Rosemann T, Lepers R. Age of peak performance in elite male and female Ironman triathletes competing in a qualifier for 'Ironman Hawaii'—Ironman Switzerland from 1995–2011. Open Access J Sports Med. 2012;3:175–82.
33. Rüst CA, Knechtle B, Knechtle P, Rosemann T, Lepers R. Participation and performance in Triple Iron ultra-triathlon—a cross-sectional and longitudinal data analysis. Asian J Sports Med. 2012;3:145–52.
34. Lepers R, Maffiuletti NA. Age and gender interactions in ultra-endurance performance: insight from triathlon. Med Sci Sports Exerc. 2011;43:134–9.
35. Hoffman MD, Ong JC, Wang G. Historical analysis of participation in 161 km ultramarathons in North America. Int J Hist Sport. 2010;27:1877–91.
36. Reaburn PRJ, Dascombe BJ, Janse de Jonge X. Body composition and gender differences in performance. In: Driskell JA, Wolinsky I. Nutritionnal assesement of athletes, 2nd ed. CRC Press, Boca Raton; 2011, p. 121–147.
37. Sparling PB. A meta-analysis of studies comparing maximal oxygen uptake in men and women. Res Q Exerc Sport. 1980;51(3):542–52.
38. Joyner MJ. Physiological limiting factors and distance running: influence of gender and age on record performances. Exerc Sport Sci Rev. 1993;21:103–33.
39. Bunc V, Heller J, Horcic J, Novotny J. Physiological profile of best Czech male and female young triathletes. J Sports Med Phys Fitness. 1996;36(4):265–70.
40. Shaskey DJ, Green GA. Sports haematology. Sports Med. 2000;29(1):27–38.
41. Wiswell RA, Hawkins SA, Jaque SV, Hyslop D, Constantino N, Tarpenning K, Marcell T, Schroeder ET. Relationship between physiological loss, performance decrement, and age in master athletes. J Gerontol A Biol Sci Med Sci. 2001;56(10):M618–26.
42. Yasuda N, Gaskill SE, Ruby BC. No gender-specific differences in mechanical efficiency during arm or leg exercise relative to ventilatory threshold. Scand J Med Sci Sports. 2008;18(2):205–12.
43. Kimber NE, Ross JJ, Mason SL, Speedy DB. Energy balance during an ironman triathlon in male and female triathletes. Int J Sport Nutr Exerc Metab. 2002;12(1):47–62.
44. Landers GJ, Ong KB, Ackland TR, Blanksby BA, Main LC, Smith D. Kinanthropometric differences between 1997 World championship junior elite and 2011 national junior elite triathletes. J Sci Med Sport. 2012. doi:10.1016/j.jsams.2012.09.006.
45. Knechtle B, Wirth A, Baumann B, Knechtle P, Rosemann T, Oliver S. Differential correlations between anthropometry, training volume, and performance in male and female Ironman triathletes. J Strength Cond Res. 2010;24(10):2785–93.
46. Knechtle B, Wirth A, Baumann B, Knechtle P, Rosemann T. Personal best time, percent body fat, and training are differently associated with race time for male and female ironman triathletes. Res Q Exerc Sport. 2010;81(1):62–8.
47. VanHeest JL, Mahoney CE. Female athletes: factors impacting successful performance. Curr Sports Med Rep. 2007;6(3):190–4.
48. Zaryski DJ. Smith, training principles and issues for ultra-endurance athletes. Curr Sports Med Rep. 2005;4(3):165–70.
49. Cheuvront SN, Carter III R, DeRuisseau KC, Moffart RJ. Running performance differences between men and women: an update. Sports Med. 2005;35:1017–24.
50. Sparkling PB, O'Donnell EM, Snow TK. The gender difference in distance running performance has plateaued: an analysis of world rankings from 1980 to 1996. Med Sci Sports Exerc. 1998;30:1725–9.
51. Speechly DP, Taylor SR, Rogers GG. Differences in ultraendurance exercise in performance-matched male and female runners. Med Sci Sports Exerc. 1996;28(3):359–65.
52. Hoffman MD. Ultramarathon trail running comparison of performance-matched men and women. Med Sci Sports Exerc. 2008;40(9):1681–6.
53. Whipp BJ, Ward SA. Will women soon outrun men? [Letter]. Nature. 1992;335(6355):25.

54. Hoffman MD. Performance trends in 161-km ultramarathons. Int J Sports Med. 2010;31(1):31–7.
55. Hoffman MD, Wegelin JA. The Western States 100-Mile Endurance Run: participation and performance trends. Med Sci Sports Exerc. 2009;41(12):2191–8.
56. Rüst CA, Knechtle B, Knechtle P, Pfeifer S, Rosemann T, Lepers R, et al. Gender difference and age-related changes in performance at the long distance duathlon World Championships. J Strength Cond Res. 2013;27(2):293–301.
57. Rüst CA, Knechtle B, Rosemann T, Lepers R. Sex difference in race performance and age of peak performance in the Ironman Triathlon World Championship from 1983 to 2012. Extreme Physiol Med. 2012;1:15.
58. Sultana F, Brisswalter J, Lepers R, Hausswirth C, Bernard T. Effects of age and gender on Olympic triathlon performances. Sci Sport. 2008;23:130–5.
59. Tanaka H, Seals DR. Age and gender interactions in physiological functional capacity: insight from swimming performance. J Appl Physiol. 1997;82(3):846–51.
60. Eichenberger E, Knechtle B, Knechtle P, Rüst CA, Rosemann T, Lepers R. Best performances by men and women open-water swimmers during the 'English Channel Swim' from 1900 to 2010. J Sports Sci. 2012;30(12):1295–301.
61. Eichenberger E, Knechtle B, Knechtle P, Rüst CA, Rosemann T, Lepers R, et al. Sex difference in open-water ultra-swim performance in the longest freshwater lake swim in Europe. J Strength Cond Res. 2013;23(1):e48–55.
62. Fischer G, Knechtle B, Rüst CA, Rosemann T. Male swimmers cross the English Channel faster than female swimmers. Epub: Scand J Med Sci Sports; 2012.
63. Toussaint HM. Differences in propelling efficiency between competitive and triathlon swimmers. Med Sci Sports Exerc. 1990;22(3):409–15.
64. Lavoie JM, Montpetit RR. Applied physiology of swimming. Sports Med. 1986;3:165–89.
65. McLean SP, Hinrichs RN. Sex differences in the centre of buoyancy location of competitive swimmers. J Sports Sci. 1998;16:373–83.
66. Pendergast DR, Di Prampero PE, Craig AB, Wilson DR, Rennie DW. Quantitative analysis of the front crawl in men and women. J Appl Physiol. 1977;43:475–9.
67. Pate RR, Sparling PB, Wilson GE, Cureton KJ, Miller BJ. Cardiorespiratory and metabolic responses to submaximal and maximal exercise in elite women distance runners. Int J Sports Med. 1987;8:91–5.
68. Zamparo P. Effects of age and gender on the propelling efficiency of the arm stroke. Eur J Appl Physiol. 2006;97(1):52–8.
69. Toussaint HM, de Groot G, Savelberg HH, Vervoorn K, Hollander AP, van Ingen Schenau GJ. Active drag related to velocity in male and female swimmers. J Biomech. 1988;21(5):435–8.
70. Schumacher YO, Mueller P, Keul J. Development of peak performance in track cycling. J Sports Med Phys Fitness. 2001;41:139–46.
71. Rüst CA, Knechtle B, Rosemann T, Lepers R. Men cross America faster than women—the 'Race Across America' (RAAM) from 1982 to 2012. Int J Sports Physiol Perform. 2013 (Epub ahead of print).
72. Abou Shoak M, Knechtle B, Knechtle P, Rüst CA, Rosemann T, Lepers R. Participation and performance trends in ultra-cycling. Open Access J Sports Med. 2013;4:41–51.
73. Levis DA, Kamon E, Hodgson JL. Physiological differences between genders: implications for sports conditioning. Sports Med. 1986;3:357–69.
74. Lucia A, Joyos H, Chicharro JL. Physiological response to professional road cycling: climbers vs time trialists. Int J Sports Med. 2000;21:505–12.
75. Impellizzeri FM, Marcora SM. The physiology of mountain biking. Sports Med. 2007;37(1):59–71.
76. Bernard T, Hausswirth C, Le Meur Y, Bignet F, Dorel S, Brisswalter J. Distribution of power output during the cycling stage of a triathlon world cup. Med Sci Sports Exerc. 2009;41(6):1296–302.
77. Le Meur Y, Hausswirth C, Dorel S, Bignet F, Brisswalter J, Bernard T. Influence of gender on pacing adopted by elite triathletes during a competition. Eur J Appl Physiol. 2009;106(4):535–45.
78. Hopkins WG, Schabort EJ, Hawley JA. Reliability of power in physical performance tests. Sports Med. 2001;31:211–34.
79. Stevenson J, Song H, Cooper JA. Age and sex differences pertaining to modes of locomotion in triathlon. Med Sci Sports Exerc. 2013;45(5):976–84.
80. Seiler S, De Koning JJ, Foster C. The fall and rise of the gender difference in elite anaerobic performance 1952–2006. Med Sci Sports Exerc. 2007;39:534–40.
81. Lepers R. Age and gender considerations in triathlon. In: Friel J, Vance J, editors. Triathlon Science. Human kinetics, Inc; 2013. p. 39–53.
82. Mayhew JL, Salm PC. Gender differences in anaerobic power tests. Eur J Appl Physiol Occup Physiol. 1990;60(2):133–8.
83. Tanaka H, Seals DR. Endurance exercise performance in Masters athletes: age-associated changes and underlying physiological mechanisms. J Physiol. 2008;586(1):55–63.
84. Ransdell LB, Vener J, Huberty J. Master athletes: an analysis of running, swimming and cycling performance by age and gender. J Exerc Sci Fit. 2009;7:S61–73.
85. Tanaka H, Seals DR. Invited review: dynamic exercise performance in Masters athletes: insight into the effects of primary human aging on physiological functional capacity. J Appl Physiol. 2003;95(5):2152–62.
86. Maharam LG, Bauman PA, Kalman D, Skolnik H, Perle SM. Masters athletes: factors affecting performance. Sports Med. 1999;28:273–85.
87. Deaner RO. Physiology does not explain all sex differences in running performance. Med Sci Sports Exerc. 2013;45(1):146–7.
88. Reaburn P, Dascombe B. Endurance performance in masters athletes. Eur Rev Aging Phys Act. 2008;5:31–42.
89. Leyk D, Erley O, Ridder D, Leurs M, Rüther T, Wunderlich M, et al. Age-related changes in marathon and half-marathon performances. Int J Sports Med. 2007;28(6):513–7.
90. Knechtle B, Rüst CA, Rosemann T, Lepers R. Age-related changes in 100-km ultra-marathon running performances. Age (Dordr). 2012;34(4):1033–45.
91. Lepers R, Stapley PJ, Cattagni T, Gremeaux V, Knechtle B. Limits in endurance performance of octogenarian athletes. J Appl Physiol. 2013;114(6):829.
92. Baker AB, Tang YQ. Aging performance for masters records in athletics, swimming, rowing, cycling, triathlon, and weightlifting. Exp Aging Res. 2010;36:453–77.
93. Balmer J, Bird S, Davison R, Lucia A. Effect of age on 16.1-km time-trial performance. J Sports Sci. 2008;26(2):197–206.
94. Balmer J, Bird S, Davison R. Indoor 16.1-km time-trial performance in cyclists aged 25–63 years. J Sports Sci. 2008;26(1):57–62.
95. Balmer J, Potter CR, Bird SR, Davison RC. Age-related changes in maximal power and maximal heart rate recorded during a ramped test in 114 cyclists age 15–73 years. JAPA. 2005;13:125–36.
96. Bernard T, Sultana F, Lepers R, Hausswirth C, Brisswalter J. Age related decline in Olympic triathlon performance: effect of locomotion mode. Exp Aging Res. 2010;36:1–15.

97. Knechtle B, Rüst CA, Knechtle P, Rosemann T, Lepers R. Age-related changes in ultra-triathlon performances. Extreme Physiol Med. 2012;1:5.
98. Lepers R, Sultana F, Bernard T, et al. Age-related changes in triathlon performances. Int J Sports Med. 2010;31(4):251–6.
99. Hunter SK, Stevens AA, Magennis K, Skelton KW, Fauth M. Is there a sex difference in the age of elite marathon runners? Med Sci Sports Exerc. 2011;43(4):656–64.
100. Kallinen M, Markku A. Aging, physical activity and sports injuries: an overview of common sports injuries in the elderly. Sports Med. 1995;20:41–52.
101. Lepers R, Maffiuletti NA, Rochette L, Brugniaux J, Millet GY. Neuromuscular fatigue during a long-duration cycling exercise. J Appl Physiol. 2002;92:1487–93.
102. Bonnefoy M, Kostka T, Arsac LM, Berthouze SE, Lacour JR. Peak anaerobic power in elderly men. Eur J Appl Physiol. 1998;77:182–8.
103. Kostka T. Quadriceps maximal power and optimal shortening velocity in 335 men aged 23–88 years. Eur J Appl Physiol. 2005;95:140–55.
104. McCrory JL, Salacinski AJ, Hunt SE, Greenspan SL. Thigh muscle strength in senior athletes and healthy controls. J Strength Cond Res. 2009;23:2430–6.
105. Stiefel M, Knechtle B, Lepers R. Master triathletes have not reached limits in their Ironman triathlon performance. Scand J Med Sci Sports (Epub 2012 May 14). doi:10.1111/j.1600-0838.2012.01473.x.
106. Medic N, Starkes JL, Young BW. Examining relative age effects on performance achievement and participation rates in Masters athletes. J Sports Sci. 2007;25(12):1377–84.
107. Gulbin JP, Gaffney PT. Ultraendurance triathlon participation: typical race preparation of lower level triathletes. J Sports Med Phys Fitness. 1999;39(1):12–5.
108. Knechtle B, Knechtle P, Rüst CA, Rosemann T. A comparison of anthropometric and training characteristics of Ironman triathletes and Triple Iron ultra-triathletes. J Sports Sci. 2011;29(13):1373–80.
109. Knechtle B, Wirth A, Rosemann T. Predictors of race time in male ironman triathletes: physical characteristics, training, or prerace experience? Percept Mot Skills. 2010;111(2):437–46.
110. Leyk D, Erley O, Gorges W, Ridder D, Rüther T, Wunderlich M, et al. Performance, training and lifestyle parameters of marathon runners aged 20–80 years: results of the PACE study. Int J Sports Med. 2009;30:360–5.
111. Wright VJ, Perricelli BC. Age-related rates of decline in performance among elite senior athletes. Am J Sports Med. 2008;36(3):443–50.
112. Gremeaux V, Gayda M, Lepers R, Sosner P, Juneau M, Nigam A. Exercise and longevity. Maturitas. 2012;73(4):312–7.
113. Trappe S, Hayes E, Galpin A, Kaminsky L, Jemiolo B, Fink W, Trappe T, Jansson A, Gustafsson T, Tesch P. New records in aerobic power among octogenarian lifelong endurance athletes. J Appl Physiol. 2013;114(1):3–10.

PLOS ONE

RESEARCH ARTICLE

Maturation-, age-, and sex-specific anthropometric and physical fitness percentiles of German elite young athletes

Melanie Lesinski[1], Alina Schmelcher[1], Michael Herz[1], Christian Puta[2], Holger Gabriel[2], Adamantios Arampatzis[3,4], Gunnar Laube[3,4], Dirk Büsch[5], Urs Granacher[1]*

1 Division of Training and Movement Sciences, University of Potsdam, Research Focus Cognition Sciences, Potsdam, Germany, **2** Department of Sports Medicine and Health Promotion, Friedrich Schiller University, Jena, Germany, **3** Department of Training and Movement Sciences, Humboldt-Universität zu Berlin, Berlin, Germany, **4** Berlin School of Movement Science, Berlin, Germany, **5** Institute of Sports Science, Carl von Ossietzky University of Oldenburg, Oldenburg, Germany

* urs.granacher@uni-potsdam.de

OPEN ACCESS

Citation: Lesinski M, Schmelcher A, Herz M, Puta C, Gabriel H, Arampatzis A, et al. (2020) Maturation-, age-, and sex-specific anthropometric and physical fitness percentiles of German elite young athletes. PLoS ONE 15(8): e0237423. https://doi.org/10.1371/journal.pone.0237423

Editor: Daniel Boullosa, Universidade Federal de Mato Grosso do Sul, BRAZIL

Received: February 10, 2020

Accepted: July 14, 2020

Published: August 13, 2020

Data Availability Statement: All relevant data are within the manuscript and its Supporting Information files.

Funding: This study is part of the research project "Resistance Training in Young athletes" that was funded by the German Federal Institute of Sport Science (ZMVI1-081901 14-18, ZMVI4-081901/20-23). The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Abstract

The aim of this study was to establish maturation-, age-, and sex-specific anthropometric and physical fitness percentile reference values of young elite athletes from various sports. Anthropometric (i.e., standing and sitting body height, body mass, body mass index) and physical fitness (i.e., countermovement jump, drop jump, change-of-direction speed [i.e., T-test], trunk muscle endurance [i.e., ventral Bourban test], dynamic lower limbs balance [i.e., Y-balance test], hand grip strength) of 703 male and female elite young athletes aged 8–18 years were collected to aggregate reference values according to maturation, age, and sex. Findings indicate that body height and mass were significantly higher ($p<0.001$; $0.95 \leq d \leq 1.74$) in more compared to less mature young athletes as well as with increasing chronological age ($p<0.05$; $0.66 \leq d \leq 3.13$). Furthermore, male young athletes were significantly taller and heavier compared to their female counterparts ($p<0.001$; $0.34 \leq d \leq 0.50$). In terms of physical fitness, post-pubertal athletes showed better countermovement jump, drop jump, change-of-direction, and handgrip strength performances ($p<0.001$; $1.57 \leq d \leq 8.72$) compared to pubertal athletes. Further, countermovement jump, drop jump, change-of-direction, and handgrip strength performances increased with increasing chronological age ($p<0.05$; $0.29 \leq d \leq 4.13$). In addition, male athletes outperformed their female counterpart in the countermovement jump, drop jump, change-of-direction, and handgrip strength ($p<0.05$; $0.17 \leq d \leq 0.76$). Significant age by sex interactions indicate that sex-specific differences were even more pronounced with increasing age. Conclusively, body height, body mass, and physical fitness increased with increasing maturational status and chronological age. Sex-specific differences appear to be larger as youth grow older. Practitioners can use the percentile values as approximate benchmarks for talent identification and development.

Competing interests: The authors have declared that no competing interests exist.

Introduction

There are primarily two pathways which are pursued to develop a gifted young child into a talented elite athlete. These are early specialization and diversification [1]. While both pathways have proven to be successful in developing high performance athletes, more recent evidence has shown that there is an increased risk with early specialization to sustain acute and/or overuse injuries which may ultimately lead to drop out from organized sports [2, 3]. Diversification on the other hand has proven to be particularly successful with cgs (centimeters, grams, seconds) sports in terms of developing successful elite athletes [4]. A premise of the diversification approach is to lay a foundation of physical fitness before developing sport-specific performance [5]. In other words, fitness development precedes sport specialization. Accordingly, reference values are needed for this first step of the diversification approach to evaluate physical fitness levels of youth athletes, irrespective of the sport they practice. To pursue this promising approach, reference data are needed of various physical fitness tests to support coaches and athletes with talent identification, selection, and development [6]. This will help to better monitor and guide the development of talented youth athletes on their early diversification pathway. For the general youth population, studies exist with large cohorts that provide age- and sex-specific percentile norm values for different anthropometric and physical fitness outcomes (e.g., hand grip strength, 1-kg ball push, standing long jump, 50-m sprint, shuttle run test) [7–9]. These studies suggest that anthropometry and physical fitness develop with increasing age in a sex-specific non-linear fashion. However, these data sets cannot be utilized with young athletes because by definition, talented youth are equal to or above the 90th percentile of the respective general population [10, 11]. Superior performance of young sporting talents is due to both, nature (i.e., genes) but also nurture (e.g., regular training over several years) [12]. Accordingly, it is timely and imperative to establish cohort specific reference values that can be used for talent identification, selection, and development. However, previous studies with young athletes reported only age-, sex-, and/or sport-specific mean values for different physical fitness outcomes [13–15]. For instance, Opstoel et al. [14] determined mean values of 620 children aged 9 to 11 who participated in at least one specific sport (i.e., in total 25 different sports) for several physical fitness outcomes (e.g., hand grip strength, countermovement jump [CMJ], standing long jump, shuttle run test). Yet, there is currently no study available that has established maturation-specific anthropometric and physical fitness percentiles of elite young athletes. Of note, maturation is a non-linear process, which is why there is often a discrepancy between chronological age and maturation among young athletes [16–18]. This is a major challenge in youth sport where competitions are mainly regulated by chronological age-groups.

To the authors' knowledge, there are no studies available that provide maturation-, age-, and sex-specific anthropometric and physical fitness percentiles for young athletes. Therefore, the purpose of this cross-sectional study was to present and discuss maturation, age, and sex-specific anthropometric (e.g., body height, body mass) and physical fitness (e.g., CMJ, drop jump [DJ], change-of-direction [CoD] speed) percentile reference values of young elite athletes from various sports. With reference to the relevant literature [16, 19, 20], we hypothesized that body height, body mass, and physical fitness increase with age and maturation in a sex-specific but non-linear fashion.

Materials and methods

Participants

A convenience sample of 703 male (♂ = 420) and female (♀ = 283) young elite athletes aged 8–18 years who originally participated in a large research project entitled "Resistance Training

in Young Athletes" (https://www.uni-potsdam.de/kraftprojekt/english.php) was used to aggregate anthropometric and physical fitness reference values. Athletes were from 18 different sports including soccer (41 ♀/ 49 ♂), volleyball (24 ♀/24 ♂), basketball (10 ♀/29 ♂), handball (29 ♀/84 ♂), judo (36 ♀/53 ♂), wrestling (3 ♀/6 ♂), boxing (3 ♀/7 ♂), canoeing (21 ♀/42 ♂), rowing (21 ♀/0 ♂), ski jumping (0 ♀/18 ♂), nordic combination (5 ♀/15 ♂), speed skating (13 ♀/8 ♂), swimming (37 ♀/26 ♂), weight lifting (3 ♀/7 ♂), badminton (13 ♀/8 ♂), gymnastics (0 ♀/18 ♂), athletics (13 ♀/11 ♂), and modern pentathlon (12 ♀/14 ♂). Participants were recruited from German elite sport schools and followed a training regime consisting of regular physical education together with their sport-specific training and competitions. On average, participating athletes practiced their sport between 2 and 12 years. All athletes were enrolled at elite sport schools and performed a minimum of 10 hours of training per week. Each participant was coded for his/her maturity status, age, and sex. Maturity was determined by calculating the time from peak-height-velocity (PHV) according to the regression equations of Mirwald et al. [21] for boys:

Maturity offset = -9.236 + (0.0002708 * leg length * sitting height)—(0.001663 * age * leg length) + (0.007216 * age* sitting height) + (0.02292 * weight by height ratio) and girls:

Maturity offset = -9.376 + (0.0001882 * leg length * sitting height) + (0.0022 * age * leg length) + (0.005841·age * sitting height)—(0.002658·age * weight) + (0.07693 * weight by height ratio).

In accordance with Faigenbaum et al. [22], maturity was classified as pre-pubertal (i.e., < -1 year before PHV), pubertal (i.e., ±1 years around PHV), and post-pubertal (i.e., > 1 year after PHV). Prior to the start of the study, all participants were informed about potential risks and benefits of the study and athletes as well as their legal guardians provided their written informed consent. The protocol was approved by local ethical commissions (University of Potsdam: submission No. 5/2014; Charité Berlin: EA2/076/15; Friedrich-Schiller-University Jena: 458510/15).

Testing procedures

Baseline data of anthropometric and physical fitness tests were used from intervention studies of a larger research project to aggregate reference values. All anthropometric and physical fitness tests were performed under strictly standardized conditions. In these studies, baseline testing always started with the assessment of anthropometrics (i.e., standing and sitting height, body mass). Tests were always conducted in the morning before fitness testing. According to Mirwald et al. [21], standing and sitting height were measured to the nearest mm. Two measurements were taken for each anthropometric variable and averaged for analysis. A third measurement was required if the first two differed by more than 4 mm for standing or sitting height. Prior to physical fitness testing, a standardized warm-up protocol (i.e., ten minutes of jumping, running and agility/change-of-direction drills) was performed. The physical fitness test battery included the assessment of vertical jump performance (i.e., DJ, CMJ), CoD speed (i.e., T-test), dynamic balance of the lower extremities (i.e., Y balance-test), trunk muscle endurance (i.e., ventral Bourban-test), and hand grip strength. Participants were familiarized with all physical fitness tests prior to data assessment. Hand and leg dominance were determined using the lateral preference inventory [23].

Assessment of hand grip strength. Grip strength of the dominant hand was measured using a hand-held dynamometer (Jamar Plus, Performance Health, Warrenville, IL, USA) which showed good test-retest (ICC > 0.80) and inter-rater reliability (ICC > 0.97) [24]. During testing, participants were seated upright, elbows by the side of the body and flexed at an angle of 90˚. Participants were instructed to press the dynamometer grip as forcefully as

possible for 5 s while maintaining their position (i.e., no additional movements from upper or lower body). Three trials were conducted and the best trial was used for further analysis.

Assessment of jump performance. CMJ and DJ performances were measured using an optoelectric cell system (Optojump, Microgate, Bolzano, Italy). Excellent test–retest reliability (intraclass correlation coefficient [ICC]) was previously reported for the estimation of vertical jump height using the Optojump photocell system (ICC = 0.98) [25]. For CMJ, athletes stood in an upright erect standing position, feet shoulder-width apart, and hands akimbo. CMJs were initiated with a countermovement which was immediately followed by a concentric explosive upward movement. For DJ, participants stood in an upright erect standing position on a 40 cm box, feet shoulder-width apart, and hands akimbo. Participants were asked to step off the box with their dominant leg, drop down to land evenly on both feet on the ground, keep ground contact time short, and jump-off the ground with a double-leg vertical jump at maximal effort. All participants jumped with shoes as well as were instructed to jump as high as possible (CMJ, DJ) and to keep ground contact as short as possible (DJ). Three trials were conducted for each jump test. The best trial in terms of maximal jump height (CMJ, DJ) was taken for further analysis. Furthermore, participants' performance index was calculated using the following formula: DJ performance index = drop jump height [m] / contact time [s]. The best trial in terms of maximal DJ performance index was taken for further analysis.

Assessment of trunk muscle endurance. Endurance of the trunk muscles was assessed using the ventral Bourban-test. The test has previously proven to be valid as well as reliable with an ICC of 0.87 [26]. During test performance, athletes are in plank position, elbows shoulder-with apart, forearms flat on the ground and legs extended. A reference rod of the alignment device touched the athlete´s lower back at the iliac crests. In this position, athletes were asked to lift their feet (2–5 cm) alternately (i.e., 1 s per foot) for as long as possible according to the beat of a metronome. If athletes lost contact with the reference rod for longer than 2–4 seconds, they received a warning from the test instructor. Test time until failure was recorded using a stopwatch and taken as dependent variable. Alternatively, test time to the third warning was used for further analysis.

Assessment of change-of-direction speed. Change-of-direction speed (CoD) was assessed using the T-test [27], which showed high test-retest reliability with an ICC of 0.98 [27]. Athletes had to complete a course, set up as a "T" using four cones, in the shortest possible time. For this purpose, they sprinted forward, performed sidesteps and ran backwards. The athletes started without a start signal and sprint time was measured using a double-light electronic gate system (WITTY; Microgate Srl, Bolzano, Italy). Following a submaximal test trial, the fastest out of two trials was taken for further analysis.

Assessment of dynamic balance. The lower quarter Y balance-test was used to assess dynamic balance [28]. According to Plisky et al. [28], ICC values for the three different movement directions ranged between 0.89 and 0.93 and showed high test-retest reliability. Athletes were barefooted and positioned in single leg stance on the Y-Balance-Test-Kit (Move2Perform, Evansville, IN, USA). They were asked to reach with the contralateral leg as far as possible into three different movement directions (i.e., ventral, posteromedial, posterolateral). Athletes always started the test while they stood on the right leg. With the left leg, participants had to reach three times in one direction before switching sides and directions. For familiarization purposes, all athletes completed three trials per leg and per movement direction before the tests started. The best performance (furthest reach) in each direction was used for further analysis. According to Filipa et al. [29], a composite score was calculated according to the equation: composite score = [(maximum anterior reach distance + maximum posteromedial reach distance + maximum posterolateral reach distance)/(leg length × 3)] × 100 and taken as dependent variable for further analysis. Of note, leg length was assessed by measuring the

distance between the anterior superior iliac spine and the most distal aspect of the medial malleolus while the athlete lies in supine position.

Statistical analyses

Data are mean values and standard deviations (SDs) with 95% confidence intervals for anthropometrics and physical fitness. After data were tested and confirmed for normal distribution (i.e., Shapiro Wilk test), an univariate ANOVA was applied with the factors sex, maturity status, and age as between subject comparators. Bonferroni corrected post-hoc tests were computed for multiple comparisons to determine outcomes according to maturation and age. The level of significance was set at $p < 0.05$ for each comparison. In addition, the classification of effect sizes was determined by calculating Cohen's d from partial eta-squared. Effect sizes constitute a means to determine whether a difference is a difference of practical concern. According to Cohen [30], effect sizes can be classified as small ($d < 0.5$), medium ($0.5 \leq d < 0.8$), or large ($d \geq 0.8$). Percentile analyses were computed separately for boys and girls according to maturation and age. The 20^{th}, 40^{th}, 50^{th}, 60^{th}, and 80^{th} percentiles were calculated. Due to the limited overall data pool of elite young athletes [31] and in accordance with other authors [32, 33], anthropometric and physical fitness differences as well as percentile reference values were only calculated if 30 participants were available within a subgroup. All analyses were conducted using IBM SPSS Statistics for Windows, Version 26.0 (IBM Corp., Armonk, NY, USA).

Results

Anthropometry and physical fitness differences by maturity status, age, and sex

Tables 1 and 2 contain sex-specific anthropometric and physical fitness values according to chronological age (Table 1) and maturity status (Table 2).

Effects of chronological age. Significant main effects of chronological age were found for all anthropometric and physical fitness test values ($p < 0.05$; $0.29 \leq d \leq 1.08$), except for the Y-balance test ($p > 0.05$; $d = 0.19$) (Table 1). Post-hoc analyses indicated significantly higher body height and mass with increasing age ($p < 0.05$; $0.66 \leq d \leq 3.13$), except for 13 and 14 years old athletes ($p > 0.05$; $0.14 \leq d \leq 0.25$). Furthermore, post-hoc analyses indicated significantly better hand grip strength, CMJ, DJ, and CoD performances with increasing age ($p < 0.05$; $0.40 \leq d \leq 4.27$). Notably, jump performance did not increase considerably between 12 and 13 year old athletes ($p > 0.05$; $0.18 \leq d \leq 0.39$). Furthermore, hand grip strength did not improve considerably in athletes aged 13 and 14 years ($p > 0.05$; $0.09 \leq d \leq 0.25$).

Effects of maturity status. Significant main effects of maturity were found for all anthropometric and physical fitness tests ($p < 0.01$; $0.26 \leq d \leq 1.57$), except for the Y-balance and the Bourban test ($p > 0.05$; $0.11 \leq d \leq 0.12$) (Table 2). Post-hoc analyses indicated that body height and mass were significantly higher ($p < 0.001$; $0.95 \leq d \leq 1.85$) in more matured young athletes (i.e., pre-pubertal < pubertal < post-pubertal). Further, post-pubertal compared to pubertal athletes showed significantly better performances in jump (i.e., CMJ height, DJ height, DJ performance index, DJ ground contact time) and CoD tests ($p < 0.001$; $1.57 \leq d \leq 3.13$).

Effects of sex. Furthermore, significant main effects of sex were found for anthropometric and physical fitness tests. More precisely, male young athletes were significantly taller and heavier compared with female young athletes ($p < 0.001$; $0.34 \leq d \leq 0.50$). Furthermore, males outperformed females in CMJ, DJ, CoD performances and hand grip strength ($p < 0.05$; $0.17 \leq d \leq 0.76$) (Tables 1 and 2).

Table 1. Anthropometric and physical fitness differences according to chronological age and sex in young athletes.

	chronological age								main/interaction effects		
	12		13		14		15		p-value (d)		
	boys	girls	boys	girls	boys	girls	boys	girls	age	sex	age x sex
	mean ± SD (95% CI) n = 55	mean ± SD (95% CI) n = 45	mean ± SD (95% CI) n = 91	mean ± SD (95% CI) n = 61	mean ± SD (95% CI) n = 69	mean ± SD (95% CI) n = 56	mean ± SD (95% CI) n = 76	mean ± SD (95% CI) n = 52	p (d)	p (d)	p (d)
ANTHROPOMETRY											
standing height [cm]	161.3 ± 11.7	162.2 ± 10.0	168.4 ± 10.7	165.7 ± 7.5	170.4 ± 10.0	166.9 ± 8.2	179.0 ± 9.9	172.2 ± 8.3	<0.001	<0.001	0.01
	(158.7–163.9)	(159.3–165.0)	(166.5–170.5)	(163.3–168.2)	(168.1–172.7)	(164.3–169.4)	(176.8–181.2)	(168.5–173.8)			-0.3
	n = 54	n = 45	n = 91	n = 59	n = 68	n = 56	n = 76	n = 52	-0.94	-0.34	
sitting height [cm]	83.2 ± 6.2	84.5 ± 4.9	86.7 ± 5.2	86.8 ± 4.1	87.4 ± 5.5	87.6 ± 3.9	91.3 ± 5.8	89.3 ± 3.8	<0.001	0.872	0.104
	(81.9–84.5.6)	(83.1–86.1)	(85.7–87.8)	(85.5–88.1)	(86.1–88.7)	(86.3–89.0)	(90.1–92.5)	(87.9–90.7)		-0.01	-0.3
	n = 55	n = 44	n = 90	n = 58	n = 60	n = 54	n = 67	n = 51	-0.86		
BMI [kg/m^2]	19.0 ± 3.0	19.3 ± 2.5	19.4 ± 2.9	20.2 ± 2.5	20.0 ± 2.7	20.9 ± 2.4	21.6 ± 3.1	20.9 ± 3.7	<0.001	0.538	0.035
	(18.1–19.9)	(18.4–20.3)	(18.8–20.2)	(19.4–21.1)	(19.2–20.8)	(20.2–21.8)	(20.9–22.4)	(19.9–21.8)		-0.06	-0.27
	n = 53	n = 45	n = 88	n = 58	n = 64	n = 54	n = 69	n = 50	-0.5		
body mass [kg]	50.2 ± 13.0	51.4 ± 10.7	55.7 ± 13.2	55.6 ± 9.4	58.4 ± 12.0	58.6 ± 8.4	69.5 ± 14.2	62.3 ± 8.5	<0.001	0.114	0.009
	(46.8–53.5)	(47.7–55.0)	(53.1–58.3)	(52.4–58.8)	(55.3–61.4)	(55.2–61.9)	(66.1–72.9)	(58.7–65.8)		-0.15	-0.31
	n = 53	n = 44	n = 88	n = 60	n = 64	n = 54	n = 69	n = 49	-0.88		
PHYSICAL FITNESS											
CMJ height [cm]	26.0 ± 5.6	23.6 ± 3.4	27.6 ± 4.3	25.0 ± 4.4	30.4 ± 7.2	26.5 ± 4.3	36.2 ± 8.8	27.4 ± 5.5	<0.001	<0.001	<0.001
	(24.2–25.3)	(21.9–25.3)	(26.4–28.8)	(23.5–26.5)	(28.8–31.9)	(25.0–28.0)	(34.7–37.6)	(25.8–29.1)			
	n = 43	n = 42	n = 86	n = 53	n = 51	n = 55	n = 61	n = 47	-0.89	-0.76	-0.45
DJ height [cm]	22.1 ± 5.4	22.5 ± 5.0	23.8 ± 4.9	22.4 ± 5.0	26.7 ± 6.6	23.4 ± 4.0	31.0 ± 7.3	24.8 ± 5.0	<0.001	<0.001	<0.001
	(20.5–23.8)	(20.9–24.2)	(22.6–24.9)	(21.0–23.9)	(25.1–28.2)	(21.9–24,8)	(29.6–32.3)	(23.3–26.3)			
	n = 43	n = 42	n = 86	n = 52	n = 51	n = 55	n = 62	n = 49	-0.77	-0.47	-0.43
DJ ground	251±101	242±53	232±61	215±42	230±67	209±29	204±35	211±35	<0.001	0.078	0.287
contact time [ms]	(220–283)	(226–258)	(218–245)	(203–227)	(218–245)	(201–217)	(195–249)	(201–221)		-0.01	-0.009
	n = 43	n = 42	n = 86	n = 52	n = 51	n = 55	n = 62	n = 49	-0.05		
DJ performance	0.97 ± 0.39	0.96 ± 0.28	1.07 ± 0.29	1.08 ± 0.30	1.27 ± 0.40	1.14 ± 0.26	1.59 ± 0.50	1.21 ± 0.36	<0.001	<0.001	<0.001
index [m/s]	(0.9–1.1)	(0.9–1.1)	(1.0–1.1)	(1.0–1.2)	(1.2–1.4)	(1.0–1.2)	(1.5–1.7)	(1.1–1.3)			
	n = 43	n = 42	n = 86	n = 52	n = 51	n = 55	n = 62	n = 49	-0.88	-0.35	-0.44
T-test [s]	11.71 ± 1.08	12.14 ± 0.71	10.98 ± 0.72	11.73 ± 0.94	10.82 ± 0.91	11.23 ± 0.96	10.19 ± 0.75	10.92 ± 0.93	<0.001	<0.001	0.309
	(11.4–12.0)	(11.9–12.4)	(10.8–11.2)	(11.5–12.0)	(10.6–11.1)	(11.0–11.5)	(10.0–10.4)	(10.7–11.2)			-0.19
	n = 34	n = 42	n = 75	n = 52	n = 51	n = 51	n = 62	n = 48	-1.08	-0.66	
Y-balance [%]	103.9 ± 8.9	101.1 ± 4.6	101.3 ± 6.9	102.6 ± 6.9	103.7 ± 11.2	104.3 ± 6.1	104.9 ± 10.0	101.9 ± 8.1	0.5	0.286	0.263
	(101.0–106.9)	(97.8–104.3)	(98.6–104.0)	(99.9–105.3)	(100.9–106.4)	(101.3–106.7)	(102.5–107.3)	(99.4–104.4)	-0.19	-0.13	-0.24
	n = 30	n = 25	n = 36	n = 36	n = 34	n = 36	n = 45	n = 42			
Bourban-test [s]	112.6 ± 50.1	102.2 ± 61.5	162.4 ± 178.2	141.0 ± 163.4	152.4 ± 108.1	163.0 ± 307.5	120.9 ± 44.2	103.2 ± 38.1	<0.001	0.515	0.869
	(65.9–159.4)	(56.5–147.8)	(129.1–195.6)	(100.3–181.7)	(110.5–194.3)	(121.4–204.5)	(82.5–159.2)	(60.0–146.4)		-0.06	-0.08
	n = 41	n = 42	n = 80	n = 53	n = 51	n = 52	n = 61	n = 48	-0.29		
Hand grip	29.2 ± 8.0	26.2 ± 4.6	33.1 ± 8.7	30.5 ± 5.2	33.5 ± 9.3	32.0 ± 4.5	41.3 ± 10.5	32.7 ± 5.9	<0.001	<0.001	0.02
strength [kg]	(26.8–31.6)	(23.8–28.5)	(31.3–34.9)	(28.3–32.7)	(31.3–35.7)	(29.5–34.4)	(39.1–43.5)	(29.9–35.4)			-0.33

(*Continued*)

Table 1. (Continued)

	chronological age								main/interaction effects		
	12		13		14		15		p-value (d)		
	boys	girls	boys	girls	boys	girls	boys	girls	age	sex	age x sex
	mean ± SD	mean ± SD	mean ± SD	mean ± SD	mean ± SD	mean ± SD	mean ± SD	mean ± SD	p (d)	p (d)	p (d)
	(95% CI)	(95% CI)	(95% CI)	(95% CI)	(95% CI)	(95% CI)	(95% CI)	(95% CI)			
	n = 55	n = 45	n = 91	n = 61	n = 69	n = 56	n = 76	n = 52			
	n = 40	n = 40	n = 74	n = 47	n = 47	n = 39	n = 48	n = 30	-0.8	-0.5	

Data were only calculated if at least 30 participants were available within a subgroup. For the subgroups of 8, 9, 10, 11, 16, 17, and 18 years old young athletes there were less than 30 participants and, thus, data were not calculated and reported.

BMI = body mass index, CMJ = countermovement jump, DJ = drop jump, SD = standard deviation.

https://doi.org/10.1371/journal.pone.0237423.t001

Table 2. Anthropometric and physical fitness differences according to maturity status and sex in young athletes.

	biological age						main/interaction effects		
	pre-pubertal		pubertal		post-pubertal		p-value (d)		
	(maturity offset: -1.98±0.71)		(maturity offset: 0.04±0.55)		(maturity offset: 2.57±1.07)				
	boys	girls	boys	girls	boys	girls	age	sex	age x sex
	mean ± SD	mean ± SD	mean ± SD	mean ± SD	mean ± SD	mean ± SD	p (d)	p (d)	p (d)
	(95% CI)	(95% CI)	(95% CI)	(95% CI)	(95% CI)	(95% CI)			
	n = 78	n = 4	n = 162	n = 54	n = 149	n = 216			
ANTHROPOMETRY									
standing height [cm]	150.9 ± 10.5	142.3 ± 9.5	167.9 ± 8.7	155.8 ± 7.3	183.5 ± 9.6	169.7 ± 7.9			
	(148.9–153.0)	(132.3–152.3)	(166.5–169.3)	(153.4–158.2)	(182.1–184.9)	(168.5–170.9)	<0.001	<0.001	0.398
	n = 77	n = 3	n = 162	n = 53	n = 149	n = 216	-1.57	-0.5	-0.11
sitting height [cm]	78.1 ± 4.8	76.1 ± 2.5	86.0 ± 4.4	81.9 ± 4.6	93.3 ± 5.7	88.4 ± 4.0			
	(77.1–79.1)	(71.5–80.7)	(85.3–86.7)	(80.6–83.1)	(92.5–94.1)	(87.8–89.0)	<0.001	<0.001	0.389
	n = 78	n = 3	n = 162	n = 54	n = 141	n = 211	-1.44	-0.34	-0.11
BMI [kg/m^2]	17.5 ± 2.1	16.0 ± 0.1	19.5 ± 2.7	18.5 ± 2.0	22.2 ± 3.1	21.1 ± 2.5			
	(17.0–18.0)	(15.7–16.3)	(19.1–19.9)	(17.9–19.0)	(21.7–22.8)	(20.7–21.4)	0.004	0.554	0.929
	n = 76	n = 3	n = 156	n = 53	n = 134	n = 204	-0.26	-0.05	-0.03
body mass [kg]	40.3 ± 8.7	32.6 ± 4.4	55.2 ± 11.5	45.0 ± 7.4	74.8 ± 14.6	60.9 ± 9.9			
	(38.3–42.2)	(21.6–43.5)	(53.4–57.1)	(43.0–47.1)	(72.3–77.4)	(59.5–62.2)	<0.001	<0.001	0.066
	n = 76	n = 4	n = 156	n = 54	n = 132	n = 203	-1.42	-0.39	-0.19
PHYSICAL FITNESS									
CMJ height [cm]	25.1 ± 4.2	NA	28.5 ± 6.1	24.3 ± 3.8	37.7 ± 8.8	26.6 ± 5.0			
	(23.6–26.6)	n = 1	(27.6–29.5)	(22.5–26.0)	(36.7–38.8)	(25.7–27.4)	<0.001	0.041	<0.001
	n = 64		n = 151	n = 48	n = 131	n = 209	-0.77	-0.17	-0.48
DJ height [cm]	21.4 ± 4.3	NA	24.6 ± 6.4	22.7 ± 5.2	31.1 ± 6.9	24.0 ± 4.9			
	(19.9–22.7)	n = 1	(23.6–25.5)	(21.1–24.3)	(30.1–32.1)	(23.3–24.8)	<0.001	0.357	0
	n = 63		n = 151	n = 48	n = 135	n = 209	-0.57	-0.08	-0.39
DJ ground contact time [ms]	213 ± 55	NA	237 ± 78	244 ± 66	207 ± 44	213 ± 38	<0.001		
	(199–227)	n = 1	(224–249)	(224–263)	(199–215)	(208–219)		0.83	0.845
	n = 63		n = 151	n = 48	n = 135	n = 209	-0.05	-0.02	-0.001
DJ performance index [m/s]	1.1 ± 0.3	NA	1.1 ± 0.4	1.0 ± 0.3	1.6 ± 0.5	1.2 ± 0.3			
	(1.0–1.1)	n = 1	(1.1–1.2)	(0.9–1.1)	(1.5–1.6)	(1.1–1.2)	<0.001	0.482	0.002
	n = 63		n = 151	n = 48	n = 135	n = 209	-0.63	-0.06	-0.29

(*Continued*)

Table 2. (Continued)

	biological age						main/interaction effects		
	pre-pubertal (maturity offset: -1.98±0.71)		pubertal (maturity offset: 0.04±0.55)		post-pubertal (maturity offset: 2.57±1.07)		p-value (d)		
	boys	girls	boys	girls	boys	girls	age	sex	age x sex
	mean ± SD (95% CI) n = 78	mean ± SD (95% CI) n = 4	mean ± SD (95% CI) n = 162	mean ± SD (95% CI) n = 54	mean ± SD (95% CI) n = 149	mean ± SD (95% CI) n = 216	p (d)	p (d)	p (d)
T-test [s]	12.0 ± 0.8	NA	11.1 ± 0.9	12.4 ± 1.0	10.1 ± 0.7	11.3 ± 1.0			
	(11.8–15.2)	n = 1	(10.9–11.2)	(12.2–12.7)	(9.9–10.2)	(11.1–11.4)	<0.001	<0.001	0.504
	n = 58		n = 131	n = 47	n = 136	n = 205	-1.04	-0.37	-0.1
Y-balance test [%]	103.7 ± 5.6	108.2 ± 13.7	103.5 ± 10.8	107.3 ± 4.2	105.8 ± 10.0	102.6 ± 7.0			
	(101.1–106.4)	(98.6–117.8)	(101.4–105.5)	(103.1–111.4)	(104.0–107.6)	(101.2–104.0)	0.567	0.378	0.014
	n = 39	n = 3	n = 64	n = 16	n = 88	n = 148	-0.11	-0.09	-0.31
Bourban-test [s]	338.1 ± 481.2	NA	153.1 ± 150.0	133.7 ± 70.0	125.4 ± 56.8	130.9 ± 180.1			
	(286.9–389.2)	n = 1	(119.3–187.0)	(74.9–192.4)	(90.4–160.3)	(103.0–159.0)	0.354	0.429	0.64
	n = 62		n = 142	n = 47	n = 132	n = 204	-0.12	-0.07	-0.08
Hand grip strength [kg]	23.1 ± 4.5	NA	32.2 ± 7.5	23.8 ± 4.3	43.9 ± 10.0	31.5 ± 5.3			
	(21.2–24.9)	n = 1	(31.0–33.4)	(21.5–26.1)	(42.5–45.3)	(30.4–32.6)	0.000	0.001	0.026
	n = 55		n = 130	n = 35	n = 99	n = 153	-1.16	-0.3	-0.25

Maturity was determined by calculating the time from peak-height-velocity (PHV) according to the equations as provided by Mirwald et al., [21]. Maturity was classified as pre-pubertal (i.e., > 1 year before PHV), pubertal (i.e., ±1 year around PHV), and post-pubertal (i.e., > 1 year after PHV).

BMI = body mass index, CMJ = countermovement jump, DJ = drop jump; NA = not applicable; SD = standard deviation.

https://doi.org/10.1371/journal.pone.0237423.t002

Interaction effects of the factors maturity, age, and sex. Our analyses showed significant sex by maturity interactions for almost all physical fitness tests ($p < 0.05$; $0.25 \leq d \leq 0.48$), except for the T-test and the Bourban-test ($p > 0.05$; $0.08 \leq d \leq 0.10$) (Table 2). Post-hoc analyses indicated more pronounced sex-specific differences (i.e., better CMJ, DJ, DJ performance index, and hand grip strength in boys compared to girls) in post-pubertal (Δ23–30%; $p < 0.001$; $1.00 \leq d \leq 1.63$) compared with pubertal athletes (Δ14–26%; $p < 0.05$; $0.34 \leq d \leq 0.99$). Due to the low number of pre-pubertal girls (n = 1–4), sex-specific differences were not computed for this cohort. Furthermore, significant sex by age interactions were found for anthropometrics, jump performance, and hand grip strength ($p < 0.05$; $0.27 \leq d \leq 0.45$) (Table 1). Post-hoc analyses indicated that 12, 13, 14, and 15 years old male athletes showed significantly better CMJ performances compared with their female counterparts (9–24%; $p < 0.05$; $0.51 \leq d \leq 1.16$). Furthermore, our analyses indicated that 14 and 15 year old males showed better DJ performance compared with females (i.e., DJ height, DJ performance index) (Δ11–24%; $p < 0.05$; $0.40 \leq d \leq 0.97$). These sex-specific differences were even more pronounced with increasing chronological age. Sex-specific differences in body mass, standing and sitting height were only found in 15 years old athletes (boys > girls; Δ4–21%; $p < 0.01$; $0.56 \leq d \leq 0.94$).

Percentile values according to maturity status, age, and sex

Tables 3 and 4 illustrate sex-specific percentile values according to chronological age (Table 3) and maturity status (Table 4) for jump tests (CMJ, DJ), change-of direction speed tests (T-test), strength tests (Bourban-test, hand grip strength test), and balance tests (Y-balance test).

Table 3. Sex- and chronological age-specific anthropometric and physical fitness percentiles of German elite young athletes.

	age (years)	n	P_{20}	P_{40}	P_{50}	P_{60}	P_{80}
Anthropometry							
			boys				
standing height [cm]	12	54	150.6	159.0	160.8	164.0	170.4
	13	91	157.9	166.4	169.0	173.0	177.8
	14	68	162.4	166.4	169.0	171.7	180.1
	15	76	171.2	176.1	178.1	181.3	187.9
			girls				
standing height [cm]	12	45	153.3	159.0	161.2	165.9	173.2
	13	59	160.1	164.3	165.5	167.0	171.3
	14	56	160.0	163.5	165.8	167.3	174.0
	15	52	165.5	169.0	170.9	172.5	176.2
			boys				
sitting height [cm]	12	55	77.7	81.5	82.9	84.3	88.0
	13	90	81.0	85.0	87.5	88.1	91.6
	14	60	82.5	85.2	86.5	87.5	93.0
	15	67	87.2	90.9	92.0	93.2	95.9
			girls				
sitting height [cm]	12	45	80.1	83.1	84.7	86.2	89.6
	13	58	83.7	86.0	87.5	88.0	90.1
	14	54	84.6	87.0	87.8	88.6	90.5
	15	51	86.6	89.1	90.0	90.4	91.7
			boys				
BMI [kg/m^2]	12	53	17.1	17.8	18.3	18.8	20.5
	13	88	17.2	18.4	18.8	19.5	21.3
	14	64	18.0	19.2	19.4	20.0	22.0
	15	68	18.7	20.3	21.4	22.1	24.0
			girls				
BMI [kg/m^2]	12	45	17.3	18.4	18.8	19.6	21.2
	13	58	18.3	19.4	19.9	20.3	21.8
	14	54	19.0	20.5	20.9	21.4	22.6
	15	49	19.1	20.5	21.0	21.6	23.0
			boys				
body mass [kg]	12	53	39.5	45.1	48.2	49.8	59.7
	13	88	44.0	50.3	53.8	56.7	67.0
	14	64	48.5	54.5	55.5	60.0	68.1
	15	68	55.5	64.0	68.3	70.4	82.7
			girls				
body mass [kg]	12	45	43.2	47.1	48.8	54.3	61.3
	13	60	49.1	52.8	54.4	55.8	62.1
	14	54	52.0	56.3	59.6	61.5	65.5
	15	49	54.7	58.7	61.6	63.7	69.1
Physical fitness							
			boys				
countermovement jump [cm]	12	43	20.6	24.4	24.8	27.3	30.1
	13	86	24.4	26.3	27.4	28.4	30.2
	14	51	25.4	27.7	28.5	30.8	33.8
	15	61	29.9	32.6	34.4	36.9	41.1
			girls				

(*Continued*)

Table 3. (Continued)

	age (years)	n	P_{20}	P_{40}	P_{50}	P_{60}	P_{80}
countermovement jump [cm]	12	42	20.1	22.3	23.5	25.3	26.8
	13	53	21.4	23.6	25.3	26.5	28.5
	14	55	23.0	25.3	26.7	27.3	30.1
	15	47	23.0	25.0	26.4	27.7	31.6
boys							
drop jump (DJ) [cm]	12	43	18.0	20.6	21.9	23.1	25.4
	13	86	19.1	22.0	32.4	24.7	28.5
	14	51	21.8	24.5	25.8	27.5	31.0
	15	62	24.3	27.8	30.4	32.3	37.0
girls							
drop jump (DJ) [cm]	12	42	18.4	20.9	22.2	23.5	25.1
	13	52	19.1	21.2	23.0	24.1	26.0
	14	55	19.6	22.0	22.8	24.1	26.3
	15	49	19.8	23.8	25.3	26.3	28.5
boys							
drop jump ground contact time [ms]	12	42	290	236	217	209	189
	13	85	262	231	218	207	186
	14	47	259	219	206	202	188
	15	58	225	206	202	188	202
girls							
drop jump ground contact time [ms]	12	43	278	240	231	222	207
	13	50	248	221	208	200	177
	14	54	237	214	203	196	187
	15	49	233	215	206	196	187
boys							
drop jump performance index [m/s]	12	42	0.61	0.83	0.91	0.99	1.20
	13	86	0.81	1.01	1.04	1.10	1.28
	14	48	0.91	1.14	1.26	1.38	1.55
	15	58	1.18	1.41	1.52	1.65	2.03
girls							
drop jump performance index [m/s]	12	43	0.78	0.87	0.93	0.95	1.17
	13	51	0.85	0.97	1.04	1.15	1.37
	14	54	0.95	1.08	1.11	1.14	1.32
	15	49	0.89	1.05	1.20	1.27	1.44
boys							
T-test [s]	12	34	12.28	11.99	11.86	11.66	10.71
	13	75	11.96	11.21	10.93	10.76	10.30
	14	51	11.38	11.03	10.71	10.50	9.98
	15	62	10.67	10.22	9.99	9.89	9.64
girls							
T-test [s]	12	42	12.79	12.47	12.32	12.04	11.46
	13	52	12.41	12.10	11.84	11.66	10.76
	14	51	11.91	11.34	11.09	10.79	10.47
	15	48	11.65	10.83	10.71	10.53	10.20
boys							
Y-balance test (dominant) [%]	12	30	96.1	102.3	103.9	106.2	108.5
	13	36	95.8	97.5	101.1	104.2	108.6
	14	34	95.0	98.1	101.0	104.1	117.1
	15	45	96.9	100.1	104.4	107.1	112.6
girls							

(*Continued*)

Table 3. (Continued)

	age (years)	n	P_{20}	P_{40}	P_{50}	P_{60}	P_{80}
Y-balance test (dominant) [%]	12	25	NA	NA	NA	NA	NA
	13	36	96.9	101.0	102.0	102.8	111.0
	14	36	99.0	104.7	105.0	106.8	108.4
	15	42	93.7	99.6	101.2	102.9	111.2
boys							
Bourban-test [s]	12	41	72	96	102	115	152
	13	80	79	100	123	150	192
	14	51	91	106	124	135	187
	15	61	77	105	119	125	148
girls							
Bourban-test [s]	12	42	57	75	83	95	142
	13	53	63	90	104	124	175
	14	52	71	88	99	109	158
	15	48	71	90	94	104	137
boys							
hand grip strength (dominant) [kg]	12	40	22.9	25.8	29.2	30.0	33.3
	13	74	25.1	29.4	30.8	33.5	40.7
	14	47	26.7	29.7	30.5	33.2	40.8
	15	48	30.9	37.0	40.1	44.8	51.7
girls							
hand grip strength (dominant) [kg]	12	40	22.6	25.3	26.4	27.5	30.1
	13	47	26.5	28.4	29.5	30.2	36.0
	14	39	27.3	31.0	32.7	33.7	36.0
	15	30	29.1	32.1	32.6	34.2	37.3

Sex-specific percentile reference values for anthropometric and physical fitness data. Data were only calculated if at least 30 participants were available within a subgroup. For the subgroups including 8, 9, 10, 11, 16, 17, and 18 year old athletes, less than 30 participants were available which is why these cells could not be filled. NA = not applicable (< 30 participants).

https://doi.org/10.1371/journal.pone.0237423.t003

Discussion

This study systematically aggregated anthropometric and physical fitness data of 703 elite young athletes aged 8–18 years from various sports and computed percentile values. Data were analyzed and expressed as percentile values according to maturity status, age, and sex. Findings indicate that anthropometry and physical fitness significantly increase with increasing maturity status and age, except for the Y-balance test. In general, male young athletes were taller, heavier, and they outperformed their female peers in CMJ, DJ, CoD, and hand grip strength performances. These sex-specific differences increase with increasing age.

Maturity-, age-, and sex-specific differences in anthropometry and physical fitness

The pathway from childhood through adolescence into adulthood inevitably leads to body growth as well as somatic and cognitive maturation. Motor development depends on and is influenced by growth and maturation and consequently affects physical fitness [19]. Unlike chronological age, maturation is not a linear process. Skeletal, sexual and somatic maturation in children differ individually in timing and tempo which is why there is often a discrepancy between chronological age and maturation among youths [16–18, 34]. Therefore, maturity

Table 4. Sex- and maturity-specific anthropometric and physical fitness percentiles of German elite young athletes.

	maturity status	n	P_{20}	P_{40}	P_{50}	P_{60}	P_{80}
Anthropometry							
	boys						
standing height [cm]	pre-pubertal (mo: -1.97 ± 0.72)	77	141.1	149.1	152.0	154.8	159.6
	pubertal (mo: 0.04 ± 0.54)	162	159.5	165.2	168.0	169.9	175.7
	post-pubertal (mo: 2.44 ± 1.11)	149	174.8	180.2	183.7	186.5	190.8
	girls						
standing height [cm]	pre-pubertal (mo: -2.18 ± 0.97)	3	NA	NA	NA	NA	NA
	pubertal (mo: 0.0 8 ± 0.58)	53	150.2	154.1	155.9	157.7	161.0
	post-pubertal (mo: 2.65 ± 1.04)	216	162.9	166.8	169.0	171.2	175.1
	boys						
sitting height [cm]	pre-pubertal (mo: -1.98 ± 0.71)	78	73.5	76.9	78.3	79.6	82.2
	Pubertal (mo: 0.04 ± 0.54)	162	82.1	84.7	86.0	87.0	89.1
	post-pubertal (mo: 2.44 ± 1.11)	141	90.2	92.7	93.7	95.3	97.0
	girls						
sitting height [cm]	pre-pubertal (mo: -2.01 ± 0.86)	4	NA	NA	NA	NA	NA
	Pubertal (mo: 0.06 ± 0.59)	54	77.4	80.7	81.5	82.6	85.8
	post-pubertal (mo: 2.65 ± 1.04)	212	85.3	87.8	88.7	89.5	91.5
	boys						
BMI [kg/m^2]	pre-pubertal (mo: -1.98 ± 0.72)	76	15.6	16.7	17.1	17.5	18.7
	pubertal (mo: 0.03 ± 0.54)	156	17.5	18.5	18.8	19.4	21.3
	post-pubertal (mo: 2.42 ± 1.14)	131	19.4	20.9	22.0	22.5	24.7
	girls						
BMI [kg/m^2]	pre-pubertal (mo: -2.18 ± 0.97)	3	NA	NA	NA	NA	NA
	pubertal (mo: 0.08 ± 0.58)	53	16.6	17.6	18.3	18.5	20.6
	post-pubertal (mo: 2.59 ± 1.00)	204	19.0	20.3	20.9	21.3	22.8
	boys						
body mass [kg]	pre-pubertal (mo: -1.98 ± 0.72)	76	32.6	38.0	39.3	40.6	46.0
	pubertal (mo: 0.03 ± 0.54)	156	46.1	50.6	53.8	55.7	62.7
	post-pubertal (mo: 2.42 ± 1.11)	131	62.1	68.7	71.4	77.6	87.1
	girls						
body mass [kg]	pre-pubertal (mo: -2.01 ± 0.86)	4	NA	NA	NA	NA	NA
	pubertal (mo: 0.06 ± 0.59)	54	37.4	44.0	45.2	47.1	52.4
	post-pubertal (mo: 2.60 ± 1.00)	204	52.3	58.0	60.0	62.6	68.7
Physical fitness							
	boys						
countermovement jump [cm]	pre-pubertal (mo: -1.96 ± 0.74)	64	21.5	23.9	24.7	25.4	28.7
	pubertal (mo: 0.03 ± 0.55)	151	24.3	26.5	27.9	28.8	32.4
	post-pubertal (mo: 2.48 ± 1.05)	131	30.6	34.4	36.7	38.3	44.3
	girls						
countermovement jump [cm]	pre-pubertal (mo: -1.64 ± NA)	1	NA	NA	NA	NA	NA
	pubertal (mo: 0.10 ± 0.56)	48	21.2	23.2	23.9	25.7	27.3
	post-pubertal (mo: 2.63 ± 1.03)	209	22.5	25.3	26.4	27.3	30.4
	boys						
drop jump [cm]	pre-pubertal (mo: -1.96 ± 0.75)	63	18.0	19.7	21.0	21.9	24.5
	pubertal (mo: 0.03 ± 0.55)	151	19.6	22.7	23.7	25.0	29.2
	post-pubertal (mo: 2.54 ± 1.11)	135	24.9	28.7	30.4	32.4	37.0
	girls						

(*Continued*)

Table 4. (Continued)

	maturity status	n	P_{20}	P_{40}	P_{50}	P_{60}	P_{80}
drop jump [cm]	pre-pubertal (mo: -1.64 ± NA)	1	NA	NA	NA	NA	NA
	pubertal (mo: 0.10 ± 0.56)	48	18.4	20.9	22.6	23.7	25.6
	post-pubertal (mo: 2.63 ± 1.03)	209	19.6	22.7	24.1	25.1	28.0
boys							
drop jump ground contact time [ms]	pre-pubertal (mo: -1.98 ± 0.75)	62	262	202	195	190	177
	pubertal (mo: 0.03 ± 0.55)	144	269	226	216	208	189
	post-pubertal (mo: 2.55 ± 1.12)	128	229	206	200	191	180
girls							
drop jump ground contact time [ms]	pre-pubertal (mo: -1.64 ± NA)	1	NA	NA	NA	NA	NA
	pubertal (mo: 0.09 ± 0.55)	47	290	241	220	214	196
	post-pubertal (mo: 2.64 ± 1.03)	207	237	219	207	200	185
boys							
drop jump performance index [m/s]	pre-pubertal (mo: -1.98 ± 0.75)	62	0.76	0.97	1.00	1.10	1.30
	pubertal (mo: 0.03 ± 0.55)	145	0.80	1.00	1.10	1.16	1.41
	post-pubertal (mo: 2.55 ± 1.12)	130	1.18	1.38	1.47	1.63	1.98
girls							
drop jump performance index [m/s]	pre-pubertal (mo: -1.64 ± NA)	1	NA	NA	NA	NA	NA
	pubertal (mo: 0.09 ± 0.55)	47	0.74	0.85	0.90	1.04	1.19
	post-pubertal (mo: 2.64 ± 1.03)	209	0.90	1.04	1.12	1.21	1.42
boys							
T-test [s]	pre-pubertal (mo: -1.99 ± 0.76)	58	12.77	12.55	12.02	11.90	11.25
	pubertal (mo: 0.04 ± 0.55)	131	11.77	11.20	10.98	10.72	10.30
	post-pubertal (mo: 2.53 ± 1.10)	136	10.60	10.12	9.99	9.82	9.56
girls							
T-test [s]	pre-pubertal (mo: -1.64 ± NA)	1	NA	NA	NA	NA	NA
	pubertal (mo: 0.12 ± 0.54)	47	13.36	12.45	12.27	12.05	11.47
	post-pubertal (mo: 2.65 ± 1.04)	205	12.14	11.47	10.96	10.76	10.41
boys							
Y-balance test (dominant) [%]	pre-pubertal (mo: -1.74 ± 0.51)	39	97.7	102.1	102.7	105.3	108.1
	pubertal (mo: 0.10 ± 0.55)	64	94.3	97.7	101.1	106.9	111.7
	post-pubertal (mo: 2.61 ± 1.29)	88	97.6	101.7	103.6	106.6	116.0
girls							
Y-balance test (dominant) [%]	pre-pubertal (mo: -2.13 ± 1.01)	3	NA	NA	NA	NA	NA
	pubertal (mo: 0.01 ± 0.66)	16	NA	NA	NA	NA	NA
	post-pubertal (mo: 2.68 ± 1.00)	148	96.9	100.8	102.3	104.4	107.6
boys							
Bourban-test [s]	pre-pubertal (mo: -1.97 ± 0.76)	62	83	124	158	181	378
	pubertal (mo: 0.04 ± 0.55)	142	79	101	124	136	183
	post-pubertal (mo: 2.53 ± 1.11)	132	84	103	114	125	150
girls							
Bourban-test [s]	pre-pubertal (mo: -1.64 ± NA)	1	NA	NA	NA	NA	NA
	pubertal (mo: 0.10 ± 0.57)	47	75	101	126	139	184
	post-pubertal (mo: 2.64 ± 1.03)	202	68	87	98	110	148
boys							
hand grip strength (dominant) [kg]	pre-pubertal (mo: -1.96 ± 0.73)	55	19.4	22.1	23.4	23.9	26.5
	pubertal (mo: 0.03 ± 0.55)	130	25.6	29.5	30.7	32.5	39.0
	post-pubertal (mo: 2.30 ± 0.92)	99	33.7	41.3	44.9	47.6	52.9

(*Continued*)

Table 4. (Continued)

	maturity status	n	P_{20}	P_{40}	P_{50}	P_{60}	P_{80}
girls							
hand grip strength (dominant) [kg]	pre-pubertal (mo: -1.64 ± NA)	1	NA	NA	NA	NA	NA
	pubertal (mo: 0.07 ± 0.59)	35	19.1	22.9	24.8	25.9	27.2
	post-pubertal (mo: 2.44 ± 0.95)	153	27.0	29.7	31.5	32.9	36.2

Maturity was determined by calculating the time from peak-height-velocity (PHV) according to the equation as provided by Mirwald et al., [21]. Maturity was classified as pre-pubertal (i.e., > 1 year before PHV), pubertal (i.e., ±1 year around PHV), and post-pubertal (i.e., > 1 year after PHV) [21].
mo = maturity offset; NA = not applicable (< 30 participants).

https://doi.org/10.1371/journal.pone.0237423.t004

and chronological age should be considered when assessing anthropometry and particularly physical fitness in young athletes. Changes in individual structural constraints through growth are temporarily very dramatic. This is evident on a whole-body level (e.g., changes in size and proportion of the whole body) as well as on a system level (e.g., skeletal system, the muscular system, and the endocrine system) [35].

As hypothesized, our findings indicate that body height and mass significantly increase with increasing maturity status. Furthermore, post-pubertal athletes performed significantly better in various physical fitness tests (i.e., CMJ height, DJ height, DJ performance index, T-test, hand grip strength) compared to pubertal athletes. Thus, it seems that increases in body size, hormones, and muscle strength, caused by puberty, can improve physical fitness. There are rarely any studies that examined the effects of maturity on anthropometrics and physical fitness in children and adolescents. Jones et al. [36] reported that self-assessed stage of sexual maturity correlated positively with objectively measured physical fitness (i.e., vertical jump, 20-m shuttle run test, hand grip strength) in untrained boys and girls. Our results regarding the effects of biological maturity should be considered preliminary and appear to be in accordance with study findings from the general population of non-athletic youth [36].

In addition, our findings indicated that athletes were taller and heavier and they performed significantly better with increasing chronological age in selected physical fitness test (i.e., CMJ height, DJ height, DJ performance index, T-test, hand grip strength). Merely, physical fitness did not improve between 12 and 13 year old athletes as well as anthropometry did not change between 13 and 14 years old athletes. Data from different studies which have previously examined non-athletic youth [9, 19, 37] and athletic youth [13, 15, 38] confirm our findings in as much as improvements in physical fitness were reported with increasing chronological age. The performance enhancements can most likely be explained by changes in body size, physique, and body composition that are important factors affecting for physical fitness in general and muscle strength in particular [19, 39].

Furthermore, our findings indicate that young male athletes are taller as well as heavier compared to female young athletes. Further, results revealed that young male athletes outperformed young female athletes in the vertical jump, CoD- and hand grip strength test. Therefore, whilst stronger and leaner than many of their non-athletic peers, young female athletes are not as tall, as strong, nor as fast as their male counterparts [18]. The sex-specific anthropometric and physical performance differences are well in line with findings from previous studies regarding sexual dimorphism in the general population of non-athletic youth [16, 19, 20] and can most likely be attributed to higher absolute and relative strength levels in boys compared to girls [17, 18, 39]. The detected sex—maturity as well as sex—age interactions indicated that differences in physical fitness outcomes increased between male and female young athletes with increasing maturity and chronological age respectively. Previous studies

highlighted that sex differences in physical fitness are rather small prior to the onset of puberty [19]. This finding could not be demonstrated in our results, due to a small sample size in pre-pubertal children, especially in girls. However, during the adolescent growth spurt, sex differences become more pronounced [15]. This can mainly be explained by hormone-dependent changes in body composition [15, 19, 39]. Boys show significant rise in growth of bone, stature and muscle mass and simultaneous loss of fat in limbs under the influence of testosterone [17, 39]. Moreover, results of several studies indicate that testosterone is responsible for improved anaerobic enzyme systems and structural development of fast twitch muscle fibers in muscles [40]. Thus, an increase in testosterone may determine the greater formation and development of fast twitch muscle fibers that positively affect the performance of explosive muscle actions [40]. Girls experience lesser increment in stature and muscle mass, but a significant accumulation of body fat [18, 39]. Thus, the beneficial effects of maturational changes are present but less evident in girls where sport-related motor performances tend to plateau from mid-adolescence. Thus, as a result of sex-related differences in growth and maturation during adolescence, the post-pubertal male athlete is stronger and has more muscle mass than the post-pubertal female athlete [35].

Physical fitness percentiles

Chronological age provides a useful point of reference when referring to growth and maturity status. However, biological processes do not progress in a linear fashion [17–19]. Therefore, youth of the same chronological age display wide variability in the development of morphological and physiological characteristics. This is a major challenge in youth sport where competitions mainly are regulated by chronological age-groups to establish equal chances of success for all athletes [16, 41]. But, within a prescribed age-group, variations in maturity status can deliver a distinct advantage not only in performance but also for talent identification. For example, boys who mature earlier are generally taller, heavier, have higher mass-to-stature ratios and, thus, are generally more prone to success in most types of exercise, particularly in those that involve strength, velocity and power [16, 17] than those who mature at a later age. Many young athletes drop out of sport or are cut from sport squads for instance due to retarded timing and tempo of their growth and maturation. For this reason, practitioners and coaches should be aware of the effects of age, growth and maturation on sports performance and should provide opportunities for all talented children irrespective of the maturational status [16, 17]. Maturity-specific reference for anthropometry and physical fitness tests for male and female young athletes are necessity to assess a youth athletes' performance adequately. Due to lack of literature that examines maturity-specific anthropometric and physical fitness percentiles for male and female youth, especially young athletes, findings are preliminary.

In terms of the established chronological age-specific percentile reference values of anthropometry and physical fitness tests for male and female young athletes, previous research mainly examined the general population of untrained children and adolescents. For juvenile non-athletes, studies with large cohorts are available [7, 9]. For instance, Tomkinson et al. [9] established sex- and age-specific percentile reference values for physical fitness (e.g., hand grip strength, bent-arm hang, standing long jump, 20-m shuttle run) in children and adolescents aged 9–17 years. However, young athletes represent a small segment of the general population with regards to their motor skill and performance levels. This is due to their genetic predisposition but also their exposure to regular training [11, 42]. Accordingly, anthropometric and physical fitness norms from the general youth population cannot be translated to young athletes [42]. Only few studies are available that provide age- and sex specific mean values for

youth from different sports [13–15, 38]. However, given that none of them established age- and sex-specific percentiles, findings are again preliminary.

Study limitations

We enrolled a convenience sample in this study consisting of 703 male and female young elite athletes from 18 different sports. The distribution of age and sex of the tested elite young athletes is not the same for each different sport. Due to the variety of sports within each sub-group the results may be affected by variations in morphology, growth, maturity status, and/or physique among athletes of different sports. Furthermore, the sample size of each sub-group is rather small and varies according to the sub-group under investigation (e.g., pre-pubertal girls are under-represented). Doing research in elite (youth) sports is always limited as to the size of the available cohort. The overall sample of young athletes is small compared to the general youth population. Consequently, it is not appropriate to compare the size of our study cohort with previous studies reporting norm values of the general youth population. In this study, we were able to enroll 703 male and female elite young athletes aged 8–18 years. While we acknowledge that the included number of participants is small when compared to studies using the general youth population, it is rather large compared to other studies who examined youth athletes [13, 43, 44]. Nonetheless, due to the small sample size in several sub-groups, the established reference values have to be interpreted with caution. Furthermore, we acknowledge that different individuals measured and tested the enrolled athletes. However, test instructions and test protocols were highly standardized and the observers were experienced exercise scientists. Moreover, we have applied the predicted maturity offset method according to Mirwald et al. [21] to estimate participants' maturity status. The application of the gold standard (x-ray exams of the left wrist) would have certainly provided higher accuracy according to the actual maturity status. However, this method is costly and causes radiation exposure which is why we decided to apply the Mirwald method based on anthropometrics (i.e., sitting and standing height). Müller et al. [45] performed a cross-validation study with young athletes using x-ray exams and the Mirwald method and concluded that the prediction equations to determine age at PHV appears to be a valid method for the assessment of biological maturity. Finally, the applicability of our observations to young athletes of other countries and races may be limited.

Conclusions

This study examined maturity-, age-, and sex-specific differences in anthropometrics and physical fitness in young athletes from various sports. Our findings indicate that body height and mass increased significantly with increasing maturity status and chronological age. Further, our results showed that physical fitness (i.e., CMJ height, DJ height, DJ performance index, DJ ground contact time, T-test, hand grip strength) was significantly better in post-pubertal compared to pubertal athletes. In addition, physical fitness outcomes (i.e., CMJ height, DJ height, DJ performance index, T-test, hand grip strength) improved with increasing chronological age (i.e., $12 = 13 < 14 < 15$ years).

Furthermore, maturity-, age- and sex-specific percentile reference values including a wide range of physical fitness outcomes (i.e., CMJ, DJ height, DJ performance index, DJ ground contact time, CoD speed, dynamic balance, trunk strength endurance, hand grip strength) were established. The percentile reference values add value to existing norms for children and adolescents for the specific sub-population of trained youth. Practitioners and coaches can use the established percentile values as approximate benchmarks to identify and develop young athletes with specific fitness characteristics for talent identification and development.

Supporting information

S1 Data.
(XLSX)

Acknowledgments

We thank the coaches, the technical staff, and foremost the young athletes who participated in the KINGS research project ("Resistance Training in Young athletes").

Author Contributions

Conceptualization: Melanie Lesinski, Christian Puta, Holger Gabriel, Adamantios Arampatzis, Gunnar Laube, Dirk Büsch, Urs Granacher.

Formal analysis: Melanie Lesinski, Alina Schmelcher, Michael Herz.

Investigation: Melanie Lesinski, Alina Schmelcher, Michael Herz, Christian Puta, Gunnar Laube.

Methodology: Melanie Lesinski, Urs Granacher.

Project administration: Melanie Lesinski, Urs Granacher.

Writing – original draft: Melanie Lesinski, Alina Schmelcher, Michael Herz.

Writing – review & editing: Melanie Lesinski, Christian Puta, Holger Gabriel, Adamantios Arampatzis, Gunnar Laube, Dirk Büsch, Urs Granacher.

References

1. Côté J, Lidor R, Hackfort D. ISSP position stand: to sample or to specialize? seven postulates about youth sport activities that lead to continued participation and elite performance. Int J Sport Exerci Psyc. 2009; 7(1):7–17.
2. DiFiori JP, Benjamin HJ, Brenner JS, Gregory A, Jayanthi N, Landry GL, et al. Overuse injuries and burnout in youth sports: a position statement from the american medical society for sports medicine. Br J Sports Med. 2014; 48(4):287–8. https://doi.org/10.1136/bjsports-2013-093299 PMID: 24463910
3. Jayanthi NA, LaBella CR, Fischer D, Pasulka J, Dugas LR. Sports-specialized intensive training and the risk of injury in young athletes: a clinical case-control study. Am J Sports Med. 2015; 43(4):794–801. https://doi.org/10.1177/0363546514567298 PMID: 25646361
4. Moesch K, Elbe AM, Hauge ML, Wikman JM. Late specialization: the key to success in centimeters, grams, or seconds (cgs) sports. Scand J Med Sci Sports. 2011; 21(6):e282–90. https://doi.org/10.1111/j.1600-0838.2010.01280.x PMID: 21401722
5. Malina R, Sławinska T, Ignasiak Z, Rożek K, Kochan K, Domaradzki J, et al. Sex differences in growth and performance of track and field athletes 11–15 Years. 2010; 24(1):79.
6. Lloyd RS, Oliver JL. The youth physical development model: a new approach to long-term athletic development. Strength Con J. 2012; 34(3):61–72.
7. Golle K, Muehlbauer T, Wick D, Granacher U. Physical fitness percentiles of german children aged 9–12 years: findings from a longitudinal study. PLoS One. 2015; 10(11):e0142393. https://doi.org/10.1371/journal.pone.0142393 PMID: 26544848
8. De Miguel-Etayo P, Gracia-Marco L, Ortega FB, Intemann T, Foraita R, Lissner L, et al. Physical fitness reference standards in european children: the IDEFICS study. Int J Obes (Lond). 2014; 38 Suppl 2: S57–66.
9. Tomkinson GR, Carver KD, Atkinson F, Daniell ND, Lewis LK, Fitzgerald JS, et al. European normative values for physical fitness in children and adolescents aged 9–17 years: results from 2 779 165 eurofit performances representing 30 countries. Br J Sports Med. 2018; 52(22):1445–14563. https://doi.org/10.1136/bjsports-2017-098253 PMID: 29191931

10. Gagné F. Constructs and models pertaining to exceptional human abilities. In: Heller KA, Mönks FJ, Passow AH, editors. International handbook of research and development of giftedness and talent. Elmsford, NY: Pergamon Press; 1993. p. 69–87.
11. Williams CA, Oliver JL, Lloyd RS. Talent development. In: Lloyd RS, Oliver JL, editors. Strength and conditioning for young athletes: science and application. 1st ed. Abingdon, Oxon: Routledge; 2014.
12. Issurin VB. Evidence-based prerequisites and precursors of athletic talent: a review. Sports Med. 2017; 47(10):1993–2010. https://doi.org/10.1007/s40279-017-0740-0 PMID: 28493064
13. Yanci J, Los Arcos A, Castillo D, Cámara J. Sprinting, Change of direction ability and horizontal jump performance in youth runners according to gender. J Hum Kinet. 2017; 60:199–207. https://doi.org/10.1515/hukin-2017-0103 PMID: 29340000
14. Opstoel K, Pion J, Elferink-Gemser M, Hartman E, Willemse B, Philippaerts R, et al. Anthropometric characteristics, physical fitness and motor coordination of 9 to 11 year old children participating in a wide range of sports. PLoS One. 2015; 10(5):e0126282. https://doi.org/10.1371/journal.pone.0126282 PMID: 25978313
15. Tonnessen E, Svendsen IS, Olsen IC, Guttormsen A, Haugen T. Performance development in adolescent track and field athletes according to age, sex and sport discipline. PLoS One. 2015; 10(6): e0129014. https://doi.org/10.1371/journal.pone.0129014 PMID: 26043192
16. Armstrong N. Development of the youth athlete. Abingdon: Routledge; 2019.
17. Armstrong N, McManus AM. Physiology of elite young male athletes. Med Sport Sci. 2011; 56:1–22. https://doi.org/10.1159/000320618 PMID: 21178364
18. McManus AM, Armstrong N. Physiology of elite young female athletes. Med Sport Sci. 2011; 56:23–46. https://doi.org/10.1159/000320626 PMID: 21178365
19. Malina RM, Bouchard C., Bar-Or O. Growth, maturation, and physical activity. 2nd ed. Champaign: Human Kinetics; 2004.
20. Reiman MP, Manske RC. Functional testing in human performance. Champaign: Human Kinetics; 2009.
21. Mirwald RL, Baxter-Jones AD, Bailey DA, Beunen GP. An assessment of maturity from anthropometric measurements. Med Sci Sports Exerc. 2002; 34(4):689–94. https://doi.org/10.1097/00005768-200204000-00020 PMID: 11932580
22. Faigenbaum AD, Lloyd DG, Oliver JL. Essentials of youth fitness. Champaign: Human Kinetice; 2020.
23. Coren S. The lateral preference inventory for measurement of handedness, footedness, eyedness, and earedness: norms for young adults. Bull Psychon Soc. 1993; 31(1):1–3.
24. Mathiowetz V, Weber K, Volland G, Kashman N. Reliability and validity of grip and pinch strength evaluations. J Hand Surg Am. 1984; 9(2):222–6. Epub 1984/03/01. https://doi.org/10.1016/s0363-5023(84)80146-x PMID: 6715829.
25. Markovic G, Dizdar D, Jukic I, Cardinale M. Reliability and factorial validity of squat and countermovement jump tests. J Strength Cond Res. 2004; 18(3):551–5. https://doi.org/10.1519/1533-4287(2004)18<551:RAFVOS>2.0.CO;2 PMID: 15320660
26. Tschopp M, Bourban P, Hübner K, Marti B. Reliability of a standardized, dynamic trunk muscle strength test: experiences with healthy male elite athletes. Swiss J of Sports Med Sports Traumatol. 2001; 49:67–72.
27. Pauole K, Madole K, Garhammer J, Lacourse M, Rozenek R. Reliability and validity of the T-test as a measure of agility, leg power, and leg speed in college-aged men and women. J Strength Cond Res. 2000; 14(4):443–50.
28. Plisky PJ, Rauh MJ, Kaminski TW, Underwood FB. Star excursion balance test as a predictor of lower extremity injury in high school basketball players. J Orthop Sports Phys Ther. 2006; 36(12):911–9. https://doi.org/10.2519/jospt.2006.2244 PMID: 17193868
29. Filipa A, Byrnes R, Paterno MV, Myer GD, Hewett TE. Neuromuscular training improves performance on the star excursion balance test in young female athletes. J Orthop Sports Phys Ther. 2010; 40(9):551–8. https://doi.org/10.2519/jospt.2010.3325 PMID: 20710094
30. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale: Erlbaum; 1988.
31. Sands W, Cardinale M, McNeal J, Murray S, Sole C, Reed J, et al. Recommendations for measurement and management of an elite athlete. Sports (Basel). 2019; 7(5).
32. Hoffman J. Norms for fitness, performance, and health. Champaign: Human Kinetics; 2006.
33. Sole CJ, Suchomel TJ, Stone MH. Preliminary scale of reference values for evaluating reactive strength index-modified in male and female NCAA division I athletes. Sports (Basel). 2018; 6(4).

34. Lloyd RS, Oliver JL, Faigenbaum AD, Myer GD, De Ste Croix MB. Chronological age vs. biological maturation: implications for exercise programming in youth. J Strength Cond Res. 2014; 28(5):1454–64. https://doi.org/10.1519/JSC.0000000000000391 PMID: 24476778

35. Haywood KM, Getchell N. Life span motor development. 4th ed. Champaign: Human Kinetics; 2005.

36. Jones MA, Hitchen PJ, Stratton G. The importance of considering biological maturity when assessing physical fitness measures in girls and boys aged 10 to 16 years. Ann Hum Biol. 2000; 27(1):57–65. https://doi.org/10.1080/030144600282389 PMID: 10673141

37. Ramos-Sepulveda JA, Ramirez-Velez R, Correa-Bautista JE, Izquierdo M, Garcia-Hermoso A. Physical fitness and anthropometric normative values among colombian-indian schoolchildren. BMC Public Health. 2016; 16:962. https://doi.org/10.1186/s12889-016-3652-2 PMID: 27619491

38. Canhadas IL, Pignataro Silva RL, Chaves CR, Portes IA. Anthropometric and physical fitness characteristics of young male soccer players Rev Bras Cineantropom Desempenho Hum. 2010; 12(4):239–45.

39. Goswami B, Roy AS, Dalui R, Bandyopadhyay A. Impact of pubertal growth on physical fitness. Am J Sports Sci Med. 2014; 2(5A):34–9.

40. Viru A, Viru M. Resistance exercise and testosterone. In: Kraemer WJ, Rogol AD, editors. The endocrine system in sports and exercise. Oxford: Blackwell; 2005. p. 319–38.

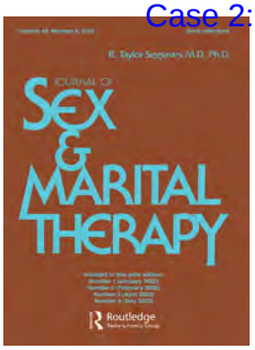
41. Malina RM, Cumming SP, Rogol AD, Coelho ESMJ, Figueiredo AJ, Konarski JM, et al. Bio-banding in youth sports: background, concept, and application. Sports Med. 2019; 49(11):1671–85. https://doi.org/10.1007/s40279-019-01166-x PMID: 31429034

42. Bergeron MF, Mountjoy M, Armstrong N, Chia M, Cote J, Emery CA, et al. International olympic committee consensus statement on youth athletic development. Br J Sports Med. 2015; 49(13):843–51. https://doi.org/10.1136/bjsports-2015-094962 PMID: 26084524

43. Karahan M. Age-related physical performance differences in male soccer players. THE ANTHROPOLOGIST. 2016; 24:724–9.

44. Bujnovky D, Maly T, Ford KR, Sugimoto D, Kunzmann E, Hank M, et al. Physical fitness characteristics of high-level youth football players: influence of playing position. Sports (Basel). 2019; 7(2).

45. Müller L, Müller E, Hildebrandt C, Kapelari K, Raschner C. The assessment of biological maturation for talent selection–which method can be used? Sportverletz Sportschaden. 2015; 29:56–63. https://doi.org/10.1055/s-0034-1399043 PMID: 25710395



Routledge
Taylor & Francis Group

Journal of Sex & Marital Therapy

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/usmt20

Reconsidering Informed Consent for Trans-Identified Children, Adolescents, and Young Adults

Stephen B. Levine, E. Abbruzzese & Julia W. Mason

To cite this article: Stephen B. Levine, E. Abbruzzese & Julia W. Mason (2022): Reconsidering Informed Consent for Trans-Identified Children, Adolescents, and Young Adults, Journal of Sex & Marital Therapy, DOI: 10.1080/0092623X.2022.2046221

To link to this article: https://doi.org/10.1080/0092623X.2022.2046221

Published online: 17 Mar 2022.

JOURNAL OF SEX & MARITAL THERAPY
https://doi.org/10.1080/0092623X.2022.2046221

Routledge
Taylor & Francis Group

Review

Reconsidering Informed Consent for Trans-Identified Children, Adolescents, and Young Adults

Stephen B. Levine[a], e. Abbruzzese[b] and Julia w. Mason[c]

[a]Department of Psychiatry, Case Western Reserve University, Cleveland, OH, USA; [b]Society for Evidence-based Gender Medicine (SEGM), Twin Falls, ID, USA; [c]Calcagno Pediatrics, Gresham, OR, USA

ABSTRACT

in less than a decade, the western world has witnessed an unprecedented rise in the numbers of children and adolescents seeking gender transition. Despite the precedent of years of gender-affirmative care, the social, medical and surgical interventions are still based on very low-quality evidence. The many risks of these interventions, including medicalizing a temporary adolescent identity, have come into a clearer focus through an awareness of detransitioners. The risks of gender-affirmative care are ethically managed through a properly conducted informed consent process. its elements—deliberate sharing of the hoped-for benefits, known risks and long-term outcomes, and alternative treatments—must be delivered in a manner that promotes comprehension. The process is limited by: erroneous professional assumptions; poor quality of the initial evaluations; and inaccurate and incomplete information shared with patients and their parents. we discuss data on suicide and present the limitations of the Dutch studies that have been the basis for interventions. Beliefs about gender-affirmative care need to be separated from the established facts. A proper informed consent process can both prepare parents and patients for the difficult choices that they must make and can ease professionals' ethical tensions. even when properly accomplished, however, some clinical circumstances exist that remain quite uncertain.

KEYWORDS
informed consent;
ethics;
gender dysphoria;
gender identity;
detransition

Introduction

Reconsideration of the meanings, purposes, indications, and processes of informed consent for transgender-identified youth is urgently needed. Parents of gender atypical children are considering social transition as early as preschool or grade school. Parents of preteens and teens are considering supporting their children's wishes to present in a new gender, take puberty blockers and cross-sex hormones, and plan for surgical alterations. College-aged youth are declaring new identities for the first time and obtaining hormones and surgery without their parents' knowledge.

When uncertain parents of children and teens consult their primary care providers, they are usually referred to specialty gender services. Parents and referring clinicians assume that specialists with "gender expertise" will undertake a thorough evaluation. However, the evaluations preceding the recommendation for gender transition are often surprisingly brief (Anderson & Edwards-Leeper, 2021) and typically lead to a recommendation for hormones and surgery, known as *gender-affirmative* treatment.

CONTACT Stephen B. Levine sbl2@case.edu Department of Psychiatry, Case Western Reserve University, 23425 Commerce Park #104, Cleveland, OH44106-7078, USA.

This article has been corrected with minor changes. These changes do not impact the academic content of the article.

Despite the widely recognized deficiencies in the evidence supporting gender-affirmative interventions (National Institute for Health & Care Excellence, 2020a; 2020b), the process of obtaining informed consent from patients and their families has no established standard. There is no consensus about the requisite elements of evaluations, nor is there unanimity about how informed consent processes should be conducted (Byne et al., 2012). These two matters are inconsistent from practitioner to practitioner, clinic to clinic, and country to country.

Social transition, hormonal interventions, and surgery have profound implications for the course of the lives of young patients and their families. It is incumbent upon professionals that these consequences be thoroughly, patiently clarified over time prior to undertaking any element of transition. The informed consent process does not preclude transition; it merely educates the family about the state of the science underpinning the decision to transition. Social transition, hormones, and surgeries are unproven in a strict scientific sense, and as such, to be ethical, require a thorough and fully informed consent process.

Ethical Concerns About Inadequate Informed Consent

The concept of informed consent in medicine has roots in both ethical theory and law. The ethical foundation is centered in the principles of beneficence, justice, and respect for autonomy, while the legal issues have to do with questions of malpractice (Katz et al., 2016).

Patients consenting to treatment must meet age-based and decisional capacity requirements (Katz et al., 2016). Minors less than the age of consent participate in decision-making by providing *assent*—an agreement with the intervention. The limited maturational cognitive capacities of minors are the key reason why parents serve as the ethical and legal surrogates for medical decision-making, tasked with signing an informed consent document (Grootens-Wiegers, Hein, van den Broek, & de Vries, 2017).

The informed consent process consists of three main elements: a disclosure of information about the nature of the condition and the proposed treatment and its alternatives; an assessment of patient and caregiver understanding of the information and capacity for medical decision-making; and obtaining the signatures that signify informed consent has been obtained (Katz et al., 2016). The current expectation that clinicians and institutions are required to thoroughly inform their patients about the benefits, risks, and uncertainties of a particular treatment, as well as about alternatives, has a long legal history in the United States (Lynch, Joffe, & Feldman, 2018).

Ethical concerns about inadequate informed consent for trans-identified youth have several potentially problematic sources, including *erroneous assumptions* held by professionals; *poor quality of the evaluation process*; and *incomplete and inaccurate information* that the patients and family members are given.

These concerns are amplified by the *dramatic growth* in demand for youth gender transition witnessed in the last several years that has led to a perfunctory informed consent process. A rushed process does not allow for a proper discussion of not only the benefits, but the profound risks and uncertainties associated with gender transition, especially when gender transition is undertaken before mature adulthood.

a. *Dramatic growth in demand for services threatens true informed consent*

Gender identity variations were thought to be extremely rare a generation ago. While the incidence in youth had not been officially estimated, in adults it was 2-14 per 100,000 (American Psychiatric Association, 2013, p. 454). However, around 2006, the incidence among youth began to rise, with a dramatic increase observed in 2015 (Aitken et al., 2015, de Graaf, Giovanardi, Zitz, & Carmichael, 2018). Currently, 2-9% of U.S. high school students identify as transgender, while in colleges, 3% of males and 5% of females identify as gender-diverse (American College Health Association, 2021; Johns et al., 2019; Kidd et al., 2021).

Whereas previously most of the affected individuals identified as the opposite sex, there is now a growing trend toward identifying as *nonbinary*: neither male nor female or both male and female (Chew et al., 2020). A recent study reported that the majority of transgender-identifying youth (63%) now have a non-binary identity (Green, DeChants, Price, & Davis, 2021). Although the incidence of natal males asserting a trans identity in adolescence has significantly increased, the dramatic increase is driven primarily by the natal females requesting services (Zucker, 2017). Many suffer from significant comorbid mental health disorders, have neurocognitive difficulties such as ADHD or autism or have a history of trauma (Becerra-Culqui et al., 2018; Kozlowska, McClure, et al., 2021).

The increase in rates of transgender identification is reflected in the numbers of youth seeking help from medical professionals. For example, according to data reported by the Tavistock gender clinic in the UK, in 2009, there were 51 requests for services (de Graaf et al., 2018); in 2019-2020, 2728 referrals were recorded—a 53-fold increase in just over a decade (Tavistock & Portman NHS Foundation Trust, 2020). The growing number of urban transgender health centers that have arisen in recent years (HRC, n.d.) reflects the increased demand for gender-related medical care among young people in North America Australia, and Europe.

This unprecedented increase has created pressure on institutions and practitioners to rapidly evaluate these youth and make recommendations about treatment. To respond to growing demand, an innovative *informed consent model of care* has been developed. Under this model, mental health evaluations are not required, and hormones can be provided after just one visit following the collection of a patient's or guardian's consent signature (Schulz, 2018). The provision of transition services under this model of care is available not just to those over 18, but for younger patients as well (Planned Parenthood League of Massachusetts, n.d.).

Although following the informed consent model of care for hormones and surgeries for youth may diminish clinicians' ethical or moral unease (Vrouenraets et al., 2020), we believe this model is the antithesis of true informed consent, as it jeopardizes the ethical foundation of patient autonomy. Autonomy is not respected when patients consenting to the treatment do not have an accurate understanding of the risks, benefits, and alternatives.

b. *Assumptions held by professionals influence the integrity of the informed consent process*

Gender-dysphoric children and teens can intensely occupy the belief that their lives will be immensely improved by transition. Clinicians who have embraced the gender-affirmative model of care operate on the assumption that children and teens know best what they need to be happy and productive (Ehrensaft, 2017). These professionals, responding to the youths' passionate pleas, see their role as validating the young person's fervent wishes for hormones and surgery and clearing the path for gender transition. In doing so, they privilege the ethical principle of respect for patient autonomy (Clark & Virani, 2021) over their obligations for beneficence and non-maleficence.

Many of the gender-affirmative clinicians subscribe to the theory of *minority stress* – the supposition that the frequently co-occurring psychiatric symptoms of gender-dysphoric individuals are a result of prejudice and discrimination brought about by gender non-conformity (Rood et al., 2016; Zucker, 2019), and that gender transition will ameliorate these symptoms. Some even claim that gender-affirmative care will successfully treat not only depression and anxiety but will also resolve neurocognitive deficits frequently present in gender-dysphoric individuals (Turban, 2018; Turban, King, Carswell, & Keuroghlian, 2020; Turban & van Schalkwyk, 2018). These latter assertions have proven controversial even among the proponents of gender-affirmative interventions (Strang et al., 2018; van der Miesen, Cohen-Kettenis, & de Vries, 2018). The minority stress theory as the sole explanatory mechanism for co-occurring mental health illness has also been questioned in light of the evidence that psychiatric symptoms frequently predate the onset of gender dysphoria (Bechard, VanderLaan, Wood, Wasserman, & Zucker, 2017; Kaltiala-Heino, Sumia, Työläjärvi, & Lindberg, 2015; Kozlowska, Chudleigh, McClure, Maguire,

& Ambler, 2021). Other clinicians recognize the limits of gender-affirmative care and are aware that youth with underlying psychiatric issues are likely to continue to struggle post-transition (Kaltiala, Heino, Työläjärvi, & Suomalainen, 2020), but, unaware of alternative approaches such as gender-exploratory psychotherapy or watchful waiting (Bonfatto & Crasnow, 2018; Churcher Clarke & Spiliadis, 2019; Spiliadis, 2019), these well-meaning professionals continue to treat youth with gender-affirmative interventions despite lingering doubts.

It is common for gender-affirmative specialists to erroneously believe that gender-affirmative interventions are a *standard of care* (Malone, D'Angelo, Beck, Mason, & Evans, 2021; Malone, Hruz, Mason, Beck, et al:, 2021). Despite the increasingly widespread professional beliefs in the safety and efficacy of pediatric gender transition, and the endorsement of this treatment pathway by a number of professional medical societies, the best available evidence suggests that the benefits of gender-affirmative interventions are of very low certainty (Clayton et al., 2021; National Institute for Health & Care Excellence, 2020a; 2020b) and must be carefully weighed against the health risks to fertility, bone, and cardiovascular health (Alzahrani et al., 2019; Biggs, 2021; Getahun et al., 2018; Hembree et al., 2017; Nota et al., 2019). Recently, emphasis has also been placed on psychosocial risks and as yet unknown medical risks (Malone, D'Angelo, et al., 2021).

Five scientific observations question and refute the assumption that an individual's experience of incongruence of sex and gender identity is best addressed by supporting the newly assumed gender identity with psychosocial and medical interventions.

1. The most foundational aspect of the diagnoses of "gender dysphoria" (DSM-5) and "gender incongruence" (ICD-11), requisite for the provision of medical treatment, is in flux, as professionals disagree on whether the presence of distress is a key diagnostic criterion, as stated in the DSM-5, or is irrelevant, as is the case according to the latest ICD-11 criteria (American Psychiatric Association, 2013; World Health Organization, 2019). Further, these diagnoses have never been properly field-tested (de Vries et al., 2021).
2. There are no randomized controlled studies demonstrating the superiority of various affirmative interventions compared to alternatives. There isn't even agreement about which outcome measures would be ideal in such studies.
3. There are few long-term follow-up studies of various interventions using predetermined outcome measures at designated intervals. Studies that have been conducted are, at best, inconsistent. Higher quality studies with longer-follow-up fail to demonstrate durable positive impacts on mental health (Bränström & Pachankis, 2020a; 2020b).
4. Rates of post-transition desistance, increased mental suffering, increased incidence of physical illness, educational failure, vocational inconstancy, and social isolation have not been established.
5. Numerous cross-sectional and prospective studies of transgender adults consistently demonstrate a high prevalence of serious mental health and social problems as well as suicide (Asscheman et al., 2011; Dhejne et al., 2011). Controversies about how to deal with trans-identified youth must consider the well described vulnerabilities of transgender adults.

It is equally important to realize that to date, research about alternative approaches, such as psychotherapy or watchful waiting, shares the scientific limitations of the research of more invasive interventions: there are no control groups, nor is there systematic follow-up at predetermined intervals with predetermined means of measurement (Bonfatto & Crasnow, 2018; Churcher Clarke & Spiliadis, 2019; Spiliadis, 2019). Parents and patients need to be informed of this as well.

Perhaps the single most problematic assumption held by some gender clinicians is that the young patients have simply been "born in the wrong body." This assumption seemingly frees clinicians from having to contend with the ethical dilemmas of recommending body-altering

interventions that are based on very low-quality evidence. Despite the principle of development that biology, psychosocial factors, and culture generate behavior, these clinicians may believe that atypical genders are created by biology. This reductionistic approach has been criticized repeatedly (Kendler, 2019).

While the origins of childhood or adolescent onset of gender incongruence have not yet been fully elucidated, brain studies of increasing technical sophistication have yet to demonstrate a distinct structure or pattern that accounts for an atypical gender identity, after statistically controlling for sexual orientation and exposure to exogenous hormones (Frigerio, Ballerini, & Valdés Hernández, 2021). Twin studies also demonstrate that while biology plays a role in one's experience of "gender incongruence," it is far from deterministic (Diamond, 2013).

A growing number of clinicians and researchers are noting that the dramatic rise of teens declaring a trans identity appears to be, at least in part, a result of peer influence (Anderson, 2022; Hutchinson, Midgen, & Spiliadis, 2020; Littman 2018; Littman, 2020; Zucker, 2019). Some have noted yet another influx of trans-identified youth emerging during the COVID lockdowns, and have hypothesized that increased isolation coupled with heavy internet exposure may be responsible (Anderson, 2022). While the research into the phenomenon of social influence as a contributor to trans identification of youth is still in its infancy, the possibility that clinicians are providing treatments with permanent consequences to address what may be transient identities in youth poses a serious ethical dilemma.

c. *Poor evaluations*

There is a growing recognition that rapid evaluations which disregard factors contributing to the development of gender dysphoria in youth are problematic. In November 2021, two-leaders of the World Professional Organization for Transgender Health (WPATH) warned the medical community that the "The mental health establishment is failing trans kids" (Anderson & Edwards-Leeper, 2021). Frequently, evaluations provided by gender clinicians may only ascertain the diagnosis of *gender dysphoria* (DSM-5) or its ICD-11 counterpart *gender incongruence*, and screen for conspicuous mental illness prior to recommending hormones and surgeries. These limited, abbreviated evaluations overlook, and as a result fail to address, the relevant issue of the forces that may have influenced the young person's current gender identity.

Confirming the young person's self-diagnosis of gender dysphoria or gender incongruence is easy. Clarifying the developmental forces that have influenced it and determining an appropriate intervention are not. Contextualizing these forces involves an understanding of child and adolescent developmental processes, childhood adversity, co-existing physical and cognitive disadvantages, unfortunate parental or family circumstances (Levine, 2021), as well as the role of social influence (Anderson, 2022; Anderson & Edwards-Leeper, 2021; Littman, 2018; 2021).

The poor quality of mental health evaluations has been a point of significant discontent for a growing number of parents of gender-dysphoric youth. Increasingly, parents have formed dozens of support groups in North America, Europe, Australia and New Zealand, united in their objections to the idea that the best or the only treatment for their gender-dysphoric children is affirmation (Genspect, 2021). These distressed parents, recognizing that their son or daughter may eventually decide to present to others as a trans person, want a psychotherapeutic investigation to understand what contributed to the development of this identity and an exploration of noninvasive treatment options. Frequently, they cannot find anyone in their community who does not recommend immediate affirmation.

The American Academy of Pediatrics' Committee of Bioethics recognizes that "parents…are better situated than others to understand the unique needs of their children and to make appropriate, caring decisions regarding their children's health care" (Katz et al., 2016). The plight of the families unable to find specialists capable of conducting thorough evaluations draws attention to the widespread acceptance of medical interventions for gender-dysphoric youth as the first line of treatment. The problem is that such care has been established through precedent rather

than through scientific demonstrations of its efficacy. We contend that parents and patients have a right to know this, and that it is the professionals' responsibility and obligation to inform them of the state of knowledge in this arena of care.

d. *Incorrect information shared*

In sharing the information with patients and families, two key areas of uncertainty must be emphasized. The first one is the uncertain permanence of a child's or an adolescent's gender identity (Littman, 2021; Ristori & Steensma, 2016; Singh, Bradley, & Zucker, 2021; Vandenbussche, 2021; Zucker, 2017). The second is the uncertain long-term physical and psychological health outcomes of gender transition (National Institute for Health & Care Excellence, 2020a; 2020b). Unfortunately, gender specialists are frequently unfamiliar with, or discount the significance of, the research in support of these two concepts. As a result, the informed consent process rarely adequately discloses this information to patients and their families.

Problematically, it is common for gender clinicians to emphasize the risk of suicide if a young person's wish to transition gender is not immediately fulfilled. There is a significant amount of misinformation surrounding the question of suicidality of trans-identified youth (Biggs, 2022). Providers of gender-affirmative care should be careful not to unwittingly propagate misinformation regarding suicide to parents and youths. They should also be reminded that any conversations about suicide should be handled with great care, due to its socially contagious nature (Bridge et al., 2020; HHS, 2021).

i. High rate of desistance/natural resolution of gender dysphoria in children is not disclosed

There have been eleven research studies to date indicating a high rate of resolution of gender incongruence in children by late adolescence or young adulthood without medical interventions (Cantor, 2020; Ristori & Steensma, 2016; Singh et al., 2021). An attempt has been made to discount the applicability of this research, suggesting that the studies were based on merely gender non-conforming, rather than truly gender-dysphoric, children (Temple Newhook et al., 2018). However, a reanalysis of the data prompted by this critique confirmed the initial finding: Among children meeting the diagnostic criteria for "Gender Identity Disorder" in DSM-IV (currently "Gender Dysphoria in DSM-5), 67% were no longer gender-dysphoric as adults; the rate of natural resolution for gender dysphoria was 93% for children whose gender dysphoria was significant but subthreshold for the DSM diagnosis (Zucker, et al., 2018). It should be noted that high resolution of childhood-onset gender dysphoria had been recorded before the practice of social transition of young children was endorsed by the American Academy of Pediatrics (Rafferty et al., 2018). It is possible that social transition will predispose a young person to persistence of transgender identity long-term (Zucker, 2020).

The information regarding the resolution of gender dysphoria among those with adolescent-onset gender dysphoria, which is currently the predominant presentation, is less clear. A growing body of evidence suggests that for many teens and young adults, a post-pubertal onset of transgender identification can be a transient phase of identity exploration, rather than a permanent identity, as evidenced by a growing number of young detransitioners (Entwistle, 2020; Littman, 2021; Vandenbussche, 2021). Previously, the rate of detransition and regret was reported to be very low, although these estimates suffered from significant limitations and were likely undercounting true regret (D'Angelo, 2018). However, in the last several years since gender-affirmative care has become popularized, the rate of detransition appears to be accelerating.

According to a recent study from a UK adult gender clinic, 6.9% of those treated with gender-affirmative interventions detransitioned within only 16 months of starting treatment, and another 3.4% had a pattern of care suggestive of detransition, yielding a rate of probable detransition in excess of 10%. Another 21.7% of patients disengaged from the clinic without completing

their treatment plan (Hall, Mitchell, & Sachdeva, 2021). While some of these individuals later reengaged with the gender service, the authors concluded, "detransitioning might be more frequent than previously reported." Another study from a UK primary care practice found that 12.2% of those who had started hormonal treatments either detransitioned or documented regret, while the total of 20% stopped the treatments for a wider range of reasons. The mean age of their presentation with gender dysphoria was 20, and the patients had been taking gender-affirming hormones for the average 5 years (17 months-10 years) prior to discontinuing.

Comparing these much higher rates of treatment discontinuation and detransition to the significantly lower rates reported by the older studies, the researchers noted: "Thus, the detransition rate found in this population is novel and questions may be raised about the phenomenon of overdiagnosis, overtreatment, or iatrogenic harm as found in other medical fields" (Boyd, Hackett, & Bewley, 2022 p.15). Indeed, given that regret may take up to 8-11 years to materialize (Dhejne, Öberg, Arver, & Landén, 2014; Wiepjes et al., 2018), many more detransitioners are likely to emerge in the coming years. Detransitioner research is still in its infancy, but two recently published studies examining detransitioner experiences report that detransitioners from the recently-transitioning cohorts feel they had been rushed to medical gender-affirmative interventions with irreversible effects, often without the benefit of appropriate, or in some instances any, psychologic exploration (Littman, 2021; Vandenbussche, 2021).

Clinicians should also disclose to patients and parents that there is no test which can accurately predict who will persist in their transgender identification upon reaching mature adulthood (Ristori & Steensma, 2016). Families should be made aware that a period of strong cross-sex identification in childhood is commonly associated with future homosexuality (Korte et al., 2008). Research in desistance confirms that the majority of youth whose gender dysphoria resolves naturally do indeed grow up to be gay, lesbian, or bisexual adults (Cantor, 2020, Appendix; Singh et al., 2021).

ii. Implications of very low-quality evidence that underlies the practice of pediatric gender transition are not explained

The evidence underlying the practice of pediatric gender transition is widely recognized to be of very low quality (Hembree et al., 2017). In 2020, the most comprehensive systematic review of evidence to date, commissioned by the UK National Health System (NHS) and conducted by the National Institute for Health and Care Excellence (NICE), concluded that the evidence for both puberty blocking and cross-sex hormones is of very low certainty (National Institute for Health & Care Excellence, 2020a; 2020b).

According to the NICE review of evidence for puberty blockers, the studies "are all small, uncontrolled observational studies, which are subject to bias and confounding, and are of very low certainty as assessed using modified GRADE [Grading of Recommendations, Assessment, Development and Evaluations]. All the included studies reported physical and mental health comorbidities and concomitant treatments very poorly" (National Institute for Health & Care Excellence, 2020a, p.13). NICE reached similar conclusions regarding the quality of the evidence for cross-sex hormones (National Institute for Health & Care Excellence, 2020b).

Problematically, the implications of administering a treatment with irreversible, life-changing consequences based on evidence that has an official designation of "very low certainty" according to modified GRADE is rarely discussed with the patients and the families. GRADE is the most widely adopted tool for grading the quality of evidence and for making treatment recommendations worldwide. GRADE has four levels of evidence, also known as certainty in evidence or quality of evidence: very low, low, moderate, and high (BMJ Best Practice, 2021). When evidence is assessed to be "very low certainty," there is a high likelihood that the patients will not experience the effects of the proposed interventions (Balshem et al., 2011).

In the context of providing puberty blockers and cross-sex hormones, the designation of "very low certainty" signals that the body of evidence asserting the benefits of these interventions is

highly unreliable. In contrast, several negative effects are quite certain. For example, puberty blockade followed by cross-sex hormones leads to infertility and sterility (Laidlaw, Van Meter, Hruz, Van Mol, & Malone, 2019). Surgeries to remove breasts or sex organs are irreversible. Other health risks, including risks to bone and cardiovascular health, are not fully understood and are uncertain, but the emerging evidence is alarming (Alzahrani et al., 2019; Biggs, 2021).

iii. The question of suicide is inappropriately handled

Suicide among trans-identified youth is significantly elevated compared to the general population of youth (Biggs, 2022; de Graaf et al., 2020). However, the "transition or die" narrative, whereby parents are told that their only choice is between a "live trans daughter or a dead son" (or vice-versa), is both factually inaccurate and ethically fraught. Disseminating such alarmist messages hurts the majority of trans-identified youth who are not at risk for suicide. It also hurts the minority who are at risk, and who, as a result of such misinformation, may forgo evidence-based suicide prevention interventions in the false hopes that transition will prevent suicide.

The notion that trans-identified youth are at alarmingly high risk of suicide usually stems from biased online samples that rely on self-report (D'Angelo et al., 2020; James et al., 2016; The Trevor Project, 2021), and frequently conflates suicidal thoughts and non-suicidal self-harm with serious suicide attempts and completed suicides. Until recently, little was known about the actual rate of suicide of trans-identified youth. However, a recent analysis of data from the biggest pediatric gender clinic in the world, the UK's Tavistock, found the rate of completed youth suicides to be 0.03% over a 10-year period, which translates into the annual rate of 13 per 100,000 (Biggs, 2022). While this rate is significantly elevated compared to the general population of teens, it is far from the epidemic of trans suicides portrayed by the media.

The "transition or die" narrative regards suicidal risk in trans-identified youth as a different phenomenon than suicidal risk among other youth. Making them an exception falsely promises the parents that immediate transition will remove the risk of suicidal self-harm. Trans patients themselves complain about the so-called "trans broken arm syndrome" – a frustrating pattern whereby physicians "blame" all the problems the patients are experiencing on their trans status, and a result, fail to perceive and respond to other sources of distress (Paine, 2021). Clinicians caring for trans-identified youth should be reminded that suicide risk in all patients is a multi-factorial phenomenon (Mars et al., 2019). To treat trans youths' suicidality as an exception is to deny them evidence-based care.

A recent study of three major youth clinics concluded that suicidality of trans-identifying teens is only somewhat elevated compared to that of youth referred for mental health issues unrelated to gender identity struggles (de Graaf et al., 2020). Another study found that transgender-identifying teens have relatively similar rates of suicidality compared to teens who are gay, lesbian and bisexual (Toomey, Syvertsen, & Shramko, 2018). Depression, eating disorders, autism spectrum conditions, and other mental health conditions commonly found in transgender-identifying youth (Kaltiala-Heino, Bergman, Työläjärvi, & Frisen, 2018; Kozlowska, McClure, et al., 2021; Morandini, Kelly, de Graaf, Carmichael, & Dar-Nimrod, 2021) are all known to independently contribute to the probability of suicide (Biggs, 2022; Simon & VonKorff, 1998; Smith, Zuromski, & Dodd, 2018).

The "transition or suicide" narrative falsely implies that transition will prevent suicides. Clinicians working with trans-identified youth should be aware that although in the short-term, gender-affirmative interventions can lead to improvements in some measures of suicidality (Kaltiala et al., 2020), neither hormones nor surgeries have been shown to reduce suicidality in the long-term (Bränström & Pachankis, 2020a; 2020b). Alarmingly, a longitudinal study from Sweden that covered more than a 30-year span found that adults who underwent surgical transition were 19 times more likely than their age-matched peers to die by suicide overall, with female-to-male participants' risk 40 times the expected rate (Dhejne et al., 2011, Table S1).

Another key longitudinal study from the Netherlands concluded that suicides occur at a similar rate at all stages of transition, from pretreatment assessment to post-transition follow-up (Wiepjes et al., 2020). The data from the Tavistock clinic also did not show a statistically significant difference between completed suicides in the "waitlist" vs. the "treated" groups (Biggs, 2022). Luckily, in both groups, completed suicides were rare events (which may have been responsible for the lack of statistical significance). Thus, we consider the "transition or die" narrative to be misinformed and ethically wrong.

In our experience working with trans-identified youth, an adolescent's suicidality can sometimes arise as a response to parental distress, resistance, skepticism, or wish to investigate the forces shaping the new gender identity before social transition and hormone therapy. When mental health professionals or other healthcare providers fail to recognize the legitimacy of parental concerns, or label the parents as transphobic, this only tends to intensify intrafamilial tension. Clinicians would be well-advised that gender transition is not an appropriate response to suicidal intent or threat, as it ignores the larger mental health and social context of the young patient's life—the entire family is often in crisis. Trans-identified adolescents should be screened for self-harm and suicidality, and if suicidal behaviors are present, an appropriate evidence-based suicide prevention plan should be put in place (de Graaf et al., 2020).

The Dutch Study: the questionable basis for the gender affirmative model of care for youth

Few practitioners of gender-affirmative interventions, and even fewer patients and families, realize that the foundation of the practice of medically transitioning minors stems from a single Dutch proof of concept study, the outcomes of which were documented in two publications (de Vries, Steensma, Doreleijers, Cohen, & Kettenis, 2011; de Vries et al., 2014). The former (de Vries et al., 2011) reported on cases who underwent puberty blockade, while the latter (de Vries et al., 2014) reported on a subset of the cases who completed surgeries.

The Dutch study subjects' high level of psychological functioning at 1.5 years after surgery, which was the study end point, was an impressive feat. However, both of the studies suffer from a high risk of bias due to their study design, which is effectively a non-randomized case series—one of the lowest levels of evidence (Mathes & Pieper, 2017; National Institute for Health & Care Excellence, 2020a). In addition, the studies suffer from limited applicability to the populations of adolescents presenting today (de Vries, 2020). The interventions described in the study are currently being applied to adolescents who were not cross-gender identified prior to puberty, who have significant mental health problems, as well as those who have non-binary identities—all of these presentations were explicitly disqualified from the Dutch protocol. Despite these limitations, the Dutch clinical experiment has become the basis for the practice of medical transition of minors worldwide and serves as the basis for the recommendations outlined in the 2017 Endocrine Society guidelines (Hembree et al., 2017).

We contend that the Dutch studies have been misunderstood and misrepresented as providing evidence of the safety and efficacy of these interventions for all youth. It is important that both the strengths and the weaknesses of these two studies are understood, as to date, the Dutch experience presents the best available evidence behind the practice of pediatric gender transition.

Rationale for pediatric transition

Prior to the 1990s, gender transitions were typically initiated in mature adults (Dhejne et al., 2011). However, it was noted that particularly for natal male patients, hormonal and surgical interventions failed to achieve satisfactory results, and patients had a "never disappearing masculine appearance" (Delemarre-van de Waal & Cohen-Kettenis, 2006). The lack of adequate cosmetic outcomes was thought to contribute to the frequently disappointing outcomes of medical

gender transition, with persistently high rates of mental illness and suicidality post-transition (Delemarre-van de Waal & Cohen-Kettenis, 2006; Dhejne et al., 2011; Ross & Need, 1989).

In the mid 1990s, a team of Dutch researchers hypothesized that by carefully selecting a subset of gender-dysphoric children who would likely be transgender-identified for the rest of their lives, and by medically intervening before puberty left an irreversible mark on their bodies, the cosmetic outcomes would be improved—and as a result, mental health outcomes might be improved (Gooren & Delemarre-van de Waal, 1996).

Mixed study findings

In 2014, the Dutch research team published a key longitudinal study of mental health outcomes of 55 youths who completed medical and surgical transition (de Vries et al., 2014). The 2014 paper (sometimes referred to as the "Dutch study") reported that for youth with severe gender dysphoria that started in early childhood and persisted into mid-adolescence, a sequence of puberty blockers, cross-sex hormones, and breast and genital surgeries (including a mandatory removal of the ovaries, uterus and testes), with ongoing extensive psychological support, was associated with positive mental health and overall function 1.5 years post-surgery.

While the Dutch reported resolution of gender dysphoria post-surgery in study subjects, the reported psychological improvements were quite modest (de Vries et al., 2014). Of the 30 psychological measurements reported, nearly half showed no statistically significant improvements, while the changes in the other half were marginally clinically significant at best (Malone, D'Angelo, et al., 2021). The scores in anxiety, depression, and anger did not improve. The change in the Children's Global Assessment Scale, which measures overall function, was one of the most impressive changes—however it too remained in the same range before and after treatment (de Vries et al., 2014).

Problematic discordance between reduced gender dysphoria and lack of meaningful improvements in psychological measures

The discordance between the marked reduction in gender dysphoria, as measured by the UGDS (Utrecht Gender Dysphoria Scale), and the lack of meaningful changes in psychological function using standard measures, warrants further examination. There are three plausible explanations for this lack of agreement. Any one of these three explanations calls into question the widely assumed notion that the medical interventions significantly improve mental health or lessen or eradicate gender dysphoria.

One possible explanation is that gender dysphoria as measured by UGDS, and psychological function as measured by most standard instruments, are not correlated. This contradicts the primary rationale for providing gender-affirmative treatments for youth (which is to improve psychological health and functioning), and if true, ethically threatens these medical interventions. The other plausible explanation stems from the high psychological function of all the subjects at baseline; the subjects were selected because they were free from significant mental health problems (de Vries et al., 2014). As a result, there was little opportunity to meaningfully improve. This explanation highlights a key limitation in applying the study's results to the majority of today's gender-dysphoric youth, who often present with a high burden of mental illness (Becerra-Culqui et al., 2018; Kozlowska, McClure, et al., 2021). The study cannot be used as evidence that these procedures have been proven to improve depression, anxiety, and suicidality.

A third possible explanation for the discordance between only minor changes in psychological outcomes but a significant drop in gender dysphoria comes from a close examination of the UGDS scale itself and how it was used by the Dutch researchers. This 12-item scale, designed by the Dutch to assess the severity of gender dysphoria and to identify candidates for hormones

and surgeries, consists of "male" (UGDS-aM) and "female" (UGDS-aF) versions (Iliadis et al., 2020). At baseline and after puberty suppression, biological females were given the "female" scale, while males were given the "male" scale. However, post-surgery, the scales were flipped: biological females were assessed using the "male" scale, while biological males were assessed on the "female" scale (de Vries et al., 2014). We maintain that this handling of the scales may have at best obscured, and at worst, severely compromised the ability to meaningfully track how gender dysphoria was affected throughout the treatment.

Consider this example. At baseline, a gender-dysphoric biological female would rate items from the "female" scale such as: "I prefer to behave like a boy" (item 1); "I feel unhappy because I have to behave like a girl" (item 6) and "I wish I had been born a boy" (item 12). Positive answers to these questions would have contributed to a high baseline gender dysphoria score. After the final surgery, however, this same patient would be asked to rate items from the "male" scale, including the following: "My life would be meaningless if I had to live as a boy" (item 1); "I hate myself because I am a boy" (item 6) and "It would be better not to live than to live as a boy" (item 12). A gender-dysphoric female would not endorse these statements (at any stage of the intervention), which would lead to a lower gender dysphoria score.

Thus, the detected drop in the gender dysphoria scores for biological males and females may have had less to do with the success of the interventions, and more to do with switching the scale from the "female" to the "male" version (and vice-versa) between the baseline and post-surgical period. This, too, may explain why no changes in gender dysphoria were noted between baseline and the puberty blockade phase, and were only recorded after the final surgery, when the scale was switched.

It must be considered that had the researchers administered the "flipped" scale earlier, at the completion of the puberty blocker stage, UGDS scale could have registered a reduction in gender dysphoria. Likewise, however, one must consider the possibility that had *both sets of scales* been administered to the same individual at baseline, a "reduction" in gender dysphoria could have been registered upon switching of the scale, *well before any interventions began*. The question here is whether the diminishment of quantitative measures of gender dysphoria is largely an artifact of what scale was used.

It must be noted that the UGDS measure has been demonstrated only to effectively differentiate between clinically referred gender-dysphoric individuals, non-clinically referred controls, and participants with disorders of sexual development, and was not designed to detect changes in gender dysphoria during treatment (Steensma, McGuire, Kreukels, et al. 2013). The presence of items such as "I dislike having erections" (item 11, UGDS-aM), which would have to be rated by birth-females, and "I hate menstruating because it makes me feel like a girl (item 10, UGDS-aF), which would be presented to birth-males, neither of which could be meaningfully rated by either at any stage of the interventions, further illustrates that UGDS has questionable validity for the purpose of detecting meaningful changes in gender dysphoria as a result of medical and surgical treatment.

The updated UGDS scale (UGDS-GS), developed by the Dutch after the publication of their seminal study, has eliminated the two-sex version of the scale in favor of a single battery of questions applicable to both sexes (McGuire et al., 2020). This change may lead to a more reliable measurement of treatment-associated changes in future research. Other gender dysphoria scales also exist (Hakeem, Črnčec, Asghari-Fard, Harte, & Eapen, 2016; Iliadis et al., 2020) and may or may not be better suited for the purposes of measuring the impact of medical interventions on underlying gender distress. Gender dysphoria, of course, may also prove to be a more complex concept than can be measured by any scale.

Other limitations

The two Dutch studies were conducted without a control group (de Vries et al., 2011; de Vires et al., 2014). Nor could the researchers control for mental health interventions, which all the

subjects received in addition to hormones and surgery. The Dutch only evaluated mental health outcomes and did not assess physical health effects of hormones and surgery. The sample size was small: the final study reported the outcomes of only 55 children, and as few as 32 were evaluated on key measures of psychological outcomes.

It is important to realize that the Dutch sample was carefully selected, which introduced a source of bias, and also challenges the study's applicability. From the 196 adolescents initially referred, 111 were considered eligible to start puberty blockers, and of this group, only the 70 most mature and mentally stable who proceeded to cross-sex hormones were included in the study (de Vries et al., 2011). Of note, 97% of the selected cases were attracted to members of their natal sex at baseline. All were cross-sex identified, with no cases of nonbinary identities. The final study only followed 55, rather than the original 70 cases, further excluding from reporting the outcomes of subjects who had experienced adverse events, including: one death from surgery-related complications and three cases of obesity and diabetes that rendered subjects ineligible for surgery. Three more subjects refused to be contacted or dropped out of care, which may mask adverse outcomes (de Vries et al., 2014).

There is no knowledge of the fate of 126 patients who did not participate in the Dutch study. Longer term outcomes of the subjects who did participate are lacking. We are aware of only one case of long-term follow-up for a female-to-male patient treated by the Dutch team in the 1990s. The case study describing the subject's functioning at the age of 33 found that the patient did not regret gender transition. However, he reported struggling with significant shame related to the appearance of his genitals and to his inability to sexually function; had problems maintaining long-term relationships; and experienced depressive symptoms (Cohen-Kettenis, Schagen, Steensma, de Vries, & Delemarre-van de Waal, 2011). Notably, these problems had not yet emerged when the same patient was assessed at the age of 20, when he reported high levels of satisfaction in general, and was "very satisfied with the results [of the metoidioplasty]" in particular (Cohen-Kettenis & van Goozen, 1998, p.248). Since the last round of psychological outcomes of the individuals in the Dutch study was obtained when the subjects were around 21 years of age (de Vries et al., 2014), it raises questions how they will fare during the decade when new developmental tasks, such as career development, forming long-term intimate relationships and friendships, or starting families come into focus.

As to the unknown outcomes of the patients rejected by the Dutch protocol, one study did report on 14 adolescents who sought gender reassignment in the same clinic, but were disqualified from treatment due to "psychological or environmental problems" (Smith, Van Goozen, & Cohen-Kettenis, 2001, p. 473). The study found that at follow-up 1-7 years after the original application, 11 of the 14 no longer wished to transition, and 2 others only slightly regretted not transitioning (Malone, D'Angelo, et al., 2021; Smith et al., 2001). This further underscores the importance of conducting research utilizing control groups and following the subjects for an extended period.

A recent attempt to replicate the results of the first Dutch study (de Vries et al., 2011) found no demonstrable psychological benefit from puberty blockade, but did find that the treatment adversely affected bone development (Carmichael et al., 2021). The final Dutch study (de Vries et al., 2014) has never been attempted to be replicated with or without a control group.

The scaling of the Dutch Protocol beyond original indications

The medical and surgical sequence of Dutch protocol has been aggressively scaled worldwide without the careful evaluations and vetting practiced by the Dutch. The protocol's original investigators have recently expressed concern that the interventions they described have been widely adopted on four continents without several of the protocol's essential discriminatory features (de Vries, 2020).

The extensive multi-year multidisciplinary evaluations of the children have been abbreviated or simply bypassed. The medical sequence is routinely used for children with post-pubertal onset of transgender identities complicated by mental health comorbidities (Kaltiala-Heino et al., 2018), and not just for those high-functioning adolescents with persistent early life cross-identifications, as was required by the Dutch protocol (de Vries & Cohen-Kettenis, 2012). Further, it has become increasingly common to socially transition children before puberty (Olson, Durwood, DeMeules, & McLaughlin, 2016), even though this was explicitly discouraged by the Dutch protocol at the time (de Vries & Cohen-Kettenis, 2012).

In addition, medical transition is frequently initiated much earlier than recommended by the original protocol (de Vries & Cohen-Kettenis, 2012). The authors of the protocol were aware that most children would have a spontaneous realignment of their gender identity with sex by going through early- to mid-stages of puberty (Cohen-Kettenis, Delemarre-van de Waal, & Gooren, 2008). The average age of initiating puberty blockade in the Dutch study was around 15. In contrast, currently the age limit has been lowered to the age of Tanner stage II, which can occur as early as 8-9 years (Hembree et al., 2017). Irreversible cross-sex hormones, initiated in the Dutch study at the average age of nearly 17, are currently commonly prescribed to 14-year-olds, and this lower age threshold has been recommended by WPATH Standard of Care 8 draft, the final version of which is due to be released in early 2022. The fact that children are transitioned before their identity is tested against the biological reality and before natural resolution of gender dysphoria has had a chance to occur is a major deviation from the original Dutch protocol. Systematic follow-up, reassessments, and tracking and publishing of outcomes are not performed.

As the lead Dutch researchers have begun to call for more research into the novel presentation of gender dysphoria in youth (de Vries, 2020; Voorzij, 2021) and question the wisdom of applying the hormonal and surgical treatment protocols to the newly presenting cases, many recently educated gender specialists mistakenly believe that the Dutch protocol proved the concept that its sequence helps all gender-dysphoric youth. Although aware of the Dutch study's importance, they seem to be unaware of its agreed upon limitations, and the Dutch clinicians' own discomfort that most new trans-identified adolescents presenting for care today significantly differ from the population the Dutch had originally studied. These facts, of course, underscore the need for a robust informed consent process.

The recommendations for informed consent process for children, adolescents, and young adults

Consent for all stages of gender transition should be explicit, not implied

Noninvasive medical care or care that carries little risk of harm does not require a signed informed consent document; rather, consent is implied through the act of a patient presenting for care. For example, when a parent brings in a child for a skin laceration or abscess, consent for sutures or simple incision and drainage is implied. Similarly, when a child presents with pneumonia and is hospitalized, consent for chest x-ray, IV fluids, and antibiotics is also implied. It is assumed that patients or their guardians agree to the interventions and understand the benefits and risks. When risks are greater, such as prior to surgery, chemotherapy, or another invasive procedure, an informed consent document is signed. Such situations require an explicit, or express informed consent.

In the context of interventions for gender dysphoria or gender incongruence, the uncertainties associated with puberty blocking, cross-sex hormones, and gender-affirmative surgeries are well-recognized (Manrique et al., 2018; National Institute for Health & Care Excellence, 2020a; 2020b; Wilson et al., 2018). In these cases, consent should be explicit rather than implied because of the complexity, uncertainty, and risks involved.

Informed consent for social transition represents a gray area. Evidence suggests that social transition is associated with the persistence of gender dysphoria (Hembree et al.,

2017; Steensma, McGuire, Kreukels, Beekman, & Cohen-Kettenis, 2013). This suggests that social gender transition is a form of a psychological intervention with potential lasting effects (Zucker, 2020). While the causality has not been proven, the possibility of iatrogenesis and the resulting exposure to the risks of future medical and surgical gender dysphoria treatments, qualifies social gender transition for explicit, rather than implied, consent.

Full unbiased disclosure of benefits, risks and alternatives is requisite

When mental health professionals are involved in evaluations and recommendations, the informed consent process begins either as part of an extended evaluation or is integrated in a psychotherapeutic process, separately or together, with the parents and patient. When pediatricians, nurse practitioners, or primary care physicians perform the initial evaluation, the informed consent process is more likely to be labeled as such in a briefer series of meetings.

In all settings, the informed consent discussions for gender-affirmative care should include three central ideas:

1. The decision to initiate gender transition may predispose the child to persist in their transgender identity long-term.
2. Many of the physical changes contemplated and undertaken are irreversible.
3. Careful long-term studies have not been done to verify that these interventions enable better physical and mental health or improved social functioning, or that they do not cause harm.

The informed consent process, culminating with a signed document, signifies that parents and patient have been educated about the short- and long-term risks, benefits and uncertainties associated with all relevant stages of the gender-affirmative interventions. The process must also inform the patients and families about the full range of alternative treatments, including the choice of not socially or medically treating the child's or adolescent's current state of gender/body incongruence.

Decisional capacity to consent needs to be assessed and family should be involved

Trans-identified youth typically present themselves as strongly desiring hormones and ultimately, surgery. It should not be assumed that their eagerness is matched with the capacity to carefully consider the consequences of their realized desires. Trans-identified youth younger than the age of consent should be part of the informed consent process, but they may not be mature enough to recognize or admit their concerns about the proposed intervention. For this reason, it is the parents who, after careful consideration, are responsible for signing an informed consent document.

The issue of the exact age at which adolescents are mature enough to consent to gender transition has proven contentious: courts have been asked to decide about competence to consent to gender-affirmative hormones for youth in the United Kingdom and Australia (Ouliaris, 2021). In the United States, the legal age for medical consent for gender-affirmative interventions varies by state.

When patients are age 18 and older, and in some jurisdictions as young as age 15 (Right to medical or dental treatment without parental consent, 2010), they do not legally require parental approval for medical procedures. But because an individual's change of gender has profound implications for parents, siblings, and other family members, it is usually prudent for clinicians to seek their input directly or indirectly during the informed consent process. This is done by requesting a meeting with the parents.

A recent study by a Dutch research team attempted to evaluate the decisional capacity of adolescents embarking on gender transition (Vrouenraets, de Vries, de Vries, van der Miesen, & Hein, 2021). The researchers administered the MacCAT-T tool, comprised of the *understanding*, *appreciating*, *reasoning*, and *expressing a choice* domains, to 74 adolescents who were 14.7 years old on average (with the minimum age of 10). They concluded that the adolescents were competent to consent to starting pubertal suppression, calling for similar research for the <12 group, particularly because "birth-assigned girls … may benefit from puberty suppression as early as 9 years of age" (Vrouenraets et al., 2021 p.7).

This study suffers from two significant limitations involving the MacCAT-T tool. It was never designed for children. Rather, it was designed to assess medical consent capacities of adults suffering from conditions such as dementia, schizophrenia, and other psychiatric disorders. There is a fundamental lack of equivalency between consenting to treatment by adults with cognitive impairments and obtaining consent from healthy children whose age-appropriate cognitive capacities are intact, but who lack the requisite life experiences to consent to profound life-changing medical interventions. We doubt, for example, whether even highly intelligent children who have not had sexual experiences can meaningfully comprehend the loss of future sexual function and reproductive abilities.

In addition, even for adults, the MacCAT-T tool has been criticized for its exclusive focus on cognitive aspects of capacity, failing to account for the non-cognitive aspects such as values, emotions and other biographic and context specific aspects inherent in the complexity of the decision process in real life (Breden & Vollmann, 2004). Children's values and emotions undergo tremendous change during the process of maturation.

The authors' conclusion about their young patients' competence to consent should be compared with what a panel of judges wrote in the challenge to the Tavistock treatment protocol (Bell v Tavistock, 2020):

> …the clinical intervention we are concerned with here is different in kind to other treatments or clinical interventions. In other cases, medical treatment is used to remedy, or alleviate the symptoms of, a diagnosed physical or mental condition, and the effects of that treatment are direct and usually apparent. The position in relation to puberty blockers would not seem to reflect that description. [para 135]

> …we consider the treatment in this case to be in entirely different territory from the type of medical treatment which is normally being considered. [para 140]

> … the combination here of lifelong and life changing treatment being given to children, with very limited knowledge of the degree to which it will or will not benefit them, is one that gives significant grounds for concern. [para 143]

It seems clear that perceptions of children as young as 10 years of age as medically competent vary by country, state, and the institution where the doctor works, and, by clinicians' beliefs about the long-term benefits of these interventions. We maintain that the claim that children can consent to extremely life-altering intervention is fundamentally a philosophical claim (Clark & Virani, 2021). Our view in this matter is that consent is primarily a parental function.

Informed consent should be viewed as a process rather than an event

Most institutions that care for transgender-identified individuals have devised obligatory consent forms that outline the risks and uncertainties of hormonal and surgical gender-affirmative interventions. However, the requisite signatures are frequently collected in a perfunctory manner (Schulz, 2018), akin to signatures collected ahead of a common surgical procedure. The purpose of such informed consent documents appears to be to protect practitioners from lawsuits, rather than attend to the primary ethical foundation of the process.

Although obtaining the signatures is important, the signed document should signify that the process of informed consent has been undertaken over an extended time period and is not simply quickly completed (Vrouenraets et al., 2021). We believe the latter approach poses an ethical concern (Levine, 2019).

The internal dynamics of the trans-identified young person and their families vary considerably. Parental capacities, their private marital and intrafamilial relationships, their cultural awareness, religious and political sensibilities all influence the amount of time necessary to undertake a thorough informed consent process. It is not prudent to suggest a specific duration for the process of informed consent, other than to emphasize that it requires a slow, patient, thoughtful question and answer period as the parents and patient contemplate the meaning of what is known and unknown and whether to embark on alternative approaches to the management of gender dysphoria before the age of full neurological maturity has been reached, mental health comorbidities have been addressed, and a true informed consent by the patient is more likely.

Final thoughts

Sixty years of experience providing medical and surgical assistance to transgender-identified persons have seen many changes in who is treated, when they are treated, and how they are treated. Today, the emphasis has shifted to the treatment of the unprecedented numbers of youth declaring a trans identity. As adolescents pursue social, medical, and surgical interventions, health care providers may experience unease about patients' cognitive and emotional capacities to make decisions with life-changing and enduring consequences. An unrushed informed consent process helps the provider, the parents, and the patient.

Three issues tend to obscure the salience of informed consent: conspicuous mental health problems, uncertainty about the minor's personal capacity to understand the irreversible nature of the interventions, and parental disagreement. Physical and psychiatric comorbidities can contribute to the formation of a new identity, develop as its consequence, or bear no connection to it. Assessing mental health and the minor's functionality is one of the reasons why rapid affirmative care may be dangerous for patients and their families. For example, when situations involve autism, learning disorders, sexual abuse, attachment problems, trauma, separation anxiety, previous depressed or anxious states, neglect, low IQ, past psychotic illness, eating disorders or parental mental illness, clinicians must choose between ignoring these potentially causative conditions and comorbidities and providing appropriate treatment before affirmative care (D'Angelo et al., 2020).

For youth less than the age of majority, informed consent via parents provides a legal route for treatment but it does not make the decision to socially transition, provide hormones, or surgically remove breasts or testes less fraught with uncertainty. The best that health professionals can do is to ensure that the consent process informs the patient and parents of the current state of science, which is sorely lacking in quality research. It is the professionals' responsibility to ensure that the benefits patients and parents seek, and the risks they are assuming, are clearly appreciated as they prepare to make this often-excruciating decision.

Young people who have reached the age of majority, but who have not reached full maturation of the brain represent a unique challenge. It is well-recognized that brain remodeling proceeds through the third decade of life, with the prefrontal cortex responsible for executive function and impulse control the last to mature (Katz et al., 2016). The growing number of detransitioners who had been old enough to legally consent to transition, but who no longer felt they were transgender upon reaching their mid-20's, raises additional concerns about this vulnerable age group (Littman, 2021; Vandenbussche, 2021).

When the clinician is uncertain whether a young person is competent to comprehend the implications of the desired treatment—that is, when informed consent cannot inform the patient—the clinician may need more time with the patient. When parents or guardians do

not agree about whether to use puberty blockers or cross-sex hormones, clinicians are in an uneasy spot (Levine, 2021). This occurs in both intact and divorced families. Australia has given legal instructions to clinicians facing these uncertainties: the court is to be asked to decide (Ouliaris, 2021). The court system in the UK has been grappling with similar issues in recent years. While it is a rare case that ends up in a courtroom, clinicians devoted to a deliberate informed consent process are still likely to encounter ethical dilemmas that they cannot resolve.

Acknowledgments

The authors wish to thank SEGM staff for their grant and bibliographic support.

Funding

This work was supported by the Society for Evidence-based Gender Medicine.

References

Aitken, M., Steensma, T. D., Blanchard, R., VanderLaan, D. P., Wood, H., Fuentes, A., Spegg, C., Wasserman, L., Ames, M., Fitzsimmons, C. L., Leef, J. H., Lishak, V., Reim, E., Takagi, A., Vinik, J., Wreford, J., Cohen-Kettenis, P. T., de Vries, A. L. C., Kreukels, B. P. C., & Zucker, K. J. (2015). Evidence for an altered sex ratio in clinic-referred adolescents with gender dysphoria. *The Journal of Sexual Medicine*, *12*(3), 756–763. doi:10.1111/jsm.12817

Alzahrani, T., Nguyen, T., Ryan, A., Dwairy, A., McCaffrey, J., Yunus, R., Forgione, J., Krepp, J., Nagy, C., Mazhari, R., & Reiner, J. (2019). Cardiovascular disease risk factors and myocardial infarction in the transgender population. *Circulation: Cardiovascular Quality and Outcomes*, *12*(4). doi:10.1161/CIRCOUTCOMES.119.005597

American College Health Association. (2021). *American College Health Association-National College Health Assessment III: Undergraduate Student Reference Group Data Report Spring 2021*. Boston: ACHA-NCHA III. https://www.acha.org/documents/ncha/NCHA-III_SPRING-2021_UNDERGRADUATE_REFERENCE_GROUP_DATA_REPORT.pdf

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed.). doi:10.1176/appi.books.9780890425596

Anderson, E. (2022, January 3). Opinion: When it comes to trans youth, we're in danger of losing our way. *The San Francisco Examiner.* Retrieved January 5, 2022, from http://www.sfexaminer.com/opinion/are-we-seeing-a-phenomenon-of-trans-youth-social-contagion/

Anderson, E., Edwards-Leeper, L. (2021, November 24). The mental health establishment is failing trans kids. Washington, DC: Washington Post. Retrieved December 20, 2021, from https://www.washingtonpost.com/outlook/2021/11/24/trans-kids-therapy-psychologist/

Asscheman, H., Giltay, E. J., Megens, J. A. J., de Ronde, W. (Pim), van Trotsenburg, M. A. A., & Gooren, L. J. G. (2011). A long-term follow-up study of mortality in transsexuals receiving treatment with cross-sex hormones. *European Journal of Endocrinology*, *164*(4), 635–642. doi:10.1530/EJE-10-1038

Balshem, H., Helfand, M., Schünemann, H. J., Oxman, A. D., Kunz, R., Brozek, J., Vist, G. E., Falck-Ytter, Y., Meerpohl, J., & Norris, S. (2011). GRADE guidelines: 3. Rating the quality of evidence. *Journal of Clinical Epidemiology*, *64*(4), 401–406. doi:10.1016/j.jclinepi.2010.07.015

Becerra-Culqui, T. A., Liu, Y., Nash, R., Cromwell, L., Flanders, W. D., Getahun, D., Giammattei, S. V., Hunkeler, E. M., Lash, T. L., Millman, A., Quinn, V. P., Robinson, B., Roblin, D., Sandberg, D. E., Silverberg, M. J., Tangpricha, V., & Goodman, M. (2018). Mental health of transgender and gender nonconforming youth compared with their peers. *Pediatrics*, *141*(5), e20173845. doi:10.1542/peds.2017-3845

Bechard, M., VanderLaan, D. P., Wood, H., Wasserman, L., & Zucker, K. J. (2017). Psychosocial and psychological vulnerability in adolescents with gender dysphoria: A "proof of principle" Study. *Journal of Sex & Marital Therapy*, *43*(7), 678–688. doi:10.1080/0092623X.2016.1232325

Bell v Tavistock and Portman NHS Foundation Trust. (2020). EWHC 3274. The High Court of Justice (2020). https://www.judiciary.uk/wp-content/uploads/2020/12/Bell-v-Tavistock-Judgment.pdf

Biggs, M. (2021). Revisiting the effect of GnRH analogue treatment on bone mineral density in young adolescents with gender dysphoria. *Journal of Pediatric Endocrinology and Metabolism*. doi:10.1515/jpem-2021-0180

Biggs, M. (2022). Suicide by clinic-referred transgender adolescents in the United Kingdom. *Archives of Sexual Behavior*.

BMJ Best Practice. (2021). What is grade? Retrieved January 1, 2022, from https://bestpractice.bmj.com/info/us/toolkit/learn-ebm/what-is-grade/

Bonfatto, M., & Crasnow, E. (2018). Gender/ed identities: An overview of our current work as child psychotherapists in the Gender Identity Development Service. *Journal of Child Psychotherapy*, *44*(1), 29–46. doi:10.1080/0075417X.2018.1443150

Boyd, I., Hackett, T., & Bewley, S. (2022). Care of transgender patients: A general practice quality improvement approach. *Healthcare*, *10*(1), 121. doi:10.3390/healthcare10010121

Bränström, R., & Pachankis, J. E. (2020a). Reduction in mental health treatment utilization among transgender individuals after gender-affirming surgeries: A total population study. *American Journal of Psychiatry*, *177*(8), 727–734. doi:10.1176/appi.ajp.2019.19010080

Bränström, R., & Pachankis, J. E. (2020b). Correction to Bränström and Pachankis. (2020). *American Journal of Psychiatry*, *177*(8), 734–734. doi:10.1176/appi.ajp.2020.1778correction

Breden, T. M., & Vollmann, J. (2004). The Cognitive Based Approach of Capacity Assessment in Psychiatry: A Philosophical Critique of the MacCAT-T. *Health Care Analysis*, *12*(4), 273–283. doi:10.1007/s10728-004-6635-x

Bridge, J. A., Greenhouse, J. B., Ruch, D., Stevens, J., Ackerman, J., Sheftall, A. H., Horowitz, L. M., Kelleher, K. J., & Campo, J. V. (2020). Association Between the Release of Netflix's 13 Reasons Why and Suicide Rates in the United States: An Interrupted Time Series Analysis. J Am Acad Child Adolesc Psychiatry, *59*(2), 236–243. doi:10.1016/j.jaac.2019.04.020

Byne, W., Bradley, S.J., Coleman, E., Eyler, A.E., Green, R., Menvielle, E.J., Meyer-Bahlburg, H.F.L., Pleak, R.R. & Tompkins, D.A. (2012). Report of the American Psychiatric Association Task Force on Treatment of Gender Identity Disorder. Archives of Sexual Behavior, *41*(4):759–796. doi:10.1007/s10508-012-9975-x

Cantor, J. M. (2020). Transgender and gender diverse children and adolescents: Fact-checking of AAP Policy. *Journal of Sex & Marital Therapy*, *46*(4), 307–313. doi:10.1080/0092623X.2019.1698481

Carmichael, P., Butler, G., Masic, U., Cole, T. J., De Stavola, B. L., Davidson, S., Skageberg, E. M., Khadr, S., & Viner, R. M. (2021). Short-term outcomes of pubertal suppression in a selected cohort of 12 to 15 year old young people with persistent gender dysphoria in the UK. *PLOS ONE*, *16*(2), e0243894. doi:10.1371/journal.pone.0243894

Chew, D., Tollit, M. A., Poulakis, Z., Zwickl, S., Cheung, A. S., & Pang, K. C. (2020). Youths with a non-binary gender identity: A review of their sociodemographic and clinical profile. *The Lancet Child & Adolescent Health*, *4*(4), 322–330. doi:10.1016/S2352-4642(19)30403-1

Churcher Clarke, A., & Spiliadis, A. (2019). 'Taking the lid off the box': The value of extended clinical assessment for adolescents presenting with gender identity difficulties. *Clinical Child Psychology and Psychiatry*, *24*(2), 338–352. doi:10.1177/1359104518825288

Clark, B. A., & Virani, A. (2021). This Wasn't a "split-second decision": An empirical ethical analysis of transgender youth capacity, rights, and authority to consent to hormone therapy. *Journal of Bioethical Inquiry*, *18*(1), 151–164. doi:10.1007/s11673-020-10086-9

Clayton, A., Malone, W. J., Clarke, P., Mason, J., & D'Angelo, R. (2021). Commentary: The signal and the noise—questioning the benefits of puberty blockers for youth with gender dysphoria—a commentary on Rew et al. (2021). *Child and Adolescent Mental Health*, *27*, camh.12533. doi:10.1111/camh.12533

Cohen-Kettenis, P. T., Delemarre-van de Waal, H. A., & Gooren, L. J. G. (2008). The treatment of adolescent transsexuals: Changing insights. *The Journal of Sexual Medicine*, *5*(8), 1892–1897. doi:10.1111/j.1743-6109.2008.00870.x

Cohen-Kettenis, P. T., Schagen, S. E. E., Steensma, T. D., de Vries, A. L. C., & Delemarre-van de Waal, H. A. (2011). Puberty suppression in a gender-dysphoric adolescent: A 22-year follow-up. *Archives of Sexual Behavior*, *40*(4), 843–847. doi:10.1007/s10508-011-9758-9

Cohen-Kettenis, P. T., & van Goozen, S. H. M. (1998). Pubertal delay as an aid in diagnosis and treatment of a transsexual adolescent. *European Child & Adolescent Psychiatry*, *7*(4), 246–248. doi:10.1007/s007870050073

D'Angelo, R. (2018). Psychiatry's ethical involvement in gender-affirming care. *Australasian Psychiatry*, *26*(5), 460–463. doi:10.1177/1039856218775216

D'Angelo, R., Syrulnik, E., Ayad, S., Marchiano, L., Kenny, D. T., & Clarke, P. (2020). One size does not fit all: In support of psychotherapy for gender dysphoria. *Archives of Sexual Behavior*, *50*, 7–16. doi:10.1007/s10508-020-01844-2

de Graaf, N. M., Giovanardi, G., Zitz, C., & Carmichael, P. (2018). Sex ratio in children and adolescents referred to the gender identity development service in the UK (2009–2016)). *Archives of Sexual Behavior*, *47*(5), 1301–1304. doi:10.1007/s10508-018-1204-9

de Graaf, N. M., Steensma, T. D., Carmichael, P., VanderLaan, D. P., Aitken, M., Cohen Kettenis, P. T., de Vries, A., Kreukels, B., Wasserman, L., Wood, H., & Zucker, K. J. (2020). Suicidality in clinic-referred transgender adolescents. *European child & adolescent psychiatry*, *31*, 67–83. doi: 10.1007/s00787-020-01663-9. Advance online publication. doi:10.1007/s00787-020-01663-9

de Vries, A. L. C. (2020). Challenges in timing puberty suppression for gender-nonconforming adolescents. *Pediatrics*, *146*(4), e2020010611. doi:10.1542/peds.2020-010611

de Vries, A. L. C., Beek, T. F., Dhondt, K., de Vet, H. C. W., Cohen-Kettenis, P. T., Steensma, T. D., & Kreukels, B. P. C. (2021). Reliability and clinical utility of gender identity-related diagnoses: comparisons between the ICD-11, ICD-10, DSM-IV, and DSM-5. *LGBT Health*, *8*(2), 133–142. doi:10.1089/lgbt.2020.0272

de Vries, A. L. C., & Cohen-Kettenis, P. T. (2012). Clinical management of gender dysphoria in children and adolescents: The Dutch approach. *Journal of Homosexuality*, *59*(3), 301–320. doi:10.1080/00918369.2012.653300

de Vries, A. L. C., McGuire, J. K., Steensma, T. D., Wagenaar, E. C. F., Doreleijers, T. A. H., & Cohen-Kettenis, P. T. (2014). Young adult psychological outcome after puberty suppression and gender reassignment. *Pediatrics*, *134*(4), 696–704. doi:10.1542/peds.2013-2958

de Vries, A. L. C., Steensma, T. D., Doreleijers, T. A. H., & Cohen-Kettenis, P. T. (2011). Puberty suppression in adolescents with gender identity disorder: A prospective follow-up study. *The Journal of Sexual Medicine*, *8*(8), 2276–2283. doi:10.1111/j.1743-6109.2010.01943.x

Delemarre-van de Waal, H. A., & Cohen-Kettenis, P. T. (2006). Clinical management of gender identity disorder in adolescents: A protocol on psychological and paediatric endocrinology aspects. *European Journal of Endocrinology*, *155*(suppl_1), S131–S137. doi:10.1530/eje.1.02231

Dhejne, C., Lichtenstein, P., Boman, M., Johansson, A. L. V., Långström, N., & Landén, M. (2011). Long-term follow-up of transsexual persons undergoing sex reassignment surgery: Cohort study in Sweden. *PLoS ONE*, *6*(2), e16885. doi:10.1371/journal.pone.0016885

Dhejne, C., Öberg, K., Arver, S., & Landén, M. (2014). An analysis of all applications for sex reassignment surgery in Sweden, 1960–2010: Prevalence, incidence, and regrets. *Archives of Sexual Behavior*, *43*(8), 1535–1545. doi:10.1007/s10508-014-0300-8

Diamond, M. (2013). Transsexuality among twins: Identity concordance, transition, rearing, and orientation. *International Journal of Transgenderism*, *14*(1), 24–38. doi:10.1080/15532739.2013.750222

Ehrensaft, D. (2017). Gender nonconforming youth: Current perspectives. *Adolescent Health, Medicine and Therapeutics*, Volume *8*, 57–67. doi:10.2147/AHMT.S110859

Entwistle, K. (2020). Debate: Reality check – Detransitioner's testimonies require us to rethink gender dysphoria. *Child and Adolescent Mental Health*, *26*, 15–16. camh.12380. doi:10.1111/camh.12380

Frigerio, A., Ballerini, L., & Valdés Hernández, M. (2021). Structural, functional, and metabolic brain differences as a function of gender identity or sexual orientation: A systematic review of the human neuroimaging literature. *Archives of sexual behavior*, *50*(8), 3329–3352. doi:10.1007/s10508-021-02005-9

Genspect (2021). Retrieved December 20, 2021, from https://genspect.org/groups/

Getahun, D., Nash, R., Flanders, W. D., Baird, T. C., Becerra-Culqui, T. A., Cromwell, L., Hunkeler, E., Lash, T. L., Millman, A., Quinn, V. P., Robinson, B., Roblin, D., Silverberg, M. J., Safer, J., Slovis, J., Tangpricha, V., & Goodman, M. (2018). Cross-sex hormones and acute cardiovascular events in transgender persons: A cohort study. *Annals of Internal Medicine*, *169*(4), 205. doi:10.7326/M17-2785

Gooren, L., & Delemarre-van de Waal, H. (1996). The feasibility of endocrine interventions in juvenile transsexuals. *Journal of Psychology & Human Sexuality*, *8*(4), 69–74. doi:10.1300/J056v08n04_05

Green, A. E., DeChants, J. P., Price, M. N., & Davis, C. K. (2021). Association of gender-affirming hormone therapy with depression, thoughts of suicide, and attempted suicide among transgender and nonbinary youth. *Journal of Adolescent Health*, S1054139X21005681. doi:10.1016/j.jadohealth.2021.10.036

Grootens-Wiegers, P., Hein, I. M., van den Broek, J. M., & de Vries, M. C. (2017). Medical decision-making in children and adolescents: Developmental and neuroscientific aspects. *BMC Pediatrics*, *17*(1), 120. doi:10.1186/s12887-017-0869-x

Hakeem, A., Črnčec, R., Asghari-Fard, M., Harte, F., & Eapen, V. (2016). Development and validation of a measure for assessing gender dysphoria in adults: The Gender Preoccupation and Stability Questionnaire. *International Journal of Transgenderism*, *17*(3–4), 131–140. doi:10.1080/15532739.2016.1217812

Hall, R., Mitchell, L., & Sachdeva, J. (2021). Access to care and frequency of detransition among a cohort discharged by a UK national adult gender identity clinic: Retrospective case-note review. *BJPsych Open*, *7*(6), e184. doi:10.1192/bjo.2021.1022

Hembree, W. C., Cohen-Kettenis, P. T., Gooren, L., Hannema, S. E., Meyer, W. J., Murad, M. H., Rosenthal, S. M., Safer, J. D., Tangpricha, V., & T'Sjoen, G. G. (2017). Endocrine treatment of gender-dysphoric/gender-incongruent persons: An endocrine society clinical practice guideline. J Clin Endocrinol Metab, *102*(11), 3869–3903. doi:10.1210/jc.2017-01658

HHS. (2021). What does "suicide contagion" mean, and what can be done to prevent it? Retrieved December 28, 2021, from https://www.hhs.gov/answers/mental-health-and-substance-abuse/what-does-suicide-contagion-mean/index.html

HRC. (n.d.). Clinical care for gender-expansive children & adolescents. Retrieved January 4, 2022, from https://www.hrc.org/resources/interactive-map-clinical-care-programs-for-gender-nonconforming-childr

Hutchinson, A., Midgen, M., & Spiliadis, A. (2020). In support of research into rapid-onset gender dysphoria. *Archives of Sexual Behavior*, *49*(1), 79–80. doi:10.1007/s10508-019-01517-9

Iliadis, S. I., Axfors, C., Friberg, A., Arinell, H., Beckman, U., Fazekas, A., Frisen, L., Sandström, L., Thelin, N., Wahlberg, J., Södersten, M., & Papadopoulos, F. C. (2020). Psychometric properties and concurrent validity

of the Transgender Congruence Scale (TCS) in the Swedish setting. *Scientific Reports*, *10*(1), 18701. doi:10.1038/s41598-020-73663-3

James, S. E., Herman, J. L., Rankin, S., Keisling, M., Mottet, L., & Anafi, M. (2016). *The report of the 2015 U.S. Transgender Survey*. Washington, DC: National Center for Transgender Equality.

Johns, M. M., Lowry, R., Andrzejewski, J., Barrios, L. C., Demissie, Z., McManus, T., Rasberry, C. N., Robin, L., & Underwood, J. M. (2019). Transgender identity and experiences of violence victimization, substance use, suicide risk, and sexual risk behaviors among high school students – 19 states and large urban school districts, 2017. *MMWR. Morbidity and Mortality Weekly Report*, *68*(3), 67–71. doi:10.15585/mmwr.mm6803a3

Kaltiala, R., Heino, E., Työläjärvi, M., & Suomalainen, L. (2020). Adolescent development and psychosocial functioning after starting cross-sex hormones for gender dysphoria. *Nordic Journal of Psychiatry*, *74*(3), 213–219. doi:10.1080/08039488.2019.1691260

Kaltiala-Heino, R., Bergman, H., Työläjärvi, M., & Frisen, L. (2018). Gender dysphoria in adolescence: Current perspectives. *Adolescent Health, Medicine and Therapeutics*, Volume *9*, 31–41. doi:10.2147/AHMT.S135432

Kaltiala-Heino, R., Sumia, M., Työläjärvi, M., & Lindberg, N. (2015). Two years of gender identity service for minors: Overrepresentation of natal girls with severe problems in adolescent development. *Child and Adolescent Psychiatry and Mental Health*, *9*(1), 9. doi:10.1186/s13034-015-0042-y

Katz, A. L., Macauley, R. C., Mercurio, M. R., Moon, M. R., Okun, A. L., Opel, D. J., & Statter, M. B. (2016). Informed consent in decision-making in pediatric practice. Committee on Bioethics. *Pediatrics*, *138*(2), e20161484. doi:10.1542/peds.2016-1484

Kendler K. S. (2019). From many to one to many-the search for causes of psychiatric illness. *JAMA psychiatry*, *76*(10), 1085–1091. doi:10.1001/jamapsychiatry.2019.1200

Kidd, K. M., Sequeira, G. M., Douglas, C., Paglisotti, T., Inwards-Breland, D. J., Miller, E., & Coulter, R. W. S. (2021). Prevalence of gender-diverse youth in an urban school district. *Pediatrics*, *147*(6), e2020049823. doi:10.1542/peds.2020-049823

Korte, A., Goecker, D., Krude, H., Lehmkuhl, U., Grüters-Kieslich, A., & Beier, K. M. (2008). Gender identity disorders in childhood and adolescence: Currently debated concepts and treatment strategies. *Deutsches Ärzteblatt International*, *105*(48), 834–841. doi:10.3238/arztebl.2008.0834

Kozlowska, K., Chudleigh, C., McClure, G., Maguire, A. M., & Ambler, G. R. (2021). Attachment patterns in children and adolescents with gender dysphoria. *Frontiers in Psychology*, *11*. doi:10.3389/fpsyg.2020.582688

Kozlowska, K., McClure, G., Chudleigh, C., Maguire, A. M., Gessler, D., Scher, S., & Ambler, G. R. (2021). Australian children and adolescents with gender dysphoria: Clinical presentations and challenges experienced by a multidisciplinary team and gender service. *Human Systems*, 26344041211010776. doi:10.1177/26344041211010777

Laidlaw, M. K., Van Meter, Q. L., Hruz, P. W., Van Mol, A., & Malone, W. J. (2019). Letter to the Editor: "Endocrine treatment of gender-dysphoric/gender-incongruent persons: An Endocrine Society Clinical Practice Guideline"." *The Journal of Clinical Endocrinology & Metabolism*, *104*(3), 686–687. doi:10.1210/jc.2018-01925

Levine, S. B. (2021). Reflections on the clinician's role with individuals who self-identify as transgender. *Archives of Sexual Behavior*, *50*, 3527–3536. doi:10.1007/s10508-021-02142-1

Levine, S.B. (2019). Informed Consent for Transgendered Patients, *Journal of Sex and Marital Therapy*, *45*(3):218–229. doi:10.1080/0092623X.2018.1518885

Littman, L. (2018). Parent reports of adolescents and young adults perceived to show signs of a rapid onset of gender dysphoria. *PLoS ONE 13*(8): e0202330. doi:10.1371/journal.pone.0202330

Littman, L. (2020). The use of methodologies in Littman (2018) is consistent with the use of methodologies in other studies contributing to the field of gender Dysphoria research: Response to Restar (2019). *Archives of Sexual Behavior*, *49*(1), 67–77. doi:10.1007/s10508-020-01631-z

Littman, L. (2021). Individuals treated for gender dysphoria with medical and/or surgical transition who subsequently detransitioned: A survey of 100 detransitioners. *Archives of Sexual Behavior, 50*, 3353–3369. doi:10.1007/s10508-021-02163-w

Lynch, H.F., Joffe, S., Feldman, E. (2018). Informed consent and the role of the treating physician. NEJM *378*:25, 435–438.

Malone, W., D'Angelo, R., Beck, S., Mason, J., & Evans, M. (2021). Puberty blockers for gender dysphoria: The science is far from settled. *The Lancet Child & Adolescent Health*, *5*(9), e33–e34. doi:10.1016/S2352-4642(21)00235-2

Malone, W. J., Hruz, P. W., Mason, J. W., & Beck, S. (2021). Letter to the editor from william j. Malone et al: "Proper care of transgender and gender-diverse persons in the setting of proposed discrimination: a policy perspective"." J Clin Endocrinol Metab, *106*(8), e3287–e3288. doi:10.1210/clinem/dgab205

Manrique, O. J., Adabi, K., Martinez-Jorge, J., Ciudad, P., Nicoli, F., & Kiranantawat, K. (2018). Complications and patient-reported outcomes in male-to-female vaginoplasty–where we are today: A systematic review and meta-analysis. *Annals of Plastic Surgery*, *80*(6), 684–691. doi:10.1097/SAP.0000000000001393

Mars, B., Heron, J., Klonsky, E. D., Moran, P., O'Connor, R. C., Tilling, K., Wilkinson, P., & Gunnell, D. (2019). Predictors of future suicide attempt among adolescents with suicidal thoughts or non-suicidal self-harm: A population-based birth cohort study. *The Lancet Psychiatry*, *6*(4), 327–337. doi:10.1016/S2215-0366(19)30030-6

Mathes, T., & Pieper, D. (2017). Clarifying the distinction between case series and cohort studies in systematic reviews of comparative studies: Potential impact on body of evidence and workload. *BMC Medical Research Methodology*, *17*(1), 107. doi:10.1186/s12874-017-0391-8

McGuire, J. K., Berg, D., Catalpa, J. M., Morrow, Q. J., Fish, J. N., Nic Rider, G., Steensma, T., Cohen-Kettenis, P. T., & Spencer, K. (2020). Utrecht Gender Dysphoria Scale - gender spectrum (UGDS-GS): Construct validity among transgender, nonbinary, and LGBQ samples. *International Journal of Transgender Health*, *21*(2), 194–208. doi:10.1080/26895269.2020.1723460

Morandini, J. S., Kelly, A., de Graaf, N. M., Carmichael, P., & Dar-Nimrod, I. (2021). Shifts in demographics and mental health co-morbidities among gender dysphoric youth referred to a specialist gender dysphoria service. *Clinical Child Psychology and Psychiatry*, 135910452110468. doi:10.1177/13591045211046813

National Institute for Health and Care Excellence. (2020a). Evidence review: Gonadotrophin releasing hormone analogues for children and adolescents with gender dysphoria. https://arms.nice.org.uk/resources/hub/1070905/attachment

National Institute for Health and Care Excellence. (2020b). Evidence review: Gender-affirming hormones for children and adolescents with gender dysphoria. https://arms.nice.org.uk/resources/hub/1070871/attachment

Nota, N. M., Wiepjes, C. M., de Blok, C. J. M., Gooren, L. J. G., Kreukels, B. P. C., & den Heijer, M. (2019). Occurrence of acute cardiovascular events in transgender individuals receiving hormone therapy: Results from a Large Cohort Study. *Circulation*, *139*(11), 1461–1462. doi:10.1161/CIRCULATIONAHA.118.038584

Olson, K. R., Durwood, L., DeMeules, M., & McLaughlin, K. A. (2016). Mental health of transgender children who are supported in their identities. *Pediatrics*, *137*(3), 1–16. iii.

Ouliaris, C. (2021). Consent for treatment of gender dysphoria in minors: evolving clinical and legal frameworks. *The Medical journal of Australia*, Advance online publication. doi:10.5694/mja2.51357

Paine, E. A. (2021). "Fat broken arm syndrome": Negotiating risk, stigma, and weight bias in LGBTQ healthcare. Soc Sci Med, *270*, 113609. doi:10.1016/j.socscimed.2020.113609

Planned Parenthood League of Massachusetts. (n.d.) Gender affirming hormone therapy. Retrieved December 26, 2021, from https://www.plannedparenthood.org/planned-parenthood-massachusetts/campaigns/gender-affirming-hormone-therapy

Rafferty, J., Committee on Psychosocial Aspects of Child and Family Health, Committee on Adolescence, & Section on Lesbian, Gay, Bisexual, and Transgender Health and Wellness. (2018). Ensuring comprehensive care and support for transgender and gender-diverse children and adolescents. *Pediatrics*, *142*(4), e20182162. doi:10.1542/peds.2018-2162

Right to medical or dental treatment without parental consent, Oregon ORS Volume 3, Title 11, 109.640 (2010). https://oregon.public.law/statutes/ors_109.640

Ristori, J., & Steensma, T. D. (2016). Gender dysphoria in childhood. *International Review of Psychiatry*, *28*(1), 13–20. doi:10.3109/09540261.2015.1115754

Rood, B. A., Reisner, S. L., Surace, F. I., Puckett, J. A., Maroney, M. R., & Pantalone, D. W. (2016). Expecting rejection: Understanding the minority stress experiences of transgender and gender-nonconforming individuals. *Transgender Health*, *1*(1), 151–164. doi:10.1089/trgh.2016.0012

Ross, M. W., & Need, J. A. (1989). Effects of adequacy of gender reassignment surgery on psychological adjustment: A follow-up of fourteen male-to-female patients. *Archives of Sexual Behavior*, *18*(2), 145–153. doi:10.1007/BF01543120

Schulz, S. L. (2018). The informed consent model of transgender care: An alternative to the diagnosis of gender dysphoria. *Journal of Humanistic Psychology*, *58*(1), 72–92. doi:10.1177/0022167817745217

Simon, G. E., & VonKorff, M. (1998). Suicide mortality among patients treated for depression in an insured population. *American Journal of Epidemiology*, *147*, 155–160. doi:10.1093/oxfordjournals.aje.a009428

Singh, D., Bradley, S. J., & Zucker, K. J. (2021). A follow-up study of boys with gender identity disorder. *Frontiers in Psychiatry*, *12*. doi:10.3389/fpsyt.2021.632784

Smith, Y. L. S., Van Goozen, S. H. M., & Cohen-Kettenis, P. T. (2001). Adolescents with gender identity disorder who were accepted or rejected for sex reassignment surgery: A prospective Follow-up Study. J Am Acad Child Adolesc Psychiatry, *40*(4), 472–481. doi:10.1097/00004583-200104000-00017

Smith, A. R., Zuromski, K. L., & Dodd, D. R. (2018). Eating disorders and suicidality: What we know, what we don't know, and suggestions for future research. *Current Opinion in Psychology*, *22*, 63–67. doi:10.1016/j.copsyc.2017.08.023

Spiliadis, A. (2019). Towards a gender exploratory model: Slowing things down, opening things up and exploring identity development. Metalogos Systemic Therapy Journal, *35*, 1–9. https://www.ohchr.org/Documents/Issues/SexualOrientation/IESOGI/Other/Rebekah_Murphy_TowardsaGenderExploratoryModelslowingthingsdownopeningthingsupandexploringidentitydevelopment.pdf

Steensma, T. D., McGuire, J. K., Kreukels, B. P. C., Beekman, A. J., & Cohen-Kettenis, P. T. (2013). Factors associated with desistence and persistence of childhood gender dysphoria: A quantitative follow-up study. *Journal of the American Academy of Child & Adolescent Psychiatry*, *52*(6), 582–590. doi:10.1016/j.jaac.2013.03.016

Strang, J. F., Janssen, l, Tishelman, A., Leibowitz, S. F., Kenworthy, L., McGuire, J. K., Edwards-Leeper, L., Mazefsky, C. A., Rofey, D., Bascom, J., Caplan, R., Gomez-Lobo, V., Berg, D., Zaks, Z., Wallace, G. L., Wimms, H., Pine-Twaddell, E., Shumer, D., Register-Brown, K., … Anthony, L. G. (2018). Revisiting the link: Evidence of

the rates of autism in studies of gender diverse individuals. *Journal of the American Academy of Child and Adolescent Psychiatry*, *57*(11), 885–887. doi:10.1016/j.jaac.2018.04.023

Tavistock and Portman NHS Foundation Trust. (2020). Gender identity development service referrals in 2019–20 same as 2018–19. https://tavistockandportman.nhs.uk/about-us/news/stories/gender-identity-development-service-referrals-2019-20-same-2018-19/

Temple Newhook, J., Pyne, J., Winters, K., Feder, S., Holmes, C., Tosh, J., Sinnott, M.-L., Jamieson, A., & Pickett, S. (2018). A critical commentary on follow-up studies and "desistance" theories about transgender and gender-nonconforming children. *International Journal of Transgenderism*, *19*(2), 212–224. doi:10.1080/15532739.2018.1456390

The Trevor Project. (2021). National Survey on LGBTQ Youth Mental Health 2021. Retrieved January 3, 2022, from https://www.thetrevorproject.org/survey-2021/?section=SuicideMentalHealth

Toomey, R. B., Syvertsen, A. K., & Shramko, M. (2018). Transgender Adolescent Suicide Behavior. *Pediatrics*, *142*(4). doi:10.1542/peds.2017-4218

Turban, J. L. (2018). Potentially reversible social deficits among transgender youth. *Journal of Autism and Developmental Disorders*, *48*(12), 4007–4009. doi:10.1007/s10803-018-3603-0

Turban, J. L., King, D., Carswell, J. M., & Keuroghlian, A. S. (2020). Pubertal suppression for transgender youth and risk of suicidal ideation. *Pediatrics*, *145*(2), e20191725. doi:10.1542/peds.2019-1725

Turban, J. L., & van Schalkwyk, G. I. (2018). "Gender Dysphoria" and autism spectrum disorder: is the link real? *Journal of the American Academy of Child & Adolescent Psychiatry*, *57*(1), 8–9.e2. doi:10.1016/j.jaac.2017.08.017

van der Miesen, A. I. R., Cohen-Kettenis, P. T., & de Vries, A. L. C. (2018). Is there a link between gender dysphoria and autism spectrum disorder? *Journal of the American Academy of Child & Adolescent Psychiatry*, *57*(11), 884–885. doi:10.1016/j.jaac.2018.04.022

Vandenbussche, E. (2021). Detransition-related needs and support: A cross-sectional online survey. *Journal of Homosexuality*, *20,* 1–19. doi:10.1080/00918369.2021.1919479

Voorzij. (2021). More research is urgently needed into transgender care for young people: "Where does the large increase of children come from?" Retrieved December 20, 2021, from https://www.voorzij.nl/more-research-is-urgently-needed-into-transgender-care-for-young-people-where-does-the-large-increase-of-children-come-from/.

Vrouenraets, L., de Vries, A., de Vries, M. C., van der Miesen, A., & Hein, I. M. (2021). Assessing medical decision-making competence in transgender youth. *Pediatrics*, *148*, e2020049643. Advance online publication. doi:10.1542/peds.2020-049643

Vrouenraets, L., Hartman, L. A., Hein, I. M., de Vries, A., de Vries, M. C., & Molewijk, B. (2020). Dealing with moral challenges in treatment of transgender children and adolescents: Evaluating the role of moral case deliberation. *Archives of sexual behavior*, *49*(7), 2619–2634. doi:10.1007/s10508-020-01762-3

Wiepjes, C. M., den Heijer, M., Bremmer, M. A., Nota, N. M., Blok, C. J. M., Coumou, B. J. G., & Steensma, T. D. (2020). Trends in suicide death risk in transgender people: Results from the Amsterdam Cohort of Gender Dysphoria Study (1972–2017)). *Acta Psychiatrica Scandinavica*, *141*(6), 486–491. doi:10.1111/acps.13164

Wiepjes, C. M., Nota, N. M., de Blok, C. J. M., Klaver, M., de Vries, A. L. C., Wensing-Kruger, S. A., de Jongh, R. T., Bouman, M.-B., Steensma, T. D., Cohen-Kettenis, P., Gooren, L. J. G., Kreukels, B. P. C., & den Heijer, M. (2018). The Amsterdam Cohort of Gender Dysphoria Study (1972–2015): Trends in Prevalence, Treatment, and Regrets. *The Journal of Sexual Medicine*, *15*(4), 582–590. doi:10.1016/j.jsxm.2018.01.016

Wilson, S. C., Morrison, S. D., Anzai, L., Massie, J. P., Poudrier, G., Motosko, C. C., & Hazen, A. (2018). Masculinizing top surgery: A systematic review of techniques and outcomes. *Annals of Plastic Surgery*, *80*(6), 679–683. doi:10.1097/SAP.0000000000001354

World Health Organization. (2019). International statistical classification of diseases and related health problems (11th ed.). https://icd.who.int/

Zucker, K. J. (2017). Epidemiology of gender dysphoria and transgender identity. *Sexual Health*, *14*(5), 404. doi:10.1071/SH17067

Zucker, K. J. (2018). The myth of persistence: Response to "A critical commentary on follow-up studies and 'desistance' theories about transgender and gender non-conforming children" by Temple Newhook et al. (2018). *International Journal of Transgenderism*, *19*(2), 231–245. doi:10.1080/15532739.2018.1468293

Zucker, K. J. (2019). Adolescents with gender dysphoria: Reflections on some contemporary clinical and research issues. *Archives of Sexual Behavior*, *48*(7), 1983–1992. doi:10.1007/s10508-019-01518-8

Zucker, K. J. (2020). Debate: Different strokes for different folks. *Child and Adolescent Mental Health*, *25*(1), 36–37. doi:10.1111/camh.12330

Chapter 2

Armstrong N, McManus AM (eds): The Elite Young Athlete.
Med Sport Sci. Basel, Karger, 2011, vol 56, pp 23–46

Physiology of Elite Young Female Athletes

Alison M. McManus[a] · Neil Armstrong[b]

[a]Institute of Human Performance, University of Hong Kong, Hong Kong, SAR, China; [b]Children's Health and Exercise Research Centre, University of Exeter, Exeter, UK

Abstract

The participation of girls in elite sport has increased exponentially over the past 30 years. Despite these increases a tradition for recruiting boys for exercise studies persists and our knowledge of the physiologic response to exercise in girls remains limited. Girls' physiology varies with age and maturation and is underpinned by a divergent hormonal milieu which begins early in foetal life. Sexual dimorphism underlies much of the physiologic response to exercise, and becomes most acute during adolescence when boys become taller, heavier, less fat and are more muscular than girls. Young girl athletes are not simply smaller, less muscular boys. The widening sex disparity in responses to exercise during puberty cannot always be accounted for by size. The woeful number of studies on girls and our prior inability to non-invasively study the complexity of the cellular metabolic response to exercise means an integrative understanding of girls' physiological responses to exercise remains elusive. Success in elite sport requires intense training, which for a long time was thought to cause disruption to normal growth and maturation. It would appear that exercise training, without other predisposing factors, is unlikely to cause aberrations to either growth or maturation. Nevertheless, there is clear evidence of a boundary between healthy and unhealthy levels of exertion when coupled with caloric limitation. Sports in which intense training is combined with the need for leanness may predispose girls to increased risk of skeletal and reproductive health problems, and ensuring risk is minimised should be a priority.

The young female athlete is unique. She stands out from her peers, who show declining levels of physical activity from early puberty [1]. In comparison to boys, she remains under-represented in competitive sport [2]. This persists to Olympic level competition with 1,704 fewer women competing at the 2008 Beijing Olympics than men [3]. Moreover, there are fewer scientific data on physiologic issues associated with exercise in girls than in boys. Traditionally, boys have been recruited for exercise studies and a search of key databases shows that this persists with comparatively fewer articles investigating physiologic responses to exercise in girls. This preference for recruiting boys reflects social constraints which can be traced back to Victorian values surrounding women, exercise and health [4]. Central to the issue of women's involvement in sport was biology, bolstered by the idea that a woman's structural and functional ability was unable to tolerate strenuous exercise, presenting considerable risk to her reproductive health. Indeed, whether or not chronic training causes less than optimal structural and functional alterations in girls remains the topic of lively debate [5–8].

Sexual dimorphism does indeed underlie much of the physiological response to acute and chronic exercise, and although a myriad of factors have been shown to influence the development of sport performance [9], structural and functional capacity represent a significant contribution to the gender differences notable. Figure 1

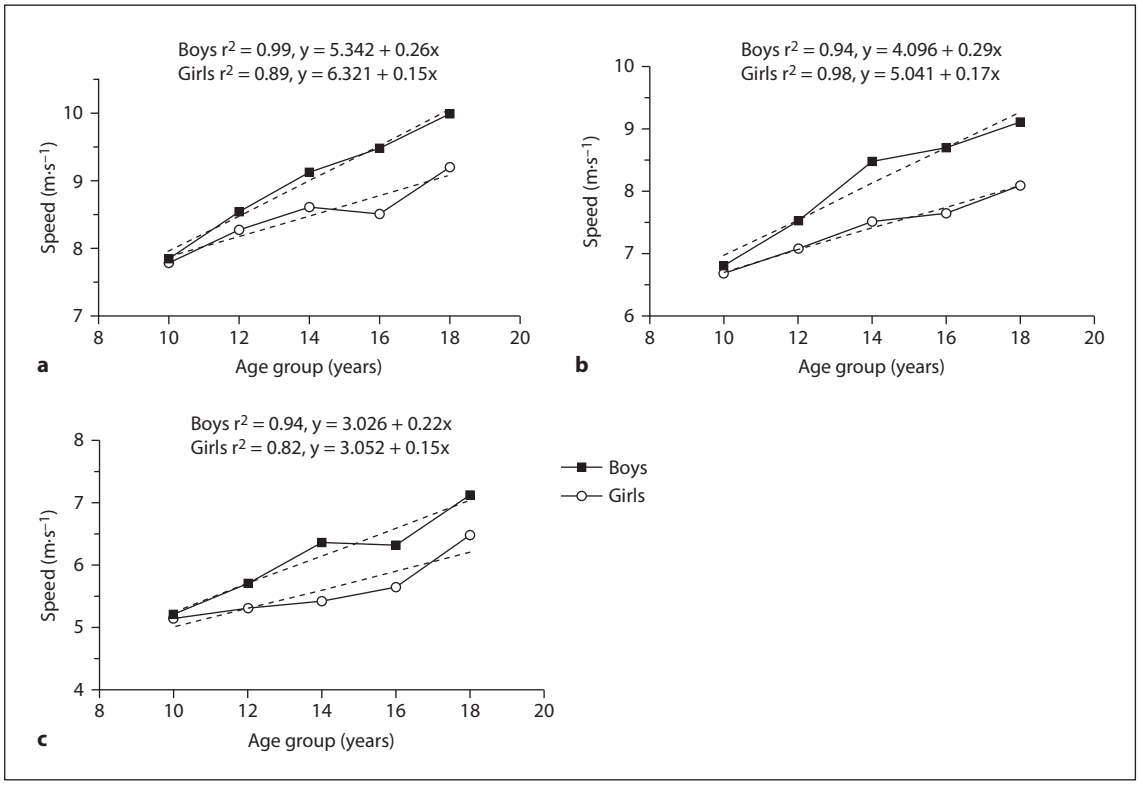


Fig. 1. Average running speed from North American age-group track and field records and world junior records over 100m (**a**), 400m (**b**) and 1,500m (**c**) for girls and boys.

illustrates the average running speed over 100, 400 and 1,500 m for girls and boys competing in age-group track and field championships in North America and world junior records [10, 11]. Analyses of these data show that regardless of distance, average running speed increases with age in both girls and boys (r^2 >0.82 for all distances). Prior to 11 years of age differences in average speed are minimal, thereafter disparity becomes more pronounced with boys 8–15% faster than girls from 13 years of age onwards. This is consistent with the 9–15% advantage in speed adult men have over women across a wide range of distances from 100 m to 200 km [12]. When the slopes of the regression lines were compared between the sexes by age, these were significantly different for 100 and 400 m with the rate of improvement greater in the boys than the girls (see fig. 1). The rate of improvement in running speed over 1,500 m was similar between boys and girls with a pooled slope of 0.19. That an interaction between distance, speed developmental trajectories and sex exists suggests intriguing physiological mechanisms.

A key question to be addressed is whether the superior performances of boys relate to qualitative discrepancies in functional capacity or alternatively, they are simply a function of the increased body size and disparate body composition that accompanies adolescent growth and maturation. Appropriate adjustment for differences in body size and composition dampen many apparent

physiological differences [13, 14]; however, young girl athletes are not simply smaller, less muscular boys. Girls' physiology varies with age and is underpinned by a divergent hormonal milieu which begins early in foetal life. There is evidence that in the two-cell stage of embryonic development, long before visual gonadal differentiation, the sex-determining region of the Y chromosome has already been transcribed [15]. Testosterone secretion commences at about 3 months in the male foetus, the absence of which in the female foetus allows maturation of the female reproductive organs, and by birth subtle differences in cardiac function and body composition already exist [16, 17]. The accentuated hormonal adjustments which occur during adolescence result in differential development and a widening sex disparity in many physiological responses to exercise, some of which cannot be accounted for by size.

The primary focus of this chapter is on physiologic issues that are associated with exercise in girls; but to illustrate sexual dimorphism, comparisons will be made with boys where relevant. The chapter begins with a brief overview of growth and maturation in girls, including a discussion of issues related to body composition and size. This is followed by a focus on the acute responses to sustained aerobic exercise, as well as short-duration high-intensity exercise in girls. The possible physiological mechanisms that underlie these responses, such as sex differences in pulmonary, cardiac, and peripheral function, as well as cellular metabolism are discussed. The chapter concludes with consideration of the contention that intensive training poses a substantial threat to the development and health of young girls.

Growth and Maturation

Growth hormone (GH), insulin-like growth factor I (IGF-I), the sex steroids and insulin are all potent anabolic hormones. Their complex interactions enable linear growth, bone mineralization, increases in muscle and metabolic adaptations during childhood and adolescence and much of this hormonal mélange is sex dependent [18]. For example, the physiological effects of the sex steroids testosterone and oestrogen differ markedly, with evidence that combined testosterone and GH administration causes increases in IGF-I concentrations, resulting in enhanced anabolism, greater increases in fat free mass and higher whole body protein synthesis in boys [19]. Oestrogen administration, on the other hand, has been shown to have no effect on whole body protein synthesis in girls [19]. Sex dimorphic growth and development is most pronounced during adolescence, which forms the primary focus of the following sections.

Stature and body mass follow a double-sigmoid growth pattern in girls and boys, with rapid gains in infancy, slower yearly gains of about 5–6 cm in stature and 2.25–2.75 kg in body mass through childhood, and a second rapid gain in adolescence [20]. Girls usually begin adolescent growth before boys and progress at a faster rate than boys [21, 22]. At the peak of the adolescent growth spurt, girls gain approximately 8–9 cm in a year in stature. Boys only gain about 3 cm more in stature during the adolescent growth spurt, but are about 11–13 cm taller by adulthood because of their extra pre-adolescent growth [23]. From onset to completion, adolescent growth in stature lasts about 4–4.5 years in girls, until rising levels of oestrogen induce epiphyseal fusion marking an end to the growth in stature, usually around a skeletal age of 15 years [18, 21]. Peak body mass velocity is in lag by some 4–6 months with peak height velocity and total body mass gains of 16 kg are usual during the adolescent growth spurt in girls [24].

Other body proportions such as sitting height, leg length, biacromial and bicristal breadths follow a similar growth pattern to stature. Leg length and sitting height differ little between girls and boys during childhood [23]. At the onset of adolescence rapid growth in leg length precedes trunk growth. Boys surpass girls in leg length by about 12 years of age and in sitting height around

14 years of age. Yet, the ratio of sitting height to stature is higher in girls than boys through adolescence indicating relatively shorter legs in girls for the same stature. Girls have a marginally wider bicristal breadth than boys from late childhood to late adolescence, when boys catch-up [23]. In contrast, boys experience much more dramatic increases in biacromial breadth compared to girls [23]. When bicristal breadth is expressed as a ratio of biacromial breadth (hip-to-shoulder ratio), in comparison to boys values are higher in girls from early childhood, with bicristal breadth approximately 72–73% of biacromial breath, remaining quite stable through adolescence. In boys a decline in this ratio is noted from about 70% at 11 years of age, to 65% by 16 years of age, which is an outcome of the disproportionately faster growth of biacromial breadth [23]. Interestingly, although stature as well as mass-to-stature ratio differ between girls involved in different competitive sports, minimal differences in other proportions such as arm span and seated height have been reported [25, 26]. Greater mass-to-stature ratios can confer performance benefit in some sports such as throwing events; however, the combined effect of broader hips and shorter legs that usually accompany a greater mass-for-stature and characterize early maturation in girls, is generally disadvantageous. Although data are sparse, young female athletes in running events or gymnastics are generally more likely to be characterized by longer legs, lower hip-to-shoulder ratios and lower mass-for-stature.

Assessment of maturity stage is vital but poses considerable challenge. Skeletal age is the biological marker of choice, but is hindered by ethical constraints related to ionizing radiation exposure. The timing and tempo of sexual maturation in girls has most commonly been described using the visual descriptive stages of secondary sexual characteristics. These were first documented by Reynolds and Wines [27], and then further refined by J.M. Tanner and are better known as Tanner stages [28]. There is a large normal variation in the timing and tempo of sexual maturation in girls, as well as clearly documented sex differences. Like linear growth, girls normally begin sexual maturation before boys and progress toward full maturity at a faster tempo than boys [22].

A recent large-scale longitudinal study of Caucasian and African-American children suggests that the average girl begins breast development at 9.8 years, whilst the average boy begins genital development at about 10.3 years [21]. Pubic hair growth usually occurs around 10.2 years in girls and around 11.3 years in boys. The onset of the initial stages of sexual maturity in these American girls is somewhat earlier than previously published data for European girls, for whom breast budding was reported to occur at about 10.5 years and Tanner Stage two for pubic hair at 10.8 years [28, 29]. This discrepancy probably reflects the different racial mix of the groups, in addition to a possible secular trend for a declining age of onset of maturation [30].

Asynchronous maturation of secondary sex characteristics in girls is common and has been defined as a difference of at least 4 months between breast and pubic hair development. About 51–66% of girls follow an asynchronous maturation pattern [21, 31]. Most (about 70%) follow a thelarchal pathway, with breast development beginning prior to pubic hair growth (adrenarche). Asynchrony usually persists into Tanner stage 3, the onset of which is on average 11.3 years for breast development, with pubic hair stage three occurring some 2 months later [31]. As the latter stages of sexual maturity are attained, development becomes more synchronous. A minority of girls (approximately 30%) follow an adrenarchal asynchronous pathway, in which pubic hair development precedes breast development. Thelarchal asynchrony is believed to result from initial advanced stimulation of gonadotropin and oestrogen, thereby enabling earlier breast development. The converse is true of adrenarchal asynchrony, where the advanced production of testosterone and adrenal hormones promotes earlier pubic hair growth. Age of onset of menses usually

occurs during Tanner stage three for breast development. For those girls following the thelarchal pathway, menses occurs at an earlier age, around 12.6 years. Those following the adrenarchal pathway usually begin menses around 13.1 years [21, 31]. Lower oestrogen levels are noted in girls following an adrenarchal maturation pathway which persists throughout adolescence. This affords a body composition advantage, characterized by a lower sum of skinfolds, percent body fat and waist-to-hip ratio [31]. Girls involved in intensive training are generally characterized by lower percent body fat, but there is little evidence to suggest that they preferentially follow an asynchronous adrenarchal pathway [32].

The presence of asynchronous sexual maturation has implications when comparing young athletes with non-athletes, as well as making comparisons between girls and boys. Whether alignment is on the basis of a single marker (e.g. pubic hair development), the creation of a composite score for both pubic hair and breast development, or on differing secondary sex characteristics (such as genitalia and breast development), the assumption is that the timing of the appearance of a particular characteristic, as well as the tempo, is homogeneous. This is clearly not always the case, and it has been suggested that alignment of sexual maturation to other biological or somatic markers of maturation (e.g. age at menarche or peak height velocity) is more appropriate [33]. Menarchal age is convenient if retrospective, otherwise, like peak height velocity, a prospective research design is necessary. Assessment of maturational stage continues to present a real methodological challenge to paediatric exercise physiologists.

Body Composition and Size

Fat

Small sex differences in fat mass and percent body fat are evident from mid-childhood, with levels rising substantially in girls during adolescence. Body fat gains by the end of puberty usually result in 26–31% body fat in the average adolescent girl [34, 35]. Young athletes are generally leaner than the average non-athletic girl, but this is dependent on the chosen sport. Values as low as 14.3% have been reported for 15-year-old rhythmic gymnasts, with gymnasts generally showing lower body fat than other athletes [34]. Body fat values from 21 to 25% have been reported for dancers, distance runners and cross-country skiers from the ages of 10–17 years [36–39]. Higher values have been reported for 13- to 17-year-old high-school athletes competing in lacrosse, soccer, softball, swimming, track and field and volleyball (mean 27.4 ± 0.7%) [40].

Sex steroids are major determinants of body fat distribution, with the increases in body fat generally subcutaneous and in the gluteal and femoral regions in girls. Fat mass, combined with a smaller leg length-to-stature ratio, lowers the centre of gravity in girls, thereby affording better balance. However, fat mass is also negatively related to heat dissipation, which may prove disadvantageous in girls during endurance events in hot environmental conditions [41].

Fat tissue has relatively uniform properties throughout life, with negligible water content and a tissue density of 0.9007 $kg \bullet l^{-1}$ [42]. Recent reference data from Wells et al. [42] have shown that in comparison lean tissue shows sex-specific chemical maturation, with decreases in water content and increases in density with increasing age. These new data have implications for the assessment of body fat since previous reference data extrapolated rather than directly assessed age specific tissue densities and hydration. Wells et al.'s [42] work provides the first comprehensive empirical data set for lean tissue properties for 4- to 23-year-old boys and girls, with lean tissue density values in girls of 1.0905 $kg \bullet l^{-1}$ at 8–9 years, rising to 1.1021 $kg \bullet l^{-1}$ at 16–17 years. Lean tissue hydration values declined with age from 75.2% at 8–9 years to 73.7% at 16–17 years

in girls. Importantly, this study has shown that these new values differ from previously simulated values. The lean tissue density values of Wells et al. [42] were consistently higher, whilst the hydration data were consistently lower than those reported by Lohman [43]. Comparisons of % fat calculated from densitometry with the Lohman [43] formula led to a between-study error of –1 to 2.5% fat in the average girl. Wells et al. [42] provide important new reference values for the assessment of body fat by both hydrometry (total body water, bioelectrical impedance) and densitometry which should ensure greater clarity in future analyses.

Muscle

At birth, boys tend to have a greater lean mass than girls. This difference remains small but detectable throughout childhood with about a 10% greater lean mass in boys than girls prior to puberty [17]. The sharp increase in muscle mass disparity between the sexes during puberty indicates a primary role of the gonadal steroidal hormones. Muscle mass in girls increases from about 25 kg at 10 years of age to about 45 kg by 18 years of age [42]. Reported values for 15- to 17-year-old female athletes are not dissimilar, ranging from 42 to 53 kg [23]. These gains in muscle tissue represent an increase of about 5% in muscle mass. The relative contribution of muscle mass to total body mass usually declines once consideration is given to the relative contribution of fat mass. In comparison, the androgen-mediated growth of muscle in boys results in muscle mass reaching about 55% of total body mass at maturity [44]. The greater overall skeletal muscle mass in adolescent boys creates a potential cascade of functional differences apparent in adults such as differing muscle fibre size, activities of metabolic enzymes, lipid content and oxidation, relative expression of myosin isoforms, and fatigability [45–50]. Maturation of these features remains poorly understood.

The use of ultrasonography and magnetic resonance imaging (MRI) are providing insight into changes in muscle architecture with growth. A recent study using ultrasonography demonstrated that muscle thickness (a marker of physiological cross-sectional area) and pennation angle were correlated with age from 4 to 10 years in both sexes [51]. Findings from MRI studies have shown similarly that muscle cross-sectional area increases with age from childhood through adolescence, and more so in boys than girls [52]. Pennation angle on the other hand has not been found to differ between the sexes [51]. Whilst muscle cross-sectional area and pennation angle are related to age, this has not been shown in muscle fibre length [51]. Muscle fibre length has been found to have high inter- and intra-individual variation, which may reflect a greater malleability in response to external stimuli such as the extent and intensity of exercise [51].

Morphological change in the muscle impacts upon function. Maximal strength, for example, is dependent on the specific joint angle (force-length relationship), contraction type, muscle cross-sectional area and velocity. The length of the muscle fibre is proportional to the absolute maximum contraction velocity, whilst the pennation angle dictates the proportion of force transmitted to the tendon. Muscle strength, expressed as torque, increases with age in children, but gains are greater in boys. This has been presumed to be an outcome of the greater muscle cross-sectional area [53]. Alternatively, there may be intrinsic sex differences in the fibre composition and fatigue characteristics of skeletal muscle that materialize during adolescence that also influence the ability to increase torque.

In adults, several studies have reported higher glycolytic enzyme activity and lower oxidative enzyme activity in men compared to women, supporting the contention that men have a lower proportion of type I fibres [50, 54]. Data on muscle fibre typing in children are limited because of the invasive nature of the biopsy methodology,

but there is evidence to show that differentiation of fibre type occurs during the first few years of life. About 10% of skeletal muscle fibres remain undifferentiated up until puberty, with no sex difference notable in the percentage of type I fibres (slow-twitch oxidative fibres) during childhood [55]. By adolescence females have a lower % of type I muscle fibres than males [45, 56] and the type II muscle fibres of young men are bigger than their type I fibres, something not evident in young women [56, 58].

Although boys gain more in strength than girls during adolescence, elite girl athletes are stronger than their less athletic peers. For example, average quadriceps and biceps isometric strength was reported to be 22% greater in elite gymnasts and swimmers and 18% greater in tennis players compared to less athletic school children [59]. Interestingly this study found no differences in strength between sports in the girls, even when co-varied for body mass. The relationship between strength and body mass, or strength-to-mass ratio, has been seen as an important predictor of sport performance particularly in gymnastics, middle- and long-distance running. Indeed, elite adult women runners such as Yvonne Murray and Greta Waitz were 17–18% below the average body mass for their stature at the peak of their running careers, which suggests relative strength was high. The work of Bencke et al. [60] has shown that 11-year-old girl gymnasts were the smallest, lightest and possessed the highest explosive strength compared to other athletes, suggesting high relative strength confers advantage in some sports in girls.

Bone

Bone characteristics differ little between boys and girls prior to puberty, but then follow two sex-divergent growth paths. During the adolescent growth spurt boys experience increases predominantly in bone diameter and cortical thickness due to periosteal apposition [17]. Girls on the other hand experience increases in cortical thickness, a decrease in medullary diameter, and little increase in periosteal diameter as a result of oestrogen inhibition of periosteal apposition [17, 61]. It should be noted that bone accretion and endocortical features appear to be site specific with data showing endocortical resorption at the mid-femur and proximal tibia in girls through puberty, but no endocortical resorption at the radial diaphysis [62, 63].

During puberty, bone mineral content (BMC) accrual rate is in lag with muscle accrual rate, suggesting that muscle enlargement, and concomitant increases in muscle force, are important for bone development [64]. Indeed, the 'functional muscle-bone unit' hypothesis suggests muscle force is a primary determinant of bone mass, structure and strength [65]. Young female runners and gymnasts have been shown to have elevated bone mass and enlarged bone size at specific sites such as the radius and lumbar spine in gymnasts [66] and the femur in runners [67], reflecting the specific mechanical-loading patterns these sports require. This has led some to conclude that muscular force alone explains the impact loading effect on bone [68–70]. On the contrary, recent research has shown that bone mass, size and strength increases in the upper extremity in gymnasts are independent of maturation, stature and muscle cross-sectional area and substantiates the hypothesis that other non-muscular loading factors may also account for skeletal adaptations [71, 72].

Puberty is the most favourable period for augmented bone mineralization, with about one quarter of adult bone being laid down. Bone mineral accrual is sex and maturity dependent and appears to be enhanced by oestrogen. It is clear that the early pubertal and pre-menarchal years are particularly important for young girls in terms of optimizing their bone mineralization and weight-bearing exercise plays a key complementary role in this process [73].

Body Size Considerations

By the time the adolescent growth spurt is complete the body size, shape and composition of boys and girls is different. Boys have become taller, have longer legs, broader shoulders, are heavier, and have less fat and more muscle than girls. The effect of these discrepancies on performance is substantial, and it is important in understanding girls' physiologic responses to exercise that we are able to effectively partition the impact of size from function. Traditionally in exercise physiology this has been achieved by expressing the physiological measure of interest (y) as a ratio of an appropriate marker of body size (x) to give the ratio y/x. Tanner suggested in 1949 that the use of such ratio standards to scale physiological measurements to size was 'theoretically fallacious and unless in exceptional circumstances, misleading' [74]. Yet this has largely been ignored with much of the comparison between men and women, or boys and girls based on ratio standards [75, 76]. An implicit assumption with the ratio standard is that the relationship is linear and the y intercept is zero. Additionally, ratio standards should only be used when the coefficient of variation (V) for body size (x), divided by the coefficient of variation (V) for the physiological variable (y), equals the Pearson product moment correlation coefficient (r) for the two variables, expressed by the equation $V_x/V_y = rx,y$. These assumptions are rarely met and the outcome is scaling distortion, which may have obscured our understanding of the physiologic responses of girls [77].

Theoretically, morphological and physiological variables are scaled according to the general allometric equation $y = ax^b$, where y is the morphological or physiological variable of interest, x is the chosen size denominator, *b* is the scaling exponent and *a* is the constant [78]. When this equation is solved the resultant power function ratio (y/x^b) is derived. Various studies have shown that with careful consideration of the denominator, alternative approaches, such as the allometric power function ratio or more complex multilevel modelling of longitudinal data, are more appropriate than ratio scaling when comparisons of various physiological outcomes between individuals of differing body size are sought [79, 80]. These alternatives should, wherever possible, be utilised.

Acute Responses to Aerobic Exercise

Peak oxygen uptake (peak $\dot{V}O_2$), the highest $\dot{V}O_2$ elicited during an exercise test to exhaustion in children, is well-established as the best single measure of aerobic fitness [81]. In comparison to boys, girls are characterised with a smaller absolute peak $\dot{V}O_2$. Predicted values range from 1.5 to 2.2 litres • min^{-1} in 10- to 16-year-old girls and are lower than boys by 11, 19, 23 and 27% at ages 10, 12, 14 and 16 years of age, respectively [82]. Peak $\dot{V}O_2$ is strongly correlated with body size and composition and thus, much of the divergence in values reflects this. When expressed as a ratio standard with body mass (ml • kg^{-1} • min^{-1}), peak $\dot{V}O_2$ shows a progressive decline in girls from 13 years of age, with values dropping from approximately 45 to 35 ml • kg^{-1} • min^{-1} [83]. In contrast, mass-related peak $\dot{V}O_2$ in young female runners has been found to be relatively constant with values of 56.3, 57.1, 56.9 and 54.3 ml • kg^{-1}·min^{-1} at ages 10, 12, 14 and 16 years, respectively [84]. Likewise, peak $\dot{V}O_2$ has been shown to be fairly stable between 11 and 16 years of age in elite girl swimmers and tennis players [85], with values of 51–52 ml • kg^{-1} • min^{-1} and 47–49 ml • kg^{-1} • min^{-1}, respectively. When multilevel modelling was used to account for mass, stature and biological age, the elite girl swimmers and tennis players showed increases in peak $\dot{V}O_2$ until late puberty when increases became non-significant [85]. Similarly, in the less athletic population when more appropriate allometric adjustment is used to partition size effects in body mass and stature, peak $\dot{V}O_2$ has been found to increase significantly from 11 to 13

years in girls, and then remain constant with no decline into adulthood evident [81].

Dramatic pubertal changes in muscle, fat and mass contribute to the widening of the sex difference in peak $\dot{V}O_2$. When a marker of body fat was included in a multilevel regression model which incorporated body mass, stature and age, the sex difference in peak $\dot{V}O_2$ was reduced, but the greater increase in boys' peak $\dot{V}O_2$ with growth compared to girls was still not fully explained [83]. Equally, longitudinal data have shown that even when differences in body mass and fat mass are controlled for allometrically, girls utilise less oxygen than boys during submaximal exercise, and this becomes more pronounced with age [86]. Understanding the physiologic mechanisms that underlie these size-independent sex differences in peak and submaximal $\dot{V}O_2$ requires consideration of the coordinated systems response, which includes pulmonary, cardiac and peripheral adjustments to the demands in muscular energy. A discussion of key features of each follows.

Pulmonary

It was generally assumed that because exercise training exerts little influence on lung structure or function that the lungs exert minimal influence on oxygen transport. However, there is evidence that lung function adaptation does occur as a consequence of exercise training in girls [87]. Moreover recent investigation of sex differences in pulmonary structure and function in adults has shown considerable effects on gas exchange and the integrated ventilatory response during exercise, in particular exercise-induced arterial hypoxia [88]. There are well-documented sex differences in anatomical aspects of the pulmonary system which occur during lung growth [89]. The consequence of sex dependent pubertal thoracic growth is a larger thoracic width in boys. When coupled with a greater muscle mass for generating lower lung function, boys have approximately 25% greater lung volumes than girls who are matched for stature [89]. By adult life, in addition to the smaller lung volumes, stature and age independent lower resting diffusion capacity (corrected for haemoglobin), lower maximal expiratory flow rates [90], and a greater occurrence of exercise-induced hypoxia has been shown in women [91]. Equally, there is also evidence that when matched for size and aerobic power women do not have reduced diffusion capacity or impaired ventilation perfusion during exercise [92].

In children, like adults, exercise pulmonary gas exchange depends on pulmonary ventilation ($\dot{V}_E$) and at maximal work rates high rates of ventilation are usual. Maximal values of 49–95 litres • min^{-1} have been recorded for girls between the ages of 9 and 16 years [93] and there is a consistent sex difference with values somewhat higher in boys (58–105 litres • min^{-1}) for the same age span. It should be noted that cross-study comparisons are difficult given the dependence of ventilation on the protocol and data such as these need to be interpreted cautiously. Maximum ventilation remains higher in boys, whether controlled for body size using a ratio standard or allometric adjustment with either stature and/or body mass [94, 95]. Thus, the higher peak $\dot{V}O_2$ in boys is indeed supported by a higher $\dot{V}_E$.

During exercise, an expiratory flow limitation is apparent in adult women but not men, resulting in a greater oxygen cost of breathing and the onset of arterial desaturation [88, 96]. Recent evidence has provided a comparison between pre-pubertal boys and girls and found no difference in the occurrence or severity of expiratory flow limitation between girls and boys and no changes in arterial saturation during exercise to maximum [97]. Others have found little evidence of exercise induced arterial hypoxaemia in pre-pubertal girls, or lower ventilatory efficiency at maximum [98, 99]. When Armstrong et al. [94] compared ventilatory parameters during submaximal exercise at the same absolute intensities they noted that girls demonstrated higher ventilatory equivalents for oxygen and carbon dioxide, i.e. poorer ventilatory

efficiency in comparison to boys. However, when they compared submaximal ventilatory efficiency during the same relative exercise intensities, values were remarkably similar between the sexes. This suggests that differences apparent at absolute submaximal exercise intensities simply reflect the higher relative percentage of maximum that girls are working at and do not denote true inefficiency.

There is little evidence that prior to puberty pulmonary structure or function limits oxygen uptake, however, considerable evidence has shown pulmonary function influences gas exchange in adult women, suggesting that maturational adjustments occur. At present however, there is little evidence to substantiate this.

Blood Volume and Haemoglobin

Assessment of blood volume in children and adolescents is complex and the variability in techniques means there are considerable discrepancies between studies. There are conflicting results regarding changes in blood volume with age. Some have shown that blood volume per unit body mass increases with age [100], others have found no change [101], whilst others report decreasing blood volume with age [102]. Likewise, data on sex differences in blood volume between girls and boys are mixed. When normalised using a ratio standard with body mass, differences between girls and boys were apparent from about 6 years of age, with values lower in the girls [103]. In contrast, when normalised using a ratio standard for lean body mass, sex differences are no longer apparent for pre-pubertal children, or at any maturational stage [103].

In boys, haemoglobin rises through adolescence to about 152 $g \bullet l^{-1}$ by 16 years of age [104]. Girls, on the other hand, usually demonstrate a plateau in haemoglobin concentration with values of about 137 $g \bullet l^{-1}$ by 16 years of age [104]. Highly trained adolescent female athletes also show lower haemoglobin concentration values compared to trained boys, with about a 7% difference [105]. Fully saturated, 1 g of haemoglobin carries 1.34 ml of oxygen, and one would presume that the smaller increase in haemoglobin in girls would result in a reduced oxygen carrying capacity in comparison to boys. However, it has been shown that haemoglobin concentration is not a significant predictor of peak $\dot{V}O_2$ in 11- to 17-year-olds once body size and composition and maturation have been controlled for [83].

Cardiac and Vascular Considerations

There are clear differences in cardiac function at rest and during exercise between girls and boys, with differences apparent even prior to puberty. The electrical conduction system is influenced by sex steroid hormones, with girls normally having higher resting heart rates than boys – somewhere in the magnitude of 90 beats per minute at around 10–12 years of age [106]. This is thought to relate to intrinsic differences in the sinus node pacemaker [107], a difference notable at birth with newborn boys displaying lower baseline heart rates than girls [16]. The higher resting heart rate in girls is often explained as an artefact of differences in cardiac dimensions, and indeed the ratio of heart mass to body mass has been found to be higher in boys than girls at birth, remaining so through adolescence [106]. Heart volume has also been found to be greater in boys with values of 342 and 403 ml for pre-pubertal girls and boys, respectively, and of 466 and 561 ml for pubertal girls and boys, respectively [108]. When adjusted for body mass these differences were found to persist through puberty (female 10.0 $ml \bullet kg^{-1}$; male 10.8 $ml \bullet kg^{-1}$). Inconsistencies in the findings, however, are present and others have found no differences in either left ventricular mass [109] or heart volume [110].

Echocardiographic studies that have shown greater left ventricular mass in boys compared

Table 1. Oxygen uptake, stroke index, cardiac index and arteriovenous oxygen difference at maximal cycle ergometer exercise

Reference	Sex	n	Age (years)	SI ($ml \cdot m^{-2}$)	HR (bpm)	CI (litres $\cdot min^{-1} \cdot m^{-2}$)	a-v O_2 difference ($ml \cdot 100\ ml^{-1}$)	Peak $\dot{V}O_2$ (litres $\cdot min^{-1} \cdot kg^{-1}$)	Peak $\dot{V}O_2$ (litres $\cdot min^{-1}$)
Cumming [111]	F	29	11.8±3.1	46±3†	174±11	8.61±8.1†	–	–	–
	M	31	12.6±3.5	56±13	170±17	10.1± 1.8	–	–	–
Rowland et al. [112]	F	24	11.7±0.5	55±9†	198±9	10.9±1.7†	12.3±1.9	40.4±5.8†	1.84±31
	M	25	12.0±0.4	62±9	199±11	12.3±2.2	12.2±1.7	47.1±6.1	1.98±28
Obert et al. [113]									
Pre-training experimental group	F	7	10.66±0.3	47±7†	204±5	9.4±1.4†	13.2±1.6	40.9±8.9†	–
	M	9	10.66±0.5	52±8	199±9	10.5±1.8	13.0±2.1	44.1±6.1	–
Pre-training control group	F	10	10.41±0.3	46±6†	202±7	9.4±1.2†	13.1±2.8†	42.4±5.6†	–
	M	9	10.5±0.3	49±5	202±7	9.7±0.8	15.6±1.5	51.5±6.3	–
Winsley et al. [114]	F	9	10.2±0.3	45±6	192±11	8.7±1.1	12.6±1.6†	–	1.23±.08†
	M	9	10.1±0.5	47±8	195±11	8.9±1.4	14.8±2.1	–	1.41±.18

SI = Stroke index; HR = heart rate; CI = cardiac index; a-v O_2 difference = arterio-venous oxygen difference; $\dot{V}O_2$ = oxygen uptake. † Significant differences noted.

to girls have suggested that the reduced cardiac mass in girls may be associated with reduced contractility, reduced pre-load or increased afterload [106]. All of these could result in a reduced stroke index (SI) and therefore reduced cardiac index (CI). Cardiac index has generally been found to be higher in boys than girls at maximal exercise (table 1) and in the absence of sex differences in maximal heart rate, it would appear that SI most likely accounts for this difference. Absolute maximal SI index has been reported to be between 7 and 13% less in girls than boys. When corrected for body fat, this difference was reduced to 5.2% [112], but remained nonetheless. Interestingly, the lower maximal SI index apparent in girls has not always been found to relate to left ventricular dimensions, which suggests sex differences may instead relate to other factors such as the peripheral pump, systemic vascular resistance or differing adrenergic responses [113, 115].

Evidence of cardiac re-modelling following training has provided some insight into the role systemic vascular resistance may play in SI differences between boys and girls [113]. Following 13 weeks of training, both pre-pubertal boys and girls increased LV end-diastolic diameter and left ventricular mass. However, only LV end-diastolic diameter was related to percent increase in SI. Percent increase in SI was also inversely related to systemic vascular resistance, suggestive of vascular adaptations in response to high-intensity training. Of note, the decrease in systemic vascular resistance was greater in the boys than the girls, which may account for the greater increase in maximal SI in the boys.

The vasoregulatory capacity of the arterial and arteriolar vessels manipulates peripheral resistance. When blood is effectively distributed to the working muscle, peripheral resistance is reduced, which unloads the heart improving the capacity of the heart to increase SI. This is achieved by improving the flow of blood to and through the muscle and both the vasculature and skeletal muscle pump are involved. Interestingly, no sex differences in arterial compliance have been noted in pre- and early-pubertal children [116], although the beneficial role of oestrogen in vasodilation is well established and female advantage in arterial compliance is apparent in adults [117].

The skeletal muscle pump utilises the rhythmic muscle contractions to empty the venous vessels, aiding blood muscle hyperaemia and venous return. There is scant information on the skeletal muscle pump in children, but evidence in boys suggests, like adults, the skeletal muscle pump is associated with improved CI [118, 119]. In Rowland et al.'s [119] study, arteriovenous oxygen (a-v O_2) difference, a composite index of the haematological components of oxygen delivery, remained constant during unloaded exercise suggesting the increases in muscle oxygen supply were met by the increasing blood volume. Conversely, as exercise intensity increased with loading, a-v O_2 difference increased indicating decreased effectiveness of the muscle pump in satisfying the metabolic demands of the working muscle.

Whilst some studies have found no differences in estimated a-v O_2 difference at maximal or submaximal intensities between pre-pubertal girls and boys [112, 120], there are conflicting findings. Data recently published from a thoracic impedance measure of peak CI and MRI markers of cardiac size [114] demonstrated that pre-pubertal boys had a 16.7% higher a-v O_2 difference than girls. This was the only distinguishing factor to explain the significantly higher peak $\dot{V}O_2$ in the boys compared to the girls and unlike other studies no difference in either CI or SI were apparent at maximal exercise. It is interesting to note that a-v O_2 difference was 16% lower in the girls of the control group (table 1) in the study of Obert et al. [113]. These findings are intriguing, but confirmatory studies are needed to help understand the inconsistencies in the extant data.

Muscle Cellular Metabolism during Moderate Intensity Exercise

Characterizing muscle metabolism during exercise is extremely challenging and for a long time hampered by the need for invasive measurement of enzymatic activity. ^{31}P magnetic resonance spectroscopy (MRS) has enabled the study of high energy phosphates non-invasively in human skeletal muscle. This technique can provide an estimation of skeletal muscle metabolic activity via examination of creatine phosphate (PCr), inorganic phosphate (P_i) and intracellular pH, and has been validated in both adults and children [121, 122]. There remain methodological challenges in the paediatric population, which have been outlined by Armstrong and Fawkner [123], but the data available are providing fascinating insight into cellular metabolic processes.

Children, like adults show high correspondence between MRS determined muscle phosphocreatine (PCr) activity and the pulmonary oxygen uptake ($p\dot{V}O_2$) kinetic response [124, 125]. This implies that $p\dot{V}O_2$ kinetics also provide a marker of energy utilization at the muscular level, one which may prove very useful in understanding the interplay between cardiopulmonary and metabolic processes during exercise. More comprehensive descriptions of oxygen uptake kinetic assessments have been provided elsewhere [126] and only the salient issues related to girls' responses are summarized here. The $p\dot{V}O_2$ kinetic response is tri-phasic, but only phases II and III pertain to muscle oxygen uptake kinetics. During moderate intensity exercise, the phase II $p\dot{V}O_2$ kinetic response involves

an exponential increase in oxygen uptake toward steady state, which signifies increases in muscle $\dot{V}O_2$. The primary response is described by a time constant (τ), representing the time taken (s) to achieve 63% of the change in $p\dot{V}O_2$. The attainment of a steady state denotes phase III. At higher workloads, i.e. those above the maximal lactate steady state, the $p\dot{V}O_2$ kinetic response alters, with phase III showing a delayed increase, eventually resulting in a $p\dot{V}O_2$ value higher than predicted on the basis of exercise intensity. This 'slow component' represents an increasing inefficiency in energy turnover and negatively correlates with increases in $\dot{V}O_2$ per unit increases in work, suggesting fatigue. To ensure confidence in the kinetic parameters estimated, the level of measurement rigour needed is high [126]. Few of the available oxygen uptake kinetics studies with children provide this and as such information on girls is very limited.

There is little evidence of a sex difference in $p\dot{V}O_2$ kinetic responses during moderate intensity exercise in children [127]. Neither have sex differences been found in boys and girls for MRS determined pH, P_i to PCr ratio (P_i/PCr) or PCr kinetic time constant at either the onset or offset of moderate intensity exercise [128]. In contrast, a study of the kinetic responses to high-intensity exercise found sex differences [129]. Results showed phase II $p\dot{V}O_2$ kinetics were approximately 20% slower in pre-pubertal girls compared to boys and the relative contribution of the $p\dot{V}O_2$ slow component to the end exercise $p\dot{V}O_2$ in the girls was about 30% greater. This is suggestive of a lower tolerance of fatigue in the girls, but the mechanisms underlying this response are not yet understood. One hypothesis suggests that these differences reflect a difference between boys and girls in the energetic profiles of the recruited muscles.

Barker et al. [130] have explored high-intensity exercise responses of the quadriceps muscle using MRS in children, but showed in accord with the $p\dot{V}O_2$ kinetic work, that girls responded with a greater anaerobic metabolic contribution than boys. These findings were partly attributed to the inequalities in maturity status, with relatively immature boys compared to the girls. Maturation of the cellular anaerobic response was noted in the girls in this study, who progressed from a response that was attenuated prior to puberty, but adult-like with ensuing maturation. This was not apparent in the boys, most likely an artefact of the narrow age range of the boys (9–12 years). Generally, high-intensity work requires the recruitment of fast twitch muscle fibres that are faster and larger, with a greater glycolytic and lower oxidative capacity. As discussed earlier, there is evidence of sex differences in muscle fibre type and size which vary with age and maturation, and clearly comparison of cellular metabolism during high-intensity exercise in girls and boys who are more closely aligned in terms of maturation is something which deserves further enquiry.

To summarise, there are differences between boys and girls in the aerobic responses to exercise which cannot be accounted for solely by size. Ventilatory parameters do not appear to influence peak $\dot{V}O_2$ in pre-pubertal children, however, there is scant information on the maturation of ventilatory responses in girls. It has been suggested that peripheral factors may be more important in defining aerobic fitness than cardiac function [131], but these are poorly understood in children and in particular in girls.

Acute Responses to High-Intensity Exercise

Most sports require short-duration bursts of high-intensity effort, which are supported by high muscle energy turnover. The direct examination of muscular energetics during short-duration high-intensity exercise is complex and instead investigations have largely concentrated on mechanical output markers of short duration exercise performance. The most commonly employed tests are the Wingate cycle ergometer test (WAnT) and cycle ergometer force-velocity tests,

both eliciting markers of leg power. Wingate test values for leg peak power in girls aged 11–16 years have ranged from 260 to 542 W [132–136], whilst comparable values between 250 and 555 W have been recorded using force-velocity tests in similarly aged girls [137–140]. Mean power values from the Wingate test have ranged from 228 to 341 W in 11- to 16-year-old girls [132, 136]. It is interesting to note that neither peak nor mean power appear to be unusually high in young girls who are engaged in elite tennis, swimming or gymnastics training [141]. Higher values have been recorded for elite handball players and elite sprinters [133, 141], which could not be fully explained by age and body composition; however, when comparison was made with published values for less athletic girls, peak and mean power for these elite girl athletes were not substantially different.

Longitudinal data have shown that leg peak power increases with age in both boys and girls, but the increases in boys are greater than in girls. In a study of 7- to 18-year-olds, peak power was shown to increase by 273% in girls from 7 to 16 years of age, and then to plateau [142]. In comparison, boys showed increases of 375% over this period with no plateau at 16 years. Armstrong et al. [132] examined changes in leg peak power from 12 to 17 years of age and noted increases of 66% in girls, whilst boys increased peak power by 120% over the same period. The increases noted by Armstrong et al. [132] are similar in magnitude to those of Martin et al. [142] when the same age range is considered. Similar age-related increases in mean power have been noted, again with increases in boys almost double those of girls between the ages of 12 and 17 years [132]. Sex differences in peak leg power do not appear to emerge until about 14 years of age [141], whilst mean power is greater in boys than girls from about 13 years of age [132]. Clearly age is an important predictor of short-term power in young people. The influence of stature and mass as predictors of peak and mean power have also been established [132, 143], highlighting the need to consider both body mass and composition when assessing short-term power. De Ste Croix and colleagues [143] have shown that in addition to the effects of body mass, sum of skinfolds and age, MRI determined thigh muscle volume exerts considerable influence on young people's short-term power output. Furthermore, De Ste Croix and colleagues [143] have shown using multilevel modelling that in addition to the effects of body mass, sum of skinfolds and age, MRI-determined thigh muscle volume has a significant impact on young people's short-term power output during cycling.

There are very few data on skeletal muscle metabolism during short-duration high-intensity exercise in girls. A MRS study of pH and P_i/PCr ratio during supramaximal plantar flexion exercise in pre-pubertal and pubertal girls found that the maturational differences in pH and P_i/PCr values were not statistically significant [144]. The authors concluded that glycolytic metabolism was not maturity dependent; rather, it was dependent on muscle cross-sectional area. A more recent study of the PCr kinetics and intracellular pH response during high-intensity exercise also showed a non significant sex difference in pH. It is worthy of note that in both studies [144, 145] there was considerable variability within small samples which may be masking biological significance. Wilcox et al. [145] did demonstrate that the PCr cost per watt was higher in the girls compared to the boys [145]. These findings suggest lower efficiency in the girls compared to the boys, which may be an outcome of differences in muscle fibre type, muscle activation patterns or leg vasodilatory response. However, this study failed to demonstrate that the differences in the slow component of the PCr response between children and adults were statistically significant, raising doubt that age-related change in muscle fibre recruitment substantially influences skeletal muscle metabolism during high-intensity exercise.

Does Intensive Training Pose a Threat to the Development and Health of Young Girls?

Growth

Many young athletes begin formal training before 10 years of age, with young elite gymnasts, swimmers and tennis players entering their respective sport between the ages of 6 and 7.5 years [146]. In a number of countries young girls are recruited into specialised sport schools as young as 5 years of age [147]. These elite young athletes train intensively all year round, for many hours, with weekly training volumes of 24 h not being unusual [148]. Whether intensive training such as this distorts normal growth and maturation remains a topic of much debate [7, 149].

Evidence of reduced or delayed growth in some young athletes, such as gymnasts, has been suggested to be a direct outcome of the intensive training these youngsters have endured [5, 50, 151]. Counter-argument contends that growth reductions or delay in young athletes simply reflect their late maturation [136, 152–154]. The Training of Young Athletes (TOYA) study found that elite young female swimmers and tennis players were generally taller than the general population throughout the growth period (close to the 75th percentile for stature), whilst gymnasts were generally smaller (below the 50th percentile for stature) from 10 to 17 years of age [146]. What was noteworthy was that by 18 years of age, the gymnasts were above the 50th percentile for stature and when aligned by biological age (years from attainment of menarche), gymnasts, swimmers and tennis players showed no significant differences in height. The catch-up growth noted in the gymnasts was indicative of late maturation and apparent in girls who are not involved in competitive training, but who mature late. Both the fathers and mothers of gymnasts have been found to be significantly shorter than the parents of other athletes and genetic predisposition for stature has been not only been shown to be preserved, but often exceeded [146, 155]. Combined, this evidence indicates that the tendency for short stature in gymnasts is not, as argued by some [151], evidence of a training-induced alteration in growth, but more likely a reflection of a genetic predisposition for later development and short stature.

Reproductive Health

Menstrual dysfunction in young athletes has also been interpreted as evidence that intensive training in young girls is harmful to reproductive health [156]. Menstrual dysfunction includes delayed menarche (onset after 16 years), luteal phase defects, oligomenorrhea and amenorrhea (table 2). Several studies have concluded that female gymnasts, swimmers and ballet dancers have delayed menarche. De Ridder et al. [29] observed that, in comparison to a control group of girls matched for maturation and fatness, girls involved in competitive gymnastics exhibited delayed menarche. An early hypothesis suggested that because these young girls had low levels of body fat they did not attain a critical level of body fat (22%) necessary for menstruation. The wide variability noted in body fat at menarche [35] has provided proof that a threshold of 22% body fat is incorrect. Additionally, there is sufficient experimental evidence in women to show that it is not body fat but caloric deprivation that affects reproductive health [157].

Genetic predisposition for late menarche in athletes has also been explored. When age of menarche in a group of elite gymnasts was correlated with maternal menarchal age, it was, on average, in lag by 0.81 years [158]. This lag was double that noted for elite swimmers and triple that for elite tennis players. These data suggest that despite a genetic predisposition for delayed menarche, this does not fully explain the extent of the delay, signalling that training may indeed delay menarche in gymnasts. However, Baxter-Jones et al. [158] went on to show that when the time period between menarchal age and retirement from the sport were

Table 2. Components of the female athlete triad, diagnosis and prevention

Component of the triad	Diagnosis	Warning signs	Prevention
Energy availability [160]	Energy availability is defined as energy intake minus exercise energy expenditure, with a threshold of 30 kcal • kg^{-1} LBM • day^{-1}.	Low body mass (>85% of ideal body mass for stature). Fatigue.	Monitor dietary intake. Monitor training volume. Focus on healthy eating and caloric balance. Educate youngsters about nutritional fads. Reinforce message that body mass is only one aspect of good performance. Educational information on nutrition and energy expenditure, e.g. http://kidshealth.org/teen/food/sports/triad.html
Eating disorders [173]	*Anorexia nervosa* Refusal to maintain body mass over a minimally normal mass for age and stature. Intense fear of gaining mass or becoming fat. Disturbed body image. Secondary amenorrhea. *Bulimia nervosa* Recurrent episodes of binge eating (eating a large amount of food in a discrete period of time and lacking control over eating during the episode). Recurrent inappropriate compensatory behaviour such as self-induced vomiting, laxatives or excessive exercise. The binge-eating and purging behaviours occur at least twice a week for 3 months.	*Anorexia* Dramatic loss in body mass. Preoccupation with food, calories and body mass. Wears baggy clothes. Fine, downy facial hair. Mood swings. Avoidance of food-related social activities. *Bulimia* Noticeable loss in body mass. Excessive worry over weight. Bathroom visits after eating. Depression. Strict dieting followed by binging. Dental erosion.	Promote healthy body image. Removal of body mass/fat monitoring by coaches. Provide opportunities for developing self-coping strategies. Deemphasize body mass and thinness. Provide opportunities for nutritional counselling.
Menstrual dysfunction [174]	Oligomenorrhea – irregular menses (length between cycles >35 days). Primary amenorrhea – absence of menstruation by 15 years in girls with secondary sexual characteristics. Secondary amenorrhea- absence of menstrual cycles for 3 cycles after onset of menses.	Irregular or absent menstrual cycle.	Ask athletes to keep a training diary and include monitoring of menstrual cycle. Help girls understand that secondary amenorrhea is not normal. Provide education on reproductive health and the link between menstruation and bone health. Provide dietary education and help girls understand the link between diet and reproductive health.

Table 2. Continued

Component of the triad	Diagnosis	Warning signs	Prevention
	Luteal phase dysfunction – shortened secretory phase of the menstrual cycle, typically less than 10 days.		
Low bone mineral density [175]	If comparison with age, gender, stature and race specific Z-scores yields values ≤2.0 this is classified as a low bone mineral density for chronological age. Osteoporosis is diagnosed if low BMD for chronological age is accompanied by one or more of the following fracture histories: long bone fracture of the lower extremities; vertebral compression fracture and two or more long–bone fractures of the upper extremities.	Secondary amenorrhea. Stress fracture. History of fractures.	Provide educational information on osteoporosis. Provide information on nutrition for bone health, particularly focusing on calcium-rich foods. Monitor diet and provide opportunities for nutritional counselling.

considered, 92% of the girls began menarche prior to retiring. The authors concluded that training was therefore unlikely to cause the delay in menarche noted in the gymnasts, instead there was simply a chronological age difference in the timing of events. It would appear that exercise training, without other predisposing factors, is unlikely to be the cause of menstrual dysfunction.

The Female Athlete Triad

The female athlete triad was established in the early 1990s as a syndrome of three separate, but inter-related conditions, namely menstrual dysfunction, disordered eating and premature osteoporosis [159]. An updated position statement from the American College of Sports Medicine (ACSM) has revised the definition of the triad as the presence of one or more of (1) low energy availability (with or without eating disorders), (2) amenorrhea, and (3) osteoporosis (table 2) [160]. Prevalence estimates of components of the female triad are very dependent on the athletic group studied, with higher rates in sports where low body mass is the norm. For instance, 25% of young women in endurance, weight class or aesthetic sports had clinical eating disorders, compared to 9% of the general population [161]. Secondary amenorrhea has been reported to be as high as 69% in dancers and less than 1% in the general population [162, 163]. Osteoporosis has been found in about 13% of female athletes, although this is not too different from the normal population [164] and in pre-menopausal women low bone mineral density for age is a more appropriate marker than osteoporosis.

Much of the available data on the female triad is on college-age or young adult athletes, with few reports targeting adolescents. Two key studies of

high school athletes have shown a considerable number of girls present with components of the female triad. In a study of 170 13- to 18-year-olds, 18% had disordered eating, 24% had oligomenorrhea or amenorrhea and 22% had low bone mass for their age based on WHO diagnostic criteria [165]. A higher rate of occurrence of low energy availability (55%) was noted in a study of 80 similarly aged young athletes, with 16% diagnosed with amenorrhea and using the same WHO diagnostic criteria, 16% presented with low bone mineral density [166]. The existence of the female athlete triad in these young girls is particularly worrying given this is a time when substantial amounts of bone should be accrued.

Energy deficiency appears to be particularly harmful when combined with excessive exercise, and leads to reduced oestrogen levels, athletic amenorrhea and bone demineralisation. Loucks et al. [167] have shown that there is an energy availability threshold of 20–25 kcal • kg^{-1} LBM • day^{-1}, below which skeletal and reproductive health is compromised. This group conducted a number of studies in which women underwent energy availability manipulations, decreasing energy availability from 45 to 20 kcal • kg^{-1} LBM • day^{-1} or from 45 to 10 kcal • kg^{-1} LBM • day^{-1}. These reductions caused blunting of LH pulsatility [168] and a de-linking of bone resorption and formation [169]. The existence of an energy availability threshold may help to explain why not all athletes develop athletic amenorrhea even when following the same training programme and provides a useful marker for nutritional health.

Recent conjecture that the triad is a 'myth' [8] has caused intense debate [164, 170, 171]. This contention stems from a number of criticisms, including flaws in the epidemiological evidence, assumptions that low energy availability implies disordered eating and a lack of experimental evidence in athletes. Much of the epidemiological evidence of the prevalence of the female triad is for individual components of the triad, rather than for the synchronous appearance of all three which, it has been argued, over-inflates the extent of the problems. When occurrence of all three components is examined, prevalence falls dramatically [172]. On the other hand, the definition of the female athlete triad states explicitly that presence of one component is sufficient for diagnosis and the revised guidelines provided by the ACSM, have removed disordered eating and replaced it with energy availability, accepting that low energy availability does not equate to a pathological eating disorder [160].

Definitive conclusions on whether elite participation causes aberrations to skeletal and reproductive health are not possible. Nevertheless, there is clear evidence of a boundary between healthy and unhealthy levels of exertion when coupled with caloric limitation. Exposure to excessive training and caloric limitation causes abnormality in skeletal and reproductive function and regardless of the magnitude of the problem, young girl athletes deserve protection. Protection likely entails athlete education, coach recognition of the triad and the monitoring of both training volume and nutritional health in young elite girls. Table 2 provides an overview of the components of the triad, with possible prevention strategies for young athletes, coaches and parents.

Conclusions

Girls have differential growth and development in comparison to boys, resulting in substantial differences in body size and composition. Whilst stronger and leaner than many of her non-athletic peers, the smaller stature, shorter legs, lower muscularity and greater relative fatness of the elite girl athlete means she is not as strong, nor as fast as her male counterpart. Some discordant responses to exercise are not solely explained by body size and/or composition and there is evidence of underlying qualitative differences which require further clarification.

A young girl's involvement in elite sport predisposes her to increased risk of skeletal and reproductive health problems, particularly in sports where intense training is coupled with the need for leanness. Ensuring girls involved in elite training are in an environment which optimises their athletic potential while minimising risk is a priority. This entails sufficient knowledge of the physiologic responses to exercise in girls, as well as thorough understanding of the female triad disorder, its aetiology and prevention. Despite the burgeoning literature in the field of paediatric exercise physiology, an integrative understanding of girls' physiological responses to exercise remains elusive. Traditional technologies have proved inadequate in providing a detailed understanding of the complexity of the cellular metabolic response to exercise and this is compounded by the need to separate qualitative changes from changes which are an artefact of a growing, maturing body. More recent application of non-invasive imaging techniques and breath-by-breath gas analysis is facilitating a more integrated understanding of the responses to exercise, but the number of studies to date with girls is woefully small. In addition, more needs to be learnt about the female triad and its antecedents in younger girls. The dearth of information on girls and emergence and availability of new technologies provides plenty of scope for future studies in paediatric exercise physiology.

References

1 Armstrong N, Welsman JR, Kirby BJ: Longitudinal changes in 11–13-year-olds' physical activity. Acta Paediatr 2000;89:775–780.

2 Women's Sport and Fitness Foundation: Participation. Factsheet November 2009:http://www.wsff.org.uk/documents/Participation_factsheet_November_09.pdf

3 Women's Sport and Fitness Foundation: WSSF viewpoint – women and the Olympics. Factsheet July 2009:http://www.wsff.org.uk/documents/Women_and_the_Olympics.pdf

4 Vertinsky P: The Eternally Wounded Woman: Women, Doctors and Exercise in the Late Nineteenth Century. Chicago, University of Illinois Press, 1994.

5 Theintz G, Howald H, Weiss U, Sizonenko P: Evidence for a reduction of growth potential in adolescent female gymnasts. J Pediatr 1993;122:306–313.

6 Damsgaard R, Bencke J, Matthiesen G, Petersen J, Müller J: Is prepubertal growth adversely affected by sport? Med Sci Sports Exerc 2000;32:1698–1703.

7 Baxter-Jones ADG, Maffulli N, Mirwald RL: Does elite competition inhibit growth and delay maturation in some gymnasts? Probably not. Pediatr Exerc Sci 2003;15:373–382.

8 DiPietro L, Stachenfeld NS: The myth of the female athlete triad. Br J Sports Med 2006;40:490–493.

9 McManus AM, Armstrong N: The elite child athlete; in Armstrong N, van Mechelen W (eds): Paediatric Exercise Science and Medicine, ed 2. Oxford, Oxford University Press, 2008, pp 489–502.

10 USA Track and Field Age Group Records. http://www.dyestat.com/rivals/227840.html#GIRLS

11 World Junior Records http://www.iaaf.org/statistics/index.html

12 Coast JR, Blevins JS, Wilson BA: Do gender differences in running performance disappear with distance? Can J Appl Physiol 2004;29:139–145.

13 George K, Sharma S, Batterham A, Whyte G, McKenna W: Allometric analysis of the association between cardiac dimensions and body size variables in 464 junior athletes. Clin Sci 2001;100:47–54.

14 Nevill AM, Holder RL: Scaling, normalizing, and per ratio standards: an allometric modeling approach. J Appl Physiol 1995;79:1027–1031.

15 Ao A, Erickson RP, Winston RML, Handysiade A: Transcription of paternal Y-linked genes in the human zygote as early as the pro-nucleate stage. Zygote 1994;2:281–287.

16 Turgeon JL, Carr MC, Maki PM, Mendelsohn ME, Wise PM: Complex actions of sex steroids in adipose tissue, the cardiovascular system and brain: insights from basic science and clinical studies. Endocr Rev 2006;27:575–605.

17 Wells JCK: Sexual dimorphism of body composition. Best Pract Res Clin Endocrinol Metab 2007;21:415–430.

18 Rogol A, Clark P, Roemmich J: Growth and pubertal development in children and adolescents: effects of diet and physical activity. Am J Clin Nutr 2000;72:S521–S528.

19 Mauras N: Growth hormone and testosterone: effects on whole body metabolism and skeletal muscle in adolescence. Horm Res 2006;66:42–48.

20 Baxter-Jones ADG: Growth and maturation; in Armstrong N, van Mechelen W (eds): Paediatric Exercise Science and Medicine, ed 2. Oxford, Oxford University Press, 2008, pp 157–168.

21 Susman EJ, Houts RM, Steinberg L, Belsky J, Cauffman E, Dehart G, Friedman SL, Roisman GI, Halpern-Felsher BL, Eunice Kennedy Shriver NICHD Early Child Care Research Network: Longitudinal development of secondary sexual characteristics in girls and boys between ages 9 1/2 and 15 1/2 years. Arch Pediatr Adolesc Med 2010; 164:166–173.

22 Sherar LB, Baxter-Jones AD, Mirwald RL: Limitations to the use of secondary sex characteristics for gender comparisons. Ann Hum Biol 2004;31:586–593.

23 Malina RM, Bouchard C, Bar-Or O (eds): Growth, Maturation and Physical Activity. Champaign, Human Kinetics, 2004, pp 39–81.

24 Sinclair D, Dangerfield P: Human Growth after Birth, ed 6. Oxford, Oxford University Press, 1988.

25 Claessons AL, Veer FM, Stijnen V, Lefevre J, Maes H, Steens G, Beunen G: Anthropometric characteristics of outstanding male and female gymnasts. J Sports Sci 1991;9:53–74.

26 Damsgaard R, Bencke J, Matthiesen G, Petersen J, Müller J: Body proportions, body composition and pubertal development of children in competitive sports. Scand J Med Sci Sports 2001;11:54–60.

27 Reynolds EL, Wines JV: Individual differences in physical changes associated with adolescence in girls. Am J Dis Child 1948;75:329–350.

28 Tanner JM: Growth at Adolescence, ed 2. Oxford, Blackwell Scientific Publications, 1962.

29 de Ridder C, Thijssen J, Bruning P, Van den Brande J, Zonderland M, Erich W: Body fat mass, body fat distribution, and pubertal development: a longitudinal study of physical and hormonal sexual maturation of girls. J Clin Endocrinol Metab 1992;75:442–446.

30 Kaplowitz P: Pubertal development in girls: secular trends. Curr Opin Obstet Gynecol 2006;18:487–491.

31 Biro FM, Lucky AW, Simbartl LA, Barton BA, Daniels SR, Striegel-Moore R, Kronsberg SS, Morrison JA: Pubertal maturation in girls and the relationship to anthropometric changes: pathways through puberty. J Pediatr 2003;142: 643–646.

32 Malina R: Physical growth and biological maturation of young athletes. Exerc Sport Sci Rev 1994;22:389–434.

33 Baxter-Jones ADG, Eisenmann JC, Sherar LB: Controlling for maturation in pediatric exercise science. Pediatr Exerc Sci 2005;17:18–30.

34 Klentrou P, Plyley M: Onset of puberty, menstrual frequency, and body fat in elite rhythmic gymnasts compared with normal controls. Br J Sports Med 2003; 37:490–494.

35 Sherar LB, Baxter-Jones AD, Mirwald RL: The relationship between body composition and onset of menarche. Ann Hum Biol 2007;34:673–677.

36 Matthews BL, Bennell KL, McKay HA, Khan KM, Baxter-Jones AD, Mirwald RL, Wark JD: The influence of dance training on growth and maturation of young females: a mixed longitudinal study. Ann Hum Biol 2006;33:342–356.

37 Barrack MT, Rauh MJ, Nichols JF: Cross-sectional evidence of suppressed bone mineral accrual among female adolescent runners. J Bone Min Res 2010;25:1850–1857.

38 Rusko H, Rahkila P, Karvinen E: Anaerobic threshold, skeletal muscle enzymes and fibre composition in young female cross-country skiers. Act Physiol Scand 1980;108:263–268.

39 Thorland WG, Johnson GO, Fagot TG, Tharp GD, Hammer RW: Body composition and somatotype characteristics of Junior Olympic athletes. Med Sci Sports Exerc 1981;13:332–338.

40 Malina RM: Regional body composition: age, sex and ethnic variation; in Roche AF, Heysnfield SB, Lohman TG (eds): Human Body Composition. Champaign, Human Kinetics, 1996, pp 217–255.

41 Falk B, Dotan R: Temperature regulation and elite young athletes; in Armstrong N, McManus AM (eds): The Elite Young Athlete. Med Sports Sci. Basel, Karger, 2011, vol 56, pp 126–149.

42 Wells JCK, Williams JE, Chomtho S, Darch T, Grijalva-Eternod C, Kennedy K, Haroun D, Wilson C, Cole TJ, Fewtrll MS: Pediatric reference data for lean tissue properties: density and hydration from age 5 to 20 y. Am J Clin Nutr 2010;91:610–618.

43 Lohman TG: Assessment of body composition in children. Pediatr Exerc Sci 1989;1:19–30.

44 Rogol AD: Growth at puberty: interaction of androgens and growth hormone. Med Sci Sports Exerc 1994;26:767–770.

45 Simoneau JA, Bouchard C: Human variation in skeletal muscle fibre-type proportion and enzyme activities. Am J Physiol 1989;257:E567–E572.

46 Staron RS, Hagerman FC, Hilida RS, Murray TF, Hostler DP, Crill MT, Ragg KE, Toma K: Fiber type composition of the vastus lateralis muscle of young men and women. J Histochem Cytochem 2002;48:623–629.

47 Schrauwen-Hinderling VB, Hesslink MKC, Schrauwen P, Kooi ME: Intramyocellular lipid content in human skeletal muscle. Obesity 2006;14:357–367.

48 Roepstorff C, Steffenson CH, Madsen M, Stallknecht B, Kanstrup L, Richter RA, Riens B: Gender differences in substrate utilization during submaximal exercise in endurance-trained subjects. Am J Physiol Endocrinol Metab 2002;282: E435–E447.

49 Hunter SK, Enoka RM: Sex differences in the fatigability of arm muscles depends on absolute force during isometric contractions. J Appl Physiol 2001;91:2686–2694.

50 Jaworowski Å, Porter MM, Holmbäck AM, Downham D, Lexell J: Enzyme activities in the tibialis anterior muscle of young moderately active men and women: relationship with body composition, muscle cross-sectional area and fibre type composition. Acta Physiol Scand 2002;176:215–225.

51 Legerlotz K, Smith HK, Hing WA: Variation and reliability of ultrasonographic quantification of the architecture of the medial gastrocnemius muscle in young children. Clin Physiol Funct Imaging 2010;30:198–205.

52 De Ste Croix MBA, Armstrong N, Welsman JR, Sharp P: Longitudinal changes in isokinetic leg strength in 10–14 year olds. Ann Hum Biol 2002;29:50–62.

53 O'Brien TD, Reeves ND, Baltzopoulos V, Jones DA, Maganaris CN: In vivo measurements of muscle specific tension in adults and children. Exp Physiol 2009;95:202–210.

54 Nygaard E: Skeletal muscle fibre characteristics in young women. Acta Physiol Scand 1981;112:299–304.

55 Vogler C, Bove KE: Morphology of skeletal muscle in children. Arch Path Lab Med 1985;109:238–242.

56 Glenmark B, Hedberg G, Jansson E: Changes in muscle fibre type from adolescence to adulthood in women and men. Acta Physiol Scand 1992;146:251–259.

57 Glenmark B, Hedberg G, Kaijster L, Jansson E: Muscle strength from adolescence to adulthood: relationship to muscle fibre types. Eur J Appl Physiol Occup Physiol 1994;68:9–19.

58 Lexell J: Human aging, muscle mass, and fibre type composition. J Gerontol 1995; 50:11–16.

59 Maffulli N, King JB, Helms P: Training in elite young athletes (the training of young athletes (TOYA study): injuries, flexibility and isometric strength. Br J Sport Med 1994;28:123–136.

60 Bencke J, Damsgaard R, Saekmose A, Jørgensen P, Jørgensen K, Klausen K: Anaerobic power and muscle strength characteristics of 11 years old elite and non-elite boys and girls from gymnastics, team handball, tennis and swimming. Scand J Med Sci Sports 2002;12: 171–178.

61 Schoenau E: Bone mass increase in puberty: what makes it happen? Horm Res 2006;65(Suppl 2):2–10.

62 Kontulainen SA, Macdonald HM, McKay HA: Change in cortical bone density and its distribution differs between boys and girls during puberty. J Clin Endocrinol Metab 2006;91:2555–2561.

63 Högler W, Blimkie CJ, Cowell CT, Kemp AF, Briody J, Wiebe P, Farpour-Lambert N, Duncan CS, Woodhead HJ: A comparison of bone geometry and cortical density at the mid-femur between prepuberty and young adulthood using magnetic resonance imaging. Bone 2003;33:771–778.

64 Schoenau E, Neu MC, Manz F: Muscle mass during childhood: relationship to skeletal development. J Musculoskelet Neuronal Interact 2004;4:105–108.

65 Schoenau E: From mechanostat theory to development of the 'Functional Muscle-Bone-Unit'. J Musculoskelet Neuronal Interact 2005;5:232–238.

66 Ward KA, Roberts SA, Adams JE, Mughal MZ: Bone geometry and density in the skeleton of pre-pubertal gymnasts and school children. Bone 2005;36: 1012–1018.

67 Duncan CS, Blimkie CJR, Cowell CT, Burke ST, Briody JN, Howman-Giles R: Bone mineral density in adolescent female athletes: relationship to exercise type and muscle strength. Med Sci Sports Exerc 2002;34:286–294.

68 Haapasalo H, Kontulainen S, Sievanen H, Kannus P, Jarvinene M, Vuori I: Exercise-induced bone gain is due to enlargement in bone size without a change in volumetric bone density: a peripheral quantitative computed tomography study of the upper arms of male tennis players. Bone 2000;27:351–357.

69 Duncan CS, Blimkie CJR, Kemp A, Higgs W, Cowell CT, Woodhead H, Briody JN, Howman-Giles R: Mid-femur geometry and biomechanical properties in 15- to 18-yr-old female athletes. Med Sci Sports Exerc 2002;34:673–681.

70 Högler W, Blimkie CJR, Cowell CT, Inglis, Rauch F, Kemp AF, Wiebe P, Duncan CS, Farpour-Lambert N, Woodhead HJ: Sex-specific developmental changes in muscle size and bone geometry at the femoral shaft. Bone 2008;42:982–989.

71 Dowthwaite JN, Kanaley JA, Spadaro JA, Hickman RM, Scerpella TA: Muscle indices do not fully account for enhanced upper extremity bone mass and strength in gymnasts. J Musculoskelet Neuronal Interact 2009;9:2–14.

72 Daly RM, Saxon L, Turner CH, Robling AG, Bass SL: The relationship between muscle size and bone geometry during growth and in response to exercise. Bone 2004;34:281–287.

73 Bailey DA, Martin AD: Physical activity and skeletal health in adolescents. Pediatr Exerc Sci 1996;6:330–347.

74 Tanner JM: Fallacy of per-weight and per-surface area standards and their relation to spurious correlation. J Appl Physiol 1949;2:1–15.

75 Guenette A, Sheel AW: Exercise-induced arterial hypoxaemia in active young women. Appl Physiol Nutr 2007;32: 1263–1273.

76 Mota J, Guerra S, Leandro C, Pinto A, Ribeiro JC, Duarte JA: Association of maturation, sex, and body fat in cardiorespiratory fitness. Am J Human Biol 2002;14:707–712.

77 Armstrong N, Welsman JR: Aerobic fitness; in Armstrong N, van Mechelen W (eds): Paediatric Exercise Science and Medicine, ed 2. Oxford, Oxford University Press, 2008, pp 97–108.

78 Nevill AM, Bate S, Holder RL: Modeling physiological and anthropometric variables known to vary with body size and other confounding variables. Year Book Phys Anthropol 2005;48:141–153.

79 Tolfrey K, Barker A, Thom JM, Morse CI, Narici MV, Batterham AM: Scaling of maximal oxygen uptake by lower leg muscle volume in boys and men. J Appl Physiol 2006;100:1851–1856.

80 Rowland T, Goff D, Martel L, Ferrone L, Kline G: Normalization of maximal cardiovascular variables for body size in premenarchal girls. Pediatr Cardiol 2000;21:429–432.

81 Armstrong N, McManus AM, Welsman JR: Aerobic fitness; in Armstrong N, van Mechelen W (eds): Paediatric Exercise Science and Medicine, ed 2. Oxford, Oxford University Press, 2008, pp 269–282.

82 Armstrong N, Welsman JR: Young People and Physical Activity. Oxford, Oxford University Press, 1997.

83 Armstrong N, Welsman JR: Peak oxygen uptake in relation to growth and maturation in 11- to 17-year-old humans. Eur J Appl Physiol 2001;85:546–551.

84 Eisenmann JC, Pivarnik JM, Malina RM: Scaling peak $\dot{V}O_2$ to body mass in young male and female distance runners. J Appl Physiol 2001;90:2172–2180.

85 Baxter-Jones A, Goldstein H, Helms P: The development of aerobic power in young athletes. J Appl Physiol 1993;75: 1160–1167.

86 Welsman JR, Armstrong N: Longitudinal changes in submaximal oxygen uptake in 11- to 13-year-olds. J Sports Sci 2000; 18:183–189.

87 Courteix D, Obert P, Lecoq AM, Guenon P, Kock G: Effect of intensive swimming training on lung volumes, airway resistances and on the maximal expiratory flow-volume relationship in prepubertal girls. Eur J Appl Physiol 1997;76:264–269.

88 Sheel AW, Richards JC, Foster GE, Guenette JA: Sex differences in respiratory exercise physiology. Sports Med 2004;34:567–579.

89 Boezen HM, Jansen DF, Postma DS: Sex and gender differences in lung development and their clinical significance. Clin Chest Med 2004;25:237–245.

90 Harms CA: Does gender affect pulmonary function and exercise capacity? Respir Physiol Neurobiol 2006;151:124–131.

91 Guenette A, Sheel AW: Exercise-induced arterial hypoxaemia in active young women. Appl Physiol Nutr 2007;32:1263–1273.

92 Olfert IM, Balouch J, Kleinsasser A, Knapp A, Wagner H, Wagner PD, Hopkins SR: Does gender affect human pulmonary gas exchange during exercise? J Physiol 2004;557:529–541.

93 Fawkner SG, Armstrong N: Pulmonary function; in Armstrong N, van Mechelen W (eds): Paediatric Exercise Science and Medicine, ed 2. Oxford, Oxford University Press, 2008, pp 243–254.

94 Armstrong N, Kirby BJ, McManus AM, Welsman JR: Prepubsecents' ventilatory responses to exercise with reference to sex and body size. Chest 1997;112:1554–1560.

95 Rowland TW, Cunningham LN: Development of ventilatory responses to exercise in normal white children. Chest 1997;111:327–332.

96 Hopkins SR, Harms CA: Gender and pulmonary gas exchange during exercise. Exerc Sport Sci Rev 2004;32:50–56.

97 Swain KE, Rosenkranz SK, Beckman B, Harms CA: Expiratory flow limitation during exercise in prepubescent boys and girls: prevalence and implications. J Appl Physiol 2010;108:1267–1274.

98 Laursen PB, Tsang GCK, Smith GJ, van Velzen MV, Ignatova BB, Sprules EB, Chu KS, Coutts KD, McKenzie DC: Incidence of exercise-induced arterial hypoxemia in prepubescent females. Pediatr Pulmonol 2002;34:37–41.

99 Guerro L, Naranjo J, Carranza MD: Influence of gender on ventilatory efficiency during exercise in young children. J Sports Sci 2008;26:1455–1457.

100 Karlberg P, Lind J: Studies of the total amount of hemoglobin and the blood volume in children. 1. Determination of total hemoglobin and blood volume in normal children. Acta Paediatr Scand 1955;44:17–34.

101 Sukarochana K, Parenzan L, Thakurdas N, Kiesewetter WB: Red cell mass determinations in infancy and childhood, with the use of radioactive chromium. J Pediatr 1961;59:903–908.

102 Linderkamp O, Versmold HT, Riegel KP, Betke K: Estimation and prediction of blood volume in infants and children. Eur J Pediatr 1977;125:227–234.

103 Raes A, Van Aken S, Craen M, Donckerwolcke R, Walle JV: A reference frame for blood volume in children and adolescents. BMC Pediatrics 2006;6:3.

104 Dallman PR, Siimes MA: Percentile curves for hemoglobin and red cell volume in infancy and childhood. Pediatrics 1979;94:26–31.

105 Boyadjiev N, Taralov Z: Red blood cell variables in highly trained pubescent athletes: a comparative analysis. Br J Sports Med 2000;34:200–204.

106 Rowland TW: Developmental Exercise Physiology. Champaign, Human Kinetics, 1996, pp 117–140.

107 Larsen JA, Kadish AH: Effects of gender on cardiac arrhythmias. J Cardiovasc Electrophysiol 1998;9:655–664.

108 Wirth A, Trager E, Scheele K, Mayer D, Diehm K, Reischle K, Weicker H: Cardiopulmonary adjustment and metabolic response to maximal and submaximal physical exercise of boys and girls at different stages of maturity. Eur J Appl Physiol 1978;39:29–39.

109 Telford RD, McDonal IG, Ellis LB, Chennells MH, Sandstrom ER, Fuller PJ: Echocardiographic dimensions in trained and untrained 12-year-old-boys and girls. J Sports Sci 1988;6:49–57.

110 Thoren CAR, Asano K: Functional capacity and cardiac function in 10-year-old-boys and girls with high and low running performance; in Ilmarinen J, Valimaki I (eds): Children and Sport: Pediatric Work Physiology. Berlin, Springer, 1984 pp 182–188.

111 Cumming GR: Hemodynamics of supine bicycle exercise in 'normal' children. Am Heart J 1977;93:617–622.

112 Rowland TW, Goff D, Martel L, Ferrone L: Influence of cardiac functional capacity on gender differences in maximal oxygen uptake in children. Chest 2000;117:629–635.

113 Obert P, Mandigout S, Nottin S, Vinet A, N'Guyen L, Lecoq A: Cardiovascular responses to endurance training in children: effect of gender. Eur J Clin Invest 2003;33:199–208.

114 Winsley RJ, Fulfold J, Roberts AC, Welsman JR, Armstrong N: Sex differences in peak oxygen uptake in prepubertal children. J Sci Med Sport 2009;12:647–651.

115 Gutin B, Owens S, Trieber F, Mensah G: Exercise hemodynamics and left ventricular parameters in children; in Armstrong N, Kirby B, Welsman J (eds): Children and Exercise. Part XIX. London, E & FN Spon, 1997, pp 460–464.

116 Osika W, Dangardt F, Montgomery SM, Volkmann R, Li MG, Friberg P: Sex differences in peripheral artery intima, media and intima media thickness in children and adolescents. Atherosclerosis 2009;203:172–177.

117 Gavin KM, Seals DR, Silver AE, Moreau KL: Vascular endothelial estrogen receptor alpha is modulated by estrogen status and related to endothelial function and endothelial nitric oxide synthase in healthy women. J Clin Endocrinol Metab 2009;94:3513–3520.

118 Nadland IH, Walloc L, Toska K: Effect of the leg muscle pump on the rise in muscle perfusion during muscle work in humans. Eur J Appl Physiol 2009;105:829–841.

119 Rowland T, Lisowski R: Hemodynamic responses to increasing cycle cadence in 11-year old boys: role of the skeletal muscle pump. Int J Sports Med 2001;22:405–409.

120 Armstrong N, Welsman JR: Cardiovascular responses to submaximal treadmill running in 11 to 13 year olds. Acta Paediatr 2002;91:125–131.

121 Constantin-Theodosiu D, Greenhaff PL, McIntyre DB, Round JM, Jones DA: Anaerobic energy production in human skeletal muscle in intense contraction: a comparison of ^{31}P magnetic resonance spectroscopy and biochemical techniques. Exp Physiol 1997;82:592–601.

122 Barker A, Welsman J, Welford D, Fulford J, Williams C, Armstrong N: Reliability of ^{31}P magnetic resonance spectroscopy during an exhaustive exercise test in children. Eur J Appl Physiol 2006;98:556–565.

123 Armstrong N, Fawkner SG: Exercise metabolism; in Armstrong N, van Mechelen W (eds): Paediatric Exercise Science and Medicine, ed 2. Oxford, Oxford University Press, 2008, pp 213–226.

124 Jones AM, Poole DC: Oxygen uptake dynamics: from muscle to mouth-an introduction to the symposium. Med Sci Sports Exerc 2005;37:1542–1550.

125 Barker AR, Welsman JR, Fulford J, Welford D, Williams CA, Armstrong N: Muscle phosphocreatine and pulmonary oxygen uptake kinetics in children at the onset and offset of moderate intensity exercise. Eur J Appl Physiol 2008;102:727–738.

126 Fawkner SG, Armstrong N: Oxygen uptake kinetic response to exercise in children. Sports Med 2003;33:651–669.

127 Fawkner SG, Armstrong N, Potter CR, Welsman JR: Oxygen uptake kinetics in children and adults after the onset of moderate-intensity exercise. J Sports Sci 2002;20:319–326.

128 Barker AR, Welsman JR, Fulford J, Welford D, Armstrong N: Muscle phosphocreatine kinetics in children and adults at the onset and offset of moderate-intensity exercise. J Appl Physiol 2008;105:446–456.

129 Fawkner SG, Armstrong N: Longitudinal changes in the kinetic response to heavy intensity exercise. J Appl Physiol 2004; 97:460–466.

130 Barker AR, Welsman JR, Fulford J, Welford D, Armstrong N: Quadriceps muscle energetic during incremental exercise in children and adults. Med Sci Sports Exerc 2011;42:1303–1313.

131 Rowland TW: Circulatory responses to exercise: are we misleading Fick? Chest 2005;127:1023–1030.

132 Armstrong N, Welsman JR, Williams CA, Kirby BJ: Longitudinal changes in young people's short-term power output. Med Sci Sports Exerc 2000;32:1140–1145.

133 Thorland WG, Johnson GO, Cisar CJ, Housh TJ, Tharp GD: Strength and anaerobic responses of elite young female sprint and distance runners. Med Sci Sports Exerc 1987;19:56–61.

134 Blimkie CJ, Roche P, Bar-Or O: The anaerobic power ratio in adolescent boys and girls; in Rutenfranz J, Mocellin R, Klimt F (eds): Children and Exercise XII. International Series on Sport Science. Champaign, Human Kinetics, 1986, pp 31–37.

135 Palgi Y, Gutin B, Young J, Alejandro D: Physiological and anthropometric factors underlying endurance performance in children. Int J Sports Med 1984;5:67–73.

136 Bar-Or O: Pediatric Sports Medicine for the Practitioner (Comprehensive Manual in Pediatrics). New York, Springer, 1983.

137 Doré E, Bedu M, França NM, Van Praagh E: Anaerobic cycling performance characteristics in prepubescent, adolescent and young adult females. Eur J Appl Physiol 2001;84:476–481.

138 Aucouturier J, Lazaar N, Doré E, Meyer M, Ratel S, Duché P: Cycling peak power in obese and lean 6–8 year-old girls and boys. Appl Physiol Nutr Metab 2007;32: 367–371.

139 Van Praagh E: Development of anaerobic function during childhood and adolescence. Pediatr Exerc Sci 1990;2:336–348.

140 Williams CA, Armstrong N: Optimised peak power output of adolescent children during maximal sprint pedalling; in Ring FJ (ed): Children in Sport. Proc 1st Bath Sports Medicine Conference. Bath, University of Bath, 1995, pp 40–44.

141 Bencke J, Damsgaard R, Saekmose A, Jørgensen P, Jørgensen K, Klausen K: Anaerobic power and muscle strength characteristics of 11 years old elite and non-elite boys and girls from gymnastics, team handball, tennis and swimming. Scand J Med Sci Sports 2002;12: 171–178.

142 Martin RJF, Doré E, Twisk J, Van Praagh E, Hautier CA, Bedu M: Longitudinal changes of maximal short-term peak power in girls and boys during growth. Med Sci Sports Exerc 2004;36:498–503.

143 De Ste Croix MBA, Armstrong N, Chia MYH, Welsman JR, Parsons G, Sharpe P: Changes in short-term power output in 10- to 12-year-olds. J Sport Sci 2001;19: 141–148.

144 Peterson SR, Gaul CA, Stanton MM, Hanstock CC: Skeletal muscle metabolism during short-term, high-intensity exercise in prepubertal and pubertal girls. J Appl Physiol 1999;87:2151–2156.

145 Wilcocks RJ, Williams CA, Barker AR, Fulford J, Armstrong N: Age- and sex-related differences in muscle phosphocreatine and oxygenation kinetics during high-intensity exercise in adolescents and adults. NMR Biomed 2010; 3:569–577.

146 Erlandson MC, Sherar LB, Mirwald RL, Maffulli N, Baxter-Jones AD: Growth and maturation of adolescent female gymnasts, swimmers, and tennis players. Med Sci Sports Exerc 2008;40:34–42.

147 Hong F: Innocence lost: child athletes in China. Sport Society 2004;7:338–354.

148 Lang M: Surveillance and conformity in competitive youth swimming. Sport Educ Society 2010;15:19–37.

149 Caine D, Bass S, Daly R: Does elite competition inhibit growth and delay maturation in some gymnasts? Quite possibly. Pediatr Exerc Sci 2003;15:360–372.

150 Klentrou P, Plyley M: Onset of puberty, menstrual frequency, and body fat in elite rhythmic gymnasts compared with normal controls. Br J Sports Med 2003;37:490–494.

151 Bass S, Bradney G, Pearce E, Hendrich E, Inge K, Stuckey S, Lo SK, Seeman E: Short stature and delayed puberty in gymnasts: influence of selection bias on leg length and the duration of training on trunk length. J Pediatr 2000;1136: 149–155.

152 Georgopoulos N, Markou K, Theodoropoulou A, Paraskevopoulou P, Varaki L, Kazantzi Z, Leglise M, Vagenakis AG: Growth and pubertal development in elite female rhythmic gymnasts. J Clin Endocrinol Metab 1999;84:4525–4530.

153 Georgopoulos NA, Markou KB, Theodoropoulou A, Vagenakis GA, Benardot D, Leglise M, Dikmopoulos JCA, Vagenakis AP: Height velocity and skeletal maturation in elite female rhythmic gymnasts. J Clin Endocrinol Metab 2001;86:5159–5164.

154 Theodoropoulou A, Markou KB, Vagenakis GA, Benardot D, Leglise M, Kourounis G, Vagenakis AP, Georgopoulos NA: Delayed but normally progressed puberty is more pronounced in artistic compared with rhythmic elite gymnasts due to the intensity of training. J Clin Endocrinol Metab 2005; 90:6022–6027.

155 Baxter-Jones ADG, Helms PJ: Effects of training at a young age: a review of the Training of Young Athletes (TOYA) Study. Pediatr Exerc Sci 1996;8:310–327.

156 Constantini NW, Warren MP: Menstrual dysfunction in swimmers: a distinct entity. J Clin Endocrinol Metab 1995;80: 2740–2744.

157 Loucks AB: Energy availability, not body fatness, regulates reproductive function in women. Exerc Sport Sci Rev 2003;31: 144–148.

158 Baxter-Jones ADG, Helms P, Baines-Preece J, Preece M: Menarche in intensively training gymnasts, swimmers and tennis players. Ann Hum Biol 1994;21: 407–415.

159 Otis CL, Drinkwater B, Johnson M, Loucks A, Wilmore J: American College of Sports Medicine position stand: the female athlete triad. Med Sci Sports Exerc 1997;29:i–ix.

160 Nattiv A, Loucks AB, Manore MM, Sanborn CF, Sundgot-Borgen J, Warren MP, American College of Sports Medicine: American College of Sports Medicine position stand: the female athlete triad. Med Sci Sports Exerc 2007;39: 1867–1882.

161 Sundgot-Borgen J, Torstveit MK: Prevalence of eating disorders in elite athletes is higher than in the general population. Clin J Sport Med 2004;14:25–32.

162 Abraham SF, Beumont PJ, Fraser IS, Llewellyn-Jones D: Body weight, exercise and menstrual status among ballet dancers in training. Br J Obstet Gynaecol 1982;89:507–510.

163 Chumlea WC, Schubert CM, Roche AF, Kulin HE, Lee PA, Himes JH, Sun SS: Age at menarche and racial comparisons in US girls. Pediatrics 2003;111:110–113.

164 Beals, KA, Meyer NL: Female athlete triad update. Clin Sports Med 2007;26:69–89.

165 Nichols JF, Rauh MJ, Lawson MJ, Ming JI, Barkai HS: Prevalence of the female athlete triad syndrome among high school athletes. Arch Pediatr Adoesc Med 2006;160:137–142.

166 Hoch AZ, Pajewski NM, Moraski L, Carrera GF, Wilson CR, Hoffmann RG, Schimke JE, Gutterman DD: Prevalence of the female athlete triad in high school athletes and sedentary students. Clin J Sport Med 2009;19:421–428.

167 Loucks AB, Heath EM: Induction of low T_3 syndrome in exercising women occurs at a threshold of energy availability. Am J Physiol 1994;226:R817–R823.

168 Loucks AB, Thurma JA: Luteinizing hormone pulsatility is disrupted at a threshold of energy availability in regularly menstruating women. J Clin Endocrinol Metab 2003;88:297–311.

169 Ihle R, Loucks AB: Dose-response relationships between energy availability and bone turnover in young exercising women. J Bone Miner Res 2004;19:1231–1240.

170 Loucks AB: Refutation of 'the myth of the female athlete triad'. Br J Sports Med 2007;41:55–57.

171 De Souza MJ, Vescovi JD, Willijmas NI, Van Heest JL, Warren MP: Correction of misinterpretation and misrepresentations of the female athlete triad. Br J Sports Med 2007;41:58–59.

172 Klungland-Torstveig M, Sundgot-Borgen J: The female athlete triad exists in both elite athletes and controls. Med Sci Sports Exerc 2005;37:1449–1459.

173 American Psychiatric Association: Eating disorders; in First M (ed): Diagnostic and Statistical Manual of Mental Disorders, ed 4. Washington, American Psychiatric Publishing, 1994 pp 539–550.

174 Practice Committee of the American Society for Reproductive Medicine: Current evaluation of amenorrhea. Fertil Steril 2004;82:266–272.

175 Bianchi ML, Baim S, Bishop NJ, Gordon CM, Hans DB, Langman CB, Leonard MB, Kalkwarf HJ: Official positions of the International Society for Clinical Densitometry (ISCD) on DXA evaluation in children and adolescents. Pediatr Nephrol 2010;25:37–47.

Dr. Alison M. McManus
Institute of Human Performance, University of Hong Kong
Hong Kong, SAR (China)
Tel. +852 2589 0582, Fax +852 2855 1712, E-Mail alimac@hku.hk

Am J Physiol Heart Circ Physiol 306: H781–H788, 2014.
First published January 10, 2014; doi:10.1152/ajpheart.00994.2013.

CALL FOR PAPERS | *Sex and Gender Differences in Cardiovascular Physiology–Back to the Basics*

Why are sex and gender important to basic physiology and translational and individualized medicine?

Virginia M. Miller

Departments of Surgery, Physiology, and Biomedical Engineering, Mayo Clinic, Rochester, Minnesota

Submitted 18 December 2013; accepted in final form 7 January 2014

Miller VM. Why are sex and gender important to basic physiology and translational and individualized medicine?. *Am J Physiol Heart Circ Physiol* 306: H781–H788, 2014. First published January 10, 2014; doi:10.1152/ajpheart.00994.2013.— Sex refers to biological differences between men and women. Although sex is a fundamental aspect of human physiology that splits the population in two approximately equal halves, this essential biological variable is rarely considered in the design of basic physiological studies, in translating findings from basic science to clinical research, or in developing personalized medical strategies. Contrary to sex, gender refers to social and cultural factors related to being a man or a woman in a particular historical and cultural context. Unfortunately, gender is often used incorrectly by scientists and clinical investigators as synonymous with sex. This article clarifies the definition of sex and gender and reviews evidence showing how sex and gender interact with each other to influence etiology, presentation of disease, and treatment outcomes. In addition, strategies to improve the inclusion of female and male human beings in preclinical and clinical studies will be presented, and the importance of embedding concepts of sex and gender into postgraduate and medical curricula will be discussed. Also, provided is a list of resources for educators. In the history of medical concepts, physiologists have provided pivotal contributions to understanding health and disease processes. In the future, physiologists should provide the evidence for advancing personalized medicine and for reducing sex and gender disparities in health care.

behavior; chromosomes; health disparities; hormones; personality

SINCE THE FATHER OF modern physiology, William Harvey, made the landmark observations in 1628 that the heart pumped blood and the blood circulated (23), physiologists have contributed fundamental and critical information for the development of modern medicine. However, physiologists like other scientists have been influenced by several important cultural trends. Possibly because science was traditionally a male-dominated profession, except for the physiology of reproduction, most human physiological studies were focused on and conducted on male human beings. Thus physiological principles contained in classical textbooks and medical curricula were based on the 70-kg healthy male or on male animals. In the United States, the opposition to this restriction of research to "male" bodies resulted in legislation, i.e., the National Institutes of Health (NIH) Revitalization Act of 1993. This law mandated the inclusion of women in research involving humans that was supported by the NIH. Other trends that have influenced our understanding of physiological principles include the rapid expansion of molecular mechanistic studies and the political pressure to reduce the use of whole animals in basic and preclinical experiments. Thus studies using isolated tissues and cultured cells (including cell lines) gained popularity and, unfortunately, without attention to the biology or phenotypic characteristics of the tissue/cell donor with the assumption that the sex of the experimental material was irrelevant. But is it?

Sex is the basic biological variable that distinguishes approximately half of the population from the other. The landmark Institute of Medicine report "Exploring the Biological Contribution of Sex" concluded that sex matters in all aspects of cellular function and physiology from "womb to tomb" (77). One must wonder, then, why is sex of experimental material so often ignored in an era of genomics and personalized medicine?

Epidemiological studies have consistently identified differences in disease incidence and prevalence between men and women. Patient health advocacy groups such as the American Heart Association, American Cancer Society, and the American Lung Association, to name a few, mount targeted campaigns to educate health care providers of sex differences in symptoms, outcomes, and mortality of specific diseases. Moreover, gender, a term that is often used incorrectly by scientists

Address for reprint requests and other correspondence: V. M. Miller, Dept. of Surgery, Medical Science 4-62, Mayo Clinic, 200 First St. SW, Rochester, MN 55905 (e-mail: miller.virginia@mayo.edu).

and clinical investigators as synonymous with sex, is considered an essential topic in medical curricula related to cultural sensitivity in health care delivery to reduce health disparities among ethnic, racial, and socioeconomic groups. What, then, is the appropriate way to consider sex and gender as research variables and in medical curricula? And why are they important to basic physiology and to translational and individualized medicine? This article will address these questions by clarifying the definitions of sex and gender and by reviewing evidence showing how sex and gender interact to influence etiology, presentation, and treatment outcomes of diseases. In addition, strategies to improve the inclusion of female and males in preclinical and clinical studies will be presented, and discussion will include the importance of embedding concepts of sex and gender into postgraduate and medical curricula with a goal to improve the health of both women and men and thus reduce disparities in health care.

Sex and Gender: Are We Talking About the Same Thing?

The Institute of Medicine defines sex as "being male or female according to reproductive organs and the functions assigned by chromosomal complement (XX for female and XY for male)" (77). That is, sex is biology. There are genetic variants in sex chromosomes, but these are rare and are not considered in this review. For example, XXY (or Klinefelter's syndrome) occurs in about 1 of 580 live male births, X monosomy (or Turner's syndrome) occurs in about 1 of 5,000 live births, XXX occurs in about 1 of 1,000 births, and XYY occurs in about 1 in 1,000 male births.

Every cell has a sex, which is determined by the presence of the complement of sex chromosomes. The nongonadal functions directed by the sex chromosomes are critical to the physiology of the organism (53, 70a) and include such important functions as coagulation (65), innate immunity (10, 12), synthesis of norepinephrine (4, 17, 49), androgen sensitivity (36), energy metabolism (adiposity) (72), blood pressure (29), and apoptosis (32). These nongonadal effects of the sex chromosomes coupled with gonadal effects through genomic actions of the sex steroids result in sex differences in gene expression in every cell (28, 31, 51, 56, 66, 76, 78, 79). Thus the sex of cells in culture, in isolated tissues, and in whole animals cannot be ignored in the design of experiments.

Whereas sex is biology, as defined by the Institute of Medicine, gender is everything else (77), including psychosocial and cultural factors. According to Ristvedt (62), the term "gender" evolved over time being originally used in linguistics. In linguistics, gender was used to designate words as masculine, feminine, or neutral. This concept was exported into "sexual science" where it was used to define masculinity/femininity (46) and then psychoanalytical writers used it to encompass "socially constructed" male/female differences (62). Thus sex and gender became incorrectly synonymous. In the 1990s, even the author of this review published papers addressing "gender differences" in vascular function when in reality the studies were about "sex differences" (1, 6–8, 43, 75). Fortunately, the Institute of Medicine report clarified the definitions, and there are steady and consistent efforts to adopt and apply the term "sex" to biological factors and "gender" to psychosocial and cultural factors.

Interaction of Sex and Gender in Translational Science

Sex differences in physiology and pathophysiology can be classified in three general categories. First are conditions or diseases unique to one sex, such as conditions associated with reproduction. Second are conditions or diseases that have greater prevalence in one sex compared with the other. Third are conditions that have different age of onset, symptomatology, or response to treatment in one sex compared with the other. Examples of conditions related to cardiovascular diseases in each of these categories are shown in Table 1. Other areas of physiology can generate similar lists that encompass diseases of the respiratory system, musculoskeletal system, immunological system, gastrointestinal system, renal/urological system, development, and behavior.

Sex is a basic variable in every cellular and integrated physiological experiment. Because sex is based on two distinct chromosomal configurations, analysis of data to determine sex differences should treat sex as a dichotomous variable.

Although gender is related to sex, gender is not a dichotomous variable even though it is often considered as such when data are collected from self-declaration of gender as male and female. Gender defines behavioral, psychological, and cultural characteristics that are expressed on a continuum. For example, masculine and feminine behaviors may be defined by sociocultural expectations (37). What may be considered a neutral behavior in one culture, for example, driving or smoking, may be considered a masculine behavior in another culture. When behaviors are classified as aggressive or stoic for males compared with nurturing or expressive for females, males and females exhibit a range of scores that are distinct but overlap, i.e., a continuous variable (62). Thus analysis of outcome data by self-reported gender alone may not provide all of the information needed to draw conclusions about behaviors that could affect disease risk factors, treatment efficacy, or outcomes if self -reporting of gender differs from the biological sex of the individual.

Table 1. *Examples of cardiovascular pathophysiological conditions showing sex differences*

	Female	Male
Conditions unique to one sex	Hypertensive disorders of pregnancy Vasomotor symptoms of menopause	Erectile dysfunction
Conditions with sex-specific prevalence	Pulmonary hypertension	Myocarditis
	Raynaud's disease Migraine Postural orthostatic tachycardia syndrome	
	Heart failure with preserved ejection fraction	Heart failure with reduced ejection fraction
	Ventricular apical ballooning (Tako Tsubo) Microvascular angina	
Conditions with sex-specific onset, symptoms, and response to treatment	Hypertension, myocardial infarction, stroke	

The question then becomes, how does gender interact with sex to influence health and disease? Both sex and behaviors influenced by gender will affect physiology and pathophysiology. The relative contribution and interaction between biological sex and external sociocultural influences of gender is an ongoing topic of scientific inquiry. Sophisticated imaging modalities coupled with computational analysis showed sex differences in distinct connectivity between cerebral hemispheres and cerebellum of adolescent males and females of various ethnicities from the Philadelphia region of the United States (27). These data provide additional support for the hypothesis that conceptual and motor behaviors develop at an early age and reflect underlying differences in biology of the sexes. This conclusion should not be surprising given that every cell has a sex and cellular function that is influenced by the genome and sex steroid milieu to which the cell is exposed (2, 3).

For physiological studies using material derived from experimental animals, the issue of "gender" is mostly irrelevant, except for specific studies of behavior (22), as the environment in which experimental animals are bred and raised is controlled. However, gender will impact results of experiments in which tissues and cells are derived from humans, clinical human studies, as well as those studies of health disparities and health outcomes in humans. For example, risk of death from cardiovascular disease was lower in men with a high "femininity" behavioral score, compared with men with a high "masculinity" behavioral score (25). Some risk factors associated with gendered behaviors, such as smoking history and alcohol consumption or occupations, are captured on inclusion criteria or evaluations for clinical studies or in the medical record by a quantitative scale such as pack years of smoking. Gene expression and cellular regulatory pathways in isolated cells obtained from such individuals may reflect epigenetic changes brought about by lifelong tobacco use or exposure or exposure to xenobiotics or other environmental pollutants (herbicides, industrial and automotive smog, radiation, industrial or occupational-related chemicals, e.g., organic solvents, dyes, aerosols, asbestos). For cells and tissues obtained from commercial sources and biobanks, such information may not be available. However, at a minimum it is critical for the accurate interpretation of experiments to acknowledge that the sex of the cells or tissues may matter and could influence the results and to consider the hormonal status of the tissue donor which if unknown, could be surmised in part by age of the donor (sexually immature child, adolescent, adult of reproductive age, or senior).

In addition to sex differences in potential exposure to disease risk factors, differences in sociocultural attitudes such as ethnicity, occupation, marital/partner status, income, years of education, etc., may affect access to medical care and compliance with treatment. Furthermore, the type of cardiovascular pathology, symptoms (Table 1) and response to treatment including response to therapies (48), i.e., aspirin (61) and statins (21) and to modifiable cardiovascular risk factors such as smoking cessation, diet, and exercise are influenced by the underlying biology, i.e., sex (13, 50, 55, 60). In an era where health outcomes will be assessed from "big data" sets, it may be appropriate to exclude the category of gender and to include the category of sex together with more specific criteria that capture continuous and quantifiable variables constituting behavioral and cultural aspects of gender that are considered risk factors for disease. In Europe and Canada, the term gender-based medicine or gender differences is used more consistently to define medical conditions and health outcomes than the term sex-based medicine. Neither of these terms sex-based or gender-based medicine incorporates the totality of parameters influencing an individual's health status. However, their use represents a critical first step in acknowledging and quantifying parameters that influence health and movement toward embedding concepts of sex and gender into all aspects of health care, i.e., "individualized medicine" that encompasses both sex and gender.

Sex, Gender, and Policy in Scientific Publication and Research Funding

As mentioned in the *Introduction*, the NIH Revitalization Act of 1993 mandated the inclusion of women in studies of humans. However, the number of women included in drug and device trials remains low (14, 15, 19, 26). The reasons for low participation of women are many (38), including that some trials investigate conditions that are more common in men, for example heart failure with reduced ejection fraction. However, even in studies of conditions occurring with similar prevalence in women and men, participation by women ranged from 25–45% (40). Some studies inadvertently exclude women because of restrictive inclusion criteria based on parameters that have a sex differential, such as size or normative blood parameters (16, 59) or that women are simply not asked to participate (57).

The percentage of women enrolled in a trial may affect interpretation of the results. For example, the incidence of type II diabetes with the use of statin therapy was found to be proportional to the percentage of women enrolled in the trial: in a woman-only trial, the risk was 42%, whereas the risk was negligible (14%) in trials enrolling only men (21). Thus, to individualize treatment options, a discussion between the health care provider and patient should include discussion of risk versus benefits that are specific to the sex of the individual. Therefore, it is critical that data from mixed sex trials be analyzed by sex and those results reported even if they are negative. It is just as critical to identify under what conditions, systems, responses, mechanistic pathways, etc., are the same between the sexes as to identify those which are not. However, reporting results by sex in clinical trials remains low (27a). It is critical that data be reported by sex rather than only accounting for sex as a confounding variable so that sex-specific outcomes from multiple studies could be evaluated by meta-analysis. Such sex-specific analyses form the basis for evidence-based medical decisions and are necessary to optimize individualized treatment strategies for men and women (48).

The inability to reproduce many basic science experiments and to translate those results to the clinical arena is a concern (74). Solutions proposed to improve reproducibility of results have focused on statistical issues in experimental design and analysis (go.nature.com/oloeip) and the lack of consistent quality systems such as reagents and assays used in different laboratories: http://gbsi.org/sites/default/files/uploads/pdf/the-_case_for_standards.pdf (77a). Methods sections of papers should provide sufficient information, including the sex of the experimental animal, so that another investigator can reproduce the experiment. Reporting of sex in basic science studies is abysmal and in many cases the sex of the material is not

Table 2. *Questions regarding sex and gender included in grant applications submitted to the Canadian Institutes of Research*

1. Are sex (biological) considerations taken into account in this study? (Y/N)
2. Are gender (socio-cultural) considerations taken into account in this study? (Y/N)
3. If YES, please describe how sex and/or gender considerations will be considered in your research design. (maximum of 2,000 characters)
4. If NO, please explain why sex and/or gender are not applicable in your research design. (maximum of 2,000 characters)

reported at all (66, 70, 81). Given the lack of reporting of the sex of experimental material, an additional consideration would be for more rigorous monitoring and requirements for reporting of sex of the biological material. Gene expression and cellular processes vary by sex as discussed above, thus analyzing data from mixed groups of males and females in human studies or mixed groups of male and female cells or genetically manipulated mice may mask a treatment effect if the particular response is upregulated or positive in one sex and downregulated or negative in the other. As with the reported results with side effects of statin medications (21), other results derived from mixed populations of males and females may be influenced by the relative proportion of each sex used in the experiment.

Studying cellular mechanistic pathways in only one sex and assuming that the same regulatory pathways apply to the other sex is common in the scientific literature. Some examples include pathways involved in vascular remodeling, atherosclerosis, oxidative stress, and immune function (11, 18, 20, 30, 33–35, 39, 45, 52, 58, 64, 67–69, 71, 73, 80). However, such assumptions are not appropriate given the present knowledge of sex differences in prevalence, symptomatology, and disease outcomes. Use of nondescripted nouns such as “mice,” “humans,” “cultured endothelial cells,” etc., does not provide the same information as “female mice,” “endothelial cells derived from a 65-yr-old male,” etc. Specification of sex may affect the interpretation and translation of results derived from basic science studies that are needed to design clinical trials. Several professional societies including the American Physiological Society (41) and the Endocrine Society (9) (see http://genderedinnovations.stanford.edu for an up-to-date list of editorial policies for other journals) have editorial policies requiring reporting of the sex of the experimental material. Although the main onus is on the authors, peer reviewers, Associate Editors, and Editors share responsibility to ensure that policies are enforced.

To reduce disparities in health outcomes between men and women, it is essential for scientists and clinicians to consider sex differences as one of the underlying physiological mechanisms of disease. More research is needed to identify regula-

Table 3. *Specialized centers of research on sex differences from 2002–2016*

Theme	Institution	Years
Genes, hormones, and environment		
Genes, androgens, and intrauterine environment in polycystic ovarian syndrome	Northwestern University	2002–2007
Excess male hormones (androgens) as the key to explaining polycystic ovarian syndrome (PCOS)	Northwestern University	2007–2011
Genes, androgens, and intrauterine environment in PCOS	Northwestern University	2012–2016
Identifying the genes that put women at risk for osteoporosis	University of Missouri, Kansas City	2007–2011
Genetic and environmental origins of adverse pregnancy outcomes	University of Pittsburgh	2002–2007
Substance Abuse/Pain		
Sex differences in pain sensitivity	University of Maryland	2002–2007
Sex and gender influences on addiction and health: a developmental perspective	University of Miami	2007–2011
Role of sex and gender differences in substance abuse relapse	Medical University of South Carolina	2002–2011
Sex and gender differences in addictions and stress response	Medical University of South Carolina	2012–2016
Sex, stress, and cocaine addiction	Yale University	2002–2007
Sex, stress, and substance use disorders	Yale University	2007–2011
Gender-sensitive treatment for tobacco dependence	Yale University	2012–2016
Sex and gender factors in the pathophysiology of irritable bowel syndrome and interstitial cystitis	University of California, Los Angeles	2002–2007
A coordinated study of stress, pain, emotion, and sexual factors underlying the pelvic visceral disorders of irritable bowel disorder and interstitial cystitis	University of California, Los Angeles	2007–2011
Center for Neurovisceral Sciences and Women’s Health (sex differences in pain)	University of California, Los Angeles	2012–2016
Sex differences and progesterone effects on impulsivity, smoking, and cocaine stress	University of Minnesota	2012–2016
Pharmacology		
Pharmacology of antiepileptic and psychotropic medications during pregnancy and lactation	Emory University	2002–2007
Mechanisms by which drug transporters alter maternal and fetal drug exposure during pregnancy	University of Washington	2002–2007
Urology/Gynecology		
Birth, muscle injury, and pelvic floor dysfunction	University of Michigan, Ann Arbor	2002–2016
Mechanisms underlying female urinary incontinence	University of California, San Francisco	2002–2007
Lower urinary tract function in women	University of California, San Francisco	2007–2011
Molecular and epidemiologic basis of acute and recurrent urinary tract infections in women	Washington University	2002–2016
Metabolism		
Sex steroids, sleep, and metabolic dysfunction in women	University of Chicago	2007–2011
Metabolic consequences of loss of gonadal function	University of Colorado	2012–2016
Musculoskeletal		
Sex differences in musculoskeletal conditions across the life span	University of California, Davis	2012–2016
Brain/Mental Health and Cardiovascular		
Fetal antecedents to sex differences in depression: a translational approach	Brigham and Women’s Hospital	2007–2011
Prepubertal stress, windows of risk and sex bias for affective disturbance	University of Pennsylvania	2012–2016
Sex-specific risk for vascular dysfunction and cognitive decline	Mayo Clinic	2012–2016

tory pathways in cells of female origin and to understand differences in integrated physiological control mechanisms in male and female experimental animals. Granting agencies, such as NIH, should consider improved mechanisms for investigators to more completely address issues of sex and gender in their grant applications. For example, the Canadian Institutes of Health Research (CIRH) requires grant applicants (http://www.cihr-irsc.gc.ca/e/32019.html) to respond to specific questions about sex and gender in research (Table 2). Currently, there are no requirements to include or address rationale for use of material from only one sex in applications funded by the NIH. Although identification of the number and sex of animals used in basic science studies and other conditions relative to their husbandry are required in the "Vertebrate Animal" section of NIH applications, it is not required to discuss sex and gender relative to outcomes of long-term goals of the research plan or to the scientific validity of choice of the experimental material. For example, would it be scientifically appropriate to develop an animal model of pulmonary hypertension only in male animals when in humans, pulmonary hypertension is more prevalent in females than males.

A typical reason cited for not including both male and female animals in preclinical and mechanistic basic studies is cost. However, the cost of identification of sex differences in discovery experiments may be minimal compared with the cost and time wasted by developing an experimental model or therapeutic approach based on one sex that does not reflect disease expression in humans, fails clinical testing, or has to be withdrawn from the market due to adverse events in a mixed study population of men and women. For example, 8 out of 10 drugs withdrawn from the market between 1997 and 2001 were due to adverse events in women (24). The first medications intended for use in both sexes to have different dosage recommendations for women and men came in May 2013 for zolpidem-containing drugs. The Federal Drug Administration approved labeling changes lowering the recommended dosage of these drugs for women due to increased side effects in women. One proposal to eliminate sex discrimination in basic research for conditions that are not sex specific would be to adopt principles of Title IX directed to eliminate sex discrimination in education to biomedical research (63).

Scientific review panels for funding agencies usually do not respond favorably toward studies that explore sex differences, classifying such studies as "descriptive." There is some limited attention to this problem by the federal government. The Office of Research on Women's Health (ORWH) developed and designed an innovative, interdisciplinary-targeted funding mechanism that integrates basic, clinical, and translational investigation into sex and gender differences: Specialized Centers of Research (P50 SCOR mechanism). The range of topics for the Centers was derived from three sources: the Institute of Medicine report "Exploring the Biological Contributions to Human Health, Does Sex Matter?"(77), the ORWH 2000 "Agenda for Research on Women's Health for the 21st Century," and the NIH Strategic Plan for Women's Health and Sex Differences. Since the inception of the program in 2002, only 33 awards have been made to 26 academic centers (Table 3). Requirements for the SCORs include a minimum of three

Table 4. *Professional resources*

NIH funded Specialized Centers of Research on Sex Differences
Available at: http://orwh.od.nih.gov/interdisciplinary/scor/index.asp
Textbooks:
· Legato M. *Principles of Gender-Specific Med*icine. Elsevier. 2011. 2nd ed.
· Oertelt-Prigione, Regitz-Zagrosek. *Clinical Aspects of Gender Specific Medicine*. Springer. 2012.
· Schenk-Gustafsson K, DeCola PR, Pfaff SW, Plsetsky DS. *Handbook of Clinical Gender Medicine*. Karger. 2012.
· Regitz-Zagrosek, V. ed. *Sex and Gender Differences in Pharmacology*. Springer-Verlag 2012.
· Mattison, DR. ed. *Clinical Pharmacology During Pregnancy*. Elsevier, 2013.
Web-based Continuing Medicine Education Courses:
· NIH ORWH *Sex and Gender Differences in Health and Behavior*. Available at: http://sexandgendercourse.od.nih.gov
· NIH ORWH *The Basic Science and the Biological Basis for Sex and Gender Differences*. Available at: http://sexandgendercourse.od.nih.gov
· TTUHSC Laura W. Bush Institute for Women's Health. *Y Does X Make A Difference*?. Available at: www.laurabushinstitute.org.
· Women's Health Info Site: *Sex and Gender Resource for Clinicians and Trainees*. Available at: http://whepducomogspot.com
· *National Association of Women's Health Medical Educators Faculty Guide* (NAWHME).
· Resource listing of various educational modalities to use in integration efforts. Available at:
http://www.drexelmed.edu/Home/OtherPrograms/WomensHealthEducationProgram/Resources.aspx
Web-based Research and Educational Resources:
· *Sex and Gender Women's Health Collaborative*. Available at: www.sgwhc.org
· *Stanford University's Gendered Innovations*. Available at: http://genderedinnovations.stanford.edu
· *Canadian Institute of Gender Health. What a Difference Sex and Gender Make*. Available at: http://www.cihr-irsc.gc.ca/e/44082.html.
Articles Outlining Experimental Design Methodology:
· Becker JB, Arnold AP, Berkley KJ, Blaustein JD, Eckel LA, Hampson E et al. Strategies and Methods for Research on Sex Differences in Brain and Behavior. *Endocrinology*. 2005;146:1650-73.
· Greenspan JD, Craft RM, LeResche L, Arendt-Nielsen L, Berkley KJ, Fillingim RB et al. Studying sex and gender differences in pain and analgesia: a consensus report. *Pain*. 2007;132:S26-S45.
· Miller VM, Kaplan JR, Schork NJ, Ouyang P, Berga SL, Wenger NK et al. Strategies and Methods to Study Sex Differences in Cardiovascular Structure and Function: A Guide for Basic Scientists. *Biol Sex Differ*. 2011;2:14. doi:10.1186/2042 −6410-2-14.
· Shah K, McCormack CE, A BN. Do you know the sex of your cells? *Am J Physiol Cell Physiol*. 2014;306:C3-C18.
Professional Membership Organizations:
· Organization for the Study of Sex Differences. Available at http://www.ossdweb.org
· International Society of Gender Medicine. Available at http://www.isogem.com
Journals
· Biology of Sex Differences. Available at http://www.bsd-journal.com
· Journal of Women's Health. Available at http://www.liebertpub.com

projects that encompass basic, clinical, and translational investigation.

ORWH is the major funder of the SCORs, but the ORWH is only a small fraction of the overall NIH budget. Although funding for SCORs is also provided by six National Institutes of Health Institutes and Centers and the Federal Drug Administration, the major portion of the NIH budget focuses predominantly on basic biomedical research in males. The SCORs represent initial steps in bringing attention to the need for research on sex differences in physiological processes and translational medicine. Clearly, there is a need for expansion of this program to reflect a broader range of diseases showing sex differences and for development of new funding sources directed at understanding basic and preclinical biological mechanisms for sex differences to advance translational, individualized medicine.

Embedding Concepts of Sex and Gender into Postgraduate and Medical Curricula

Evidenced-based medicine is built on results from basic, human, and clinical studies. Results of these studies need to be included in postgraduate and medical curricula to ensure development of sex-based evidence for individualized medical decision making. In a small case study conducted at Mayo Clinic, although fourth year medical students were aware of some sex differences in drug metabolism, they reported that such information was not translated into patient care (42). A consensus statement from thought-leaders in medical education agrees that concepts of sex and gender will best be embedded throughout the postgraduate training and not as discrete units (44). The challenge is to provide sufficient mechanistic data and outcome data in all aspects of training which is a work in progress. Educators are reluctant to begin a time and labor-intensive process of developing entirely new programs and materials. However, resources are beginning to accumulate that can be incorporated and customized to fit into existing programs (Table 4). Efforts to expand these resources are underway and should hasten curricula change.

Conclusion

Individualized medicine requires viewing the patient through a sex and gender lens as a first step toward personalizing care and potentially improving outcomes. However, the evidence upon which to base sex-specific decisions needs to be improved. The first step toward accomplishing this goal is for basic scientists to provide more data regarding mechanistic and regulatory processes which are similar and which differ between males and females. In addition, there is continued need for more scientific journal editors to institute editorial policies and for those who do, to enforce these policies that require reporting and analysis of data by sex and gender. Furthermore, increased support of sex difference research by funding agencies is needed, and funding agencies should develop requirements for inclusion of female animals in basic science and women in translational studies and clinical trials. Physiologists have consistently made pivotal contributions to understanding regulatory processes in health and disease. In the future, their contributions will continue to form the basis for advancing personalized medicine. As educators, they will pass on their discoveries to future basic and clinical researchers and health care providers by embedding concepts of sex and gender differences throughout revised curricula. When sex and gender are included as essentials for scientific excellence in research and education, health disparities between women and men will be reduced.

ACKNOWLEDGMENTS

V. M. Miller is immediate past president of the Organization for the Study of Sex Differences.

GRANTS

V. M. Miller's research program is funded by National Institute of Health Grants AG-44170, HD-65987, HL-83947, and HL-90639.

DISCLOSURES

My funding and association with the Organization for the Study of Sex Differences could be considered conflicts of interest since they are focused on sex differences. However, all are reported above, and I will leave it to the reader to determine whether three is a conflict. I do not receive personal financial benefit from any of these activities.

AUTHOR CONTRIBUTIONS

V.M.M. drafted, edited, revised, and approved final version of manuscript.

REFERENCES

1. **Antoniucci D, Miller VM, Sieck GC, Fitzpatrick LA.** Gender-related differences in proliferative responses of vascular smooth muscle cells to endothelin-1. *Endothelium* 8: 137–145, 2001.
2. **Arnold AP.** Sex chromosomes and brain gender. *Nat Rev Neurosci* 5: 701–708, 2004.
3. **Arnold AP, Burgoyne PS.** Are XX and XY brain cells intrinsically different? *Trends Endocrinol Metab* 15: 6–11, 2004.
4. **Arnold AP, Chen X.** What does the "four core genotypes" mouse model tell us about sex differences in the brain and other tissues? *Front Neuroendocrinol* 30: 1–9, 2009.
6. **Barber DA, Burnett JC Jr, Fitzpatrick LA, Sieck GC, Miller VM.** Gender and relaxation to C-Type natriuretic peptide in porcine coronary arteries. *J Cardiovasc Pharmacol* 32: 5–11, 1998.
7. **Barber DA, Burnett JC Jr, Miller VM.** Gender differences in relaxations evoked by C-Type natriuretic peptide in porcine coronary arteries. *Endothelium* 2: s2, 1995.
8. **Barber DA, Miller VM.** Gender differences in endothelium-dependent relaxations do not involve NO in porcine coronary arteries. *Am J Physiol Heart Circ Physiol* 273: H2325–H2332, 1997.
9. **Blaustein JD.** Animals have a sex, and so should titles and methods sections of articles in Endocrinology. *Endocrinology* 153: 2539–2540, 2012.
10. **Bloomer LD, Nelson CP, Eales J, Denniff M, Christofidou P, Debiec R, Moore J, Consortium C, Zukowska-Szczechowska E, Goodall AH, Thompson J, Samani NJ, Charchar FJ, Tomaszewski M.** Male-specific region of the Y chromosome and cardiovascular risk: phylogenetic analysis and gene expression studies. *Arterioscler Thromb Vasc Biol* 33: 1722–1727, 2013.
11. **Bulckaen H, Prevost G, Boulanger E, Robitaille G, Roquet V, Gaxatte C, Garcon G, Corman B, Gosset P, Shirali P, Creusy C, Puisieux F.** Low-dose aspirin prevents age-related endothelial dysfunction in a mouse model of physiological aging. *Am J Physiol Heart Circ Physiol* 294: H1562–H1570, 2008.
12. **Charchar FJ, Bloomer LD, Barnes TA, Cowley MJ, Nelson CP, Wang Y, Denniff M, Debiec R, Christofidou P, Nankervis S, Dominiczak AF, Bani-Mustafa A, Balmforth AJ, Hall AS, Erdmann J, Cambien F, Deloukas P, Hengstenberg C, Packard C, Schunkert H, Ouwehand WH, Ford I, Goodall AH, Jobling MA, Samani NJ, Tomaszewski M.** Inheritance of coronary artery disease in men: an analysis of the role of the Y chromosome. *Lancet* 379: 915–922, 2012.
13. **Clifton PM, Nestel PJ.** Influence of gender, body mass index, and age on response of plasma lipids to dietary fat plus cholesterol. *Arterioscler Thromb* 12: 955–962, 1992.

14. **Dhruva SS, Bero LA, Redberg RF.** Gender bias in studies for Food and Drug Administration premarket approval of cardiovascular devices. *Circ Cardiovasc Qual Outcomes* 4: 165–171, 2011.
15. **Dhruva SS, Redberg RF.** Evaluating sex differences in medical device clinical trials: time for action. *JAMA* 307: 1145–1146, 2012.
16. **Donawa M.** Successful recruitment for medical device clinical studies. *Med Device Technol* 15: 25–27, 2004.
17. **Ely D, Milsted A, Dunphy G, Boehme S, Dunmire J, Hart M, Toot J, Martins A, Turner M.** Delivery of sry1, but not sry2, to the kidney increases blood pressure and sns indices in normotensive wky rats. *BMC Physiol* 9: 10, 2009.
18. **Feletou M, Vanhoutte PM.** Endothelial dysfunction: a multifaceted disorder (The Wiggers Award Lecture). *Am J Physiol Heart Circ Physiol* 291: H985–H1002, 2006.
19. **Geller SE, Koch A, Pellettieri B, Carnes M.** Inclusion, analysis and reporting of sex and race/ethnicity in clinical trials: have we made progress? *J Womens Health* 20: 315–320, 2011.
20. **Gibson FC 3rd, Hong C, Chou HH, Yumoto H, Chen J, Lien E, Wong J, Genco CA.** Innate immune recognition of invasive bacteria accelerates atherosclerosis in apolipoprotein E-deficient mice. *Circulation* 109: 2801–2806, 2004.
21. **Goodarzi MO, Li X, Krauss RM, Otter JI, Chien YI.** Relationship of sex to diabetes risk in statin trials. *Diabetes Care* 36: e100–e101, 2013.
22. **Hamm TE Jr, Kaplan JR, Clarkson TB, Bullock BC.** Effects of gender and social behavior on the development of coronary artery atherosclerosis in cynomolgus macaques. *Atherosclerosis* 48: 221–233, 1983.
23. **Harvey W.** *Exercitatio Anatomica de Motu Cordis et Sanguinis in Animalibus*, Springfield, IL: Thomas, 1928.
24. **Heinrich J.** General Accounting Office, GAO-01-286R Drugs Withdrawn From Market, 2001. http://www.gao.gov/new.items/d01286r.pdf.
25. **Hunt K, Lewars H, Emslie C, Batty GD.** Decreased risk of death from coronary heart disease amongst men with higher 'femininity' scores: a general population cohort study. *Int J Epidemiol* 36: 612–620, 2007.
26. **Inacio MC, Ake CF, Paxton EW, Khatod M, Wang C, Gross TP, Kaczmarek RG, Marinac-Dabic D, Sedrakyan A.** Sex and risk of hip implant failure: assessing total hip arthroplasty outcomes in the United States. *JAMA Intern Med* 173: 435–441, 2013.
27. **Ingalhalikar M, Smith A, Parker D, Satterthwaite TD, Elliott MA, Ruparel K, Hakonarson H, Gur RE, Gur RC, Verma R.** Sex differences in the structural connectome of the human brain. *Proc Natl Acad Sci USA* 111: 823–828, 2014.
27a. **Institute of Medicine (US) Board on Population Health and Public Health Practice.** *Sex-Specific Reporting of Scientific Research: A Workshop Summary*. Washington, DC: National Academic Press, 2012.
28. **Isensee J, Ruiz Noppinger P.** Sexually dimorphic gene expression in mammalian somatic tissue. *Gend Med* 4, *Suppl B: S75–S95, 2007*.
29. **Ji H, Zheng W, Wu X, Liu J, Ecelbarger CM, Watkins R, Arnold AP, Sandberg K.** Sex chromosome effects unmasked in angiotensin II-induced hypertension. *Hypertension* 55: 1275–1282, 2010.
30. **Kuroki Y, Tsuchida K, Go I, Aoyama M, Naganuma T, Takemoto Y, Nakatani T.** A study of innate immunity in patients with end-stage renal disease: special reference to toll-like receptor-2 and -4 expression in peripheral blood monocytes of hemodialysis patients. *Int J Mol Med* 19: 783–790, 2007.
31. **Kwekel JC, Desai VG, Moland CL, Vijay V, Fuscoe JC.** Sex differences in kidney gene expression during the life cycle of F344 rats. *Biol Sex Differ* 4: 14, 2013.
32. **Lagace M, Xuan JY, Young SS, McRoberts C, Maier J, Rajcan-Separovic E, Korneluk RG.** Genomic organization of the X-linked inhibitor of apoptosis and identification of a novel testis-specific transcript. *Genomics* 77: 181–188, 2001.
33. **Lenaz G, Bovina C, D'Aurelio M, Fato R, Formiggini G, Genova ML, Giuliano G, Merlo Pich M, Paolucci U, Parenti Castelli G, Ventura B.** Role of mitochondria in oxidative stress and aging. *Ann NY Acad Sci* 959: 199–213, 2002.
34. **Li Z, Yang F, Dunn S, Gross AK, Smyth SS.** Platelets as immune mediators: Their role in Host defense responses and sepsis. *Thromb Res* 127: 184–186, 2011.
35. **Libby P.** Fat fuels the flame: triglyceride-rich lipoproteins and arterial inflammation. *Circ Res* 100: 299–301, 2007.
36. **Lin TH, Yeh S, Chang C.** Tissue-specific knockout of androgen receptor in mice. *Methods Mol Biol* 776: 275–293, 2011.
37. **Mahalik JR, Locke BD, Ludlow LH, Diemer MA, Scott RP, Gottfried M, Freitas G.** Development of the conformity to masculine norms inventory. *Psychol Men Masc*: 3–25, 2003.
38. **Martin SS, Ou FS, Newby LK, Sutton V, Adams P, Felker GM, Wang TY.** Patient- and trial-specific barriers to participation in cardiovascular randomized clinical trials. *J Am Coll Cardiol* 61: 762–769, 2013.
39. **McIntyre TM, Prescott SM, Weyrich AS, Zimmerman GA.** Cell-cell interactions: leukocyte-endothelial interactions. *Curr Opin Hematol* 10: 150–158, 2003.
40. **Melloni C, Berger JS, Wang TY, Gunes F, Stebbins A, Pieper KS, Dolor RJ, Douglas PS, Mark DB, Newby LK.** Representation of women in randomized clinical trials of cardiovascular disease prevention. *Circ Cardiovasc Qual Outcomes* 3: 135–142, 2010.
41. **Miller VM.** In pursuit of scientific excellence: sex matters. *Am J Physiol Heart Circ Physiol* 302: H1771–H1772, 2012.
42. **Miller VM, Flynn PM, Lindor KD.** Evaluating sex and gender competencies in the medical curriculum: a case study. *Gend Med* 9: 180–186, e183, 2012.
43. **Miller VM, Lewis DA, Barber DA.** Gender differences and endothelium- and platelet-derived factors in the coronary circulation. *Clin Exp Pharmacol Physiol* 26: 132–136, 1999.
44. **Miller VM, Rice M, Schiebinger L, Jenkins MR, Werbinski J, Nunez A, Wood S, Viggiano TR, Shuster LT.** Embedding concepts of sex and gender health differences into medical curricula. *J Womens Health* 22: 194–202, 2013.
45. **Misra MK, Sarwat M, Bhakuni P, Tuteja R, Tuteja N.** Oxidative stress and ischemic myocardial syndromes. *Med Sci Monit* 15: RA209–RA219, 2009.
46. **Money J.** Gender: history, theory and usage of the term in sexology and its relationship to nature/nurture. *J Sex Marital Ther* 11: 71–79, 1985.
48. **Mosca L, Benjamin EJ, Berra K, Bezanson JL, Dolor RJ, Lloyd-Jones DM, Newby LK, Pina IL, Roger VL, Shaw LJ, Zhao D, Beckie TM, Bushnell C, D'Armiento J, Kris-Etherton PM, Fang J, Ganiats TG, Gomes AS, Gracia CR, Haan CK, Jackson EA, Judelson DR, Kelepouris E, Lavie CJ, Moore A, Nussmeier NA, Ofili E, Oparil S, Ouyang P, Pinn VW, Sherif K, Smith SC Jr, Sopko G, Chandra-Strobos N, Urbina EM, Vaccarino V, and Wenger NK.** Effectiveness-Based Guidelines for the Prevention of Cardiovascular Disease in Women–2011 Update: A Guideline from the American Heart Association. *Circulation* 123: 1243–1262, 2011.
49. **Negrin CD, McBride MW, Carswell HV, Graham D, Carr FJ, Clark JS, Jeffs B, Anderson NH, Macrae M, Dominiczak AF.** Reciprocal consomic strains to evaluate Y chromosome effects. *Hypertension* 37: 391–397, 2001.
50. **Njolstad I, Arnesen E, Lund-Larsen PG.** Smoking, serum lipids, blood pressure, and sex differences in myocardial infarction. A 12-year follow-up of the Finnmark study. *Circulation* 93: 450–456, 1995.
51. **Ober C, Loisel DA, Gilad Y.** Sex-specific genetic architecture of human disease. *Nat Rev Genet* 9: 911–922, 2008.
52. **Ogawa S, Lozach J, Benner C, Pascual G, Tangirala RK, Westin S, Hoffmann A, Subramaniam S, David M, Rosenfeld MG, Glass CK.** Molecular determinants of crosstalk between nuclear receptors and toll-like receptors. *Cell* 122: 707–721, 2005.
53. **Ohno S.** *Sex chromosomes and sex-linked genes***: Springer-Verlag, Berlin, Heidelberg, New York**, 1967.
55. **Pierce GL, Eskurza I, Walker AE, Fay TN, Seals DR.** Sex-specific effects of habitual aerobic exercise on brachial artery flow-mediated dilation in middle-aged and older adults. *Clin Sci (Lond)* 120: 13–23, 2011.
56. **Pierce JP, Kievits J, Graustein B, Speth RC, Iadecola C, Milner TA.** Sex differences in the subcellular distribution of angiotensin type 1 receptors and NADPH oxidase subunits in the dendrites of C1 neurons in the rat rostral ventrolateral medulla. *Neuroscience* 163: 329–338, 2009.
57. **Radecki Breitkopf C, Parker M, Balls-Berry J, Halyard M, Pinn V, Hayes S.** African-American women's perceptions and attitudes regarding participation in health research.. *J Womens Health* 22: 38–39, 2013.
58. **Raffetto JD, Khalil RA.** Matrix metalloproteinases and their inhibitors in vascular remodeling and vascular disease. *Biochem Pharmacol* 75: 346–359, 2008.
59. **Redfield MM, Rodeheffer RJ, Jacobsen SJ, Mahoney DW, Bailey KR, Burnett JC Jr.** Plasma brain natriuretic peptide concentration: Impact of age and gender. *J Am Coll Cardiol* 40: 976–982, 2002.

60. **Reilly SL, Kottke BA, Sing CF.** The effects of generation and gender on the joint distributions of lipid and apolipoprotein phenotypes in the population at large. *J Clin Epidemiol* 43: 921–940, 1990.
61. **Ridker PM, Cook NR, Lee IM, Gordon D, Gaziano JM, Manson JE, Hennekens CH, Buring JE.** A randomized trial of low-dose aspirin in the primary prevention of cardiovascular disease in women.[see comment]. *N Engl J Med* 352: 1293–1304, 2005.
62. **Ristvedt SL.** The Evolution of Gender. *JAMA Psychiatry* 71: 13–14, 2014.
63. **Sandberg K, Verbalis JG.** Sex and the basic scientist: is it time to embrace Title IX? *Biol Sex Differ* 4: 13, 2013.
64. **Schmidt EP, Lee WL, Zemans RL, Yamashita C, Downey GP.** On, around, and through: neutrophil-endothelial interactions in innate immunity. *Physiology (Bethesda)* 26: 334–347, 2011.
65. **Seremetis SV.** Sex-related differences in hemostasis and thrombosis. *J Gend Specif Med* 4: 59–64, 2001.
66. **Shah K, McCormack CE, ABN.** Do you know the sex of your cells? *Am J Physiol Cell Physiol* 306: C3–C18, 2014.
67. **Shaul PW, Wells LB, Horning KM.** Acute and prolonged hypoxia attenuated endothelial nitric oxide production in rat pulmonary arteries by different mechanisms. *J Cardiovasc Pharmacol* 22: 819–827, 1993.
68. **Smulyan H.** Nitrates, arterial function, wave reflections and coronary heart disease. *Adv Cardiol* 44: 302–314, 2007.
69. **Strehlow K, Rotter S, Wassmann S, Adam O, Grohe C, Laufs K, Bohm M, Nickenig G.** Modulation of antioxidant enzyme expression and function by estrogen. *Circ Res* 93: 170–177, 2003.
70. **Taylor KE, Vallejo-Giraldo C, Schaible NS, Zakeri R, Miller VM.** Reporting of sex as a variable in cardiovascular studies using cultured cells. *Biol Sex Differ* 2: 11, 2011.
70a. **Turner ME, Jenkins C, Milsted A, Ely DL.** Sex chromosomes. *Adv Mol Cell Biol* 34: 1–13, 2004.
71. **van Heerebeek L, Franssen CP, Hamdani N, Verheugt FW, Somsen GA, Paulus WJ.** Molecular and cellular basis for diastolic dysfunction. *Curr Heart Fail Rep* 9: 293–302, 2012.
72. **Van PL, Bakalov VK, Zinn AR, Bondy CA.** Maternal X chromosome, visceral adiposity, and lipid profile. *JAMA* 295: 1373–1374, 2006.
73. **Voghel G, Thorin-Trescases N, Farhat N, Nguyen A, Villeneuve L, Mamarbachi AM, Fortier A, Perrault LP, Carrier M, Thorin E.** Cellular senescence in endothelial cells from atherosclerotic patients is accelerated by oxidative stress associated with cardiovascular risk factors. *Mech Ageing Dev* 128: 662–671, 2007.
74. **Wadman M.** NIH mulls rules for validating key results. *Nature* 500: 14–16, 2013.
75. **Wang X, Barber DA, Lewis DA, McGregor CGA, Sieck GA, Fitzpatrick LA, Miller VM.** Gender and transcriptional regulation of NO synthase and ET-1 in porcine aortic endothelial cells. *Am J Physiol Heart Circ Physiol* 273: H1962–H1967, 1997.
76. **Waxman DJ, Holloway MG.** Sex differences in the expression of hepatic drug metabolizing enzymes. *Mol Pharmacol* 76: 215–228, 2009.
77. **Wizemann TM, Pardue ML.** *Exploring the Biological Contributions to Human Health: Does Sex Matter? Board on Health Sciences Policy*. Washington, DC: Institute of Medicine, 2001.
77a. **World Health Organization.** *WHO Expert Committee on Biological Standardization*. **Thirty-seventh Report. WHO Health Organization Technical Report Series**, 1987.
78. **Xu J, Disteche CM.** Sex differences in brain expression of X- and Y-linked genes. *Brain Res* 1126: 50–55, 2006.
79. **Yang X, Schadt EE, Wang S, Wang H, Arnold AP, Ingram-Drake L, Drake TA, Lusis AJ.** Tissue-specific expression and regulation of sexually dimorphic genes in mice. *Genome Res* 16: 995–1004, 2006.
80. **Yasunari K, Maeda K, Minami M, Yoshikawa J.** HMG-CoA reductase inhibitors prevent migration of human coronary smooth muscle cells through suppression of increase in oxidative stress. *Arterioscler Thromb Vasc Biol* 21: 937–942, 2001.
81. **Zucker I, Beery AK.** Males still dominate animal studies. *Nature* 465: 690, 2010.

Journal of Athletic Training 2019;54(5):472–482
doi: 10.4085/1062-6050-407-16

www.natajournals.org

Anterior Cruciate Ligament Injury Risk in Sport: A Systematic Review and Meta-Analysis of Injury Incidence by Sex and Sport Classification

Alicia M. Montalvo, PhD, ATC, CSCS*; Daniel K. Schneider, MD†; Kate E. Webster, PhD‡; Laura Yut, BS*; Marc T. Galloway, MD§; Robert S. Heidt Jr, MD§; Christopher C. Kaeding, MD¶; Timothy E. Kremcheck, MD||; Robert A. Magnussen, MD, MPH¶; Shital N. Parikh, MD#; Denver T. Stanfield, MD§; Eric J. Wall, MD#; Gregory D. Myer, PhD, CSCS*D, FACSM**

*Department of Athletic Training, Florida International University, Miami; †Riverside Methodist Hospital, Columbus, OH; ‡School of Allied Health, La Trobe University, Melbourne, Australia; §Mercy Health, Cincinnati, OH; ¶Department of Orthopaedics, Sports Medicine Institute, The Ohio State University, Columbus; ||Beacon Orthopaedics, Cincinnati, OH; #University of Cincinnati, OH; **Cincinnati Children's Hospital Medical Center, OH

Objective: To evaluate sex differences in incidence rates (IRs) of anterior cruciate ligament (ACL) injury by sport type (collision, contact, limited contact, and noncontact).

Data Sources: A systematic review was performed using the electronic databases PubMed (1969–January 20, 2017) and EBSCOhost (CINAHL, SPORTDiscus; 1969–January 20, 2017) and the search terms *anterior cruciate ligament* AND *injury* AND (*incidence* OR *prevalence* OR *epidemiology*).

Study Selection: Studies were included if they provided the number of ACL injuries and the number of athlete-exposures (AEs) by sex or enough information to allow the number of ACL injuries by sex to be calculated. Studies were excluded if they were analyses of previously reported data or were not written in English.

Data Extraction: Data on sport classification, number of ACL injuries by sex, person-time in AEs for each sex, year of publication, sport, sport type, and level of play were extracted for analysis.

Data Synthesis: We conducted IR and IR ratio (IRR) meta-analyses, weighted for study size and calculated. Female and male athletes had similar ACL injury IRs for the following sport types: collision (2.10/10 000 versus 1.12/10 000 AEs, IRR = 1.14, $P = .63$), limited contact (0.71/10 000 versus 0.29/10 000 AEs, IRR = 1.21, $P = .77$), and noncontact (0.36/10 000 versus 0.21/10 000 AEs, IRR = 1.49, $P = .22$) sports. For contact sports, female athletes had a greater risk of injury than male athletes did (1.88/10 000 versus 0.87/10 000 AEs, IRR = 3.00, $P < .001$). Gymnastics and obstacle-course races were outliers with respect to IR, so we created a sport category of fixed-object, high-impact rotational landing (HIRL). For this sport type, female athletes had a greater risk of ACL injury than male athletes did (4.80/10 000 versus 1.75/10 000 AEs, IRR = 5.51, $P < .001$), and the overall IRs of ACL injury were greater than all IRs in all other sport categories.

Conclusions: Fixed-object HIRL sports had the highest IRs of ACL injury for both sexes. Female athletes were at greater risk of ACL injury than male athletes in contact and fixed-object HIRL sports.

Key Words: epidemiology, knee, sprain, athletes

Anterior cruciate ligament (ACL) injury is a common and debilitating injury among athletes. It can occur from both contact and noncontact mechanisms[1,2] and has a relatively high incidence in sports involving deliberate contact.[1] The relationship between the amount of inherent contact in a sport and the risk of injury to the ACL is unclear, especially when including sex as a variable. In the United States, collision sports, such as football, rugby, and wrestling, are male dominated. Females play collision sports such as ice hockey and rugby, but contact sports such as soccer and basketball are more commonly cited when comparing ACL injury risk by sex. Whereas the rate of ACL injury in females playing soccer was among the highest, it was also high in limited-contact and noncontact sports, including alpine skiing and gymnastics, respectively.[1,3] Hootman et al[1] found some of the highest rates of ACL injury among males in collision sports (spring and fall football and wrestling). Conversely, in females, gymnastics (noncontact), followed by soccer and basketball, resulted in the highest rates of ACL injury.[1]

Deliberate contact during sport is believed to contribute to increased rates of ACL injury.[4] However, given that many ACL injuries result from noncontact mechanisms, the role of sport type in ACL injury is uncertain. Moreover, it is unclear if a sex difference in ACL injury incidence exists when stratifying by sport type (eg, collision, full contact, limited contact, and noncontact). Therefore, the purpose of our systematic review and meta-analysis was to compare the incidence rates (IRs) of ACL injury of male and female athletes in each of the following sport types: collision, contact, limited contact, and noncontact.

METHODS

We followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses[5] (PRISMA) guidelines when conducting and reporting this systematic review and meta-analysis.

Literature Search

A systematic review of the current literature was performed using the electronic databases PubMed (1969–January 20, 2017) and EBSCOhost (CINAHL and SPORTDiscus; 1969–January 20, 2017) and the following search terms: *anterior cruciate ligament* AND *injury* AND (*incidence* OR *prevalence* OR *epidemiology*). Results were further limited to peer-reviewed articles written in English.

In addition to the electronic search, we contacted experts in the field for further suggestions and examined references cited in review papers to identify any other relevant articles for potential inclusion. Publication details from all studies identified in the literature search were exported to bibliographic software (Endnote X7; Clarivate Analytics, Philadelphia, PA).

Selection Criteria

Given the large number of identified studies, a single author (A.M.M.) performed the initial screening of articles for inclusion. Any gray areas were discussed with the second author (D.K.S.), and any disagreements were decided by the senior author (G.D.M.). Articles were screened first by title, second by abstract, and third by full text according to the inclusion and exclusion criteria. We included articles in which the total number of ACL injuries and the total number of athlete-exposures (AEs) were reported by sex and the data were provided in such a way that the number of ACL injuries by sex could be calculated. We excluded articles that included further analyses on previously reported prospective studies, were written in languages other than English, or were review papers. Full texts were retrieved when the title or abstract met the selection criteria or when the status could not be determined from the title and abstract alone.

Data Extraction and Analysis

The primary variables extracted were the sport classification, number of ACL injuries for each sex, and person-time in AEs for each sex. Sports were classified as follows: *collision* (contact with an opponent or object is inherent), *contact* (contact with an opponent or object is acceptable), *limited contact* (contact with an opponent or object is discouraged), and *noncontact* (contact with an opponent or object is unexpected; Table 1). For each sport classification, we calculated the overall ACL injury rate and separate IRs for men and women. The IR ratio (IRR) between men and women was subsequently calculated using only data from studies in which injury-risk data were reported for both men and women to allow direct comparisons. Additional extracted data included year of publication, sport, sport type, and level of play. One author (A.M.M.) recorded all pertinent data from the included articles, and another author (D.K.S.) independently reviewed those data for accuracy and completeness.

Table 1. Sport Classification Key

Classification	Sport
Collision	Boxing
	Boys'/men's lacrosse
	Close-quarters combat
	Football
	Handball
	Ice hockey
	Rugby
	Wrestling
Contact	Basketball
	Field hockey
	Girls'/women's lacrosse
	Judo
	Soccer
Limited contact	Baseball
	Cheerleading
	Fencing
	Flickerball
	Floorball
	Frisbee
	Softball
	Volleyball
Noncontact	Alpine skiing
	Dance/ballet
	Running/track
Fixed-object high-impact rotational landing	Gymnastics
	Indoor obstacle-course test
	Obstacle-course race

The reported person-time unit was not uniform across studies. Therefore, to establish a common metric, we tabulated AEs. When the number of player-hours was reported, the number of AEs was estimated by dividing player-hours by 2. The assumption for converting player-hours to AEs was that each AE (1 game or 1 practice) on average would last about 2 hours. In addition, not all authors reported the number of ACL injuries by sex; instead, they provided IRs by sex. For these studies, the number of AEs and the reported IRs were used to calculate the number of ACL injuries by sex (number of ACL injuries by sex = total AEs by sex × the rate numerator by sex/the rate denominator by sex).[6–8] For studies in which the number of ACL injuries by sex could not be estimated, we e-mailed the authors to gather those data. If they did not have access to the information or did not respond, the study was excluded from the meta-analysis.[9–15]

Risk of Bias Assessment

Included studies were critically appraised independently by 2 authors (A.M.M., D.K.S.). Given that most included articles described observational cohort studies that did not include an intervention, traditional checklists were not appropriate. After a thorough search for tools to appraise observational cohort studies, we decided that the tool best suited to be used quantitatively was the Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies.[16] This tool, available through the National Institutes of Health (Bethesda, MD), assesses criteria such as participation rate, whether exposure data were collected before the outcome, whether the time frame was sufficient to allow for the outcome to occur, and the number of participants lost to follow-up after baseline. If a criterion was met, the item was scored as 1. If it was absent or not reported, the item

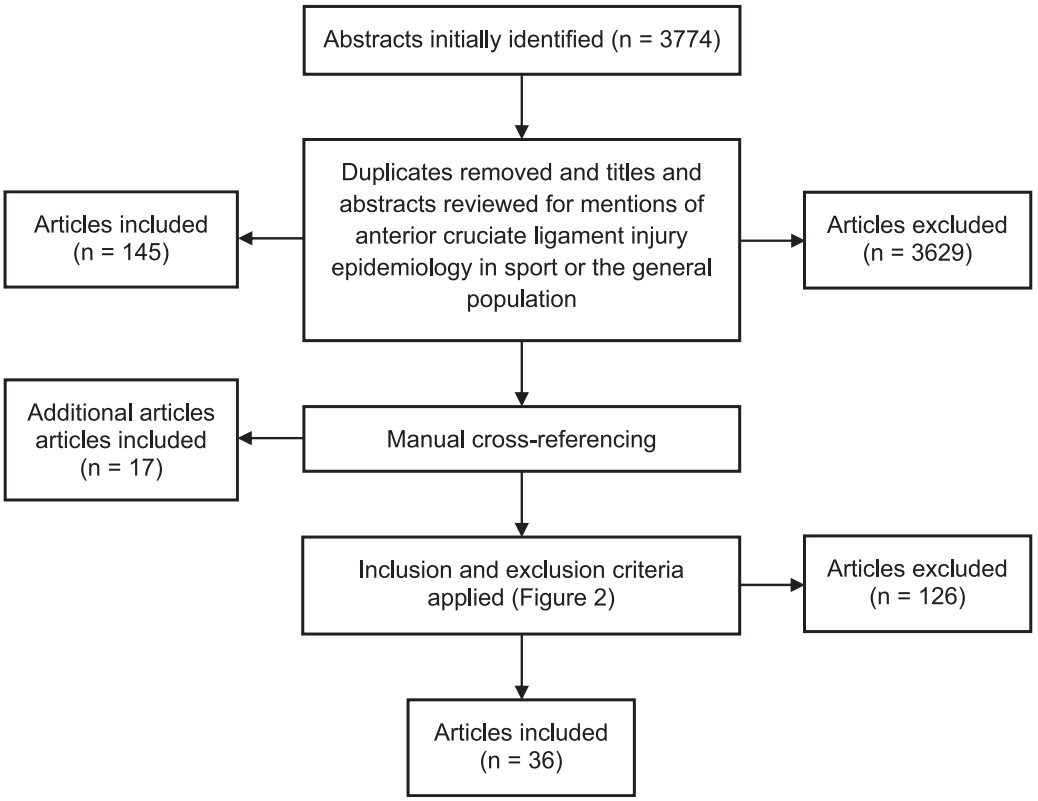


Figure 1. Flow chart of the literature review process.

was scored as zero. The maximum score possible was 14. Items were scored independently by 2 authors (A.M.M., D.K.S.). These authors discussed any discrepancies in scoring. For discrepancies that could not be resolved, a third author (G.D.M.) was consulted for arbitration. Given that the included studies with interventions were treated as cohort studies in the analyses, they were assessed with the same tool, which allowed for quality comparisons across all included studies.

Statistical Analysis

The number of included studies per analysis varied. For the total IR, any study in which authors reported the rate of either sex was included. For the IR by sex, any study in which the authors reported female or male rates was included for the respective analyses. Only studies that included both female and male athletes were used to calculate ratios. The ACL injury IR in noncontact sports comprised sports with marked differences in ACL injury IRs. Given that several outliers were present, we subdivided the category into sports that did and sports that did not include a fixed-object and high-impact landing. These latter sports were removed from the noncontact category, and a new fixed-object, high-impact rotational-landing (HIRL) category was created. *Fixed-object HIRL sports* were defined as noncontact sports that included high-impact landings from fixed objects, such as beams, vaults, and obstacles. Injury IRs for the individual studies were summarized in forest plots for the following groups by total, female, and male IRs: collision, contact, limited-contact, noncontact, and fixed-object HIRL sports. These rates were multiplied to calculate ACL injury IRs per 10 000 AEs in each respective group. Incidence rate ratios for women versus men were calculated for each group and summarized in forest plots.

Injury data were analyzed using R (version 3.3.2; R Foundation for Statistical Computing, Vienna, Austria) and the R packages meta and metafor with the functions metarate for IR and metainc for IRR weighted for individual study size. When AEs but no events (ACL injuries) were present, a continuity correction was applied. The default value for the continuity correction, 0.5, was used to calculate individual point estimates and the 95% confidence interval (CI) and to conduct a meta-analysis based on the inverse variance method. We set the α level at .05.

RESULTS

The electronic literature search yielded 3774 abstracts for initial review. After duplicates were removed, a total of 1300 unique titles remained. We screened the titles and abstracts and removed 1155 articles for lack of relevance to the research. The remaining 145 articles were manually cross-referenced, and experts were consulted to identify additional relevant articles, resulting in the inclusion of 17 more articles. Full texts of these 162 articles were obtained and assessed for the inclusion and exclusion criteria. We contacted the corresponding authors of the included articles for additional information as needed. At the end of the search, 36 articles were included in the study.[1,6–8,17–48] An outline of the literature review process is presented in Figure 1. The data that were extracted for each analysis and can be used to determine which studies were included in each analysis are shown in Table 2.

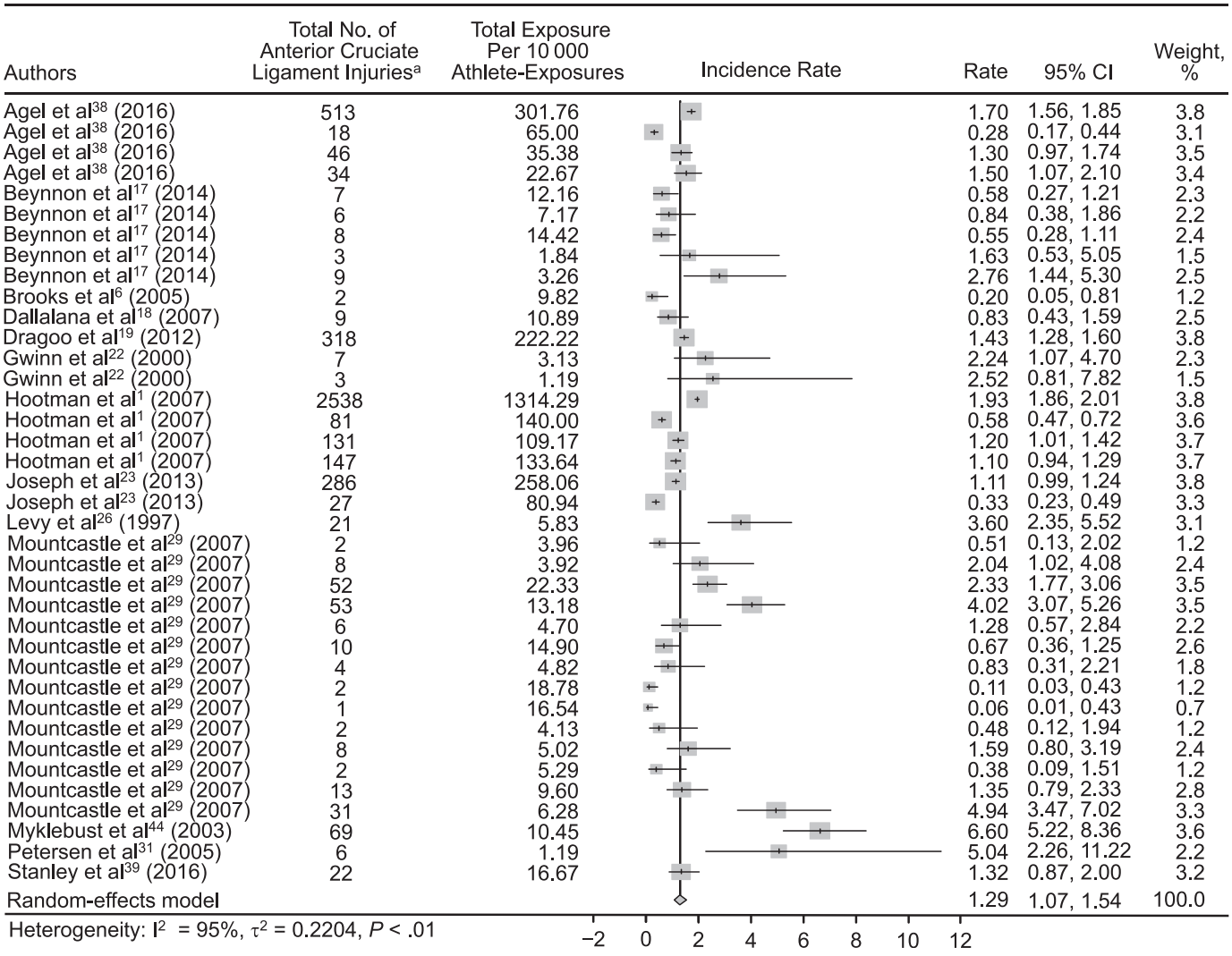


Figure 2. Forest plot for the total incidence rate of anterior cruciate ligament injury in male and female collision-sport athletes combined.
[a] Sports are provided in Table 2. Abbreviation: CI, confidence interval.

Incidence Rates for Collision Sports by Sex

In collision sports, the total IR of ACL injury among female and male athletes combined was 1.29/10 000 AEs (95% CI = 1.07, 1.54; $P < .01$, $I^2 = 95.0\%$; Figure 2). The injury IR among female athletes was 2.10/10 000 AEs (95% CI = 1.12, 3.96; $P < .01$, $I^2 = 84.0\%$; Figure 3) and among male athletes was 1.12/10 000 AEs (95% CI = 0.94, 1.33; $P < .01$, $I^2 = 93.0\%$; see Supplemental Figure 1, available online at http://dx.doi.org/10.4085/1062-6050-407-16.S1). We observed no difference between sexes for the ACL injury IR (IRR = 1.14; 95% CI = 0.68, 1.92, $P = .63$; $I^2 = 0\%$; see Supplemental Figure 2).

Incidence Rates for Contact Sports by Sex

The total IR of ACL injury in contact sports was 1.51/10 000 AEs (95% CI = 1.31, 1.75; $P < .01$, $I^2 = 90.0\%$; see Supplemental Figure 3). The injury IR was greater among female (1.88/10 000 AEs; 95% CI = 1.61, 2.20; $P < .01$, $I^2 = 88.0\%$; see Supplemental Figure 4) than among male (0.87/10 000 AEs; 95% CI = 0.69, 1.11; $P < .01$, $I^2 = 84.0\%$; see Supplemental Figure 5) athletes. We observed a difference between sexes for the ACL injury IR (IRR = 3.00; 95% CI = 2.70, 3.34; $P < .001$, $I^2 = 4.0\%$; see Supplemental Figure 6).

Incidence Rates for Limited-Contact Sports by Sex

In limited-contact sports, the total IR of ACL injury was 0.48/10 000 AEs (95% CI = 0.33, 0.70; $P < .01$, $I^2 = 91.0\%$; see Supplemental Figure 7). The injury IR in female athletes was 0.71/10 000 AEs (95% CI = 0.50, 1.01; $P < .01$, $I^2 = 84.0\%$; see Supplemental Figure 8) and in male athletes was 0.29/10 000 AEs (95% CI = 0.18, 0.48; $P < .01$, $I^2 = 63.0\%$; see Supplemental Figure 9). The IRR was calculated using only data from Mountcastle et al,[29] as data comparing injury rates among women and men in this sport type were not available. We observed no difference between sexes for the ACL injury IR (IRR = 1.21; 95% CI = 0.35, 4.20; $P = .77$, $I^2 = 0\%$; see Supplemental Figure 10).

Incidence Rates for Noncontact Sports by Sex

The total IR of ACL injury in noncontact sports was 0.25/10 000 AEs (95% CI = 0.10, 0.65; $P < .01$, $I^2 = 85.0\%$; see Supplemental Figure 11). The ACL injury IR among female athletes was 0.36/10 000 AEs (95% CI = 0.14, 0.96;

Table 2. Data Extracted From Each Included Study Continued on Next Page

				Anterior Cruciate Ligament Injuries			Athlete-Exposures	
Article (y)	Sport	Classification	Level	Female	Male	Total	Female	Male
Agel et al[38] (2016)	Football	Collision	Collegiate	0	513	513	0	3 017 647
Agel et al[38] (2016)	Ice hockey	Collision	Collegiate	3	15	18	150 000	500 000
Agel et al[38] (2016)	Lacrosse	Collision	Collegiate	0	46	46	0	353 846
Agel et al[38] (2016)	Wrestling	Collision	Collegiate	0	34	34	0	226 667
Beynnon et al[17] (2014)	Lacrosse	Collision	High school	0	7	7	0	121 583
Beynnon et al[17] (2014)	Lacrosse	Collision	Collegiate	0	6	6	0	71 731
Beynnon et al[17] (2014)	Football	Collision	High school	0	8	8	0	144 233
Beynnon et al[17] (2014)	Football	Collision	Collegiate	0	3	3	0	18 417
Beynnon et al[17] (2014)	Rugby	Collision	Collegiate	6	3	9	14 723	17 886
Brooks et al[6] (2005)	Rugby	Collision	Professional	0	2	2	0	98 205
Dallalana et al[18] (2007)	Rugby	Collision	Professional	0	9	9	0	108 920
Dragoo et al[19] (2012)	Football	Collision	Collegiate	0	318	318	0	2 222 155
Gwinn et al[22] (2000)	Rugby	Collision	Collegiate	3	4	7	8475	22 788
Gwinn et al[22] (2000)	Instructional wrestling	Collision	Amateur	1	2	3	1306	10 582
Hootman et al[1] (2007)	Football	Collision	Collegiate	0	2538	2538	0	13 142 929
Hootman et al[1] (2007)	Ice hockey	Collision	Collegiate	3	78	81	100 000	1 300 000
Hootman et al[1] (2007)	Lacrosse	Collision	Collegiate	0	131	131	0	1 091 667
Hootman et al[1] (2007)	Wrestling	Collision	Collegiate	0	147	147	0	1 336 364
Joseph et al[23] (2013)	Football	Collision	High school	0	286	286	0	2 580 637
Joseph et al[23] (2013)	Wrestling	Collision	High school	0	27	27	0	809 430
Levy et al[26] (1997)	Rugby	Collision	Collegiate	21	0	21	58 296	0
Mountcastle et al[29] (2007)	Ice hockey	Collision	Collegiate	0	2	2	0	39 587
Mountcastle et al[29] (2007)	Lacrosse	Collision	Collegiate	0	8	8	0	39 204
Mountcastle et al[29] (2007)	Football	Collision	Collegiate	0	52	52	0	223 307
Mountcastle et al[29] (2007)	Football	Collision	Amateur	1	52	53	1828	129 956
Mountcastle et al[29] (2007)	Wrestling	Collision	Collegiate	0	6	6	0	47 039
Mountcastle et al[29] (2007)	Wrestling	Collision	Amateur	0	10	10	0	149 022
Mountcastle et al[29] (2007)	Wrestling	Collision	Amateur	0	4	4	0	48 203
Mountcastle et al[29] (2007)	Close-quarters combat	Collision	Amateur	0	2	2	37 184	150 606
Mountcastle et al[29] (2007)	Boxing	Collision	Amateur	0	1	1	0	165 376
Mountcastle et al[29] (2007)	Boxing	Collision	Amateur	0	2	2	0	41 270
Mountcastle et al[29] (2007)	Handball	Collision	Amateur	4	4	8	25 090	25 090
Mountcastle et al[29] (2007)	Handball	Collision	Amateur	0	2	2	13 564	39 348
Mountcastle et al[29] (2007)	Rugby	Collision	Amateur	0	13	13	770	95 200
Mountcastle et al[29] (2007)	Rugby	Collision	Amateur	0	31	31	0	62 785
Myklebust et al[44] (2003)	Handball	Collision	Elite, subelite	69	0	69	104 468	0
Petersen et al[31] (2005)	Handball	Collision	Semiprofessional, amateur	6	0	6	11 905	0
Stanley et al[39] (2016)	Lacrosse	Collision	High school	0	22	22	0	166 667
Agel et al[38] (2016)	Basketball	Contact	Collegiate	162	70	232	736 364	875 000
Agel et al[38] (2016)	Field hockey	Contact	Collegiate	20	0	20	181 818	0
Agel et al[38] (2016)	Lacrosse	Contact	Collegiate	59	0	59	256 522	0
Agel et al[38] (2016)	Soccer	Contact	Collegiate	71	26	97	710 000	650 000
Beynnon et al[17] (2014)	Basketball	Contact	High school	6	4	10	98 296	108 622
Beynnon et al[17] (2014)	Basketball	Contact	Collegiate	5	2	7	34 882	38 927
Beynnon et al[17] (2014)	Soccer	Contact	High school	15	3	18	114 077	117 140
Beynnon et al[17] (2014)	Soccer	Contact	Collegiate	11	6	17	28 115	30 241
Beynnon et al[17] (2014)	Field hockey	Contact	Collegiate	1	0	1	25 993	0
Beynnon et al[17] (2014)	Field hockey	Contact	High school	4	0	4	82 946	0
Beynnon et al[17] (2014)	Lacrosse	Contact	High school	6	0	6	86 160	0
Beynnon et al[17] (2014)	Lacrosse	Contact	Collegiate	4	0	4	37 567	0
Faude et al[40] (2005)	Soccer	Contact	Elite	11	0	11	17 655	0
Gilchrist et al[20] (2008)	Soccer	Contact	Collegiate	25	0	25	88 139	0
Giza et al[48] (2005)	Soccer	Contact	Professional	8	0	8	177 778	0
Gomez et al[21] (1996)	Basketball	Contact	High school	11	0	11	60 376	0
Gwinn et al[22] (2000)	Basketball	Contact	Collegiate	5	1	6	10 452	11 282
Gwinn et al[22] (2000)	Soccer	Contact	Collegiate	5	1	6	6508	12 408
Gwinn et al[22] (2000)	Basketball	Contact	Amateur	0	5	5	1360	33 866
Gwinn et al[22] (2000)	Soccer	Contact	Amateur	2	10	12	742	25 462

Table 2. Continued From Previous Page and Continued on Next Page

Article (y)	Sport	Classification	Level	Anterior Cruciate Ligament Injuries			Athlete-Exposures	
				Female	Male	Total	Female	Male
Hägglund et al[41] (2009)	Soccer	Contact	Elite	8	8	16	27 078	35 681
Hootman et al[1] (2007)	Basketball	Contact	Collegiate	498	167	665	2 165 217	2 385 714
Hootman et al[1] (2007)	Field hockey	Contact	Collegiate	53	0	53	757 143	0
Hootman et al[1] (2007)	Lacrosse	Contact	Collegiate	145	0	145	852 941	0
Hootman et al[1] (2007)	Soccer	Contact	Collegiate	411	168	579	1 467 857	1 866 667
Joseph et al[23] (2013)	Soccer	Contact	High school	96	44	140	643 206	914 551
Joseph et al[23] (2013)	Basketball	Contact	High school	92	25	117	894 391	1 106 060
Kiani et al[24] (2010)	Soccer	Contact	Amateur	5	0	5	66 505	0
Krutsch et al[42] (2016)	Soccer	Contact	Professional and amateur	0	16	16	0	75 312
LaBella et al[25] (2011)	Soccer, basketball	Contact	High school	12	0	12	22 925	0
Le Gall et al[43] (2008)	Soccer	Contact	Elite, youth	12	0	12	48 359	0
Mandelbaum et al[7] (2005)	Soccer	Contact	Amateur	73	0	73	205 308	0
Messina et al[28] (1999)	Basketball	Contact	High school	0	4	4	0	84 943
Mountcastle et al[29] (2007)	Basketball	Contact	Collegiate	6	0	6	15 300	14 273
Mountcastle et al[29] (2007)	Basketball	Contact	Amateur	1	2	3	3438	19 483
Mountcastle et al[29] (2007)	Basketball	Contact	Amateur	2	12	14	16 896	100 409
Mountcastle et al[29] (2007)	Soccer	Contact	Collegiate	4	5	9	23 080	34 192
Mountcastle et al[29] (2007)	Soccer	Contact	Amateur	0	1	1	1810	10 261
Mountcastle et al[29] (2007)	Soccer	Contact	Amateur	1	13	14	14 382	80 124
Mountcastle et al[29] (2007)	Judo	Contact	Amateur	1	5	6	4600	29 900
Nagano et al[8] (2011)	Basketball	Contact	Elite	23	0	23	254 831	0
Östenberg and Roos[45] (2000)	Soccer	Contact	Elite	3	0	3	4839	0
Pfeiffer et al[32] (2006)	Basketball	Contact	High school	5	0	5	24 378	0
Pfeiffer et al[32] (2006)	Soccer	Contact	High school	1	0	1	15 270	0
Söderman et al[46] (2000)	Soccer	Contact	Elite	5	0	5	7017	0
Stanley et al[39] (2016)	Basketball	Contact	High school	35	12	47	289 256	363 636
Stanley et al[39] (2016)	Lacrosse	Contact	High school	32	0	32	101 266	0
Stanley et al[39] (2016)	Soccer	Contact	High school	31	19	50	173 184	208 791
Steffen et al[33] (2008)	Soccer	Contact	Amateur	9	0	9	66 574	0
Tegnander et al[34] (2008)	Soccer	Contact	Elite	2	0	2	14 810	0
Trojian and Collins[35] (2006)	Basketball	Contact	Professional	9	0	9	45 036	0
Waldén et al[37] (2012)	Soccer	Contact	Amateur	21	0	21	139 149	0
Waldén et al[47] (2011)	Soccer	Contact	Professional	15	20	35	52 389	164 923
Agel et al[38] (2016)	Baseball	Limited contact	Collegiate	0	12	12	0	600 000
Agel et al[38] (2016)	Softball	Limited contact	Collegiate	33	0	33	550 000	0
Agel et al[38] (2016)	Volleyball	Limited contact	Collegiate	30	0	30	500 000	0
Beynnon et al[17] (2014)	Volleyball	Limited contact	Collegiate	1	0	1	2237	0
Hootman et al[1] (2007)	Baseball	Limited contact	Collegiate	0	56	56	0	2 800 000
Hootman et al[1] (2007)	Softball	Limited contact	Collegiate	129	0	129	1 612 500	0
Hootman et al[1] (2007)	Volleyball	Limited contact	Collegiate	142	0	142	1 577 778	0
Joseph et al[23] (2013)	Volleyball	Limited contact	High school	20	0	20	841 608	0
Joseph et al[23] (2013)	Baseball	Limited contact	High school	0	6	6	0	861 964
Joseph et al[23] (2013)	Softball	Limited contact	High school	21	0	21	657 246	0
Mountcastle et al[29] (2007)	Baseball	Limited contact	Collegiate	0	1	1	0	27 674
Mountcastle et al[29] (2007)	Volleyball	Limited contact	Collegiate	2	0	2	19 357	0
Mountcastle et al[29] (2007)	Volleyball	Limited contact	Amateur	0	2	2	6856	38 849
Mountcastle et al[29] (2007)	Fencing	Limited contact	Amateur	0	1	1	12 148	16 964
Mountcastle et al[29] (2007)	Cheerleading	Limited contact	Amateur	2	2	4	16 780	16 780
Mountcastle et al[29] (2007)	Flickerball	Limited contact	Amateur	0	2	2	5845	31 896
Mountcastle et al[29] (2007)	Frisbee	Limited contact	Amateur	0	1	1	925	4829
Pasanen et al[30] (2008)	Floorball	Limited contact	Elite	10	0	10	28 679	0
Stanley et al[39] (2016)	Softball	Limited contact	High school	1	0	1	142 857	0
Stanley et al[39] (2016)	Baseball	Limited contact	High school	0	5	5	0	208 333
Liederbach et al[27] (2008)	Dance	Noncontact	Elite	10	2	12	873 067	545 266
Mountcastle et al[29] (2007)	Track	Noncontact	Collegiate	2	0	2	76 542	114 409
Mountcastle et al[29] (2007)	Skiing	Noncontact	Amateur	1	1	2	3586	20 361
Mountcastle et al[29] (2007)	Parachute	Noncontact	Amateur	0	2	2	8402	42 300
Viola et al[36] (1999)	Alpine skiing	Noncontact	Professional	10	21	31	227 766	499 070

Table 2. Continued From Previous Page

Article (y)	Sport	Classification	Level	Anterior Cruciate Ligament Injuries			Athlete-Exposures	
				Female	Male	Total	Female	Male
Agel et al[38] (2016)	Gymnastics	High-impact rotational landing	Collegiate	24	0	24	100 000	0
Gwinn et al[22] (2000)	Obstacle-course race	High-impact rotational landing	Amateur	4	3	7	650	5289
Hootman et al[1] (2007)	Gymnastics	High-impact rotational landing	Collegiate	134	0	134	406 061	0
Mountcastle et al[29] (2007)	Gymnastics	High-impact rotational landing	Collegiate	1	0	1	14 317	0
Mountcastle et al[29] (2007)	Gymnastics	High-impact rotational landing	Amateur	7	7	14	29 304	166 054
Mountcastle et al[29] (2007)	Obstacle-course race	High-impact rotational landing	Amateur	5	9	14	5323	35 630

$P < .01$, $I^2 = 74.0\%$; see Supplemental Figure 12) and among male athletes was 0.21/10 000 AEs (95% CI = 0.07, 0.62; $P < .01$, $I^2 = 70.0\%$; see Supplemental Figure 13). We observed no difference between sexes (IRR = 1.49; 95% CI = 0.79, 2.79; $P = .22$, $I^2 = 0\%$; see Supplemental Figure 14).

Incidence Rates for Fixed-Object HIRL Sports by Sex

In fixed-object HIRL sports, the total IR of ACL injury was 2.62/10 000 AEs (95% CI = 1.44, 4.75; $P < .01$, $I^2 = 89.0\%$; see Supplemental Figure 15). The ACL injury IR among female athletes was 4.80/10 000 AEs (95% CI = 2.37, 9.70; $P < .01$, $I^2 = 89.0\%$; see Supplemental Figure 16) and among male athletes was 1.75/10 000 AEs (95% CI = 0.41, 7.48; $P < .01$, $I^2 = 89.0\%$; see Supplemental Figure 17). We observed a difference between sexes (IRR = 5.51; 95% CI = 2.80, 10.82; $P < .001$, $I^2 = 0\%$; see Supplemental Figure 18).

Risk of Bias Assessment

Most studies were of moderate quality (Table 3). Three studies fulfilled 75% or more of the criteria, and 33 studies fulfilled 50% or more of the criteria. The remaining 3 studies fulfilled fewer than 50% of the criteria and were deemed to be of low quality. Studies may have received lower scores for lack of reporting information, such as the total number of eligible individuals, how outcomes were measured, and attrition. Overall, the risk of bias was deemed to be moderate.

DISCUSSION

The purpose of our study was to quantify sex differences in ACL injury risk for sports with various amounts of contact. Female athletes participating in contact and fixed-object HIRL sports had greater ACL injury IRs than their male counterparts. In contrast, the ACL injury IRs for collision, limited-contact, and noncontact sports did not differ between sexes. The findings from this meta-analysis support a previous report[4] indicating that the amount of athlete-to-athlete contact inherent to a sport was correlated with the rate of ACL injury in both male and female athletes. However, adding the fixed-object HIRL category suggested that sports such as gymnastics and obstacle-course races may result in the highest rates of ACL injury.

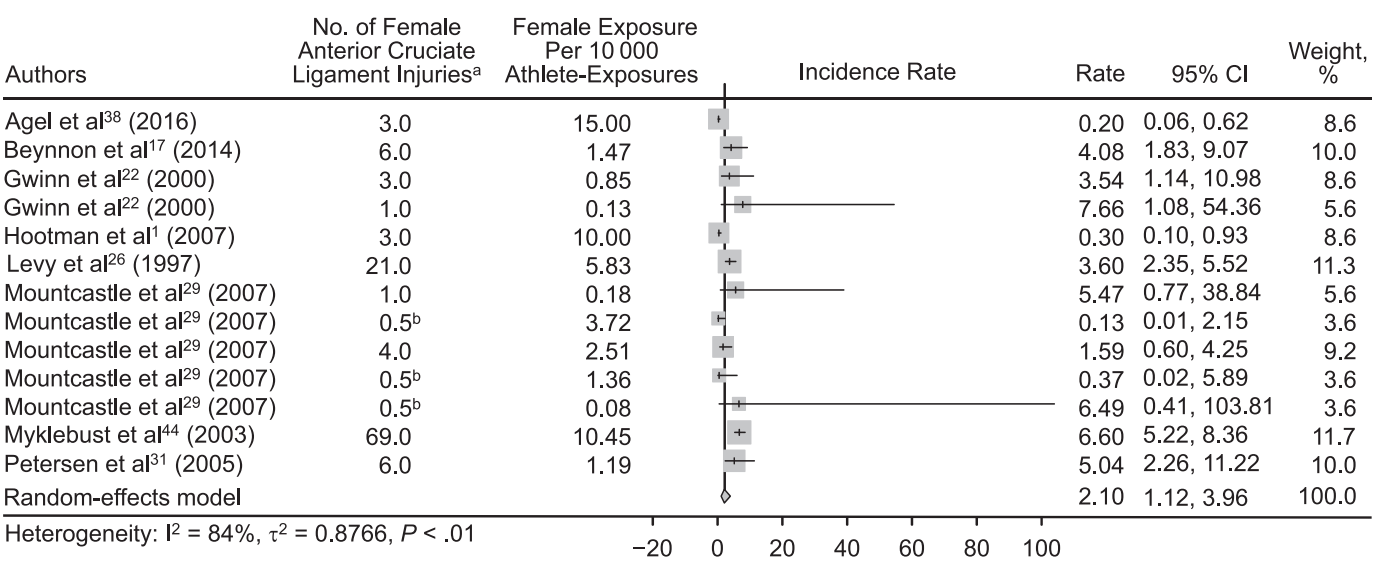


Figure 3. Forest plot for the incidence rate of anterior cruciate ligament injury in female collision-sport athletes. [a] Sports are provided in Table 2. [b] We substituted 0.1 for 0 to estimate an extremely low rate that could be used in the analysis. Abbreviation: CI, confidence interval.

Table 3. Results of Risk of Bias Assessment Using the Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies[a]

Study (y)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	Total Present
Agel et al[38] (2016)	1	1	0	1	0	1	1	1	1	1	1	0	0	0	9
Beynnon et al[17] (2014)	1	1	0	1	0	1	1	1	1	0	1	0	0	1	9
Brooks et al[6] (2005)	1	1	1	1	0	1	1	1	1	1	0	0	0	0	9
Dallalana et al[18] (2007)	1	1	1	1	0	1	1	1	1	1	0	0	0	0	9
Dragoo et al[19] (2012)	1	1	0	1	0	1	1	1	1	0	0	0	0	0	7
Faude et al[40] (2005)	1	1	0	1	0	1	1	1	1	1	0	0	0	0	8
Gilchrist et al[20] (2008)	1	1	0	1	0	1	1	1	1	1	1	1	0	0	10
Giza et al[48] (2005)	1	1	1	1	0	1	1	1	1	0	0	0	0	0	8
Gomez et al[21] (1996)	1	1	0	1	0	1	1	1	1	0	0	0	0	0	7
Gwinn et al[22] (2000)	1	1	1	1	0	1	1	1	1	0	1	0	0	0	9
Hägglund et al[41] (2009)	1	1	1	1	0	1	1	0	1	1	0	0	1	0	9
Hootman et al[1] (2007)	1	1	0	1	0	1	1	1	1	1	0	0	0	0	8
Joseph et al[23] (2013)	1	1	1	1	0	0	1	0	0	0	1	0	1	1	8
Kiani et al[24] (2010)	1	1	1	1	0	1	1	1	1	1	1	0	1	1	12
Krutsch et al[42] (2016)	1	1	1	1	0	1	1	0	0	1	1	0	0	0	8
LaBella et al[25] (2011)	1	1	0	1	0	1	1	1	1	1	1	0	1	0	10
Le Gall et al[43] (2008)	1	1	1	1	0	0	1	0	0	0	0	0	1	0	6
Levy et al[26] (1997)	1	1	1	1	0	1	1	1	0	0	1	0	1	0	9
Liederbach et al[27] (2008)	1	1	0	0	0	1	1	1	1	1	1	0	0	0	8
Mandelbaum et al[7] (2005)	1	1	1	1	0	1	1	1	1	1	1	0	0	0	10
Messina et al[28] (1999)	1	0	0	1	0	1	1	1	0	0	0	0	0	0	5
Mountcastle et al[29] (2007)	1	1	1	1	0	1	1	1	1	0	1	0	0	0	9
Myklebust et al[44] (2003)	1	1	1	1	0	1	1	1	1	0	1	0	0	0	9
Nagano et al[8] (2011)	1	0	0	1	0	0	0	1	0	0	0	0	0	0	3
Östenberg and Roos[45] (2000)	1	1	1	1	0	1	1	0	0	1	0	0	1	0	8
Pasanen et al[30] (2008)	1	0	0	1	1	1	1	1	1	1	0	1	1	0	10
Petersen et al[31] (2005)	1	1	0	1	0	1	1	1	1	1	1	0	0	0	9
Pfeiffer et al[32] (2006)	1	1	0	1	0	1	1	1	1	1	1	0	0	0	9
Söderman et al[46] (2000)	1	1	1	1	0	1	1	1	1	0	0	0	0	0	8
Stanley et al[39] (2016)	1	1	0	1	0	1	1	1	1	1	0	0	0	0	8
Steffen et al[33] (2008)	1	1	1	1	1	1	1	1	1	1	1	1	0	1	13
Tegnander et al[34] (2008)	1	1	1	1	0	1	1	1	1	0	0	0	0	0	8
Trojian and Collins[35] (2006)	1	1	1	1	0	1	1	1	1	0	1	0	0	0	9
Viola et al[36] (1999)	1	1	1	1	0	1	1	1	1	0	0	0	0	0	8
Waldén et al[37] (2012)	1	1	1	1	0	1	1	1	1	1	1	1	0	1	12
Waldén et al[47] (2011)	1	1	0	1	0	1	1	1	1	0	0	0	0	0	7

[a] 0 = *criterion was absent or not reported*; 1 = *criterion was present*.

Identifying the ACL injury IR associated with fixed-object HIRL sports is especially relevant as it pertains to military training and activities. Over a 7-year period, the IR of ACL injury in US military members of all services was 3.09/1000 person-years for men and 2.29/1000 person-years for women.[49] Investigators[49] noted that service members were at 10 times greater risk of ACL injury than the general population. This increased risk may be partially explained by participation in fixed-object HIRL activities. In contrast to our findings, Owens et al[49] did not find women to be at greater risk of ACL injury than men; however, they reported person-years because they did not have exposure information. In addition, men outnumbered women in their study[49] and, thus, had higher rates of ACL injury. Military service members, especially those participating in regular training that includes fixed-object HIRLs, may benefit from integrative neuromuscular training to mitigate their risk of ACL injury.

In addition to the military application, our findings related to fixed-object HIRL sports are also relevant considering the advent of recreational obstacle-course races (eg, Tough Mudder, Spartan, BattleFrog). These races are based on military training obstacle courses. Currently, no information about the rates of ACL injury associated with these races is available, but our results suggest that participants should exercise caution. For gymnastics, our findings indicated that the unique demands of the sport, including both implement-based activity and high-impact landings after full-body rotation, distinguish the sport from other noncontact sports regarding the ACL injury risk. Hootman et al[1] found that football, a collision sport, resulted in the greatest incidence of ACL injuries in collegiate male athletes. Our findings indicated that fixed-object HIRL sports resulted in ACL injury IRs that were similar to those of collision sports in men (1.75/10 000 versus 1.12/10 000 AEs). The ACL injury IR was more than 3 times greater among women than among men for fixed-object HIRL sports. Considering the likely mechanisms of injury (landing with rotation, stiff-legged landing), this disparity highlights the neuromuscular deficits typically demonstrated by female athletes.[50] Therefore, female athletes participating in fixed-object HIRL sports may benefit the most from preventive strategies.

We also found that female athletes participating in contact sports sustained ACL injuries at 3 times the rate of male athletes in these same sports (IRR = 3.00). These findings are similar to IRRs previously reported[4] for male and female collegiate basketball and soccer players, which

were approximately 3.6 and 2.8, respectively. However, ACL injury IRs did not differ between women and men for collision sports. The lack of a difference in ACL injury IRs between women and men in collision sports and between women in collision and contact sports may be partially explained by the lack of collision-sport participation by women. When participation was equal among women and men (contact sports), the greater ACL injury IR among women was evident. It is possible that not enough studies were available in which researchers investigated ACL injury incidence among both female and male collision athletes to detect a difference in the rates where one truly exists (ie, low statistical power).

In contrast, the ACL injury IRs for men in collision and contact sports differed (1.12/10 000 and 0.87/10 000 AEs, respectively). The sports included in these categories are similar because they require cutting and pivoting, which are dynamic maneuvers known to contribute to noncontact ACL injury mechanisms. Again, these combined findings further support the idea that neuromuscular deficits may contribute to the greater ACL injury IR among women. Although speculative, it was also possible that the men's decreased IR in collision sports compared with contact sports was due to direct-contact blows to the knee based on the nature of the sports.

Whereas our research may provide a robust estimate of sex differences in ACL injury IRs among sport types, it had limitations. The common metric of AE had to be estimated in some cases when exposure was provided in player-hours. This was necessary to include the maximum amount of data possible. As mentioned, we assumed that 2 player-hours were approximately equal to 1 AE, and we used this assumption to generate estimates of AEs. This assumption may have resulted in the overestimation or underestimation of exposure, depending on the sport. We used broad inclusion criteria to capture the greatest amount of information for generating these estimates. The included articles ranged in study quality, and the estimates are only as strong as the evidence on which they are based. However, we believed it was important to capture a wider range of studies to obtain a truer, more robust picture of ACL injury incidence among athletes. In addition, heterogeneity was relatively high (>75%) for the point estimates, indicating that populations that were grouped together may actually have differed. However, this was expected, as different sports have different demands that change the risk of sustaining an ACL injury. Moreover, heterogeneity for the rate ratios was low, and in some cases was 0%, indicating that the results were consistent and potentially generalizable. Given that female participation in collision sports was less prevalent than male participation, we included relatively few studies in which differences in ACL injury IRs between sexes were investigated. We could not control for variables known to contribute to ACL injury, including surface type, anticipation (anticipated event versus unanticipated event), or mechanism of injury (contact versus noncontact) because of a lack of information. Finally, we did not stratify by age or level of play, as those were not aims of this study.

To address these limitations, future researchers should report their findings in the most accurate units possible (player-hours) or should make both player-hours and AEs available to provide the opportunity for meta-analysis. Given that prospective designs allow for real-time data capture, investigators conducting future research in injury epidemiology should use prospective designs. Developing a standard and comprehensive checklist for criteria that should be met when performing or designing a prospective observational cohort study would provide a guide for researchers to achieve maximum study quality. This meta-analysis should be repeated in the future when more ACL injury data are available to permit comparisons of incidence rates among female and male athletes participating in collision and limited-contact sports. Finally, researchers should establish ACL injury IRs within each sport type while controlling for confounding variables, including age and level of play.

CONCLUSIONS

The incidence of ACL injury is associated with the nature of player-to-player contact inherent in the sport. Female athletes had greater ACL injury IRs than male athletes in contact and fixed-object HIRL sports. The latter sports category had the highest ACL injury IRs for both sexes, which might suggest the need for a new sport type to identify athletes at the highest risk of ACL injury. Future strategies aimed at reducing the risk of ACL injury may benefit from considering and integrating sport-related perturbation that mimics contact exposure to better equip athletes with preprepatory and avoidance techniques.

REFERENCES

1. Hootman JM, Dick R, Agel J. Epidemiology of collegiate injuries for 15 sports: summary and recommendations for injury prevention initiatives. *J Athl Train*. 2007;42(2):311–319.
2. Arendt E, Dick R. Knee injury patterns among men and women in collegiate basketball and soccer: NCAA data and review of literature. *Am J Sports Med*. 1995;23(6):694–701.
3. Prodromos CC, Han Y, Rogowski J, Joyce B, Shi K. A meta-analysis of the incidence of anterior cruciate ligament tears as a function of gender, sport, and a knee injury–reduction regimen. *Arthroscopy*. 2007;23(12):1320–1325.
4. Agel J, Arendt EA, Bershadsky B. Anterior cruciate ligament injury in National Collegiate Athletic Association basketball and soccer: a 13-year review. *Am J Sports Med*. 2005;33(4):524–530.
5. Moher D, Liberati A, Tetzlaff J, Altman DG; The PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med*. 2009;6(7):e1000097.
6. Brooks JHM, Fuller CW, Kemp SPT, Reddin DB. Epidemiology of injuries in English professional rugby union: part 2 training injuries. *Br J Sports Med*. 2005;39(10):767–775.
7. Mandelbaum BR, Silvers HJ, Watanabe DS, et al. Effectiveness of a neuromuscular and proprioceptive training program in preventing anterior cruciate ligament injuries in female athletes: 2-year follow-up. *Am J Sports Med*. 2005;33(7):1003–1010.
8. Nagano Y, Miki H, Tsuda K, Shimizu Y, Fukubayashi T. Prevention of anterior cruciate ligament injuries in female basketball players in Japan: an intervention study over four seasons [abstract]. *Br J Sports Med*. 2011;45(4):365.
9. Bjordal JM, Arnly F, Hannestad B, Strand T. Epidemiology of anterior cruciate ligament injuries in soccer. *Am J Sports Med*. 1997;25(3):341–345.
10. Mufty S, Bollars P, Vanlommel L, Van Crombrugge K, Corten K, Bellemans J. Injuries in male versus female soccer players:

epidemiology of a nationwide study. *Acta Orthop Belg*. 2015;81(2):289–295.
11. Ristolainen L, Heinonen A, Waller B, Kujala UM, Kettunen JA. Gender differences in sport injury risk and types of injuries: a retrospective twelve-month study on cross-country skiers, swimmers, long-distance runners and soccer players. *J Sports Sci Med*. 2009;8(3):443–451.
12. Moroder P, Runer A, Hoffelner T, Frick N, Resch H, Tauber M. A prospective study of snowkiting injuries. *Am J Sports Med*. 2011;39(7):1534–1540.
13. Hershman EB, Anderson R, Bergfeld JA, et al. An analysis of specific lower extremity injury rates on grass and FieldTurf playing surfaces in National Football League games: 2000–2009 seasons. *Am J Sports Med*. 2012;40(10):2200–2205.
14. Powell JW, Schootman M. A multivariate risk analysis of selected playing surfaces in the National Football League: 1980 to 1989. An epidemiologic study of knee injuries. *Am J Sports Med*. 1992;20(6):686–694.
15. Rauh MJ, Macera CA, Ji M, Wiksten DL. Subsequent injury patterns in girls' high school sports. *J Athl Train*. 2007;42(4):486–494.
16. Quality assessment tool for observational cohort and cross-sectional studies. National Heart, Lung, and Blood Institute Web site. http://www.nhlbi.nih.gov/health-pro/guidelines/in-develop/cardiovascular-risk-reduction/tools/cohort. Accessed April 7, 2018.
17. Beynnon BD, Vacek PM, Newell MK, et al. The effects of level of competition, sport, and sex on the incidence of first-time noncontact anterior cruciate ligament injury. *Am J Sports Med*. 2014;42(8):1806–1812.
18. Dallalana RJ, Brooks JH, Kemp SPT, Williams AM. The epidemiology of knee injuries in English Professional Rugby Union. *Am J Sports Med*. 2007;35(5):818–830.
19. Dragoo JL, Braun HJ, Durham JL, Chen MR, Harris AH. Incidence and risk factors for injuries to the anterior cruciate ligament in National Collegiate Athletic Association football: data from the 2004–2005 through 2008–2009 National Collegiate Athletic Association Injury Surveillance System. *Am J Sports Med*. 2012;40(5):990–995.
20. Gilchrist J, Mandelbaum BR, Melancon H, et al. A randomized controlled trial to prevent noncontact anterior cruciate ligament injury in female collegiate soccer players. *Am J Sports Med*. 2008;36(8):1476–1483.
21. Gomez E, DeLee JC, Farney WC. Incidence of injury in Texas girls' high school basketball. *Am J Sports Med*. 1996;24(5):684–687.
22. Gwinn DE, Wilckens JH, McDevitt ER, Ross G, Kao TC. The relative incidence of anterior cruciate ligament injury in men and women at the United States Naval Academy. *Am J Sports Med*. 2000;28(1):98–102.
23. Joseph AM, Collins CL, Henke NM, Yard EE, Fields SK, Comstock RD. A multisport epidemiologic comparison of anterior cruciate ligament injuries in high school athletics. *J Athl Train*. 2013;48(6):810–817.
24. Kiani A, Hellquist E, Ahlqvist K, Gedeborg R, Michaelsson K, Byberg L. Prevention of soccer-related knee injuries in teenaged girls. *Arch Intern Med*. 2010;170(1):43–49.
25. LaBella CR, Huxford MR, Grissom J, Kim KY, Peng J, Christoffel KK. Effect of neuromuscular warm-up on injuries in female soccer and basketball athletes in urban public high schools: cluster randomized controlled trial. *Arch Pediatr Adolesc Med*. 2011;165(11):1033–1040.
26. Levy AS, Wetzler MJ, Lewars M, Laughlin W. Knee injuries in women collegiate rugby players. *Am J Sports Med*. 1997;25(3):360–362.
27. Liederbach M, Dilgen FE, Rose DJ. Incidence of anterior cruciate ligament injuries among elite ballet and modern dancers: a 5-year prospective study. *Am J Sports Med*. 2008;36(9):1779–1788.
28. Messina DF, Farney WC, DeLee JC. The incidence of injury in Texas high school basketball: a prospective study among male and female athletes. *Am J Sports Med*. 1999;27(3):294–299.
29. Mountcastle SB, Posner M, Kragh JF Jr, Taylor DC. Gender differences in anterior cruciate ligament injury vary with activity: epidemiology of anterior cruciate ligament injuries in a young, athletic population. *Am J Sports Med*. 2007;35(10):1635–1642.
30. Pasanen K, Parkkari J, Pasanen M, et al. Neuromuscular training and the risk of leg injuries in female floorball players: cluster randomised controlled study. *BMJ*. 2008;337:a295.
31. Petersen W, Braun C, Bock W, et al. A controlled prospective case control study of a prevention training program in female team handball players: the German experience. *Arch Orthop Trauma Surg*. 2005;125(9):614–621.
32. Pfeiffer RP, Shea KG, Roberts D, Grandstrand S, Bond L. Lack of effect of a knee ligament injury prevention program on the incidence of noncontact anterior cruciate ligament injury. *J Bone Joint Surg Am*. 2006;88(8):1769–1774.
33. Steffen K, Myklebust G, Olsen OE, Holme I, Bahr R. Preventing injuries in female youth football: a cluster-randomized controlled trial. *Scand J Med Sci Sports*. 2008;18(5):605–614.
34. Tegnander A, Olsen OE, Moholdt TT, Engebretsen L, Bahr R. Injuries in Norwegian female elite soccer: a prospective one-season cohort study. *Knee Surg Sports Traumatol Arthrosc*. 2008;16(2):194–198.
35. Trojian TH, Collins S. The anterior cruciate ligament tear rate varies by race in professional women's basketball. *Am J Sports Med*. 2006;34(6):895–898.
36. Viola RW, Steadman JR, Mair SD, Briggs KK, Sterett WI. Anterior cruciate ligament injury incidence among male and female professional alpine skiers. *Am J Sports Med*. 1999;27(6):792–795.
37. Waldén M, Atroshi I, Magnusson H, Wagner P, Hägglund M. Prevention of acute knee injuries in adolescent female football players: cluster randomised controlled trial. *BMJ*. 2012;344:e3042.
38. Agel J, Rockwood T, Klossner D. Collegiate ACL injury rates across 15 sports: National Collegiate Athletic Association Injury Surveillance System data update (2004–2005 through 2012–2013). *Clin J Sport Med*. 2016;26(6):518–523.
39. Stanley LE, Kerr ZY, Dompier TP, Padua DA. Sex differences in the incidence of anterior cruciate ligament, medial collateral ligament, and meniscal injuries in collegiate and high school sports: 2009–2010 through 2013–2014. *Am J Sports Med*. 2016;44(6):1565–1572.
40. Faude O, Junge A, Kindermann W, Dvorak J. Injuries in female soccer players: a prospective study in the German national league. *Am J Sports Med*. 2005;33(11):1694–1700.
41. Hägglund M, Waldén M, Ekstrand J. Injuries among male and female elite football players. *Scand J Med Sci Sports*. 2009;19(6):819–827.
42. Krutsch W, Zeman F, Zellner J, Pfeifer C, Nerlich M, Angele P. Increase in ACL and PCL injuries after implementation of a new professional football league. *Knee Surg Sports Traumatol Arthrosc*. 2016;24(7):2271–2279.
43. Le Gall F, Carling C, Reilly T. Injuries in young elite female soccer players: an 8-season prospective study. *Am J Sports Med*. 2008;36(2):276–284.
44. Myklebust G, Engebretsen L, Braekken IH, Skjolberg A, Olsen OE, Bahr R. Prevention of anterior cruciate ligament injuries in female team handball players: a prospective intervention study over three seasons [abstract]. *Scand J Med Sci Sports*. 2003;13(4):272.
45. Östenberg A, Roos H. Injury risk factors in female European football: a prospective study of 123 players during one season. *Scand J Med Sci Sports*. 2000;10(5):279–285.
46. Söderman K, Werner S, Pietilä T, Engström B, Alfredson H. Balance board training: prevention of traumatic injuries of the lower extremities in female soccer players? A prospective randomized

intervention study. *Knee Surg Sports Traumatol Arthrosc*. 2000;8(6):356–363.
47. Waldén M, Hägglund M, Magnusson H, Ekstrand J. Anterior cruciate ligament injury in elite football: a prospective three-cohort study. *Knee Surg Sports Traumatol Arthrosc*. 2011;19(1):11–19.
48. Giza E, Mithöfer K, Farrell L, Zarins B, Gill T. Injuries in women's professional soccer. *Br J Sports Med*. 2005;39(4):212–216.
49. Owens BD, Mountcastle SB, Dunn WR, DeBerardino TM, Taylor DC. Incidence of anterior cruciate ligament injury among active duty US military servicemen and servicewomen. *Mil Med*. 2007;172(1):90–91.
50. Hewett TE, Myer GD, Ford KR, et al. Biomechanical measures of neuromuscular control and valgus loading of the knee predict anterior cruciate ligament injury risk in female athletes: a prospective study. *Am J Sports Med*. 2005;33(4):492–501.

SUPPLEMENTAL MATERIAL

Supplemental Figures. Series of forest plots for incidence rate ratios.

Found at DOI: http://dx.doi.org/10.4085/1062-6050-407-16.S1

*Address correspondence to Gregory D. Myer, PhD, CSCS*D, FACSM, Cincinnati Children's Hospital Medical Center, 3333 Burnet Avenue, MLC 10001, Cincinnati, OH 45229. Address e-mail to greg.myer@cchmc.org.*

Vertical Jump and Leg Power Normative Data for Colombian Schoolchildren Aged 9–17.9 Years: The FUPRECOL Study

Robinson Ramírez-Vélez,[1] Jorge E. Correa-Bautista,[1] Felipe Lobelo,[2] Eduardo L. Cadore,[3] Alicia M. Alonso-Martinez,[4] and Mikel Izquierdo[4,5]

[1]Center of Studies in Physical Activity Measurements, School of Medicine and Health Sciences, University of Rosario, Bogotá, DC, Colombia; [2]Hubert Department of Global Health, Rollins School of Public Health, Emory University, Atlanta, Georgia; [3]Exercise Research Laboratory, School of Physical Education, Federal University of Rio Grande do Sul, Porto Alegre, Brazil; [4]Department of Health Sciences, Public University of Navarra, Pamplona, Spain; and [5]Faculty of Physical Culture, Sport and Recreation, Santo Tomás University, Bogotá, DC, Colombia

Abstract

Ramírez-Vélez, R, Correa-Bautista, JE, Lobelo, F, Cadore, EL, Alonso-Martinez, AM, and Izquierdo, M. Vertical jump and leg power normative data for Colombian schoolchildren aged 9–17.9 years: the FUPRECOL study. *J Strength Cond Res* 31(4): 990–998, 2017—The aims of the present study were to generate normative vertical jump height and predicted peak power (P_{peak}) data for 9- to 17.9-year-olds and to investigate between-sex and age group differences in these measures. This was a cross-sectional study of 7,614 healthy schoolchildren (boys $n = 3,258$ and girls $n = 4,356$, mean [*SD*] age 12.8 [2.3] years). Each participant performed 2 countermovement jumps; jump height was calculated using a Takei 5414 Jump-DF Digital Vertical (Takei Scientific Instruments Co., Ltd.). The highest jump was used for analysis and in the calculation of predicted P_{peak}. Centile smoothed curves, percentiles, and tables for the 3rd, 10th, 25th, 50th, 75th, 90th, and 97th percentiles were calculated using Cole's LMS (L [curve Box-Cox], M [curve median], and S [curve coefficient of variation]) method. The 2-way analysis of variance tests showed that maximum jump height (in centimeters) and predicted P_{peak} (in watts) were higher in boys than in girls ($p < 0.01$). Post hoc analyses within sexes showed yearly increases in jump height and P_{peak} in all ages. In boys, the maximum jump height and predicted P_{peak} 50th percentile ranged from 24.0 to 38.0 cm and from 845.5 to 3061.6 W, respectively. In girls, the 50th percentile for jump height ranged from 22.3 to 27.0 cm, and the predicted P_{peak} was 710.1–2036.4 W. For girls, jump height increased yearly from 9 to 17.9 years old. Our results provide, for the first time, sex- and age-specific vertical jump height and predicted P_{peak} reference standards for Colombian schoolchildren aged 9–17.9 years.

Key Words strength, biomechanics, adolescent, percentile, jumping, muscular power

Introduction

Anaerobic power refers to the ability to perform high-intensity exercise for a fraction of a second to several minutes (e.g., distance jumped, weight lifted) (21). Typically, this energy system is assessed by testing maximal effort in cycling, running, or jumping (7). Jump testing is potentially very useful because many sports and work activities involve either jumping or movements similar to jumping, such as lifting heavy objects ballistically. The vertical jump (VJ) test is a simple method for calculating peak leg power using prediction equations based on jump height and body mass (29). This test measures the vertical displacement of the center of mass between standing on the ground and at the apex of a jump (31). Jumping primarily involves the gluteal and quadriceps muscles, which are instrumental in many sport and work activities (2,11).

Raw measures of jump height may have some use in performance appraisal, but ideally, an estimate of peak leg power should accompany jump height normative data because this provides insight beyond the outcome of the jump itself (29). For example, 2 individuals of different body weight might be able to jump vertically the exact same distance, therefore calculating power values could provide additional information to that given by jump height alone (11). However, the heavier individuals' jump would show the ability to generate greater power, which could provide an advantage in activities which involve manipulation of mass outside the body (e.g., helping to identify talented athletes).

Typically, VJ can be measured using relatively inexpensive, portable, and easy-to-use tools (23,26,28) and is a reliable (2,11,22) and valid (3) method for strength assessment.

Address correspondence to Mikel Izquierdo, mikel.izquierdo@gmail.com.

31(4)/990–998

Journal of Strength and Conditioning Research

TABLE 1. Descriptive statistics for anthropometric and jump height and peak power in 9- to 17.9-year-old Colombian schoolchildren.*†

Sex	n	Body mass (kg)	Height (cm)	Body mass index ($kg \cdot m^{-2}$)	Maximum jump height (cm)	Difference		P_{peak} (W)
						D (cm)	%	
Boys								
9–9.9	236	32.1 (7.5)	133.5 (6.5)	17.8 (3.1)	24.0 (4.9)‡			868.0 (462.4)§
10–10.9	440	34.5 (8.5)	137.3 (7.4)§	18.1 (3.3)	25.4 (5.0)‡	1.4	6.1	1,057.4 (489.7)§
11–11.9	414	37.2 (8.8)§	141.9 (8.2)§	18.3 (3.2)	26.9 (4.7)‡	2.9	6.5	1,277.1 (465.8)§
12–12.9	368	41.3 (9.1)§	147.1 (8.2)§	18.9 (3.2)	27.9 (5.1)‡	3.9	6.7	1,529.7 (520.1)§
13–13.9	379	46.0 (9.8)§	153.5 (9.3)§	19.4 (3.3)‡	30.5 (6.3)‡	6.5	7.3	1,879.8 (588.1)‡
14–14.9	415	50.0 (9.7)§	158.9 (9.1)‡	19.7 (3.0)‡	32.2 (6.6)‡	8.2	7.7	2,174.5 (646.4)‡
15–15.9	407	54.4 (9.7)§	163.3 (8.9)‡	20.3 (3.0)‡	35.1 (6.6)‡	11.1	8.4	2,544.1 (651.6)‡
16–16.9	358	57.7 (8.7)‡	166.7 (7.2)‡	20.8 (2.9)‡	36.2 (6.8)‡	12.2	8.7	2,761.1 (571.4)‡
17–17.9	241	60.8 (10.3)‡	168.1 (7.4)‡	21.5 (3.3)‡	37.6 (7.2)‡	13.6	9.0	2,986.3 (658.5)‡
Total	3,258	45.5 (13.0)§	151.9 (14.1)‡	19.4 (3.3)‡	30.5 (7.4)‡			1,877.5 (887.0)‡
Girls								
9–9.9	308	32.1 (7.4)	134.6 (7.6)	17.6 (3.0)	22.3 (5.0)			773.1 (467.2)
10–10.9	672	35.0 (7.9)	138.4 (7.6)	18.1 (3.0)	24.0 (4.6)	1.7	5.4	995.7 (456.8)
11–11.9	619	38.3 (7.9)	143.7 (7.5)	18.4 (2.9)	24.9 (4.6)	2.6	5.6	1,198.3 (429.4)
12–12.9	506	42.8 (8.6)	148.5 (7.3)	19.3 (3.0)	25.5 (4.6)	3.2	5.7	1,431.0 (483.3)
13–13.9	456	47.4 (9.0)	152.4 (6.3)	20.3 (3.2)	25.5 (5.3)	3.2	5.7	1,628.1 (504.3)
14–14.9	582	51.0 (8.9)	154.6 (6.5)	21.3 (3.3)	26.1 (5.2)	3.8	5.8	1,841.0 (512.8)
15–15.9	493	52.7 (8.6)	155.7 (6.8)	21.7 (3.1)	26.4 (5.1)	4.1	5.9	1,934.4 (483.6)
16–16.9	433	53.9 (8.6)	156.4 (5.8)	22.0 (3.1)	26.7 (6.9)	4.4	6.0	2,010.2 (579.4)
17–17.9	287	55.1 (9.3)	156.8 (6.5)	22.4 (3.6)	27.6 (5.7)	5.3	6.2	2,115.4 (541.6)
Total	4,356	44.8 (11.5)	148.7 (10.1)	20.0 (3.5)	25.4 (5.3)			1,527.7 (647.6)

*Data values are reported as mean (±) (*SD*).
†D = between-year differences (9–9.9, 10–10.9, 11–11.9, 12–12.9, 13–13.9, 14–14.9, 15–15.9, 16–16.9, 17–17.9 years) in absolute scores; % = relative change in jump height.
‡Significant difference between boys and girls within the same age group: $p < 0.0001$.
§Significant difference between boys and girls within the same age group: $p < 0.01$.

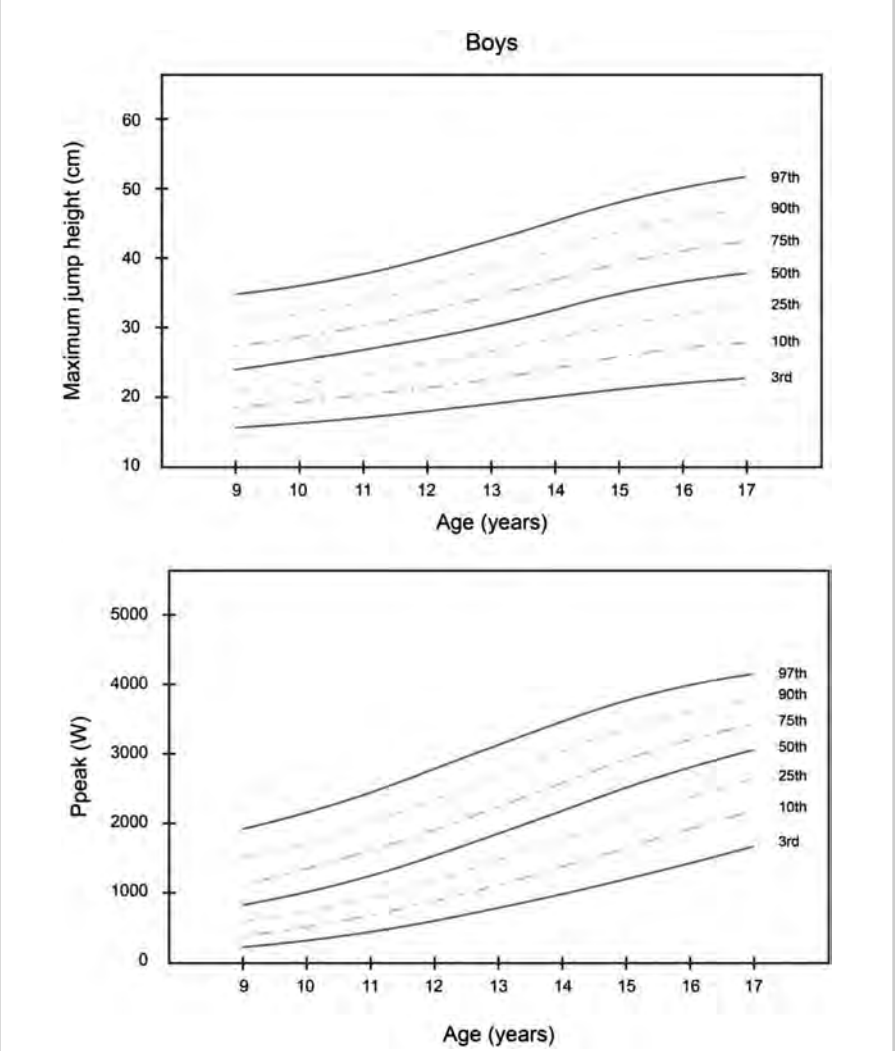


Figure 1. Centile curves for vertical jump height (in centimeters) and predicted peak power (in watts) in boys and age.

Collective VJ can be assessed using various strength performance tests, such as the countermovement jump, standing broad jump test, Abalakov jump test, and Sargent jump and leg press (12,26). Sex- and age-specific normative values for VJ in young people have been published (18,19). For children, only 2 studies attempted to develop normative data tables for the VJ (1,29). For example, the sample used by Bovet et al. (1) was African descent, with minorities of Caucasian, Indian, Chinese, and mixed origins, and it is not clear whether their lower-body explosive power is comparable with the broad range of jump abilities seen in the wider pediatric population. Thus, there is a need to refine and improve the methods used to estimate jumping performance, currently available to practitioners and coaches. To date, predicted peak leg power for jumping has not been reported for Latin American schoolchildren.

However, the majority of published VJ reference values are for schoolchildren from high-income countries in North America (18), Canada (19), Europe (29), and Africa (1). In contrast, there is a scarcity of reference values for children using harmonized measures of fitness in Latin America (9,22) and other low-to-middle-income countries undergoing nutritional transitions (17,22), making it impossible to evaluate secular trends within these regions. There are no such data available for school-aged Colombian children and adolescents.

Therefore, the aims of the present study were to generate normative VJ and predicted peak power data for 9- to 17.9-year-olds and to investigate between-sex and age group differences in these measures.

Methods

Experimental Approach to the Problem

During the 2014–2015 school year, we conducted a cross-sectional component of the FUPRECOL study (in Spanish, ASOCIACIÓN DE LA FUERZA PRENSIL CON MANIFESTACIONES DE RIESGO CARDIOVASCULAR TEMPRANAS EN NIÑOS Y ADOLESCENTES COLOMBIANOS). Briefly, this study aimed to examine the relationships between physical fitness levels in children and adolescents with cardiometabolic risk factors and healthy habits. These data were used to evaluate their health status (20,22,24) and to establish reference values for anthropometric, metabolic, and physical fitness among children and adolescents aged 9–17.9 years in Bogota, Colombia.

A convenience sample of volunteers was included and grouped by sex and age in 1-year increments (a total of 9 groups). Power calculations were based on the mean of maximum jump height (in centimeters) from the first 200 participants in the ongoing data collection (range, 20–35 cm), with a group *SD* of approximately 5.2 cm. The significance level was set to 0.05, and the required power was set to at least 0.80. The sample size was estimated to be

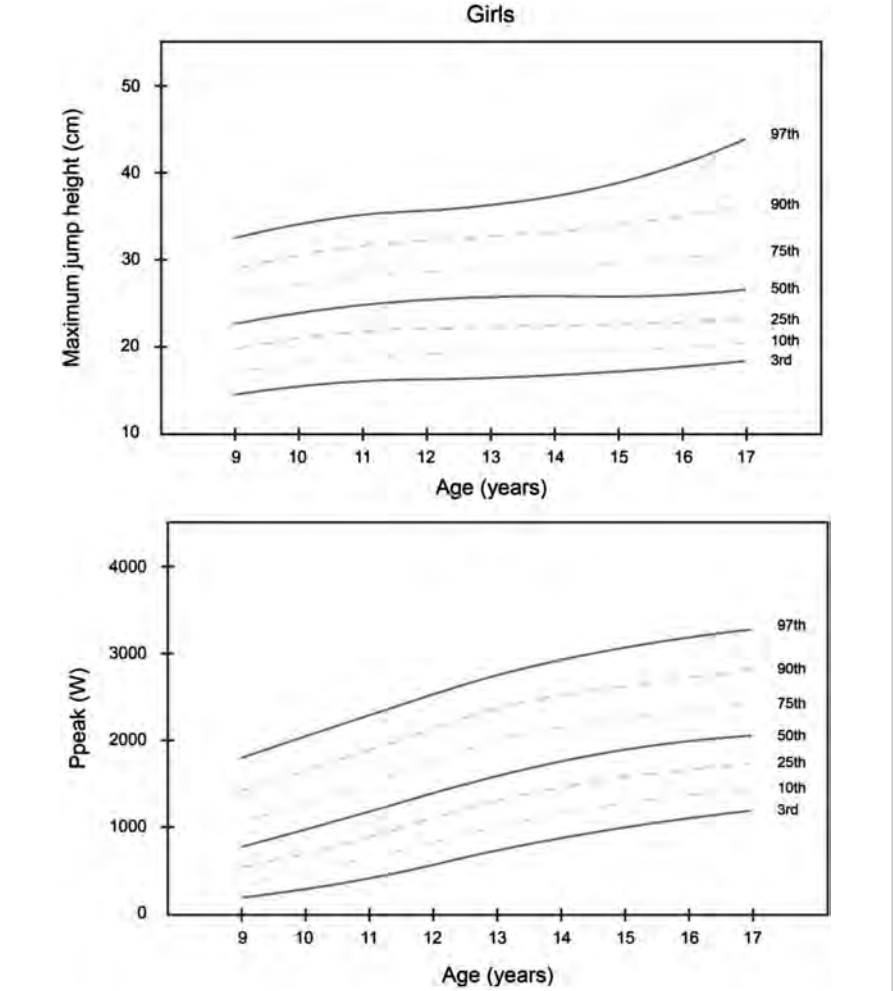


Figure 2. Centile curves for vertical jump height (in centimeters) and predicted peak power (in watts) in girls and age.

approximately 200–400 participants per group. Exclusion factors included a clinical diagnosis of cardiovascular disease, diabetes mellitus type 1 or 2, pregnancy, the use of alcohol or drugs, and, in general, the presence of any disease not directly associated with nutrition. Exclusion from the study was made effective a posteriori and without the students' knowledge to avoid any undesired situations.

Subjects

The sample comprised 7,614 healthy Colombian schoolchildren (boys $n = 3{,}258$ and girls $n = 4{,}356$, mean ± *SD* age 12.8 [2.3] years, body mass 45.1 [12.1] kg, height 1.50 [0.1] m, body mass index (BMI) 19.7 [3.4] $kg \cdot m^{-2}$). The schoolchildren were of low-to-middle socioeconomic status (1–3 as defined by the Colombian government) and enrolled in public elementary and high schools (between grades 5 and 11) in the capital district of Bogota in a municipality in the Cundinamarca Department in the Andean region. This region is located at approximately 4°35′56″N 74°04′51″W and at an elevation of approximately 2,625 m (min: 2,500 m, max: 3,250 m) above sea level. Bogota is considered an urban area with 7,862,277 inhabitants (6).

The study was approved by the institutional review board for use of human subject research in addition to the Rosario University Board (Code Nº CEI-ABN026-000262). Potential subjects and their parents or guardian (s) were informed of the purpose, benefits, and potential risks of the study, and then provided written informed consent to participate. The protocol was in accordance with the latest revision of the Declaration of Helsinki (as revised in Hong Kong in 1989 and in Edinburgh, Scotland, in 2000) and current Colombian laws governing clinical research on human subjects (Resolution 008430/1993 Ministry of Health).

Procedures

Anthropometric variables were measured by a level 2 anthropometrist certified by the International Society for the Advancement of Kinanthropometry (ISAK), in accordance with the ISAK guidelines (15), at the same time (7:00–10:00 AM) in the morning following an overnight fast. Body weight was measured with subjects wearing their underwear and without shoes using electronic scales (Tanita BC544; Tanita, Tokyo, Japan) with a low technical error of measurement (TEM = 0.510%). Height was measured using a mechanical stadiometer platform (Seca 274; Seca Hamburg, Germany; TEM = 0.019%). Body mass index was calculated as the body weight in kilograms divided by the square of height in meters. The data were recorded on paper by the FUPRECOL evaluators (22).

The VJ (in centimeters) was measured using a standard mat (Takei 5414 Jump-DF digital vertical; Takei Scientific Instruments Co., Ltd., Niigata, Japan) and then performed a countermovement jump with arms on waist. Each

Table 2. Vertical jump height (in centimeters) and predicted peak power (in watts) percentiles in boys and in girls by age.*

	n	M	*SD*	P_3	P_{10}	P_{25}	P_{50}	P_{75}	P_{90}	P_{97}
Boys										
Jump height										
9–9.9	236	24.0	4.9	16.6	18.0	20.5	24.0	26.5	29.5	35.3
10–10.9	440	25.4	5.0	16.0	19.5	22.0	25.0	28.5	32.0	35.0
11–11.9	414	26.9	4.7	18.0	21.0	24.0	27.0	30.5	32.5	35.8
12–12.9	368	27.9	5.1	19.0	22.0	24.5	27.5	31.5	34.5	38.0
13–13.9	379	30.5	6.3	19.7	23.0	26.5	30.5	34.5	39.0	44.0
14–14.9	415	32.2	6.6	20.0	23.5	28.0	32.0	36.5	41.5	44.0
15–15.9	407	35.1	6.6	22.5	26.0	31.0	35.5	40.0	43.0	47.0
16–16.9	358	36.2	6.8	23.5	28.0	31.5	36.5	40.5	45.1	49.1
17–17.9	241	37.6	7.2	22.5	28.0	33.0	38.0	42.0	47.0	50.0
Total	3,258	30.5	7.4	18.4	21.5	25.0	30.0	35.5	41.0	45.5
Peak power										
9–9.9	236	868.0	462.4	197.3	349.7	576.4	845.5	1,054.3	1,450.9	1,898.0
10–10.9	440	1,057.4	489.7	249.9	489.2	747.0	1,017.6	1,298.9	1,632.6	2,099.2
11–11.9	414	1,277.1	465.8	464.6	727.5	945.2	1,221.7	1,569.0	1,887.6	2,266.1
12–12.9	368	1,529.7	520.1	638.1	893.7	1,146.3	1,501.0	1,845.7	2,244.8	2,603.6
13–13.9	379	1,879.8	588.1	914.5	1,115.8	1,458.0	1,846.6	2,286.6	2,627.5	2,990.5
14–14.9	415	2,174.5	646.4	987.9	1,362.9	1,758.7	2,140.6	2,595.4	3,066.3	3,388.0
15–15.9	407	2,544.1	651.6	1,246.0	1,682.5	2,125.1	2,576.2	2,971.1	3,332.0	3,785.5
16–16.9	358	2,761.1	571.4	1,597.4	2,006.8	2,438.3	2,756.9	3,170.4	3,495.7	3,770.7
17–17.9	241	2,986.3	658.5	1,399.8	2,167.3	2,525.4	3,061.6	3,396.8	3,728.1	4,226.6
Total	3,258	1,877.5	887.0	474.8	797.2	1,147.9	1,775.7	2,553.0	3,108.1	3,573.6
Girls										
Jump height										
9–9.9	308	22.3	5.0	14.0	16.0	19.0	22.3	25.0	29.0	32.0
10–10.9	672	24.0	4.6	15.5	18.0	20.5	24.0	27.0	29.5	33.0
11–11.9	619	24.9	4.6	16.5	19.5	22.0	25.0	27.5	31.0	33.5
12–12.9	506	25.5	4.6	17.0	20.0	22.5	25.5	28.5	31.5	34.5
13–13.9	456	25.5	5.3	15.5	19.0	22.0	25.5	28.5	32.0	36.1
14–14.9	582	26.1	5.2	17.0	20.0	22.5	25.5	29.5	32.5	36.5
15–15.9	493	26.4	5.1	17.0	20.2	23.0	26.0	29.5	32.5	36.5
16–16.9	433	26.7	6.9	17.0	20.5	23.0	26.5	30.0	33.0	37.0
17–17.9	287	27.6	5.7	19.3	21.5	23.5	27.0	30.5	35.0	41.0
Total	4,356	25.4	5.3	16.0	19.0	22.0	25.0	28.5	31.5	35.5
Peak power										
9–9.9	308	773.1	467.2	114.1	250.1	422.9	710.1	1,018.6	1,358.3	1,806.1
10–10.9	672	995.7	456.8	259.6	443.6	675.9	949.7	1,249.4	1,567.0	2,024.2
11–11.9	619	1,198.3	429.4	455.9	680.2	898.1	1,181.2	1,449.4	1,722.5	2,150.6
12–12.9	506	1,431.0	483.3	655.3	890.4	1,100.1	1,381.9	1,715.7	2,031.0	2,460.9
13–13.9	456	1,628.1	504.3	807.5	1,004.4	1,261.6	1,612.4	1,914.8	2,274.5	2,683.6
14–14.9	582	1,841.0	512.8	962.9	1,211.2	1,477.2	1,792.5	2,191.1	2,529.3	2,905.0
15–15.9	493	1,934.4	483.6	1,059.4	1,319.7	1,624.1	1,942.2	2,226.3	2,559.6	2,898.0
16–16.9	433	2,010.2	579.4	1,117.8	1,357.3	1,670.4	1,998.4	2,288.3	2,674.6	3,106.5
17–17.9	287	2,115.4	541.6	1,254.6	1,505.5	1,737.5	2,036.4	2,387.5	2,908.0	3,369.5
Total	4,356	1,527.7	647.6	397.2	711.7	1,059.0	1,506.0	1,963.9	2,334.3	2,780.8

*M = mean; *SD* = standard deviation; P = percentile.

individual performed 1–3 submaximal practice jumps, then jumped for maximal height 2 times with 1 minute allowed for recovery between attempts. The jump began from a standing position, with plantigrade foot and the leg vertically aligned (i.e., knee angle approximately 180°). On instruction, the countermovement was performed and the knees flexed to approximately 90° before rapid extension and take-off. Landing (initial contact with the jump mat) and the knee angle extended at approximately 180°. If these criteria were not met, the jump was performed again (29). All the personnel were trained in testing and calibration procedures and maintained a calibration log. Two assessors

Table 3. Reference values (50th percentile) for vertical jump height (in centimeters) from cited studies.

Sex and age	FUPRECOL study (n = 7,614)	England (n = 1,845)	Republic of Seychelles* (n = 4,599)
Boys			
9–9.9	24		
10–10.9	25	21	
11–11.9	27	27	
12–12.9	27	30	30
13–13.9	30	32	33
14–14.9	32	36	36
15–15.9	35	37	39
16–16.9	36		
17–17.9	38		
Girls			
9–9.9	22		
10–10.9	24	22	
11–11.9	25	25	
12–12.9	25	27	28
13–13.9	25	26	29
14–14.9	25	28	30
15–15.9	26	28	30
16–16.9	26		
17–17.9	27		

*From Indian Ocean and African region.

were trained in the use of the VJ mat and the implementation of the protocol, which they practiced before the assessments. The VJ measurements in a subsample (n = 229, median age = 12.8 ± 2.4 years, 46.2 ± 12.4 kg, 1.50 ± 0.1 m, 19.9 ± 3.1 kg·m^{-2}) were recorded to ensure reproducibility on the day of the study. The reproducibility of our data was R = 0.88. Intrarater reliability was assessed by determining the intraclass correlation coefficient (intraclass correlation coefficient = 0.85, 95% confidence interval [CI] = 0.75–0.93). The systematic error when the VJ assessments were performed twice was −1.171 (*SD* 10.148) cm (95% CI = −21.063 to 18.720; n = 207) (22).

Each child was allowed 2 jumps using the correct technique. The best jump height was recorded and incorporated into the power prediction equation. The prediction equation of Sayers et al. (28) was used to predict peak leg power: P_{peak} (W): 60.7 × (VJ [cm]) + 45.3 × (body mass [kg]) − 2,055. To date, no prediction equation has been validated for children, thus the Sayers equation was used, which incorporates jump height and the participant's mass. This equation, developed from jumps performed on a force plate, has a reported difference in adults of 2.7% with power calculated from the force plate (28). The Sayers equation is an improvement on the Lewis formula (P_{peak} [W] : $\sqrt{4.9\times\text{body mass[kg]}}\times\sqrt{\text{VJ[m]}\times 9.8}$), which has been reported to underestimate predicted peak power by 70% (12). It has also been recommended as a replacement for the Lewis formula for physical assessment appraisals (19,29).

Statistical Analyses

Anthropometric and VJ characteristics from the study sample are presented as the mean with *SD*. Normality for selected variables was verified using histograms and Q-Q plots. Data were then split by sex; a 2-way analysis of variance (ANOVA) with post hoc tests (Bonferroni) was used to identify differences between age groups within sexes. The LMS method assumes that the outcome variable has a normal distribution after a Box–Cox power transformation is applied, using the LMS method implemented in LMSChartMaker Pro Version 2.54 (Medical Research Council, London, United Kingdom, http://www.healthforallchildren.com/shop-base/software/lmschartmaker-light/). Smoothed and specific curves for each age were obtained via a penalized maximum likelihood with the following abbreviations: M (median), L (Box–Cox transformation), and S (coefficient of variation) (5). The appropriate number of degrees of freedom was selected on the basis of the deviance, Q tests, and worm plots, following the suggestions of Royston and Wright (25). The 3rd, 10th, 25th, 50th, 75th, 90th, and 97th smoothing centiles were chosen as age- and gender-specific reference values. We used SPSS V. 21.0 software for Windows (SPSS, Chicago, IL, USA) for all but the LMS method calculations. Statistical significance was set at $p \leq 0.05$.

Results

Descriptive Characteristics

Descriptive statistics by gender are shown in Table 1. All the anthropometric variables, except the BMI (aged 9–12.9 years), were higher in boys than in girls ($p < 0.01$). The 2-way ANOVA tests showed that maximum jump height (in centimeters) and predicted P_{peak} (in watts) were higher in boys than in girls ($p < 0.01$). Post hoc analyses within sexes showed yearly increases in jump height and P_{peak} in all ages.

Centile Curves and Reference Values

Smoothed LMS curves (3rd, 10th, 25th, 50th, 75th, 90th, and 97th percentile) for boys and girls of the maximum jump height (in centimeters) and predicted P_{peak} (in watts) are illustrated in Figures 1 and 2. The equivalent numerical values are available in Table 2. Together, these data show that boys performed better on the tests at all ages compared with girls. In boys, the maximum jump height and predicted P_{peak} 50th percentile ranged from 24.0 to 38.0 cm and from 845.5 to 3061.6 W, respectively. In girls, the 50th percentile for jump height was 22.3–27.0 cm, and the predicted P_{peak} was 710.1–2,036.4 W. For girls, jump height increased yearly from 9 to 17.9 years of age.

Vertical Jump Height Differences: Comparisons With Previous Research

Finally, comparisons between the 50th percentile values for VJ height (in centimeters) from this study are presented in Table 3. We found that Colombian schoolchildren have lower values than children and adolescents from England and the Republic of Seychelles.

Discussion

This study aimed to generate normative VJ and predicted peak power (P_{peak}) data for 9- to 17.9-year-olds from Bogota, Colombia, and to investigate between-sex and age group differences in these outcomes. The main findings of the present study were that maximum jump height (in centimeters) and predicted P_{peak} (in watts) gradually increased in all ages and were higher in boys than in girls. These results are important because they provide normative values for anaerobic power for Colombian children and adolescents; this variable is strongly associated with functional status and motor performance in youth.

Our data are based on samples of 200–600 schoolchildren of each sex by age group and thus may better describe the patterns of VJ in both genders. England (27) and Republic of Seychelles (1) studies have used large samples, comprising 1,845 (10–16 years old) and 4,599 subjects (12–16 years old), respectively, but contain no data regarding Colombian children and adolescents. We observed moderate but significant differences (5%) between the sexes in 16- to 17.9-years-olds, which increased to 13% (boys) and 5% (girls) by ages 17–17.9 years. In adolescents (aged 14– 16.9 year), the latter magnitude of between-gender differences is similar to subjects from Republic of Seychelles (8–12%) but lower than other England sample (2–3%). In children (aged 10–12.9 years), we observed small but not significant differences (1–3%), similar to findings reported in UK schoolchildren (4%). The differences may reflect higher anaerobic fitness among international samples, fundamental differences in testing protocols, body composition or stage maturation, or some combination of explanations. In the context of VJ, in particular, Taylor (29) and Malina et al. (14) have noted that neglecting the body size effect in such movements may result in inconsistent or incorrect conclusions being drawn from such data.

The data reported here showing yearly increases in jump height are in agreement with the results reported in studies assessing VJ performance in children from primary and secondary school (10–15 years) from the East of England (29) and in secondary school children (12–15 years) from the Republic of Seychelles (1). In addition, other studies assessing children through different jump tests, such as the standing broad jump, have also shown gradual increases in jump performance in all ages (27,30). In the present study, there was an increase in jump height through maturation, especially in boys, whereas girls showed a trend toward a plateau in jump height between 14 and 16 years of age and increased further in all percentiles between 15 and 17 years. This plateau in jump height is consistent with the findings of Taylor et al. (29), who found a VJ plateau after 12 years, and also with the findings of Fortier et al. (8), who reported an increase in standing long jump performance for girls up to the age of 12 years, followed by a plateau.

Although we observed a plateau in jump height for girls between 14 and 16 years, the same was not observed in P_{peak}, which is most likely associated with gradual increases in the body mass for all ages in boys and girls (Table 1). Body mass and jump height are incorporated into the Sayers equation to predict peak. Either component results in improved P_{peak} performance. This interaction among explosive leg power, dynamic leg strength, stretch-shortening cycle ability, and arm–leg coordination (13,16). Therefore, all these factors are expected to improve during biological maturation, which explains the yearly increases in P_{peak} in boys and girls observed in the present study.

In the present study, boys jumped significantly higher than girls at all ages, with greater differences observed after the age of 13. This difference between boys and girls has been shown previously for VJ performance (1,4,29) and also for other lower-limb power tests (1,4,10,22). The marked differences observed especially after the age of 13 may be explained by greater increases in the lean mass in boys and greater increases in adiposity in girls because of puberty (14).

The VJ height of the Colombian children assessed in the present study was lower at most ages when compared with children from England (29) and the Republic of Seychelles (1). It is not clear why such differences exist, but factors such as local physical activity levels, nutritional outcomes, and jump assessment criteria could partially explain these differences. However, the influence of these factors remains speculative and should be investigated further.

This is the first study to provide measures of anaerobic power values America. It should be highlighted that VJ height and power are outcomes associated with several other neuromuscular outcomes in children (4). Moreover, poor anaerobic power is associated with cardiometabolic health in youth and with the risk of future functional decline, morbidity, and mortality (2,11).

This study had some limitations. First, it included participants from only a single region in Colombia; therefore, inferences to all Colombian children and adolescents should be made cautiously. Second, we have not considered the potential impact of recognized determinants of VJ, such as diet, physical activity patterns, sex hormone levels, sexual maturation, and ethnic factors that modulate growth and levels of anaerobic capacity. However, because our study was cross-sectional, a cohort effect may have occurred, and as a consequence, our estimations of jumping power levels could not be extrapolated from previous cohorts (5,6,10,12,29). This is an

area for future research. Nevertheless, such limitations did not compromise the results obtained here because they were similar regarding total score by gender and similar to those reported in other studies carried out in Colombian children (10).

Practical Applications

The VJ has become one of the most convenient tests used to evaluate anaerobic capacity and the effectiveness of anaerobic training programs for a variety of power sports. However, its use and interpretation as an evaluative measurement in physical activity tests are limited because there are few published reference values available for children and adolescents. Our results provide, for the first time, sex- and age-specific VJ height and predicted P_{peak} reference standards for Colombian schoolchildren aged 9–17.9 years. The data provided in the present study will be useful in the assessment of risk of poor health outcomes in youth and in the identification of schoolchildren who are well suited for anaerobic performance.

Acknowledgments

The FUPRECOL Study was carried out with the financial support of Instituto Colombiano para el Desarrollo de la Ciencia y la Tecnología "Francisco José de Caldas" COLCIENCIAS (Contract Nº 671-2014; Code 122265743978). The authors are grateful to the Bogota District Education Secretary for the data used in this study. The authors also thank the participating Bogota District students, teachers, schools, and staff. The authors disclose funding received for this work from any of the following organizations: National Institutes of Health, Welcome Trust, Howard Hughes Medical Institute, and other(s).

References

1. Bovet, P, Auguste, R, and Burdette, H. Strong inverse association between physical fitness and overweight in adolescents: A large school-based survey. *Int J Behav Nutr Phys Act* 4: 24, 2007.
2. Bui, HT, Farinas, MI, Fortin, AM, Comtois, AS, and Leone, M. Comparison and analysis of three different methods to evaluate vertical jump height. *Clin Physiol Funct Imaging* 35: 203–209, 2015.
3. Casartelli, N, Müller, R, and Maffiuletti, NA. Validity and reliability of the Myotest accelerometric system for the assessment of vertical jump height. *J Strength Cond Res* 24: 3186–3193, 2010.
4. Castro-Piñero, J, Artero, EG, España-Romero, V, Ortega, FB, Sjöström, M, Suni, J, and Ruiz, JR. Criterion-related validity of field-based fitness tests in youth: A systematic review. *Br J Sports Med* 44: 934–943, 2010.
5. Cole, TJ and Green, PJ. Smoothing reference centile curves: The LMS method and penalized likelihood. *Stat Med* 11: 1305–1319, 1992.
6. The National Administrative Department of Statistics (DANE). Analysis of the results by ethnic group in 2005 [in Spanish]. Bogotá, Colombia: The National Administrative Department of Statistics (DANE), 2007.
7. Driss, T and Vandewalle, H. The measurement of maximal (anaerobic) power output on a cycle ergometer: A critical review. *Biomed Res Int* 2013: 589361, 2013.
8. Fortier, MD, Katzmarzyk, PT, Malina, RM, and Bouchard, C. Seven-year stability of physical activity and musculoskeletal fitness in the Canadian population. *Med Sci Sports Exerc* 33: 1905–1911, 2001.
9. González, SA, Sarmiento, OL, Cohen, DD, Camargo, DM, Correa, JE, Páez, DC, and Ramírez-Vélez, R. Results from Colombia's 2014 report card on physical activity for children and youth. *J Phys Act Health* 11: S33–S44, 2014.
10. Gulías-González, R, Sánchez-López, M, Olivas-Bravo, Á, Solera-Martínez, M, and Martínez-Vizcaíno, V. Physical fitness in Spanish schoolchildren aged 6-12 years: reference values of the battery EUROFIT and associated cardiovascular risk. *J Sch Health* 84: 625–635, 2014.
11. Gutiérrez-Davila, M, Campos, J, and Navarro, E. A comparison of two landing styles in a two-foot vertical jump. *J Strength Cond Res* 23: 325–331, 2009.
12. Harman, EA, Rosenstein, MT, Frykman, PN, Rosenstein, RM, and Kramer, WJ. Estimates of human power output from vertical jump. *J Appl Sport Sci Res* 5: 116–120, 1991.
13. Leard, JS, Cirillo, MA, Katsnelson, E, Kimiatek, DA, Miller, TW, Trebincevic, K, and Garbalosa, JC. Validity of two alternative systems for measuring vertical jump height. *J Strength Cond Res* 21: 1296–1299, 2007.
14. Malina, RM, Beunen, GP, Classens, AL, Lefevre, J, Vanden Eynde, BV, Renson, R, Vanreusel, B, and Simons, J. Fatness and physical fitness of girls 7 to 17 years. *Obes Res* 3: 221–231, 1995.
15. Marfell-Jones, M, Olds, T, and Stewart, A. *International standards for anthropometric assessment*. Potchefstroom, South Africa: ISAK, 2006, pp 44–109.
16. Markovic, G, Dizdar, D, Jukic, I, and Cardinale, M. Reliability and factorial validity of squat and countermovement jump tests. *J Strength Cond Res* 18: 551–555, 2004.
17. Parra, DC, Iannotti, L, Gomez, LF, Pachón, H, Haire-Joshu, D, Sarmiento, OL, Kuhlmann, AS, and Brownson, RC. The nutrition transition in Colombia over a decade: A novel household classification system of anthropometric measures. *Arch Public Health* 73: 12, 2015.
18. Patterson, DD and Peterson, DF. Vertical jump and leg power norms for young adults. *Meas Phys Educ Exerc Sci* 81: 33–41, 2004.
19. Payne, N, Gledhill, N, Katzmarzyk, PT, Jamnik, VK, and Keir, PJ. Canadian Musculoskeletal fitness norms. *Can J Appl Physiol* 25: 430–442, 2000.
20. Prieto-Benavides, DH, Correa-Bautista, JE, and Ramírez-Vélez, R. Physical activity levels, physical fitness and screen time among children and adolescents from Bogotá, Colombia [in Spanish]: Estudio FUPRECOL. *Nutr Hosp* 32: 2184–2192, 2015.
21. Ramírez-Vélez, R, López-Albán, CA, La Rotta-Villamizar, DR, Romero-García, JA, and Izquierdo, M. Wingate anaerobic test percentile norms in colombian healthy adults. *J Strength Cond Res* 30: 217–225, 2016.
22. Ramírez-Vélez, R, Rodrigues-Bezerra, D, Correa-Bautista, JE, Izquierdo, M, and Lobelo, F. Reliability of Health-Related Physical Fitness Tests among Colombian Children and Adolescents: The FUPRECOL Study. *PLoS One* 10: e0140875, 2015.
23. Reiser, RF, Rocheford, EC, and Armstrong, CJ. Building a better understanding of basic mechanical principles through analysis of the vertical jump. *J Strength Cond Res* 28: 70–80, 2006.
24. Rodriguez-Bautista, YP, Correa-Bautista, JE, Gonzalez-Jimenez, E, Schmidt RioValle, J, and Ramirez Velez, R. Values of waist/hip ratio among children and adolescents from Bogotá, Colombia: The FUPRECOL study [in Spanish]. *Nutr Hosp* 32: 2054–2061, 2015.
25. Royston, P and Wright, EM. Goodness-of-fit statistics for age-specific reference intervals. *Stat Med* 19: 2943–2962, 2000.
26. Ruiz, JR, Castro-Piñero, J, España-Romero, V, Artero, EG, Ortega, FB, Cuenca, MM, Jimenez-Pavón, D, Chillón, P, Girela-Rejón, MJ, Gutiérrez, A, Suni, J, Sjöström, M, and Castillo, MJ. Field-based fitness assessment in young people: The ALPHA health-related fitness test battery for children and adolescents. *Br J Sports Med* 45: 518–524, 2011.
27. Saint-Maurice, PF, Laurson, KR, Kaj, M, and Csányi, T. Establishing normative reference values for standing broad jump among Hungarian youth. *Res Q Exerc Sport* 86: S37–S44, 2015.

28. Sayers, SP, Harackiewicz, DV, Harman, EA, Frykman, PN, and Rosenstein, MT. Cross-validation of three jump power equations. *Med Sci Sports Exerc* 31: 572–577, 1999.

29. Taylor, MJ, Cohen, D, Voss, C, and Sandercock, GR. Vertical jumping and leg power normative data for English school children aged 10–15 years. *J Sports Sci* 28: 867–872, 2010.

30. Yan-Chung, LM, Yu-Chow, LP, and Yee-Chung, JW. Normative reference of standing long jump indicates gender difference in lower muscular strength of pubertal growth. *Health* 5: 6–11, 2013.

31. Zainal-Abidin, N and Bakri-Adam, M. Prediction of vertical jump height from anthropometric factors in male and female martial arts athletes. *Malays J Med Sci* 20: 39–45, 2013.

Health and Care Utilization of Transgender and Gender Nonconforming Youth: A Population-Based Study

G. Nicole Rider, PhD,[a] Barbara J. McMorris, PhD,[b] Amy L. Gower, PhD,[c] Eli Coleman, PhD,[a] Marla E. Eisenberg, ScD, MPH[c]

abstract

BACKGROUND: Transgender and gender nonconforming (TGNC) adolescents have difficulty accessing and receiving health care compared with cisgender youth, yet research is limited by a reliance on small and nonrepresentative samples. This study's purpose was to examine mental and physical health characteristics and care utilization between youth who are TGNC and cisgender and across perceived gender expressions within the TGNC sample.

METHODS: Data came from the 2016 Minnesota Student Survey, which consisted of 80 929 students in ninth and 11th grade (n = 2168 TGNC, 2.7%). Students self-reported gender identity, perceived gender expression, 4 health status measures, and 3 care utilization measures. Chi-squares and multiple analysis of covariance tests (controlling for demographic covariates) were used to compare groups.

RESULTS: We found that students who are TGNC reported significantly poorer health, lower rates of preventive health checkups, and more nurse office visits than cisgender youth. For example, 62.1% of youth who are TGNC reported their general health as poor, fair, or good versus very good or excellent, compared with 33.1% of cisgender youth (χ^2 = 763.7, P < .001). Among the TGNC sample, those whose gender presentation was perceived as very congruent with their birth-assigned sex were less likely to report poorer health and long-term mental health problems compared with those with other gender presentations.

CONCLUSIONS: Health care utilization differs between TGNC versus cisgender youth and across gender presentations within TGNC youth. With our results, we suggest that health care providers should screen for health risks and identify barriers to care for TGNC youth while promoting and bolstering wellness within this community.

[a]*Program in Human Sexuality, Department of Family Medicine and Community Health,* [c]*Division of General Pediatrics and Adolescent Health, Department of Pediatrics, School of Medicine, and* [b]*School of Nursing, University of Minnesota, Minneapolis, Minnesota*

Dr Rider assisted with conceptualizing and designing this study, conducted data analyses and interpretation, drafted the initial manuscript, and revised the manuscript; Drs McMorris, Gower, Coleman, and Eisenberg assisted with conceptualizing and designing the study and interpreting the data and reviewed and revised the manuscript; and all authors approved the final manuscript as submitted and agree to be accountable for all aspects of the work.

Preliminary findings from this study were summarized in a symposium presentation at the Rainbow Health Initiative's Opportunity Conference; February 27, 2017; Saint Paul, MN.

DOI: https://doi.org/10.1542/peds.2017-1683

Accepted for publication Nov 30, 2017

WHAT'S KNOWN ON THIS SUBJECT: Transgender and gender nonconforming (TGNC) adolescents are significantly affected by mental health disparities and have difficulty accessing and receiving health care compared with cisgender youth. Previous research in this field is limited by reliance on small, nonrepresentative, and adult samples.

WHAT THIS STUDY ADDS: TGNC adolescents reported poorer health, fewer health checkups, and more nurse visits than their cisgender peers. TGNC adolescents whose gender expression strongly matched their birth-assigned sex had better health and fewer long-term mental health problems compared with other gender presentations.

To cite: Rider GN, McMorris BJ, Gower AL, et al. Health and Care Utilization of Transgender and Gender Nonconforming Youth: A Population-Based Study. *Pediatrics.* 2018;141(3): e20171683

Youth who are transgender have a gender identity and/or expression differing from societal expectations based on their birth-assigned sex, whereas youth who are cisgender have a gender identity aligning with their birth-assigned sex. Gender nonconforming describes individuals whose gender expression does not follow stereotypical conventions of masculinity and femininity and who may or may not identify as transgender.[1] Although research on youth who are transgender and gender nonconforming (TGNC) is in its nascence, studies indicate that adolescents who identify as TGNC versus cisgender experience significant mental health disparities.[2] Additional studies are needed to better understand other health risks, disparities, and access to health care among youth who are TGNC.

A paucity of health research examines TGNC adolescents' perceived gender expression (ie, the way others interpret a person's gender presentation along a spectrum from feminine to masculine). The authors of the gender minority stress and resilience model[3] theorize that misperceptions of a person's gender expression may result in a young individual feeling as if their gendered experience is negated or not affirmed. The young individual may also be placed at an elevated risk for harassment and victimization, which in turn may contribute to a heightened risk for negative health outcomes, such as depressive symptoms, self-harm, posttraumatic stress, disordered eating, and suicidal ideation and attempts.[4–9] For example, Roberts et al[10] found that youth who reported childhood gender nonconformity were at heightened risk for depressive symptoms during adolescence and early adulthood compared with those reporting childhood gender conformity. Birth-assigned males who reported childhood gender nonconformity were at the greatest risk for bullying victimization and depressive symptoms.

This vulnerability for poorer health outcomes reveals the importance of access to affordable, competent health care services for youth who are TGNC. However, historical marginalization in health care settings and a lack of competent providers create barriers to treatment and contribute to delayed access to care and longer-term health consequences.[11–18] For example, Gordon et al[19] found that gender nonconformity was associated with an increased risk for problems with mobility, usual activities, pain or discomfort, anxiety, and depression. Health scores were lower for participants with moderate gender conformity and lowest for those with low gender conformity when compared with participants reporting high gender conformity. Given the limited research on perceived gender nonconformity and health outcomes, Wylie et al[20] emphasized the importance of assessing perceived gender expression as a determinant of health disparities, particularly in population-based studies.

Previously, researchers have most often dichotomized gender into binary categories (exclusively masculine [man or boy] or feminine [woman or girl],[13] which minimizes the complexity of TGNC identities.[21] In a recent study, researchers found that 41% of a Canadian TGNC sample identified as gender nonbinary (ie, identifying as both, neither, or somewhere between masculine and feminine), which illustrates the importance of investigating the heterogeneity of gender identities and expressions among this group.[22] Health researchers who do not incorporate options to indicate nonbinary gender identities and expressions are at risk for having categories that misclassify or exclude certain gender diverse participants.[21] This categorical invisibility and erasure of diverse gender identities and expressions contribute to a lack of knowledge and training for health care providers and thereby place youth who are TGNC at risk for poorer health outcomes. In the current study, we address these concerns and illuminate health-related disparities in this underserved youth population.

Limitations in the extant research include reliance on samples of adults,[2] convenience samples, and small sample sizes. Population-based studies with large samples of adolescents are needed to generalize findings and make accurate comparisons between gender identity groups (TGNC versus cisgender). Our purpose in this study was to examine the prevalence of mental and physical health concerns and health care utilization among youth who identify as TGNC versus cisgender and across perceived gender expressions within our TGNC sample, using a large-scale, population-based sample.

METHODS

Data Source and Study Design

Data are from the Minnesota Student Survey (MSS), a statewide surveillance system coordinated by the Departments of Education, Health, Human Services, and Public Safety that is used to assess health and well-being among select grades of public school students. In 2016, 85% of the state's school districts participated. Passive parental consent procedures were used in accordance with federal laws. The analytic sample was composed of 80 929 students in ninth and 11th grade who were asked about their gender identity. The University of Minnesota's Institutional Review Board determined that this secondary analysis of existing anonymous data was exempt from review.

Survey Measures

Gender identity was assessed by using a modified version of the validated 2-item approach recommended by transgender health experts.[23–26] Birth-assigned sex was assessed by the question, "What is your biological sex?" (male or female), followed by gender identity: "Do you consider yourself transgender, genderqueer, genderfluid, or unsure about your gender identity?" (yes or no). Adolescents who provided an affirmative response to the gender identity measure comprised the TGNC group. Perceived gender expression was measured by combining 2 items validated with young adults[20] to create the following item: "A person's appearance, style, dress, or the way they walk or talk may affect how people describe them. How do you think other people at school would describe you?" (response options: very or mostly feminine, somewhat feminine, equally feminine and masculine, somewhat masculine, or very or mostly masculine).

Dependent variables included health status (general health, long-term physical health problems, long-term mental health problems, and staying home sick from school) and care utilization (nurse office visits and preventive medical and dental checkups). A description of these measures is presented in Table 1. Notably, response options for general health were dichotomized into "very good or excellent" versus "poor, fair, or good" because of a skewed distribution.

Demographics and personal characteristics included 4 variables. Students were asked their grade and to endorse 1 or more of 5 racial groups and whether they self-identified with a Hispanic ethnicity. Responses were combined to create a race and/or ethnicity variable (Hispanic or Latino; American Indian or Alaskan Native [non-Hispanic]; Asian [non-Hispanic]; Black, African, or African American [non-Hispanic]; Native Hawaiian or other Pacific Islander [non-Hispanic]; White [non-Hispanic]; and multiple race [non-Hispanic]). An indicator of poverty included whether students received free or reduced-price lunch at school. School location was coded as within or outside the 7-county Minneapolis and St. Paul metropolitan area.

TABLE 1 MSS Health Status and Health Care Utilization Measures

Measure	Survey Item	Dichotomized Responses
Health status		
General health	How would you describe your health in general?	1 = poor, fair, or good[a] 0 = very good or excellent
Long-term physical disabilities or health problems	Do you have any physical disabilities, or long-term health problems (such as asthma, cancer, diabetes, epilepsy, or something else)? Long-term means lasting 6 months or more	1 = yes 0 = no
Long-term mental health problems	Do you have any long-term mental health, behavioral, or emotional problems? Long-term means lasting 6 months or more	1 = yes 0 = no
Stayed home sick (last 30 days)	During the last 30 days, how many times have you. . .stayed home because you were sick?	1 = 1+ times 0 = none
Health care utilization		
Nurse office visits (last 30 days)	During the last 30 days, how many times have you. . .gone to the nurse's office?	1 = 1+ times 0 = none
Preventive medical checkup	When was the last time you saw a doctor or nurse for a checkup or physical examination when you were not sick or injured?	1 = during the last year[b] 0 = not in the last year
Preventive dental checkup	When was the last time you saw a dentist or dental hygienist for a regular checkup, examination, teeth cleaning, or other dental work?	1 = during the last year[c] 0 = not in the last year

[a] Response options for general health were dichotomized because of a skewed distribution.
[b] Following recommendations by the American Academy of Pediatrics for wellness checkups.[27]
[c] Following recommendations by the American Academy of Pediatric Dentistry for regular checkups.[28]

Data Analysis

Analyses were conducted by using IBM SPSS version 23 (IBM Corporation, Armonk, NY). First, χ^2 tests were used to compare demographic characteristics, health status, and care utilization measures between students who are TGNC and cisgender. A 2-sided significance level of .001 was selected to reduce type I error rate because of the large sample. Second, multiple analysis of covariance (MANCOVA) models were used to estimate least squares means of the 4 health status variables simultaneously and then the 3 care utilization variables simultaneously for TGNC students by their perceived gender expression while controlling for grade, free or reduced-price lunch, race and/or ethnicity, and school location. Pillai's trace value statistic was used to assess the significant effects of perceived gender expression and control variables on the dependent variables. For dichotomous dependent variables, adjusted least squares means can be interpreted as predicted probabilities. Analyses were conducted separately for birth-assigned male and birth-assigned female adolescents who are TGNC by using an α level of .05. Bonferroni tests were used to correct α for all post hoc comparisons between perceived gender expression groups.

RESULTS

Sample Characteristics

Participants included 2168 (2.7%) students who identified as TGNC

and 78 761 (97.3%) students who identified as cisgender. As shown in Table 2, the TGNC sample included a higher proportion of those assigned female at birth, youth of color, and those receiving free or reduced-price lunch than the cisgender sample. No significant differences emerged between students in metropolitan versus nonmetropolitan locations.

Health Statuses and Care Utilization Between Adolescents Who Are TGNC Versus Cisgender

Almost two-thirds (62.1%) of youth who are TGNC reported their general health as poor, fair, or good as opposed to very good or excellent, which is nearly twice the rate among youth who identify as cisgender (33.1%, $P < .001$; Table 3). Over half (59.3%) of youth who are TGNC also endorsed having long-term mental health problems compared with 17.4% of cisgender youth ($P < .001$). Over half (51.5%) of youth who are TGNC reported staying home from school because of illness at least once in the past month (versus 42.6% of youth who are cisgender; $P < .001$). Youth who are TGNC visited the nurse's office more often and reported lower rates of preventive medical and dental checkups during the last year than their cisgender peers.

Health Status and Care Utilization by Birth-Assigned Sex and Perceived Gender Expression

Perceived gender expression among youth who are TGNC is shown in Table 4. We found that youth who are TGNC varied across perceived gender expressions. Notably, the prevalence of TGNC adolescents with an equally feminine and masculine perceived gender expression was highest for both those assigned male (29.3%) and assigned female (41.2%) at birth compared with other perceived gender presentations. In Table 5, we present predicted probabilities and pairwise comparisons for youth who are TGNC by perceived gender expression, stratified by birth-assigned sex. As indicated by Pillai's trace, there was a significant effect of perceived gender expression for students who are TGNC and assigned male at birth on health status measures ($P < .001$) after controlling for covariates. Statistically significant differences between at least 2 groups on general health and long-term mental health problems were indicated in our results.

TABLE 2 Demographic Characteristics of MSS Participants by Gender Identity ($N = 80\,929$)

	TGNC, *n* (%)	Cisgender, *n* (%)	*P*[a]
Birth-assigned sex			<.001
Male	684 (31.9)	40 014 (50.9)	
Female	1457 (68.1)	38 639 (49.1)	
Grade			.001
Ninth	1271 (58.6)	43 368 (55.1)	
11th	897 (41.4)	35 393 (44.9)	
Race and/or ethnicity			.001
American Indian or Alaskan Native, NH	44 (2.1)	805 (1.0)	
Asian, NH	181 (8.5)	4677 (6.0)	
Black, African, or African American, NH	140 (6.5)	4545 (5.8)	
Native Hawaiian or other Pacific Islander, NH	11 (0.5)	117 (0.1)	
White, NH	1257 (58.7)	55 962 (71.5)	
Multiple race, NH	252 (11.8)	5319 (6.8)	
Hispanic or Latino	255 (11.9)	6816 (8.7)	
Free or reduced-price lunch			.001
Yes	834 (38.8)	20 936 (26.8)	
No	1315 (61.2)	57 226 (73.2)	
Location			.148
Twin Cities metropolitan area	1188 (54.8)	41 921 (53.2)	
Nonmetropolitan	980 (45.2)	36840 (46.8)	

NH, non-Hispanic.
[a] χ^2 tests of associations were used to examine differences in demographic factors.

TABLE 3 Health Status and Care Utilization of MSS Participants by Gender Identity ($N = 80\,929$)

	TGNC ($n = 2168$), *n* (%)	Cisgender ($n = 78\,761$), *n* (%)	*P*[a]
Health status			
General health			<.001
Poor, fair, or good	1299 (62.1)	25 496 (33.1)	
Very good or excellent	793 (37.9)	51 504 (66.9)	
Long-term physical disabilities or health problems			<.001
Yes	522 (25.2)	11 633 (15.2)	
No	1551 (74.8)	65 050 (84.8)	
Long-term mental health problems			<.001
Yes	1220 (59.3)	13 304 (17.4)	
No	838 (40.7)	63 096 (82.6)	
Stayed home sick (last 30 days)			<.001
1+ times	1096 (51.5)	33 367 (42.6)	
None	1031 (48.5)	44 871 (57.4)	
Care utilization			
Nurse office visits (last 30 days)			<.001
1+ times	877 (41.2)	20 298 (25.9)	
None	1252 (58.8)	57 954 (74.1)	
Preventive medical check-up			<.001
During the last year	1248 (60.0)	49 570 (64.7)	
Not in the last year	832 (40.0)	27 052 (35.3)	
Preventive dental check-up			<.001
During the last year	1477 (71.1)	62 854 (82.0)	
Not in the last year	601 (28.9)	13 803 (18.0)	

[a] χ^2 tests of associations were used to examine differences in health status and care utilization.

Pairwise comparisons revealed that participants perceived as equally feminine or masculine (49.2%) or somewhat masculine (57.5%) were significantly more likely to report poorer general health than those with a very masculine (32.1%) perceived gender expression. When compared with those with a very masculine perceived presentation (15.8%), all other perceived gender expression groups were more likely to report long-term mental health problems (range: 40.7%–45.7%). Although a significant effect for long-term physical disability or health problems ($P = .048$) was indicated in our results, no statistically significant between-group comparisons were found. By using Pillai's trace, a statistically significant effect of perceived gender expression on care utilization measures ($P = .52$) was not indicated after controlling for covariates.

For adolescents who are TGNC and were assigned female at birth, a significant effect of perceived gender expression on health status measures ($P = .001$) was indicated by using Pillai's trace after controlling for covariates. Pairwise comparisons revealed that participants with a somewhat feminine (69.5%), equally feminine and masculine (70.4%), or somewhat masculine (71.7%) perceived gender expression were significantly more likely to report poorer general health than those with a very feminine (54.0%) perceived gender expression. Compared with participants assigned female with a very feminine perceived gender expression (55.4%), participants with all other perceived gender expressions were more likely to report long-term mental health problems (range: 68.1%–76.7%). No other pairwise comparisons were statistically significant. The effect of perceived gender expression on care utilization measures was also not statistically significant.

TABLE 4 Perceived Gender Expression of TGNC Students by Birth-Assigned Sex ($n = 2095$)

	TGNC Students	
	Assigned Male at Birth ($n = 661$), *n* (%)	Assigned Female at Birth ($n = 1434$), *n* (%)
Perceived gender expression		
Very feminine	104 (15.7)	177 (12.3)
Somewhat feminine	100 (15.1)	327 (22.8)
Equally feminine and masculine	194 (29.3)	591 (41.2)
Somewhat masculine	132 (20.0)	243 (16.9)
Very masculine	131 (19.8)	96 (6.7)

Data for birth-assigned sex or perceived gender expression were missing for 73 cases.

DISCUSSION

Population-based research in which both binary and nonbinary gender categories are examined is essential for a more comprehensive understanding of health disparities and health care needs of adolescents who are TGNC.[13] In this study, we address research gaps related to health status and care utilization of youth who are TGNC by describing a variety of physical and mental health indicators in a large, population-based sample of adolescents and identifying perceived gender expression as an important factor in understanding health disparities for this understudied group.

We found that students who are TGNC reported significantly poorer health status, lower rates of preventive health checkups, and more visits to the nurse's office than their cisgender peers. Although youth who are TGNC reported an overall worse health status compared with their cisgender peers, nearly three-quarters of youth who are TGNC did not experience long-term physical disabilities or health problems, which is consistent with previous findings that this group typically does not struggle with chronic physical health concerns.[2] Over half of adolescents who are TGNC have received preventive medical and dental care; these rates are slightly lower than those reported previously, but with this information, we can continue to highlight the importance of health care providers addressing health risk while promoting wellness within this community.[29,30]

Among the TGNC sample, important differences emerged across perceived gender expressions by birth-assigned sex. Youth who are TGNC with perceived gender expressions that are incongruent or that somewhat deviate from societal expectations for their birth-assigned sex were at higher risk for poorer health outcomes. This is not surprising given the social pressures to conform to gender roles and stereotypes associated with one's birth-assigned sex and is consistent with previous studies.[10] Comparisons between perceived gender expression groups were not significantly different for any care utilization measure.

School nurses are uniquely positioned to promote, educate, and advocate for optimal health for students who are TGNC. For example, school nurses can promote antibullying policies and clubs such as Gay-Straight Alliances to improve school climate.[31,32] School nurses can assist adolescents by providing resources and information about gender identity and expression. When appropriate, school nurses can also discuss with parents of self-disclosing adolescents who are seeking support that family support is protective and rejection is potentially detrimental to health. In addition, nurses can assist with health promotion by referring to appropriate resources.[32]

TABLE 5 Predicted Probabilities of TGNC Students' Health Status and Health Care Utilization, Stratified by Perceived Gender Expression and Birth-Assigned Sex (n = 2095)

	TGNC Assigned Male at Birth (n = 661)		TGNC Assigned Female at Birth (n = 1434)	
	Responses (n)	Predicted Probability (%)	Responses (n)	Predicted Probability (%)
Health status				
General health is poor, fair, or good	P = .003[a]		P = .001[b]	
Very feminine	82	49.2	163	54.0
Somewhat feminine	90	50.8	314	69.5
Equally feminine and masculine	177	49.2	564	70.4
Somewhat masculine	123	57.5	231	71.7
Very masculine	113	32.1	85	68.6
Long-term physical disability or health problems	P = .048		P = .418	
Very feminine	82	19.5	163	22.7
Somewhat feminine	90	35.8	314	24.9
Equally feminine and masculine	177	27.1	564	25.0
Somewhat masculine	123	29.0	231	25.5
Very masculine	113	18.8	85	33.9
Long-term mental health problems	P < .001[c]		P < .001[d]	
Very feminine	82	40.7	163	55.4
Somewhat feminine	90	44.6	314	68.1
Equally feminine and masculine	177	45.7	564	69.8
Somewhat masculine	123	42.8	231	76.7
Very masculine	113	15.8	85	73.2
Stayed home sick (last 30 days)	P = .210		P = .265	
Very feminine	82	50.9	163	48.3
Somewhat feminine	90	57.5	314	55.1
Equally feminine and masculine	177	48.6	564	55.4
Somewhat masculine	123	48.2	231	48.6
Very masculine	113	40.7	85	55.0
Care utilization				
Nurse office visits (last 30 days)	P = .947		P = .688	
Very feminine	81	34.0	163	47.4
Somewhat feminine	90	34.2	321	42.1
Equally feminine and masculine	178	34.5	571	42.5
Somewhat masculine	123	31.6	235	46.2
Very masculine	113	30.4	86	45.4
Preventive medical checkup	P = .558		P = .175	
Very feminine	81	51.5	163	69.4
Somewhat feminine	90	54.0	321	61.8
Equally feminine and masculine	178	57.6	571	61.3
Somewhat masculine	123	62.4	235	57.2
Very masculine	113	59.6	86	64.3
Preventive dental checkup	P = .059		P = .225	
Very feminine	81	63.8	163	74.7
Somewhat feminine	90	63.9	321	75.2
Equally feminine and masculine	178	64.9	571	71.9
Somewhat masculine	123	74.3	235	67.2
Very masculine	113	77.0	86	76.1

In these analyses, we controlled for free and/or reduced-price lunch, race and/or ethnicity, grade, and school location. α level set at .05. Post hoc tests used Bonferroni's correction to adjust α for all pairwise comparisons. Numbers do not sum to the sample size because of missing data on at least 1 variable.

[a] Post hoc analysis indicated significant differences between equally feminine and masculine and very masculine perceived gender expressions as well as somewhat masculine and very masculine perceived gender expressions.

[b] Post hoc analysis indicated significant differences between somewhat feminine and very feminine perceived gender expressions, equally feminine and masculine and very feminine perceived gender expressions, as well as somewhat masculine and very feminine perceived gender expressions.

[c] Post hoc analysis indicated significant differences between very masculine and all other perceived gender expression groups.

[d] Post hoc analysis indicated significant differences between very feminine and all other perceived gender expression groups.

Consistent with gender minority and resilience theory, individuals perceived as gender nonconforming may be vulnerable to discrimination and have difficulty accessing and receiving health care compared with their cisgender peers.[14,29,30,33] Perceived gender nonconformity may be a risk factor for minority stressors (eg, nonaffirmation, victimization, discrimination, or rejection), which may in turn elevate adverse health outcomes for these youth.[3,13] Youth who are perceived or identify as gender nonconforming or nonbinary must also overcome unique barriers to accessing affirming health care compared with other TGNC adolescents, such as mistrust of health care providers because of fear of the youth's own gender identity or expression being misunderstood.[34,35] These barriers contribute to delays in seeking services, which may result in poorer health outcomes. More research with a focus on differences across gender identities and expressions is needed to better understand associations contributing to health risk disparities among youth who are TGNC.

To our knowledge, this is the first large, population-based study of TGNC adolescents in the United States conducted to describe prevalence rates of health status and care utilization compared with cisgender youth and to explore perceived gender expression. Because of the census-like recruitment strategy in which all schools in the state were invited to participate, findings are more generalizable than results from previous studies in which convenience samples were used. The numerous measures of health status (including both mental and physical health) and care utilization are considerable strengths of the survey, with which we address a gap in the literature for youth who are TGNC.

Although valuable information about health status and care utilization

for youth who are TGNC is provided in our results, it is important to note limitations. First, asking about biological sex may be confusing for some students more accustomed to the phrase "sex assigned at birth," which is commonly used in this population. Likewise, the measure of gender identity does not allow for differentiation of students who identify as transgender, genderqueer, or unsure. We were also unable to assess whether youth were interested in being perceived as a different gender, had received any gender-affirming medical interventions (ie, puberty blockers, gender-affirming hormones), or had socially transitioned to their affirmed gender, which may impact how their gender expression is perceived and how they feel about particular perceptions of their gender. Furthermore, we lack a measure of actual gender expression (ie, how youth perceive and present themselves in society through dress, mannerisms, and personal style). Instead, students were asked about how they think others at school perceive them, which might be interpreted as a question more reflective of gender affirmation than personal gender expression and/or presentation. Missing data (whether due to nonresponse or missed opportunities because of school absence) may result in an underestimation of TGNC identity and health status. Lastly, youth who are TGNC often use the bathroom in the nurse's office[36]; thus, students may have overreported the frequency of nurse office visits.

CONCLUSIONS

Health status and care utilization differ between youth who are TGNC versus cisgender and across perceived gender presentations. With our results, we suggest that health care providers should screen for health risks and identify barriers to care for youth who are TGNC while promoting and bolstering wellness within this community. Although youth who are TGNC generally appear healthy and many are using health care services, continued research and advocacy are needed to decrease barriers to care and improve health outcomes for these young people, particularly those whose perceived gender expressions transgress societal expectations. As such, it is important that providers develop competency to work with adolescents with diverse gender identities and expressions because health needs may differ across and within gender groups.

ACKNOWLEDGMENTS

MSS data were provided by public school students in Minnesota via local public school districts and are managed by the MSS Interagency Team.

ABBREVIATIONS

MSS: Minnesota Student Survey
TGNC: transgender and gender nonconforming

Address correspondence to G. Nicole Rider, PhD, Program in Human Sexuality, Department of Family Medicine and Community Health, University of Minnesota Medical School, 1300 S 2nd St, Suite 180, Minneapolis, MN 55454. E-mail: gnrider@umn.edu

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275).

FINANCIAL DISCLOSURE: The authors have indicated they have no financial relationships relevant to this article to disclose.

FUNDING: Supported by the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (award R21HD088757). The content of this article is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Funded by the National Institutes of Health (NIH).

POTENTIAL CONFLICT OF INTEREST: The authors have indicated they have no potential conflicts of interest to disclose.

COMPANION PAPER: A companion to this article can be found online at www.pediatrics.org/cgi/doi/10.1542/peds.2017-4079.

REFERENCES

1. Trans Student Educational Resources. LGBTQ+ definitions. Available at: www.transstudent.org/definitions. Accessed July 15, 2017
2. Institute of Medicine (US) Committee on Lesbian, Gay, Bisexual, and Transgender Health Issues and Research Gaps and Opportunities. *The Health of Lesbian, Gay, Bisexual, and Transgender People: Building a Foundation for Better Understanding*. Washington, DC: National Academies Press; 2011
3. Testa RJ, Habarth J, Peta J, Balsam K, Bockting W. Development of the gender minority stress and resilience measure. *Psychol Sex Orientat Gend Divers*. 2015;2(1):65–77
4. Hill DB, Menvielle E, Sica KM, Johnson A. An affirmative intervention for families with gender variant children: parental ratings of child mental health and gender. *J Sex Marital Ther*. 2010;36(1):6–23
5. Hidalgo MA, Ehrensaft D, Tishelman AC, et al. The gender affirmative model: what we know and what we aim to learn. *Hum Dev*. 2013;56(5):285–290
6. Garofalo R, Deleon J, Osmer E, Doll M, Harper GW. Overlooked, misunderstood and at-risk: exploring the lives and HIV

risk of ethnic minority male-to-female transgender youth. *J Adolesc Health*. 2006;38(3):230–236

7. Roberts AL, Rosario M, Corliss HL, Koenen KC, Austin SB. Childhood gender nonconformity: a risk indicator for childhood abuse and posttraumatic stress in youth. *Pediatrics*. 2012;129(3):410–417
8. Testa RJ, Rider GN, Haug NA, Balsam KF. Gender confirming medical interventions and eating disorder symptoms among transgender individuals. *Health Psychol*. 2017;36(10):927–936
9. Toomey RB, Ryan C, Diaz RM, Card NA, Russell ST. Gender-nonconforming lesbian, gay, bisexual, and transgender youth: school victimization and young adult psychosocial adjustment. *Dev Psychol*. 2010;46(6):1580–1589
10. Roberts AL, Rosario M, Slopen N, Calzo JP, Austin SB. Childhood gender nonconformity, bullying victimization, and depressive symptoms across adolescence and early adulthood: an 11-year longitudinal study. *J Am Acad Child Adolesc Psychiatry*. 2013;52(2):143–152
11. Maragh-Bass AC, Torain M, Adler R, et al. Is it okay to ask: transgender patient perspectives on sexual orientation and gender identity collection in healthcare. *Acad Emerg Med*. 2017;24(6):655–667
12. Roberts TK, Fantz CR. Barriers to quality health care for the transgender population. *Clin Biochem*. 2014;47(10–11):983–987
13. Frohard-Dourlent H, Dobson S, Clark BA, Doull M, Saewyc EM. "I would have preferred more options": accounting for non-binary youth in health research [published online ahead of print August 8, 2011]. *Nurs Inq*. 2017;24(1): e12150
14. Grant JM, Mottet LA, Tanis J, Herman JL, Harrison J, Keisling M. *National Transgender Discrimination Survey Report on Health and Health Care: Findings of a Study by the National Center for Transgender Equality and the National Gay and Lesbian Task Force*. Washington, DC: The National Center for Transgender Equality and The National Gay and Lesbian Task Force; 2010
15. Schuster MA, Reisner SL, Onorato SE. Beyond bathrooms–meeting the health needs of transgender people. *N Engl J Med*. 2016;375(2):101–103
16. Poteat T, German D, Kerrigan D. Managing uncertainty: a grounded theory of stigma in transgender health care encounters. *Soc Sci Med*. 2013;84:22–29
17. Giffort DM, Underman K. The relationship between medical education and trans health disparities: a call to research. *Sociol Compass*. 2016;10(11):999–1013
18. Safer JD, Coleman E, Feldman J, et al. Barriers to healthcare for transgender individuals. *Curr Opin Endocrinol Diabetes Obes*. 2016;23(2):168–171
19. Gordon AR, Krieger N, Okechukwu CA, et al. Decrements in health-related quality of life associated with gender nonconformity among US adolescents and young adults. *Qual Life Res*. 2017;26(8):2129–2138
20. Wylie SA, Corliss HL, Boulanger V, Prokop LA, Austin SB. Socially assigned gender nonconformity: a brief measure for use in surveillance and investigation of health disparities. *Sex Roles*. 2010;63(3–4):264–276
21. McPhail BA. Questioning gender and sexuality binaries. *J Gay Lesbian Soc Serv*. 2004;17(1):3–21
22. Veale JF, Watson RJ, Peter T, Saewyc EM. Mental health disparities among Canadian transgender youth. *J Adolesc Health*. 2017;60(1):44–49
23. Reisner SL, Biello K, Rosenberger JG, et al. Using a two-step method to measure transgender identity in Latin America/the Caribbean, Portugal, and Spain. *Arch Sex Behav*. 2014;43(8):1503–1514
24. Reisner SL, Conron KJ, Tardiff LA, Jarvi S, Gordon AR, Austin SB. Monitoring the health of transgender and other gender minority populations: validity of natal sex and gender identity survey items in a US national cohort of young adults. *BMC Public Health*. 2014;14:1224
25. Reisner SL, Deutsch MB, Bhasin S, et al. Advancing methods for US transgender health research. *Curr Opin Endocrinol Diabetes Obes*. 2016;23(2):198–207
26. The Williams Institute. *Gender-Related Measures Overview*. Los Angeles, CA: UCLA School of Law; 2013
27. Bright Futures; American Academy of Pediatrics. Recommendations for preventive pediatric health care. Available at: https://www.aap.org/en-us/documents/periodicity_schedule.pdf. Accessed April 15, 2017
28. American Academy on Pediatric Dentistry Clinical Affairs Committee; American Academy on Pediatric Dentistry Council on Clinical Affairs. Guideline on periodicity of examination, preventive dental services, anticipatory guidance/counseling, and oral treatment for infants, children, and adolescents.*Pediatr Dent*. 2008–2009;30(7 suppl):112–118
29. Enhancing transgender health care. *Am J Public Health*. 2017;107(2):230–231
30. Hoffman ND, Freeman K, Swann S. Healthcare preferences of lesbian, gay, bisexual, transgender and questioning youth. *J Adolesc Health*. 2009;45(3):222–229
31. Gower AL, Forster M, Gloppen K, et al. School practices to foster LGBT-supportive climate: associations with adolescent bullying involvement [published online ahead of print October 14, 2017]. *Prev Sci*. doi:10.1007/s11121-017-0847-4
32. National Association of School Nurses. LGBTQ students: the role of the school nurse (revised 2016). Available at: https://schoolnursenet.nasn.org/blogs/nasn-profile/2017/03/13/lgbtq-students-the-role-of-the-school-nurse. Accessed October 12, 2017
33. Pusch RS. Objects of curiosity: transgender college students' perceptions of the reactions of others. *J Gay Lesbian Issues Educ*. 2005;3(1):45–61
34. National LGBT Health Education Center, A Program of the Fenway Institute. Providing affirmative care for patients with non-binary gender identities. Available at: https://www.lgbthealtheducation.org/wp-content/uploads/2017/02/Providing-Affirmative-Care-for-People-with-Non-Binary-Gender-Identities.pdf. Accessed March 24, 2016
35. National LGBT Health Education Center. Providing mental health care for youth with non-binary gender identities. Available at: https://www.lgbthealtheducation.org/lgbt-education/online-courses/continuing-education/?y=162. Accessed March 24, 2016
36. Porta CM, Gower AL, Mehus CJ, Yu X, Saewyc EM, Eisenberg ME. "Kicked out": LGBTQ youths' bathroom experiences and preferences. *J Adolesc*. 2017;56:107–112

Journal of Sex Research

ISSN: 0022-4499 (Print) 1559-8519 (Online) Journal homepage: https://www.tandfonline.com/loi/hjsr20

How common is Intersex? A response to Anne Fausto-Sterling

Leonard Sax

To cite this article: Leonard Sax (2002) How common is Intersex? A response to Anne Fausto-Sterling, Journal of Sex Research, 39:3, 174-178, DOI: 10.1080/00224490209552139

To link to this article: https://doi.org/10.1080/00224490209552139

Published online: 11 Jan 2010.

How Common is Intersex? A Response to Anne Fausto-Sterling

Leonard Sax
The Montgomery Center for Research in Child and Adolescent Development, Maryland

Anne Fausto-Sterling's suggestion that the prevalence of intersex might be as high as 1.7% has attracted wide attention in both the scholarly press and the popular media. Many reviewers are not aware that this figure includes conditions which most clinicians do not recognize as intersex, such as Klinefelter syndrome, Turner syndrome, and late-onset adrenal hyperplasia. If the term intersex *is to retain any meaning, the term should be restricted to those conditions in which chromosomal sex is inconsistent with phenotypic sex, or in which the phenotype is not classifiable as either male or female. Applying this more precise definition, the true prevalence of intersex is seen to be about 0.018%, almost 100 times lower than Fausto-Sterling's estimate of 1.7%.*

Sometimes a child is born with genitalia which cannot be classified as female or male. A genetically female child (i.e., with XX chromosomes) may be born with external genitalia which appear to be those of a normal male. Or, a genetically male child (XY chromosomes) may be born with female-appearing external genitalia. In very rare cases, a child may be born with both female and male genitalia. Because these conditions are in some sense "in-between" the two sexes, they are collectively referred to as *intersex*.

How common is intersex? In her 1993 essay, biologist Anne Fausto-Sterling acknowledged that "it is extremely difficult to estimate the frequency of intersexuality" (Fausto-Sterling, 1993, p. 21). In this paper we will focus on establishing how often intersexual conditions occur, and what conditions should be considered intersexual.

In her most recent book, *Sexing the Body: Gender Politics and the Construction of Sexuality* (Fausto-Sterling, 2000), Fausto-Sterling maintains that human sexuality is best understood not as a dichotomy but as a continuum. She bases this assertion on her beliefs regarding intersex conditions. A chapter subtitled "The Sexual Continuum" begins with the case of Levi Suydam, an intersexual living in the 1840s who menstruated regularly but who also had a penis and testicles. Fausto-Sterling writes:

> While male and female stand on the extreme ends of a biological continuum, there are many bodies, bodies such as Suydam's, that evidently mix together anatomical components conventionally attributed to both males and females. The implications of my argument for a sexual continuum are profound. If nature really offers us more than two sexes, then it follows that our current notions of masculinity and femininity are cultural conceits.
>
> . . . Modern surgical techniques help maintain the two-sex system. Today children who are born "either/or-neither/both"—a fairly common phenomenon —usually disappear from view because doctors "correct" them right away with surgery. (Fausto-Sterling, 2000, p. 31)

Address correspondence to Leonard Sax, The Montgomery Center for Research in Child and Adolescent Development, P. O. Box 108, 19710 Fisher Avenue, Suite J, Poolesville, MD 20837; e-mail: leonardsax@prodigy.net.

This work received no outside financial support.

Fausto-Sterling asserts that 1.7% of human births are intersex. This figure was widely quoted in the aftermath of the book's publication. "Instead of viewing intersexuality as a genetic hiccup," wrote Courtney Weaver for the *Washington Post*, "[Fausto-Sterling] points out that its frequency mandates a fresher look. In one study, intersexuality typically constitute 1.7% of a community" (Weaver, 2000). The *New England Journal of Medicine* applauded Fausto-Sterling's "careful and insightful book. . . . She [Fausto-Sterling] points out that intersexual newborns are not rare (they may account for 1.7% of births), so a review of our attitudes about these children is overdue. . ." (Breedlove, 2000). "Most people believe that there are only two sex categories," went the review in *American Scientist*. "Yet 17 out of every 1,000 people fail to meet our assumption that everyone is either male or female. This is the approximate incidence of intersexuals: individuals with XY chromosomes and female anatomy, XX chromosomes and male anatomy, or anatomy that is half male and half female." (Moore, 2000, p. 545)

This reviewer assumed that Fausto-Sterling was using the term *intersex* in the usual way, the same way in which Fausto-Sterling herself used the term in her 1993 essay, "The Five Sexes" (Fausto-Sterling, 1993): to refer either to individuals who have XY chromosomes with predominantly female anatomy, XX chromosomes with predominantly male anatomy, or ambiguous or mixed genitalia. This assumption is reasonable, because all the case histories presented in her book *Sexing the Body* describe individuals who meet these criteria (Fausto-Sterling, 2000). However, as we shall see, the 1.7% statistic is based on a much broader definition of intersex.

Fausto-Sterling herself has encouraged the belief that a significant fraction of the population is neither male nor female, but intersex. In an interview with *The New York Times*, she said that "I did some research and we found that maybe 1 to 2 percent of all births do not fall strictly within the tight definition of all-male or all-female. . . .there is greater human variation than supposed. . . [We should] lighten up about what it means to be male or female. We

should definitely lighten up on those who fall in between because there are a lot of them" (Dreifus, 2001).

Is Human Sexuality a Dichotomy or a Continuum?

Fausto-Sterling's argument that human sexuality is a continuum, not a dichotomy, rests in large measure on her claim that intersex births are a fairly common phenomenon. Specifically, Fausto-Sterling computes the incidence of intersexual births to be 1.7 per 100 live births, or 1.7%. To arrive at that figure, she defines as *intersex* any "individual who deviates from the Platonic ideal of physical dimorphism at the chromosomal, genital, gonadal, or hormonal levels" (Blackless et al., 2000, p. 161).

This definition is too broad. Fausto-Sterling and her associates acknowledge that some of the individuals thus categorized as intersex "are undiagnosed because they present no symptoms" (Blackless et al., 2000, p. 152). A definition of intersex which encompasses individuals who are phenotypically indistinguishable from normal is likely to confuse both clinicians and patients.

John Wiener, a urologist, has suggested defining intersex simply as "a discordance between phenotypic sex and chromosomal sex" (Wiener, 1999). While this definition would cover most true intersex patients, there are some rare conditions which are clearly intersex which are not captured by this definition. For example, some people are *mosaics*: Different cells in their body have different chromosomes. A 46,XY/46,XX mosaic is an individual in whom some cells have the male chromosomal complement (XY) and some cells have the female chromosomal complement (XX). If such an individual has both a penis and a vagina, then there is no mismatch between phenotypic sex and genotypic sex: Both the phenotype and the genotype are intersexual. Yet according to Wiener's definition, such an individual would *not* be intersex. A more comprehensive, yet still clinically useful definition of intersex would include those conditions in which (a) the phenotype is not classifiable as either male or female, or (b) chromosomal sex is inconsistent with phenotypic sex.

This definition is of course more clinically focussed than the definition employed by Fausto-Sterling. Using her definition of intersex as "any deviation from the Platonic ideal" (Blackless et al., 2000, p. 161), she lists all the following conditions as intersex, and she provides the following estimates of incidence for each condition (number of births per 100 live births): (a) late-onset congenital adrenal hyperplasia (LOCAH), 1.5/100; (b) Klinefelter (XXY), 0.0922/100; (c) other non-XX, non-XY, excluding Turner and Klinefelter, 0.0639/100; (d) Turner syndrome (XO), 0.0369/100; (e) vaginal agenesis, 0.0169/100; (f) classic congenital adrenal hyperplasia, 0.00779/10; (g) complete androgen insensitivity, 0.0076/100; (h) true hermaphrodites, 0.0012/100; (i) idiopathic, 0.0009/100; and (j) partial androgen insensitivity, 0.00076/100. The chief problem with this list is that the five most common conditions listed are not intersex conditions. If we examine these five conditions in more detail, we will see that there is no meaningful clinical sense in which these conditions can be considered intersex. "Deviation from the Platonic ideal" is, as we will see, not a clinically useful criterion for defining a medical condition such as intersex.

The second problem with this list is the neglect of the five most common of these conditions in Fausto-Sterling's book *Sexing the Body* (Fausto-Sterling, 2000). In her book, Fausto-Sterling draws her case histories exclusively from the ranks of individuals who are unambiguously intersex. However, using Fausto-Sterling's own figures, such individuals account for less than 0.02% of the general population. None of her case histories are drawn from the five most common conditions in her table, even though these five conditions constitute roughly 99% of the population she defines as intersex. Without these five conditions, intersex becomes a rare occurrence, occurring in fewer than 2 out of every 10,000 live births.

Classic Intersex Conditions

Among classic intersex conditions, the most common are congenital adrenal hyperplasia (CAH) and complete androgen insensitivity syndrome. According to Fausto-Sterling's figures, these two conditions occur with roughly the same frequency: about 0.008/100, or 8 births out of every 100,000. There is no dispute that these conditions are indeed intersex conditions. We discuss them here because some understanding of these conditions is essential in order to perceive how these conditions differ from the other syndromes which Fausto-Sterling includes in the category of intersex.

Complete Androgen Insensitivity Syndrome

These individuals are genetically male (XY), but owing to a defect in the androgen receptor, their cells do not respond to testosterone or other androgens (Boehmer et al., 2001). As a result, these individuals do not form male genitalia. Genetically male (XY) babies with this condition typically are born with a vaginal opening and clitoris indistinguishable from those seen in normal female (XX) babies. In almost all cases, the diagnosis is not suspected until puberty, when these "girls" are brought to medical attention because they have never menstruated. Investigation at that point will invariably reveal that these "girls" are in fact genetically male, that they have undescended testicles, and that neither the uterus nor the ovaries are present. These individuals are genotypically male, but phenotypically female.

Congenital Adrenal Hyperplasia

In this syndrome, a defect in an enzyme involved in the synthesis of adrenal hormones leads to a blockage in one synthetic pathway, giving rise to excessive production of androgenic hormones in a different pathway (White, 2001). These androgens will masculinize a female (XX) fetus *in utero*. At birth, the girl's genitalia may appear completely masculine, or, more commonly, the genitalia

will be ambiguous—neither completely male nor completely female but somewhere in between.

Are These Other Conditions Intersex?

Late-Onset Congenital Adrenal Hyperplasia

In late-onset congenital adrenal hyperplasia, the defect in the enzymatic pathway typically does not manifest itself until late childhood, adolescence, or later, and the degree of disruption is much less than in classic congenital adrenal hypertrophy. Reviewing the list of conditions which Fausto-Sterling considers to be intersex, we find that this one condition—late-onset congenital adrenal hyperplasia (LOCAH)—accounts for 88% of all those patients whom Fausto-Sterling classifies as intersex (1.5/1.7 = 88%).

From a clinician's perspective, however, LOCAH is not an intersex condition. The genitalia of these babies are normal at birth, and consonant with their chromosomes: XY males have normal male genitalia, and XX females have normal female genitalia. The average woman with this condition does not present until about 24 years of age (Speiser et al., 2000). Men with LOCAH present later, if ever: Many go through life undetected or are discovered only incidentally (Holler et al., 1985). For example, if a daughter is discovered to have classic congenital adrenal hyperplasia, the parents often will be tested for evidence of overproduction of adrenal androgens, and one parent thereby may be discovered to have LOCAH. The most common presenting symptom of LOCAH in men is thinning of scalp hair, but even this symptom is seen in only 50% of men with LOCAH under 50 years of age (Dumic et al., 1985).

Fausto-Sterling recognizes that if her definition of the intersexual as "an individual who deviates from the Platonic ideal of physical dimorphism" (Blackless et al., 2000, p. 161) is to have any clinical relevance, then at least some patients with LOCAH must occasionally have problems which are intersexual in nature. Accordingly, she asserts that "when late-onset CAH occurs in childhood or adolescence and causes significant clitoral growth, it is quite possible that surgical intervention will ensue." (Blackless et al., 2000, p. 161) The only reference given in support of this statement is a first-person account in the woman's magazine *Mademoiselle* (Moreno & Goodwin, 1998). However, the article in *Mademoiselle* describes a phenotypically female but genotypically male (46,XY) individual with androgen insensitivity: in other words, a case of true intersexuality. LOCAH is never mentioned.

In a large-scale investigation of the natural history of LOCAH in women, the chief complaints of symptomatic women were one or more of the following: oligomenorrhea, hirsutism, infertility, or acne. These investigators noted that "in some cases, affected girls have shown mild clitoromegaly, but not true genital ambiguity" (Speiser et al., 2000, p. 527). Many women have no symptoms at all: "Probably many affected individuals are asymptomatic," notes another recent review (White, 2001, p. 25). A recent study of 220 women with LOCAH found mild clitoromegaly in only 10%; moderate or severe clitoromegaly was not reported (Moran et al., 2000).

Sex Chromosome Aneuploidies

Fausto-Sterling defines all sex chromosome complements other than XX or XY as intersex. Specifically, Fausto-Sterling includes Klinefelter syndrome, Turner syndrome, and all other non-XX, non-XY chromosomal variations in the intersex category.

Klinefelter syndrome. Babies born with Klinefelter syndrome (47,XXY) have normal male genitalia. Male secondary sexual characteristics develop normally in puberty, although the testicles typically are small. Erection and ejaculation are not impaired. Most men with Klinefelter syndrome are infertile, but an unknown proportion are fertile (Warburg, 1963). Because Klinefelter syndrome is most often discovered in the course of infertility evaluation, fertile men with Klinefelter syndrome are likely to go completely undetected. Abramsky and Chapple (1997) have suggested that many men with Klinefelter syndrome are never diagnosed because they are phenotypically indistinguishable from normal (46,XY) men.

Turner syndrome. Among the most salient features of Turner syndrome (45,X) are infertility and short stature: Women with Turner syndrome who are not treated with growth hormone typically will be about 16 centimeters shorter than their predicted adult height based on parental heights (Holl, Kunze, Etzrodt, Teller, & Heinze, 1994). Sas et al. (1999) have demonstrated that girls with Turner syndrome can achieve normal adult heights if daily doses of growth hormone are administered. Although most women with Turner syndrome cannot conceive a child, they can carry a child to term if a donated embryo or oocyte is implanted (Hovatta, Foudila, & Söderström-Anttila, 2000). Girls with Turner syndrome do not have ambiguous external genitalia (e.g., no clitoromegaly), nor do they typically experience confusion regarding their sexual identity. "A consistent feature documented in Turner's syndrome is the unambiguous identification with the female sex," according to a recent review in *The Lancet* (Ranke & Saenger, 2001, p. 310).

Other chromosomal variants (non-XX and non-XY, excluding Turner's and Klinefelter's). This category includes a variety of sex chromosome complements, such as XXX, XYY, and other less frequent arrangements. Fausto-Sterling considers all such conditions to be intersex. Men with an extra Y chromosome (47,XYY) are not distinguishable from normal (46,XY) men, although the average intelligence of men with this aneuploidy is lower than normal. Their fertility usually is not impaired. They are most commonly discovered in the course of evaluation for mild mental retardation or behavior problems (Fryns, Kleczkowska, Kubien, & Van den Berghe, 1995). Likewise, women with an extra X chromosome ("triple X," 47,XXX) are fertile, although the mean intelligence of women with this aneuploidy is also probably below aver-

age (Bender, Linden, & Harmon, 2001). None of these chromosomal variants are associated with ambiguous genitalia, or with any confusion regarding sexual identity. There is therefore no clinical sense in which these individuals are intersex.

Vaginal Agenesis

Fausto-Sterling estimates that about 0.0169 births per 100 are characterized by vaginal agenesis (also known as vaginal atresia), a condition in which the distal third of the vagina fails to develop and is replaced by about 2 cm of fibrous tissue (Simpson, 1999). According to the definition which I have proposed, vaginal agenesis is not an intersex condition. Girls born with this condition have an XX genotype and normal ovaries. In the majority of cases, vaginoplasty restores normal female vaginal anatomy (Robson & Oliver, 2000). Women who have undergone vaginoplasty can and do go on to have successful term pregnancies (Moura, Navarro, & Nogueira, 2000). Nosologically, vaginal agenesis is to genital anatomy as cleft palate is to maxillofacial anatomy. Surgical correction for vaginal agenesis is conceptually no different from surgical correction for cleft palate.

How Common Is Intersex?

Subtracting these five categories—LOCAH, vaginal agenesis, Turner's syndrome, Klinefelter's syndrome, and other non-XX and non-XY aneuploidies—the incidence of intersex drops to 0.018%, almost 100 times lower than the estimate provided by Fausto-Sterling. This figure of 0.018% suggests that there are currently about 50,000 true intersexuals living in the United States. These individuals are of course entitled to the same expert care and consideration that all patients deserve. Nothing is gained, however, by pretending that there are 5,000,000 such individuals.

Is Intersex a Normal Variant or a Pathological Condition?

The most original feature of Fausto-Sterling's book is her reluctance to classify true intersex conditions as pathological. Regarding babies born with both a penis and a vagina, she writes: "Perhaps we will come to view such children as especially blessed or lucky. It is not so far-fetched to think that some can become the most desirable of all possible mates, able to pleasure their partners in a variety of ways" (Fausto-Sterling, 2000, p. 113). Fausto-Sterling (2000) strongly affirms her belief that all possible combinations of sexual anatomy must be considered normal:

> Complete maleness and complete femaleness represent the extreme ends of a spectrum of possible body types. That these extreme ends are the most frequent has lent credence to the idea that they are not only natural (that is, produced by nature) but normal (that is, they represent both a statistical and a social ideal). Knowledge of biological variation, however, allows us to conceptualize the less frequent middle spaces as natural, although statistically unusual. (p. 76)

Nosological Confusion

Nosology is the science of the classification of diseases. The first principle of nosology is the distinction between the normal and the pathological. This principle poses real difficulties for Fausto-Sterling. She often uses the word *natural* synonymously with *normal* (for an example, see the previous paragraph). However, *natural* and *normal* are not synonyms. A cow may give birth to a two-headed or Siamese calf by natural processes, natural being understood as per Fausto-Sterling's definition as "produced by nature." Nevertheless, that two-headed calf unarguably manifests an *abnormal* condition.

Fausto-Sterling's insistence that all combinations of sexual anatomy be regarded as normal is reminiscent of Szasz's view of mental illness (Szasz, 1974). Szasz insisted that mental illness is not a real biological phenomenon but merely an invention of society. Like Fausto-Sterling, Szasz was suspicious of the distinction between normal and pathological. Fausto-Sterling follows the example set by Szasz in her belief that classifications of normal and abnormal sexual anatomy are mere social conventions, prejudices which can and should be set aside by an enlightened intelligentsia.

This type of extreme social constructionism is confusing and is not helpful to clinicians, to their patients, or to their patients' families. Diluting the term *intersex* to include "any deviation from the Platonic ideal of sexual dimorphism" (Blackless et al., 2000, p. 152), as Fausto-Sterling suggests, deprives the term of any clinically useful meaning.

Conclusions

The available data support the conclusion that human sexuality is a dichotomy, not a continuum. More than 99.98% of humans are either male or female. If the term *intersex* is to retain any clinical meaning, the use of this term should be restricted to those conditions in which chromosomal sex is inconsistent with phenotypic sex, or in which the phenotype is not classifiable as either male or female.

The birth of an intersex child, far from being "a fairly common phenomenon," is actually a rare event, occurring in fewer than 2 out of every 10,000 births.

References

Abramsky, L., & Chapple, J. (1997). 47,XXY (Klinefelter syndrome) and 47,XYY: Estimated rates of and indication for postnatal diagnosis with implications for prenatal counseling. *Prenatal Diagnosis, 17,* 363-368.

Bender, B. G., Linden, M. G., & Harmon, R. J. (2001). Neuropsychological and functional cognitive skills of 35 unselected adults with sex chromosome abnormalities. *American Journal of Medical Genetics, 102,* 309-313.

Blackless, M., Charuvastra, A., Derryck, A., Fausto-Sterling, A., Lauzanne, K., & Lee, E. (2000). How sexually dimorphic are we? Review and synthesis. *American Journal of Human Biology, 12,* 151-166.

Boehmer, A. L., Bruggenwirth, H., van Assendelft, C., Otten, B. J., Verleun-Mooijman, M. C., Niermeijer, M. F., et al. (2001). Genotype versus phenotype in families with androgen insensitivity syndrome. *Journal of Clinical Endocrinology & Metabolism, 86,* 4151-4160.

Breedlove, M. (2000). Sexing the body. *New England Journal of Medicine, 343,* 668.

Dreifus, C. (2001, January 2). A conversation with Anne Fausto-Sterling:

Exploring what makes us male or female. *The New York Times*, F3.
Dumic, M., Brkljacic, L., Mardesic, D., Plavsic, V., Lukenda, M., & Kastelan, A. (1985). Cryptic form of congenital adrenal hyperplasia due to 21-hydroxylase deficiency in the Yugoslav population. *Acta Endocrinologica, 109,* 386-392.
Fausto-Sterling, A. (1993). The five sexes: Why male and female are not enough. *The Sciences, 33*(2), 20-25.
Fausto-Sterling, A. (2000). *Sexing the body: Gender politics and the construction of sexuality.* New York: Basic Books.
Fryns, J. P., Kleczkowska, A., Kubien, E., & Van den Berghe, H. (1995). XYY syndrome and other Y chromosome polysomies: Mental status and psychosocial functioning. *Genetic Counseling, 6,* 197-206.
Holl, R. W., Kunze, D., Etzrodt, H., Teller, W., & Heinze, E. (1994). Turner's syndrome: Final height, glucose tolerance, bone density and psychosocial status in 25 adult patients. *European Journal of Pediatrics, 153,* 11-16.
Holler, W., Scholz, S., Knorr, D., Bidlingmaier, F., Keller, E., & Albert, E. D. (1985). Genetic differences between the salt-wasting, simple virilizing, and nonclassical types of congenital adrenal hyperplasia. *Journal of Clinical Endocrinology and Metabolism, 60,* 757-763.
Hovatta, O., Foudila, T., & Söderström-Anttila, V. (2000). Assisted reproductive techniques in Turner syndrome. In P. Saenger & A. M. Pasquino (Eds.), *Optimizing health care for Turner patients in the 21st century* (pp. 247-53). New York: Elsevier.
Moore, C. (2000). Sorting by Sex. *American Scientist, 88,* 545-546.
Moran, C., Azziz, R., Carmina, E., Dewailly, D., Fruzzetti, F., Ibanez, L., et al. (2000). 21-Hydroxylase deficient nonclassic adrenal hyperplasia is a progressive disorder: A multicenter study. *American Journal of Obstetrics and Gynecology, 183,* 1468-1474.
Moreno, A., & Goodwin, J. (1998, March). Am I a woman or a man? *Mademoiselle,* 178-181, 208.
Moura, M. D., Navarro, P. A., & Nogueira, A. A. (2000). Pregnancy and term delivery after neo-vaginoplasty in a patient with vaginal agenesis. *International Journal of Gynaecology and Obstetrics, 71,* 215-216.
Ranke, M. B., & Saenger, P. (2001). Turner's syndrome. *Lancet, 358,* 309-314.
Robson, S., & Oliver, G. D. (2000). Management of vaginal agenesis: Review of 10 years practice at a tertiary referral centre. *Australian and New Zealand Journal of Obstetrics and Gynaecology, 40,* 430-433.
Sas, T. C., Keizer-Schrama, S. M., Stijnen, T., Jansen, M., Otten, B. J., Hoorweg-Nijman, J. J., et al.(1999). Normalization of height in girls with Turner syndrome after long-term growth hormone treatment: Results of a randomized dose-response trial. *Journal of Clinical Endocrinology and Metabolism, 84,* 4607-4612.
Simpson, J. L. (1999). Genetics of the female reproductive ducts. *American Journal of Medical Genetics, 89,* 224-239.
Speiser, P. W., Knochenhauer, E. S., Dewailly, D., Fruzzetti, F., Marcondes, J. A. M., & Azziz, R. (2000). A multicenter study of women with nonclassical congenital adrenal hyperplasia: Relationship between genotype and phenotype. *Molecular Genetics and Metabolism, 71,* 527-534.
Szasz, T. S. (1974). *The myth of mental illness.* New York: Harper & Row.
Warburg, E. (1963). A fertile patient with Klinefelter's syndrome. *Acta Endocrinologica, 43,* 12-26.
Weaver, C. (2000, March 26). Birds do it. *Washington Post Book World,* p. 6.
White, P. C. (2001). Congenital adrenal hyperplasia. *Clinical Endocrinology and Metabolism, 15,* 17-41.
Wiener, J. S. (1999). Insights into the causes of sexual ambiguity. *Current Opinion in Urology, 9,* 507-511.

Manuscript accepted February 5, 2002

Endocrine CONNECTIONS | M Scharff, C M Wiepjes *et al.* | Grip strength change in trans people | **8**:7 | **1020**–1028

RESEARCH

Change in grip strength in trans people and its association with lean body mass and bone density

Miranda Scharff[1,*], Chantal Maria Wiepjes[1,*], Maartje Klaver[1], Thomas Schreiner[2], Guy T'Sjoen[3] and Martin den Heijer[1]

[1]Department of Endocrinology and Center of Expertise on Gender Dysphoria, Amsterdam University Medical Center, Vrije Universiteit, Amsterdam, the Netherlands
[2]Department of Endocrinology, Oslo University Hospital, Oslo, Norway
[3]Department of Endocrinology & Center for Sexology and Gender, Ghent University Hospital, Ghent, Belgium

Correspondence should be addressed to M den Heijer: m.denheijer@amsterdamumc.nl

*(M Scharff and C M Wiepjes contributed equally to this work)

Abstract

Objective: Gender-affirming hormonal treatment (HT) in trans people changes physical appearance. Muscle mass and strength are important aspects of physical appearance, but few data exist on the effect of HT on grip strength and muscle mass. This study aimed to investigate the change in grip strength in trans people during the first year of HT and to study the possible determinants of this change and the associations between changes in grip strength, lean body mass and bone mineral density (BMD).
Design and methods: A multicenter, prospective study was performed, including 249 transwomen and 278 transmen. Grip strength, lean body mass and BMD were measured at baseline and after 1 year.
Results: After 1 year of HT, grip strength decreased with −1.8 kg (95% CI −2.6; −1.0) in transwomen and increased with +6.1 kg (95% CI +5.5; +6.7) in transmen. No differences in grip strength change was found between age groups, BMI groups, hormonal administration routes or hormone concentrations. In transmen, increase in grip strength was associated with increase in lean body mass (per kg increase in grip strength: +0.010 kg, 95% CI +0.003; +0.017), while this was not found in transwomen (per kg increase in grip strength: +0.004 kg, 95% CI −0.000; +0.009). Change in grip strength was not associated with change in BMD in transwomen and transmen.
Conclusions: After 1 year of HT, grip strength decreased in transwomen, and increased in transmen. In transmen only, change in grip strength was associated with change in lean body mass.

Key Words
- transgender
- gender-affirming hormonal treatment
- grip strength
- muscle mass

Endocrine Connections (2019) **8**, 1020–1028

Introduction

According to the DSM criteria, gender dysphoria (GD) is defined as the incongruence between a person's assigned sex at birth and the experienced gender (1). GD often has a strong impact on the person's psychological wellbeing. Therefore, most trans people prefer treatment to change physical appearance, for example gender-affirming hormonal treatment (HT) with or without gender-affirming surgery.

Grip strength and muscle mass are good indications of the masculinity of the body. A decrease in grip strength and muscle mass could lead to a more feminine body for transwomen (male-to-female trans people),

https://ec.bioscientifica.com
https://doi.org/10.1530/EC-19-0196

while an increase in grip strength and muscle mass indicates a more masculine body for transmen (female-to-male trans people). Besides the importance of physical appearance, a change in grip strength and muscle mass also might be important in the prevention of sarcopenia (the age-related loss of muscle mass) (2) and dynapenia (the age-related loss of muscle strength) (3).

Earlier studies on effects of HT in trans people focused on bone mineral density (BMD) and body composition. These studies found an increase in lean body mass in transmen and a decrease in lean body mass in transwomen (4, 5). BMD increased in both transwomen and transmen (6). A study in cis gender people found grip strength to be a predictor of bone mass (7). Possibly, a change in BMD in trans people is associated with a change in grip strength. Some studies described the change in grip strength (8, 9); however, sample sizes of these studies were small and possible influences on these changes, for example age, BMI, administration routes of HT and sex hormone concentrations during HT, have not been studied.

The aim of this study is to investigate the time course of change in grip strength in trans people in the first year of HT, to study possible determinants of this change and to study the reciprocal associations between changes in grip strength, lean body mass (as an approximation of muscle mass) and BMD.

Materials and methods

Study design and study population

This study is part of the European Network for the Investigation of Gender Incongruence (ENIGI) study, a multicenter prospective cohort study, including treatment centers in Amsterdam, Ghent, Oslo and Florence using the same treatment protocol. The study design is published previously (10, 11) and the study is registered at https://clinicaltrials.gov/ct2/show/NCT01072825. In short, people were included from 2010 until April 2016. The included people were 18 years and older and gave informed consent. People could participate in this study when they started with HT, if they did not use gender-affirming hormones before the start of HT, and if they spoke the native language. During the first year of treatment, the people were seen every 3 months. For the current study, only data from Amsterdam, Ghent and Oslo were analyzed, as grip strength was not assessed in Florence. People were included if their grip strength was measured at baseline and after 12 months. For the analyses on lean body mass and BMD, people were included if a dual-energy X-ray absorptiometry (DXA) was performed at baseline (range 4 months before to 4 months after baseline) and after 12 months (range 10–14 months) of HT. People from Oslo were excluded from the analyses on lean body mass and BMD, because a different type of DXA scanner was used (Oslo: Lunar (GE Lunar, Madison, WI, USA); Amsterdam and Ghent: Hologic Discovery A (Hologic Inc, Bedford, MA, USA)). In total, 1017 participants were included in the overall study. After exclusions due to unknown grip strength values at baseline or 12 months ($n=60$), a follow-up of less than 1 year ($n=363$) or lost to follow-up ($n=57$), a total of 249 transwomen and 278 transmen were included in our analyses (Fig. 1).

Transwomen were treated with the anti-androgen cyproterone acetate (CPA) 50 mg daily, in combination with 2–4 mg oral estradiol valerate a day or 100 µg/24 h estradiol patch twice a week. People older than 40 years were advised to be treated with transdermal estrogens, because of thrombosis risk (12). Transmen were treated with testosterone. They could choose between testosterone gel (50 mg daily), testosterone esters (250 mg intramuscular every 2–3 weeks) or testosterone undecanoate (1000 mg intramuscular every 12 weeks).

The Medical Ethics Review Committee of Ghent approved the study protocol. Local Ethical Review Committees approved participations in the other centers.

Clinical data collection

Grip strength was measured in kilograms (kg), using an adjustable hand-held standard grip device, a Jamar Dynamometer (10, 13). The grip strength of the dominant hand was measured twice and the highest value was noted. These measurements were performed at baseline and after 3 months, 6 months, 9 months and 12 months of HT. However, from Oslo only data at baseline and after 12 months were available.

Body weight and height were measured at every visit. People were measured without shoes and in light indoor clothes. BMI was calculated by weight divided by the square of body height.

DXA

A whole-body DXA was performed to measure lean body mass of the arms and the legs. Outcomes of body composition were determined using manufacturer-supplied algorithms. The specific arm and leg regions were defined with software from Hologic.

Absolute BMD values were obtained for lumbar spine (L1–L4, LS), total hip (TH) and femoral neck (FN).

https://ec.bioscientifica.com
https://doi.org/10.1530/EC-19-0196

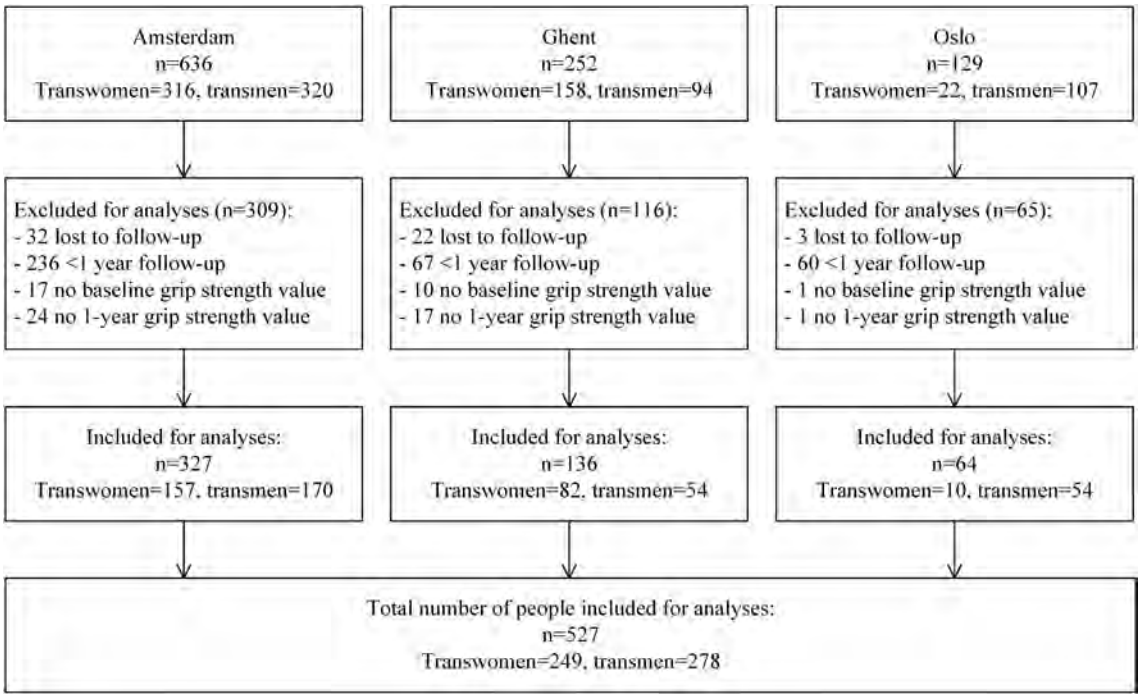


Figure 1
Inclusion flowchart.

The same type of DXA device was used in Amsterdam and Ghent (Hologic Discovery A, Hologic Inc., MA, USA) and software version 13.5.3 was used.

Laboratory measurements

Venous blood samples were obtained in the morning at baseline, after 3 months of HT and after 12 months of HT. As the results were also used during clinical care, assays with higher quality were implemented when they were available. In order to compare the new and old values likewise, conversion formulas were generated by the endocrine laboratory.

In Oslo, testosterone and estradiol were determined using a competitive immunoassay (ECLIA, Roche Diagnostic) with a lower limit of quantitation (LOQ) of 0.1 nmol/L (2.6 ng/dL) and 18.4 pmol/L (5.0 pg/mL), and a coefficient of variation (CV) of 5 and 7%, respectively. In Ghent, an E170 Modular (E2 Gen II, Roche Diagnostics) was used for testosterone (LOQ 0.4 nmol/L (11.5 ng/dL), CV 2.6%) and estradiol (LOQ 92 pmol/L (25.1 pg/mL), CV 3.2%). For estradiol, it was updated in March 2015 to an E170 Modular (E2 Gen III, Roche Diagnostics) with a conversion formula of Gen III = 6.687940 + 0.834495 × Gen II. In Amsterdam, estradiol was measured using a competitive immunoassay (Delfia, PerkinElmer, Wallac Oy, Turku, Finland) with a LOQ of 20 pmol/L (5.4 pg/mL) and a CV of <13% until July 2014. Thereafter, an LC-MS/MS (VUmc, Amsterdam, the Netherlands; LOQ 20 pmol/L (5.4 pg/mL), CV <7%) was used, with a conversion formula of LC-MS/MS = 1.60 × Delfia-29. Testosterone was measured using a radioimmunoassay (RIA, Coat-A-Count, Siemens, Los Angeles, CA, USA; LOQ 1 nmol/L (28.8 ng/dL), CV <10%) until January 2013. After that, it was measured using a competitive immunoassay (Gen III, Architect, Abbott, Abbott Park, IL, USA) with a LOQ of 0.1 nmol/L (2.9 ng/dL) and a CV <10%. Two conversion formulas were generated: <8 nmol/L: Architect = 1.1 × RIA + 0.2; >8 nmol/L: Architect = 1.34 × RIA-1.65.

Statistical analyses

Results are presented as mean with standard deviation (S.D.) in case of normal distribution, or median with interquartile range (IQR) for non-normally distributed data, and percentages. In case of non-normally distributed data, a log transformation was performed before further analyses.

To examine the course of the change in grip strength during the first year, linear mixed model analyses were performed. The influence of age, differences in BMI, different administration routes of HT and serum hormone concentrations, were analyzed. Age was divided into groups (<25 years, 25–40 years, ≥40 years), to stratify for accrual of grip strength, peak grip strength and

https://ec.bioscientifica.com
https://doi.org/10.1530/EC-19-0196

age-related decrease of grip strength. BMI was defined as underweight (<18.5 kg/m^2), normal weight (18.5–25 kg/m^2) and overweight (≥25 kg/m^2). For analyses between differences in the administration route of HT, people who used the same administration route for at least 9 months were included. Mean concentrations of estradiol and testosterone during HT were calculated by averaging the results of the measurements after 3 and after 12 months of HT. As different assays were used in the centers to determine estradiol and testosterone and no conversion formulas between the centers were available, it was not possible to analyze the estradiol and testosterone concentrations as absolute values. Therefore, center-specific tertiles were created for estradiol and testosterone and were thereafter analyzed together. The center-specific mean estradiol concentrations in transwomen were first tertile 122 pmol/L (33 pg/mL, Amsterdam), 159 pmol/L (43 pg/mL, Ghent) and 167 pmol/L (46 pg/mL, Oslo); second tertile 233 pmol/L (63 pg/mL, Amsterdam), 248 pmol/L (68 pg/mL, Ghent) and 317 pmol/L (83 pg/mL, Oslo); and third tertile 405 pmol/L (110 pg/mL, Amsterdam), 572 pmol/L (156 pg/mL, Ghent) and 398 pmol/L (108 pg/mL, Oslo). The center-specific mean testosterone concentrations for transmen were first tertile: 16 nmol/L (461 ng/dL, Amsterdam), 10 nmol/L (288 ng/dL, Ghent) and 15 nmol/L (432 ng/dL, Oslo); second tertile: 28 nmol/L (806 ng/dL, Amsterdam), 16 nmol/L (461 ng/dL, Ghent) and 21 nmol/L (605 ng/dL, Oslo); and third tertile: 54 nmol/L (1555 ng/dL, Amsterdam), 24 nmol/L (691 ng/dL, Ghent) and 33 nmol/L (950 ng/dL, Oslo).

The average of the right and left arm was calculated for the analyses on the lean body mass of the arms, and the average of the right and left leg was calculated to analyze the lean body mass of the legs. To evaluate the change of grip strength, lean body mass, BMD, and serum creatinine after 1 year, the absolute and percentage difference between the baseline values and the values after 12 months were calculated. As these were normal distributed variables, linear regression analyses were performed to calculate the mean and 95% CI.

To study the association between changes in grip strength and changes in lean body mass, BMD or creatinine concentrations, linear regression analyses were performed.

All analyses were performed separately for transwomen and transmen. Data were analyzed using STATA Statistical Software (Statacorp, version 15.1). *P* values <0.050 were considered statistically significant.

Results

Characteristics

The characteristics are shown in Table 1. No differences were found between included and excluded people in

Table 1 Characteristics of the study population.

	Transwomen (*n* = 249)		**Transmen** (*n* = 278)	
	Baseline	During HT	Baseline	During HT
Age, years	28 (23–40)		23 (20–30)	
BMI, kg/m^2	23.8 (4.5)	24.5 (4.4)	25.5 (5.6)	25.9 (4.8)
Smoking (% yes)	24.7	14.7	29.9	19.9
Alcohol (% yes)	44.8	44.6	55.3	55.8
Grip strength (kg)	41.8 (8.9)	40.0 (8.9)	33.1 (6.5)	39.2 (6.8)
Creatinine (μmol/L)	78.5 (10.8)	73.1 (10.6)	66.0 (9.0)	77.4 (10.6)
Hormone administration route[a]				
Estradiol oral	–	134	–	–
Estradiol transdermal	–	86	–	–
Testosterone gel	–	–	–	52
Testosterone undecanoate i.m.	–	–	–	130
Testosterone esters i.m.	–	–	–	81
Estradiol concentrations, pmol/L				
Amsterdam	105 (83–131)	225 (141–329)	142 (59–371)	180 (138–242)
Ghent	109 (88–130)	246 (181–342)	166 (116–431)	134 (113–169)
Oslo	120 (100–140)	257 (195–325)	280 (180–480)	168 (140–210)
Testosterone concentrations, nmol/L				
Amsterdam	18.5 (14.0–23.0)	0.8 (0.6–0.9)	1.3 (1.0–1.7)	27.8 (19.5–39.0)
Ghent	17.9 (13.5–21.5)	0.7 (0.5–1.1)	1.0 (0.7–1.3)	16.3 (11.2–21.1)
Oslo	18.1 (12.2–20.4)	0.4 (0.4–0.6)	1.0 (0.8–1.3)	21.4 (16.7–28.5)

Data are presented as median with inter quartile range, mean with standard deviation, percentages or absolute numbers.
[a]Only in people who used the same hormone administration route for >75% of the follow-up time.

https://ec.bioscientifica.com
https://doi.org/10.1530/EC-19-0196

baseline grip strength, age, BMI, estradiol concentrations, testosterone concentrations or smoking habits (Supplementary Table 1, see section on supplementary data given at the end of this article).

For the analyses on lean body mass, 171 transwomen and 154 transmen were included. For these analyses, excluded transwomen were younger (median age 26 years, IQR 21–34) than included transwomen (median age 31 years, IQR 23–42). Excluded transmen were younger (median age 22 years, IQR 19–25) and had a lower BMI (mean 24.3 kg/m^2, S.D. 5.2) than included transmen (median age 28 years, IQR 20–33, and mean BMI 26.0 kg/m^2, S.D. 5.5). No differences were found between included and excluded people in baseline grip strength, estradiol concentrations, testosterone concentrations or smoking habits.

For the analyses on LS BMD, 207 transwomen and 229 transmen were included. For TH and FN, 206 transwomen and 216 transmen were included. No differences were found between included and excluded transwomen. Excluded transmen were younger (median age 21 years, IQR 19–25) than included transmen (median age 24 years, IQR 21–31).

Grip strength

In transwomen, grip strength decreased with −1.8 kg (95% CI −2.6; −1.0), while in transmen, grip strength increased with +6.1 kg (95% CI +5.5; +6.7). The course of the grip strength change over time is shown in Fig. 2. For transwomen, 66% of the decrease in grip strength (−1.2 kg) occurred in the last 3 months, while in transmen 49% of the increase (+3.0 kg) occurred in the first 3 months. Change in grip strength did not vary between different age groups (Fig. 3A), different BMI groups (Fig. 3B) and different administration routes of HT (Fig. 3C), for both transwomen and transmen. Change in grip strength did not vary between different hormone concentrations, for either transwomen or transmen (Fig. 3D). No analyses on testosterone concentrations in transwomen could be performed, as testosterone was suppressed (<2 nmol/L, <58 ng/dL) in the majority of the transwomen (94%).

Grip strength in relation to lean body mass

As reported previously in this study population (4), a decrease in lean body mass was observed in transwomen and an increase in lean body mass was observed in transmen during the first year of HT.

Associations of change in grip strength with change in lean body mass are presented in Table 2. In transwomen, no associations between change in grip strength and change in arm or leg lean body mass was seen. In transmen, increase in grip strength was associated with an increase in arm lean body mass (per kg increase in grip strength: +0.010 kg, 95% CI +0.003; +0.017), but not with change in leg lean body mass.

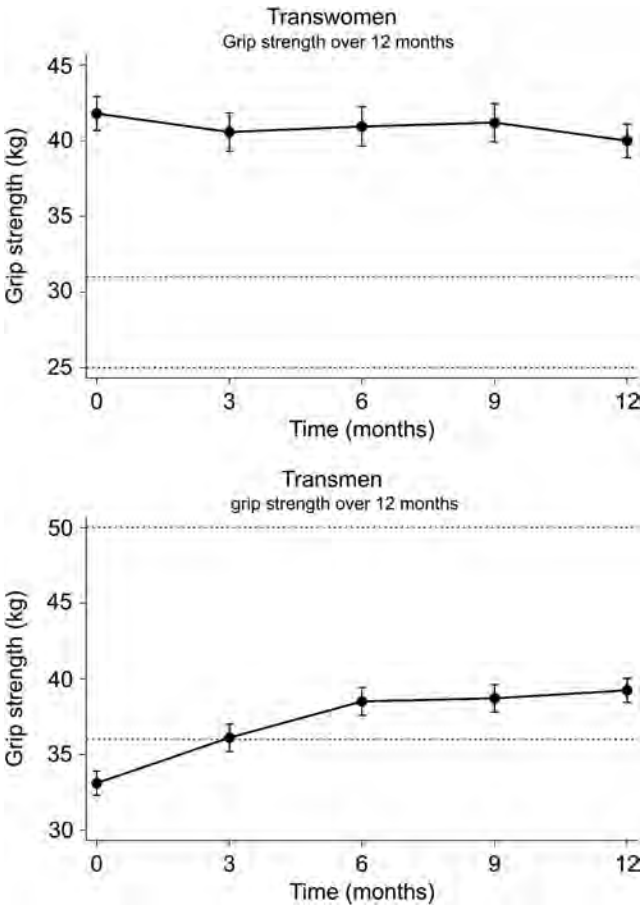


Figure 2
Change in grip strength during the first 12 months of gender-affirming hormonal treatment in transwomen and transmen. Data are presented as means with 95% CI. The 25th and 75th percentiles of the reference populations are shown with dashed lines. For transwomen the 25th percentile is 25 kg and the 75th percentile is 31 kg. For transmen the 25th percentile is 36 kg and the 75th percentile is 50 kg (18).

Grip strength in relation to BMD

As reported previously in this study population (6), increases in LS BMD, TH BMD and FN BMD were observed in transwomen, and increases in LS BMD and TH BMD, but not FN BMD, were observed in transmen during the first year of HT.

The associations of change in grip strength with change in BMD are shown in Table 2.

https://ec.bioscientifica.com
https://doi.org/10.1530/EC-19-0196

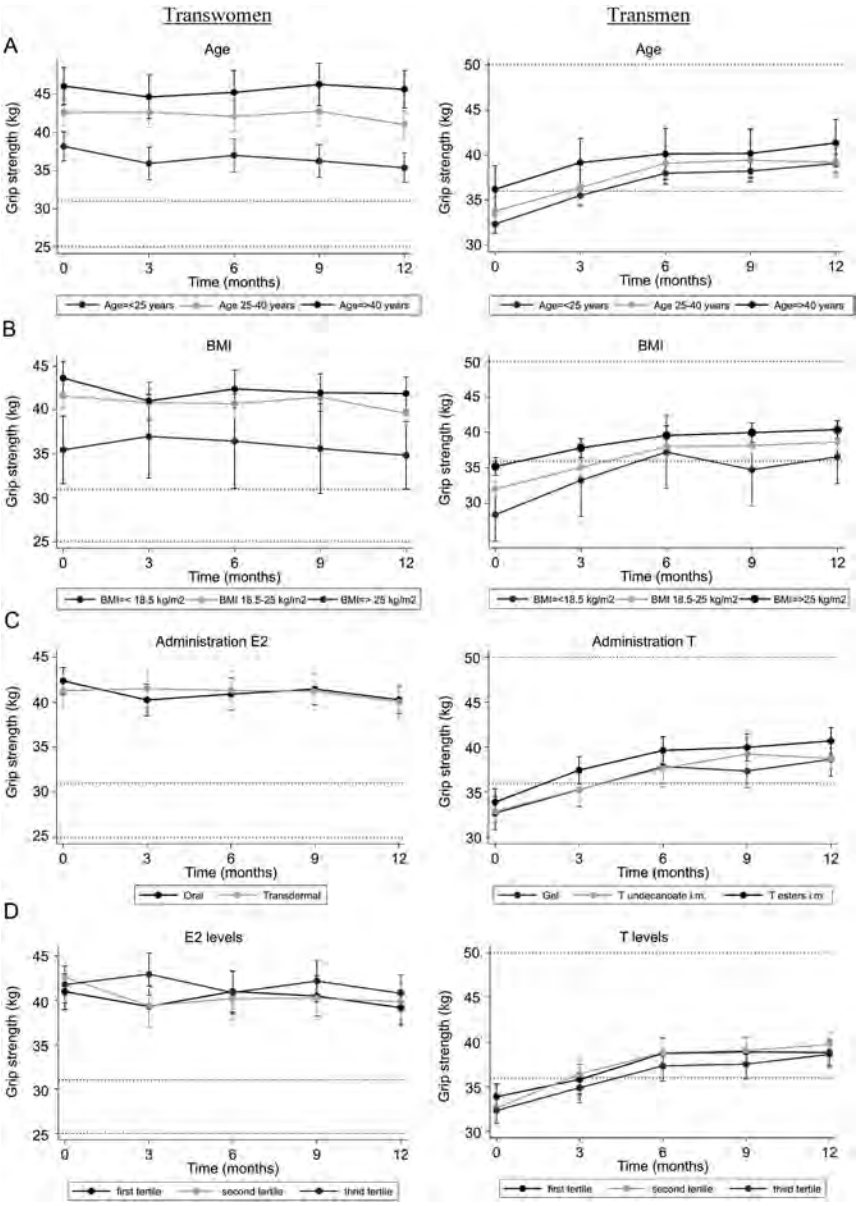


Figure 3
Differences in change in grip strength for age groups, BMI groups, routes of hormonal administration, estradiol concentrations and testosterone concentrations. Data are presented as means with 95% CIs. Reference values (25th and 75th percentile) are based on the mean age of this population. For transwomen the 25th percentile is 25 kg and the 75th percentile is 31 kg. For transmen the 25th percentile is 36 kg and the 75th percentile is 50 kg (18). (A) Grip strength change over 12 months between age, adjusted for administration route. (B) Grip strength change over 12 months between BMI groups. (C) Grip strength change over 12 months between different administration routes, adjusted for age. (D) Grip strength change over 12 months between different hormone concentrations.

Table 2 Associations between change in grip strength (per kg increase) and lean body mass, bone mineral density and creatinine, separately for transwomen and transmen.

	Transwomen		Transmen	
	Mean change (95% CI)	*P* value	Mean change (95% CI)	*P* value
Lean body mass				
Arm	+0.004 kg (–0.000; +0.009)	0.079	+0.010 kg (+0.003; +0.017)	0.003
Leg	+0.009 kg (–0.003; +0.021)	0.161	+0.014 kg (–0.002; +0.030)	0.078
Bone mineral density				
Lumbar spine	+0.02% (–0.05; +0.10)	0.556	+0.01% (–0.08; +0.09)	0.900
Total hip	+0.05% (–0.01; +0.10)	0.112	–0.02% (–0.09; +0.05)	0.630
Femoral neck	+0.05% (–0.03; +0.12)	0.201	+0.04% (–0.06; +0.15)	0.401
Creatinine	–0.1 µmol/L (–0.2; +0.1)[a]	0.535	+0.2 µmol/L (+0.0; +0.4)	0.035

[a]Data shown is per kg decrease in grip strength.

https://ec.bioscientifica.com
https://doi.org/10.1530/EC-19-0196

Change in grip strength was not associated with change in BMD in both transwomen and transmen.

Grip strength in relation to creatinine

Serum creatinine concentrations decreased with −5.0 µmol/L (95% CI −6.2; −3.8) in transwomen and increased with +11.1 µmol/L (95% CI +10.1; +12.2) in transmen. The associations of change in grip strength with change in creatinine are described in Table 2. In transmen, the increase in creatinine was associated with an increase in grip strength (per kg increase in grip strength: +0.2 µmol/L, 95% CI +0.0; +0.4). In transwomen, a decrease in creatinine tended to be associated with a decrease in grip strength (per kg decrease in grip strength: −0.1 µmol/L, 95% CI −0.2; +0.1).

Discussion

In transwomen, a decrease in grip strength after 1 year of HT was found, while in transmen, an increase in grip strength after 1 year of HT was found. These changes were also found in the serum creatinine concentrations. Grip strength change did not vary between age groups, BMI groups, administration routes and different hormone concentrations, for neither transwomen nor transmen. Change in grip strength was associated with change in lean body mass in transmen but not in transwomen. Change in grip strength was not associated with a change in BMD.

In our study, transwomen decreased in grip strength, possibly due to the lack of testosterone. This is in agreement with a study in young men using gonadotropin-releasing hormone agonists, which found that muscle mass decreased by approximately 1 kg after 10 weeks (14). The finding of an increase in grip strength in transmen is consistent with results from studies about testosterone replacement in hypogonadal men (15). Testosterone has an effect on myoblast proliferation and myoblast differentiation, and testosterone increases the number of satellite cells, which promotes protein synthesis of muscle mass (16). Thus, testosterone plays an important role in muscle mass and muscle strength. Accordingly, testosterone is not only important for muscle in men, but also for muscle in women. This is among others important for the prevention of dynapenia in older people.

The largest decrease in grip strength for transwomen took place in the last 3 months of HT, while the largest change for transmen took place in the first 3 months. The change in transmen is in line with a randomized controlled trial that shows an increase in strength in the first 6 months, after the use of testosterone in healthy men ≥60 years old (17). Interestingly, the time difference of the slow loss of grip strength in transwomen and the fast increase in transmen is opposite of what is observed for strength with training and detraining. It might possibly take longer for muscles to decrease in proteins due to lack of testosterone, than it is to increase due to administration of testosterone. However, to fully understand the difference, more research is necessary. After 12 months, the median grip strength of transwomen still falls into the 95th percentile for age-matched females. The median grip strength of transmen after 12 months falls into the 25th percentile of age-matched males (18). Thus, transwomen are still stronger than average females and transmen are still weaker than average males. However, as this study is a follow-up of 12 months, transwomen and transmen might attain a grip strength value closer to the reference values of, respectively, females and males after a longer duration of HT. No difference was found for grip strength change between age groups, BMI groups and administration routes. A larger decrease in transwomen and a smaller increase in transmen was expected at age ≥40 years, because of the age-related decrease of grip strength in combination with the lack of testosterone. However, our study population of ≥40 years was not very old (mean age 49 years for transwomen and 47 years for transmen). Furthermore, our study suggests that change in grip strength was not related to mean serum hormone concentrations during the first year of HT. This finding is line with a study on endocrine determinants in sarcopenia in 518 men aged 40–79 years, which found no association between total and bioavailable testosterone and the annual change in grip strength in men, after a follow-up of 4.3 years (19). The finding that concentrations of testosterone and estradiol were not associated with the change in grip strength can be used to inform trans people that a higher dosage does not lead to a larger increase or decrease in grip strength. Also, because no relationship was found between change in grip strength and administration routes, trans people can make a more well-informed decision when choosing between administration routes.

In transwomen, no association was found between change in lean body mass and change in grip strength, while it was only weakly associated in transmen. This is in contrast to a large cross-sectional and longitudinal perspective study of 847 participants aged 20–100 years, which found a strong correlation between grip strength and lean body mass (20). However, this correlation was

age dependent, and it was the strongest at 60 years old. People younger than 60 years appeared to be stronger than predicted with muscle mass and people older than 60 years appeared to be weaker. They also showed a stronger correlation between age and grip strength compared to age and muscle mass. Our population has a mean age of 25 years, this is significantly younger than 60 years, thus the correlation might be lower. Furthermore, change in grip strength was not associated to change in BMD. Therefore, change in BMD cannot be predicted by change in grip strength.

This is a large, multicenter, prospective study. Standardized measurements were used to measure grip strength, and there is a widespread age range. To our knowledge, this is the first study that describes the change in grip strength over the course of a year and that examines possible influences on this change. However, our study also had some limitations. First, data about physical activity were only available in smaller subgroup of the study population and was therefore not analyzed. Therefore, we could miss data that can partially explain the change in grip strength. However, in previous studies, no change in physical activity was found in transwomen (8) and transmen (9) during the first year of HT. Second, possibly a true impact of testosterone concentrations on grip strength cannot be detected, since circulating testosterone concentrations were evaluated, instead of available testosterone in muscle cells (21). In addition, the laboratory measurements were performed differently at the study centers and changed in two sites during follow-up. Although conversion formulas were generated and tertiles were used instead of absolute values, it might be that the variability in hormone concentrations affected the results. Lastly, it is questionable whether a decrease in grip strength in transwomen of almost 2 kg is clinically relevant. Nonetheless, transwomen possibly feel more feminine due to the decrease in muscle mass (22, 23). One study reported on the desired effects of HT in transwomen in Indonesia, including reduced muscle mass (23). A study in associations among masculinity, strength and attractiveness shows a positive correlation between masculinity and grip strength, and a negative correlation between masculinity and attractiveness in young women (24). This might indicate that a decrease in grip strength in women could influence the feeling of attractiveness. However, this may be dissimilar in transwomen. A questionnaire for transwomen and transmen would be of interest to investigate a possible satisfaction with the change in grip strength.

In conclusion, grip strength decreases in transwomen and increases in transmen, after 12 months of HT. This is interesting for the prevention of sarcopenia and dynapenia, since testosterone has a positive effect on muscle mass and grip strength. In addition, knowing the effects of HT on grip strength, its association with hormone concentrations and hormonal administration routes, and its relation with change in muscle mass and BMD, can help care providers and trans people with their expectations of HT. For further research, it would be interesting to evaluate the subjective satisfaction related to changes in grip strength and to evaluate the change of grip strength after a longer follow-up to see if the grip strength will reach the grip strength reference values for women or men, for transwomen and transmen, respectively.

Supplementary data

This is linked to the online version of the paper at https://doi.org/10.1530/EC-19-0196.

Declaration of interest

The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

Funding

This research did not receive any specific grant from any funding agency in the public, commercial or not-for-profit sector.

References

1 Cohen-Kettenis PT & Pfafflin F. The DSM diagnostic criteria for gender identity disorder in adolescents and adults. *Archives of Sexual Behavior* 2010 **39** 499–513. (https://doi.org/10.1007/s10508-009-9562-y)

2 Walston JD. Sarcopenia in older adults. *Current Opinion in Rheumatology* 2012 **24** 623–627. (https://doi.org/10.1097/BOR.0b013e328358d59b)

3 Clark BC & Manini TM. What is dynapenia? *Nutrition* 2012 **28** 495–503. (https://doi.org/10.1016/j.nut.2011.12.002)

4 Klaver M, de Blok CJM, Wiepjes CM, Nota NM, Dekker MJHJ, de Mutsert R, Schreiner T, Fisher AD, T'Sjoen G & den Heijer M. Changes in regional body fat, lean body mass and body shape in trans persons using cross-sex hormonal therapy: results from a multicenter prospective study. *European Journal of Endocrinology* 2018 **178** 163–171. (https://doi.org/10.1530/EJE-17-0496)

5 Wierckx K, Van Caenegem E, Schreiner T, Haraldsen I, Fisher AD, Toye K, Kaufman JM & T'Sjoen G. Cross-sex hormone therapy in trans persons is safe and effective at short-time follow-up: results from the European network for the investigation of gender incongruence. *Journal of Sexual Medicine* 2014 **11** 1999–2011. (https://doi.org/10.1111/jsm.12571)

6 Wiepjes CM, Vlot MC, Klaver M, Nota NM, de Blok CJ, de Jongh RT, Lips P, Heijboer AC, Fisher AD, Schreiner T, *et al*. Bone mineral density increases in trans persons after 1 year of hormonal treatment: a multicenter prospective observational study. *Journal of Bone and Mineral Research* 2017 **32** 1252–1260. (https://doi.org/10.1002/jbmr.3102)

7 Chan DC, Lee WT, Lo DH, Leung JC, Kwok AW & Leung PC. Relationship between grip strength and bone mineral density in healthy Hong Kong adolescents. *Osteoporosis International* 2008 **19** 1485–1495. (https://doi.org/10.1007/s00198-008-0595-1)

8 Van Caenegem E, Wierckx K, Taes Y, Schreiner T, Vandewalle S, Toye K, Kaufman JM & T'Sjoen G. Preservation of volumetric bone density and geometry in trans women during cross-sex hormonal therapy: a prospective observational study. *Osteoporosis International* 2015 **26** 35–47. (https://doi.org/10.1007/s00198-014-2805-3)

9 Van Caenegem E, Wierckx K, Taes Y, Schreiner T, Vandewalle S, Toye K, Lapauw B, Kaufman JM & T'Sjoen G. Body composition, bone turnover, and bone mass in trans men during testosterone treatment: 1-year follow-up data from a prospective case-controlled study (ENIGI). *European Journal of Endocrinology* 2015 **172** 163–171. (https://doi.org/10.1530/EJE-14-0586)

10 Dekker MJHJ, Wierckx K, Van Caenegem E, Klaver M, Kreukels BP, Elaut E, Fisher AD, van Trotsenburg MAA, Schreiner T, den Heijer M, *et al*. A European Network for the investigation of gender incongruence: endocrine part. *Journal of Sexual Medicine* 2016 **13** 994–999. (https://doi.org/10.1016/j.jsxm.2016.03.371)

11 Kreukels BP, Haraldsen IR, De Cuypere G, Richter-Appelt H, Gijs L & Cohen-Kettenis PT. A European network for the investigation of gender incongruence: the ENIGI initiative. *European Psychiatry* 2012 **27** 445–450. (https://doi.org/10.1016/j.eurpsy.2010.04.009)

12 Asscheman H, Gooren LJG & Eklund PLE. Mortality and morbidity in transsexual patients with cross-gender hormone treatment. *Metabolism: Clinical and Experimental* 1989 **38** 869–873. (https://doi.org/10.1016/0026-0495(89)90233-3)

13 Muhldorfer-Fodor M, Ziegler S, Harms C, Neumann J, Cristalli A, Kalpen A, Kundt G, Mittlmeier T & Prommersberger KJ. Grip force monitoring on the hand: manugraphy system versus Jamar dynamometer. *Archives of Orthopaedic and Trauma Surgery* 2014 **134** 1179–1188. (https://doi.org/10.1007/s00402-014-2027-3)

14 Bhasin S, Woodhouse L, Casaburi R, Singh AB, Bhasin D, Berman N, Chen X, Yarasheski KE, Magliano L, Dzekov C, *et al*. Testosterone dose-response relationships in healthy young men. *American Journal of Physiology: Endocrinology and Metabolism* 2001 **281** E1172–E1181. (https://doi.org/10.1152/ajpendo.2001.281.6.E1172)

15 Bhasin S, Storer TW, Berman N, Yarasheski KE, Clevenger B, Phillips J, Lee WP, Bunnell TJ & Casaburi R. Testosterone replacement increases fat-free mass and muscle size in hypogonadal men. *Journal of Clinical Endocrinology and Metabolism* 1997 **82** 407–413. (https://doi.org/10.1210/jcem.82.2.3733)

16 Brown M. Skeletal muscle and bone: effect of sex steroids and aging. *Advances in Physiology Education* 2008 **32** 120–126. (https://doi.org/10.1152/advan.90111.2008)

17 Storer TW, Basaria S, Traustadottir T, Harman SM, Pencina K, Li Z, Travison TG, Miciek R, Tsitouras P, Hally K, *et al*. Effects of testosterone supplementation for 3 years on muscle performance and physical function in older men. *Journal of Clinical Endocrinology and Metabolism* 2017 **102** 583–593. (https://doi.org/10.1210/jc.2016-2771)

18 Wong SL. Grip strength reference values for Canadians aged 6 to 79: Canadian Health Measures Survey, 2007 to 2013. *Health Reports* 2016 **27** 3–10.

19 Gielen E, O'Neill TW, Pye SR, Adams JE, Wu FC, Laurent MR, Claessens F, Ward KA, Boonen S, Bouillon R, *et al*. Endocrine determinants of incident sarcopenia in middle-aged and elderly European men. *Journal of Cachexia, Sarcopenia and Muscle* 2015 **6** 242–252. (https://doi.org/10.1002/jcsm.12030)

20 Kallman DA, Plato CC & Tobin JD. The role of muscle loss in the age-related decline of grip strength: cross-sectional and longitudinal perspectives. *Journal of Gerontology* 1990 **45** M82–M88. (https://doi.org/10.1093/geronj/45.3.M82)

21 Roy TA, Blackman MR, Harman SM, Tobin JD, Schrager M & Metter EJ. Interrelationships of serum testosterone and free testosterone index with FFM and strength in aging men. *American Journal of Physiology: Endocrinology and Metabolism* 2002 **283** E284–E294. (https://doi.org/10.1152/ajpendo.00334.2001)

22 Gorin-Lazard A, Baumstarck K, Boyer L, Maquigneau A, Gebleux S, Penochet JC, Pringuey D, Albarel F, Morange I, Loundou A, *et al*. Is hormonal therapy associated with better quality of life in transsexuals? A cross-sectional study. *Journal of Sexual Medicine* 2012 **9** 531–541. (https://doi.org/10.1111/j.1743-6109.2011.02564.x)

23 Idrus NI & Hymans TD. Balancing benefits and harm: chemical use and bodily transformation among Indonesia's transgender waria. *International Journal on Drug Policy* 2014 **25** 789–797. (https://doi.org/10.1016/j.drugpo.2014.06.012)

24 Van Dongen S. Associations among facial masculinity, physical strength, fluctuating asymmetry and attractiveness in young men and women. *Annals of Human Biology* 2014 **41** 205–213. (https://doi.org/10.3109/03014460.2013.847120)

Received in final form 18 June 2019
Accepted 26 June 2019
Accepted Preprint published online 26 June 2019

Aaron Baum, PhD
Mark D. Schwartz, MD

Author Affiliations: Department of Health System Design and Global Health, Icahn School of Medicine at Mount Sinai, New York, New York (Baum); Department of Population Health, New York University School of Medicine, New York (Schwartz).

Corresponding Author: Aaron Baum, PhD, Department of Health System Design and Global Health, Icahn School of Medicine at Mount Sinai, One Gustave L. Levy Place, New York, NY 10029 (aaron.baum@mssm.edu).

Accepted for Publication: May 22, 2020.

Published Online: June 5, 2020. doi:10.1001/jama.2020.9972

Author Contributions: Dr Baum had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.
Concept and design: All authors.
Acquisition, analysis, or interpretation of data: Baum.
Drafting of the manuscript: Baum.
Critical revision of the manuscript for important intellectual content: All authors.
Statistical analysis: Baum.
Administrative, technical, or material support: Schwartz.
Supervision: Schwartz.

Conflict of Interest Disclosures: Dr Baum reported receiving consulting fees from the American Board of Family Medicine outside the submitted work. Dr Schwartz reported being employed by the US Department of Veterans Affairs.

1. Rosenbaum L. The untold toll—the pandemic's effects on patients without Covid-19. *N Engl J Med*. Published online April 17, 2020. doi:10.1056/NEJMms2009984

2. COVID-19 response plan—incident-specific annex to the VHA high consequence infection (HCI) base plan. Published March 23, 2020. Accessed May 12, 2020. https://www.va.gov/opa/docs/VHA_COVID_19_03232020_vF_1.pdf

3. Fihn SD, Francis J, Clancy C, et al. Insights from advanced analytics at the Veterans Health Administration. *Health Aff (Millwood)*. 2014;33(7):1203-1211. doi:10.1377/hlthaff.2014.0054

Divergence in Timing and Magnitude of Testosterone Levels Between Male and Female Youths

Data on testosterone levels in children and adolescents segregated by sex are scarce and based on convenience samples or assays with limited sensitivity and accuracy. Such data would be useful in evaluating children with pubertal or androgen disorders and dichotomizing male and female youths participating in sport. Thus, we analyzed the timing of the onset and magnitude of the divergence in testosterone in youths aged 6 to 20 years by sex using a highly accurate assay.

Methods | Testosterone concentrations from separate cohorts of male and female youths collected during 2 periods of the National Health and Nutrition Examination Survey (NHANES; 2013-2014 and 2015-2016) were pooled into 1 data set for analyses. Briefly, NHANES uses a multistage probability design to randomly sample US residents from all 50 states. The overall response rate in the 2 data collection cycles was 70.4% in youths, and 80% of those responders elected to participate in the collection of biospecimens for testosterone analyses. All procedures accessed public, deidentified information and did not require ethical review as determined by the Mayo Clinic Institutional Review Board.

As described previously,[1] testosterone was quantified via isotope dilution liquid chromatography tandem mass spectrometry, which demonstrates a broad analytical measurement range (0.75-1400 ng/dL), excellent precision across a wide range (<3% coefficient of variation) and high accuracy (-0.7% mean bias for a 2-year period), confirmed using reference materials from the National Institute for Standards and Technology.

Full factorial analysis of variance was used to examine the change in testosterone concentration from ages 6 to 20 years by sex, focusing on the age of divergence of testosterone and the overlap at the extremes. Two-tailed post hoc analyses (Scheffe test) were used to test for differences between pairs with Bonferroni-corrected P values ($P < .025$). For all other analyses, significance was determined at $P < .05$. All analyses were performed with R software, version 3.4.2 (R Foundation).

Results | The data set included 4495 youth samples—2293 male and 2202 female—with diverse racial representation including Hispanic (36%), white (26.6%), black (23.0%), Asian (8.8%), and multiracial (6.1%). No statistical differences of race (effects or interactions) were noted.

The median testosterone concentration increased for female youths from age 6 to 20 years from 2.4 ng/dL to 29.5 ng/dL ($P < .001$), with a plateau beginning at age 14 years (**Table**). Over the same age range, the median testosterone concentration increased considerably more for male youths compared with female youths (age × sex; $P < .001$), from 1.9 ng/dL at age 6 years to 516 ng/dL at age 20 years ($P < .001$), with a plateau beginning at age 17 years. Testosterone concentration was not different between the sexes from age 6 to 10 years; however, male youths had greater testosterone concentrations than female youths from age 11 to 20 years (**Figure**).

Among youths aged 12 years or older, there was no overlap of the interquartile range of testosterone between male and female youths. After cessation of the age-related increase in testosterone for female youths (at 14 years), there was an intersection of testosterone concentration distributions between the lowest (first) percentile of male youths and the uppermost (99th) percentile of female youths (≥100 ng/dL), which includes 8 of 949 samples (<1%) for female youths.

Discussion | These data demonstrated the following: (1) the sex-related divergence of testosterone initiated at 11 years of age on average; (2) clear and distinct distributions of serum testosterone between the sexes after 11 years of age; and (3) the distribution of testosterone within male youths was much larger in magnitude and spread than the distribution of testosterone within female youths. At the population level, serum testosterone created a clear dichotomy between male and female youths, and the presented age-adjusted distributions may be useful in evaluation of pubertal and androgenic disorders in youths.

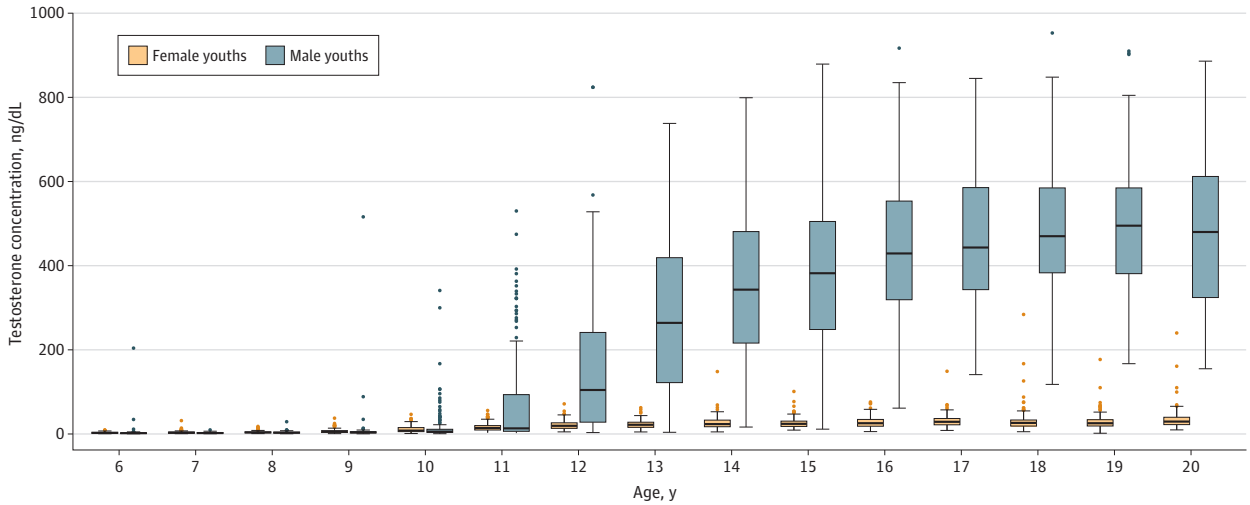
A testosterone value of 100 ng/dL distinctly separated the sexes with minimal overlap, which may have broad implications for athletic competition, as serum testosterone has been demonstrated to be strongly associated with sex differences in athletic performance.[2,3] Potential testosterone thresholds for eligibility in sports may need to be adjusted based on further information on outliers and direction of error accepted.

Table. Age-Adjusted Testosterone Concentration Percentiles

	Male youths						Female youths						
		Testosterone concentration, median, ng/dL, by percentile						Testosterone concentration, median, ng/dL, by percentile					
Age, y	No.	5th	25th	50th	75th	95th	No.	5th	25th	50th	75th	95th	*P* value[a]
6	157	0.5	1.1	1.9	2.7	4.9	166	0.8	1.6	2.4	4.0	6.5	.64
7	187	0.5	1.4	2.2	3.5	5.4	144	1.2	2.1	3.0	4.2	8.4	.16
8	171	1.2	2.0	3.0	4.4	7.6	158	1.3	2.6	3.7	5.0	9.4	.17
9	161	1.1	2.3	3.7	5.4	9.1	161	1.7	3.4	5.3	7.9	18.6	.66
10	177	1.8	3.7	5.6	11.2	76.1	168	3.0	5.5	8.0	15.2	26.2	.35
11	171	3.1	6.2	13.3	95.7	327	185	4.9	9.4	14.2	20.0	38.5	<.001
12	148	7.9	27.5	105	250	497	144	7.1	13.3	19.5	26.6	40.3	<.001
13	157	16.4	121	264	424	620	127	8.0	15.7	21.9	27.9	42.8	<.001
14	165	64.0	216	343	482	699	157	9.5	17.1	23.4	33.0	47.0	<.001
15	158	140	246	382	506	745	133	12.1	17.9	24.1	30.5	51.4	<.001
16	151	149	319	429	554	695	178	11.4	18.3	25.5	34.6	56.5	<.001
17	140	220	343	443	589	779	131	13.0	21.6	28.9	36.9	56.8	<.001
18	142	265	380	470	587	737	145	12.1	18.6	26.3	33.2	62.0	<.001
19	120	235	381	496	595	804	125	10.9	18.8	25.4	34.0	65.0	<.001
20	88	188	326	516	632	862	80	10.2	21.8	29.5	40.0	98.1	<.001

[a] Comparing male youths vs female youths at the 50th percentile.

Figure. Total Testosterone Concentrations of the US Population Aged 6 to 20 Years



The horizontal line in the middle of each box indicates the median; top and bottom borders, 75th and 25th percentiles; whiskers above and below the box, 90th and 10th percentiles; and circles beyond the whiskers, outliers beyond the 90th or 10th percentiles.

These analyses were limited by a lack of information on pubertal stages and history of androgenic disorders and by self- or parental report of male/female sex.

Jonathon W. Senefeld, PhD
Doriane Lambelet Coleman, JD
Patrick W. Johnson
Rickey E. Carter, PhD
Andrew J. Clayburn
Michael J. Joyner, MD

Author Affiliations: Department of Anesthesiology and Perioperative Medicine, Mayo Clinic, Rochester, Minnesota (Senefeld, Clayburn, Joyner); Duke University School of Law, Durham, North Carolina (Lambelet Coleman); Department of Health Sciences Research, Mayo Clinic, Jacksonville, Florida (Johnson, Carter).

Corresponding Author: Jonathon W. Senefeld, PhD, Department of Anesthesiology and Perioperative Medicine, Mayo Clinic, 200 First St SW, Rochester, MN 55905 (senefeld.jonathon@mayo.edu).

Accepted for Publication: March 28, 2020.

Author Contributions: Drs Senefeld and Joyner had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.
Concept and design: Senefeld, Lambelet Coleman, Carter, Clayburn, Joyner.
Acquisition, analysis, or interpretation of data: Senefeld, Johnson, Carter, Clayburn, Joyner.
Drafting of the manuscript: Senefeld, Carter, Clayburn.
Critical revision of the manuscript for important intellectual content: All authors.

Statistical analysis: Senefeld, Johnson, Carter.
Obtained funding: Joyner.
Administrative, technical, or material support: Lambelet Coleman, Joyner.
Supervision: Lambelet Coleman, Joyner.

Conflict of Interest Disclosures: Ms Lambelet Coleman reported receipt of personal fees from World Athletics. No other disclosures were reported.

1. Zhou H, Wang Y, Gatcombe M, et al. Simultaneous measurement of total estradiol and testosterone in human serum by isotope dilution liquid chromatography tandem mass spectrometry. *Anal Bioanal Chem*. 2017;409(25):5943-5954. doi:10.1007/s00216-017-0529-x

2. Senefeld JW, Clayburn AJ, Baker SE, Carter RE, Johnson PW, Joyner MJ. Sex differences in youth elite swimming. *PLoS One*. 2019;14(11):e0225724. doi:10.1371/journal.pone.0225724

3. Handelsman DJ. Sex differences in athletic performance emerge coinciding with the onset of male puberty. *Clin Endocrinol (Oxf)*. 2017;87(1):68-72. doi:10.1111/cen.13350

COMMENT & RESPONSE

Stress Ulcer Prophylaxis for ICU Patients

To the Editor The trial comparing proton pump inhibitors and histamine-2 receptor blockers for stress ulcer prophylaxis in patients in the intensive care unit (ICU) receiving mechanical ventilation[1] found that even though proton pump inhibitors were more effective in reducing the risk of gastrointestinal bleeding, there was no significant difference between proton pump inhibitors and histamine-2 receptor blockers in relation to 90-day all-cause mortality, *Clostridioides difficile* infection, or ICU and hospital lengths of stay. The study indirectly measured the effect of gastric acid suppression on the stated outcomes. The conclusion requires that all proton pump inhibitors be uniformly effective.

However, the data show great differences in the relative potency of proton pump inhibitors in terms of suppression of gastric acidity.[2] The most commonly used proton pump inhibitor in this and most other similar studies was pantoprazole,[1,3] which is the least effective proton pump inhibitor in terms of gastric acid suppression (40 mg of pantoprazole is similar in potency to approximately 9 mg of omeprazole).[3] Because of the low potency of pantoprazole, the potential benefits and harms related to acid suppression cannot be representative of all proton pump inhibitors. The concept that proton pump inhibitors are interchangeable is incorrect.

The unanswered questions regarding acid suppression in stress ulcer prophylaxis are whether the benefits of acid suppression and reduced gastrointestinal bleeding are outweighed by potential harms such as increased length of stay, morbidity (eg, infections), or all-cause mortality. In this study, weak acid suppression produced weak benefits.

It may be time to test the relative benefits and risks of highly effective acid suppression using one of the more potent proton pump inhibitors, such as esomeprazole or rabeprazole.[2] A negative result would obviate further trials, whereas a positive result would prompt the search for the minimal reliably effective dose and duration of therapy. Studies with the least potent proton pump inhibitors cannot be used to impute the risks and benefits of acid suppression for an entire class of drug.

Aylin Tansel, MD, MPH
David Y. Graham, MD

Author Affiliations: Division of Gastroenterology and Hepatology, Medical University of South Carolina, Charleston (Tansel); Department of Medicine, Baylor College of Medicine, Houston, Texas (Graham).

Corresponding Author: Aylin Tansel, MD, MPH, Division of Gastroenterology and Hepatology, Medical University of South Carolina, 30 Courtenay Dr, STB Ste 249, MSC 702, Charleston, SC 29425 (atansel@gmail.com).

Conflict of Interest Disclosures: Dr Graham reported receiving support from Phantom and receiving grants from RedHill Biopharma and the National Institute of Diabetes and Digestive and Kidney Diseases. No other disclosures were reported.

1. Young PJ, Bagshaw SM, Forbes AB, et al; PEPTIC Investigators for the Australian and New Zealand Intensive Care Society Clinical Trials Group, Alberta Health Services Critical Care Strategic Clinical Network, and the Irish Critical Care Trials Group. Effect of stress ulcer prophylaxis with proton pump inhibitors vs histamine-2 receptor blockers on in-hospital mortality among ICU patients receiving invasive mechanical ventilation: the PEPTIC randomized clinical trial. *JAMA*. 2020;323(7):616-626. doi:10.1001/jama.2019.22190

2. Graham DY, Tansel A. Interchangeable use of proton pump inhibitors based on relative potency. *Clin Gastroenterol Hepatol*. 2018;16(6):800-808.e7. doi:10.1016/j.cgh.2017.09.033

3. Krag M, Marker S, Perner A, et al; SUP-ICU Trial Group. Pantoprazole in patients at risk for gastrointestinal bleeding in the ICU. *N Engl J Med*. 2018;379(23):2199-2208. doi:10.1056/NEJMoa1714919

To the Editor In the study on the effect of stress ulcer prophylaxis with proton pump inhibitors vs histamine-2 receptor blockers on in-hospital mortality among ICU patients receiving invasive mechanical ventilation,[1] no restriction was imposed on the regimen of histamine-2 receptor blockers or proton pump inhibitors, and the dosing regimen was not reported. Ranitidine, which was the major histamine-2 receptor blocker used across the various study sites, is a small molecule that undergoes renal elimination primarily and has a short half-life and low protein binding. For ICU patients who often display augmented clearance or require continuous renal replacement therapy, whether the usual dose of ranitidine (50 mg every 6-8 hours administered intravenously or 150 mg twice daily enterally) provides an adequate serum concentration to confer protection against stress ulcer is unknown.

Moreover, being a competitive inhibitor means that the effect of gastric acid suppression correlates with serum concentration.[2] Bolus dosing of ranitidine may lead to fluctuations in serum concentration and hence fluctuation of gastric pH, which could compromise its protective effect against upper gastrointestinal tract bleeding. In contrast, proton pump inhibitors are highly protein bound and primarily undergo hepatic metabolism. They also have an additional advantage for being irreversible inhibitors of the H^+/K+-adenosine triphosphatase system in the parietal cells, so the effects of proton pump inhibitors last until the regeneration of new proton pumps in the parietal cells regardless of fluctuation in the serum concentration.[3]

Studies have shown ranitidine to be more effective in increasing gastric pH compared with proton pump inhibitors after single or double doses.[4] It might be worthwhile to explore the comparability of adequately dosed continuous infusion of histamine-2 receptor blockers vs proton pump inhibitors. Continuous infusion of histamine-2 receptor blockers maintains a steady serum concentration and hence a constant effect on gastric acid suppression compared with

AMERICAN JOURNAL OF PHYSIOLOGY
CELL PHYSIOLOGY

Am J Physiol Cell Physiol. 2014 Jan 1; 306(1): C3–C18.
Published online 2013 Nov 6. doi: 10.1152/ajpcell.00281.2013: 10.1152/ajpcell.00281.2013

PMCID: PMC3919971
PMID: 24196532

Do you know the sex of your cells?

Kalpit Shah, Charles E. McCormack, and Neil A. Bradbury*

Department of Physiology and Biophysics, Chicago Medical School, Rosalind Franklin University of Medicine and Science, North Chicago, Illinois
*Corresponding author.
Address for reprint requests and other correspondence: N. A. Bradbury, Dept. of Physiology and Biophysics, Chicago Medical School, 3333 Green Bay Road, North Chicago, IL 60064 (e-mail: neil.bradbury@rosalindfranklin.edu).

Received 2013 Sep 11; Accepted 2013 Oct 31.

Abstract

Do you know the sex of your cells? Not a question that is frequently heard around the lab bench, yet thanks to recent research is probably one that should be asked. It is self-evident that cervical epithelial cells would be derived from female tissue and prostate cells from a male subject (exemplified by HeLa and LnCaP, respectively), yet beyond these obvious examples, it would be true to say that the sex of cell lines derived from non-reproductive tissue, such as lung, intestine, kidney, for example, is given minimal if any thought. After all, what possible impact could the presence of a Y chromosome have on the biochemistry and cell biology of tissues such as the exocrine pancreatic acini? Intriguingly, recent evidence has suggested that far from being irrelevant, genes expressed on the sex chromosomes can have a marked impact on the biology of such diverse tissues as neurons and renal cells. It is also policy of *AJP-Cell Physiology* that the source of all cells utilized (species, sex, etc.) should be clearly indicated when submitting an article for publication, an instruction that is rarely followed (http://www.the-aps.org/mm/Publications/Info-For-Authors/Composition). In this review we discuss recent data arguing that the sex of cells being used in experiments can impact the cell's biology, and we provide a table outlining the sex of cell lines that have appeared in *AJP-Cell Physiology* over the past decade.

Keywords: amelogenin, cell line, sex, X chromosome, Y chromosome

IN 2001, THE INSTITUTE OF MEDICINE published a significant report highlighting the importance of sex as a variable in human and experimental studies (278). Over a decade later, the recommendations of this report have received meager acceptance. Most researchers acknowledge the importance of describing the sex of animals used in studies. In many cases only male animals will be used, in order to obviate any "complications" that may arise from hormonal differences in female animals during their reproductive cycles. Sex selection is obviously important in some studies, however. For example, it is obvious that a research study on milk production and lactation would utilize only female animals, whereas studies on spermatogenesis would be confined to male subjects. Despite the clear importance of knowing the sex when using whole animals, such sex assignments are paid scant attention when studies are performed using cell lines (Fig. 1). After all, cells derived from male and female organisms display the same general characteristics. Cells derived from both sexes support metabolic processes, proliferate, and undergo differentiation. Cells, whether they are obtained from a male or female, possess a nucleus, mitochondria, endoplasmic reticulum, Golgi apparatus, and other cellular organelles. The assumption is made that, because there is really no difference in architecture or function between cells from male and female organisms, the Instructions to Authors (when submitting to the *AJP-Cell Physiology*), which state that the source of all cells utilized (species, sex, etc.) should be clearly indicated, can be happily ignored. A survey of a recent issue of *AJP-Cell Physiology* revealed that only two articles referenced the sex of the animal used, and none referenced the sex of the cell lines employed. Even when including a larger sample size, 75% of all recent publications in *AJP-Cell Physiology* did not discuss the sex of cell lines or animals used in the investigations (Fig. 1). Such omissions are not peculiar to *AJP-Cell Physiology* though. A recent review of publications describing the use of cultured cells in cardiovascular studies found a similar paucity of information on the sex of the cell lines utilized (260). Why is the sex of cell lines used in studies so often omitted from the final published article? It is likely that the sex of the cells being used was simply not known by the investigators, who, like most of us, simply regard the sex of our cells as irrelevant. The utility of cultured cells

in identifying biological mechanisms, pathways, and processes is beyond doubt. Indeed, the results from such studies are often the basis for the development of new diagnostic and therapeutic interventions in human medicine. However, only half of the population may have a sex the same as the cell line on which the diagnostic test or treatment was developed. Since all cell lines have a "sex" (278), the complement of sex chromosomes has the potential to influence biochemical pathways and cell physiology (161). In this review, we provide a setting for the basis of differences between male and female cells and highlight why these differences will likely provide novel insights into the roles of the X and Y chromosomes. Throughout this review, we have avoided the use of the word "gender," specifically referring to the "sex" of cells. According to Institute of Medicine, "sex" is a biological construct dictated by the presence of sex chromosomes and in animals and humans the presence of functional reproductive organs. On the other hand, "gender" is a cultural concept referring to behaviors that might be directed by specific stimuli (visual, olfactory) or by psychosocial expectations that result from assigned or perceived sex and therefore can influence biological outcomes (161, 278). This definition has now been accepted as a new policy for sex and gender in reporting research in all APS journals (http://www.the-aps.org/mm/hp/Audiences/Public-Press/For-the-Press/releases/12/9.html). Information on the sex of cell lines routinely used by authors of publications in *AJP-Cell Physiology* is also presented. Finally, we pose several questions that we hope will guide the scientific community with regard to the potential role of sex in studies using cell lines and at least cause researchers to consider the impact of the sex of a cell on the interpretation of experimental results.[1]

Males and Females Are Different

The first question to be asked is "is there any evidence of sex differences between male and female non-sexual tissue that cannot be explained by hormonal differences?" As physiologists, we all accept that there are obvious differences between males and females. In vertebrates, sex differences are usually attributed to the effects of embryonic and post pubertal hormones. Indeed, while many of the more obvious differences between male and female vertebrates are clearly dependent on hormones, the role of hormones in other tissues is much less certain. Aristotle, the ancient Greek philosopher and polymath, more than 2,000 years ago is purported to have articulated the notion that sexual dimorphism exists at the earliest stages of embryonic growth. He believed that male embryos became "animated" 40 days post conception, whereas female embryos required a further 50 days before becoming "animated" (4). Intriguingly, recent studies tend to support the notion of early differences between male and female embryos. For example, male embryos created through in vitro fertilization grow faster prior to implantation than female embryos (6, 199, 284). Importantly, these findings suggest that genetic cellular differences between sexes exist before the onset of hormonal exposure. Moreover, even in adults, hormonal ablation or supplementation does not completely eliminate or recreate sexual differences observed in the progression of certain tumors from male and female patients (38). Furthermore, pathologies that display a sex disparity, such as neurodegenerative (242, 299), cardiovascular (266), and autoimmune (16, 82, 140) disease, differ in frequency but not severity, a difference not readily explained by hormonal differences. Thus, it is clear that not every difference observed between male and female cells can be attributed to differences in exposure to sex hormones. Fundamental to the replication of chromosomes is the telomere, that short region of repetitive nucleotides at the end of each chromatid that protects the chromosome from deterioration or fusion with other chromosomes. The length of the telomere is shorter in older males compared with females (7, 270), leading to the postulation that differences in replicative rates affect telomere shortening and aging (253), and may explain why males die younger than females (236, 274).

Males Have a Y Chromosome

On a simplistic level, differences between male and female cells are entrenched in differences in genetic content, as expressed by the presence of sex chromosomes; two X chromosomes in female cells, and one X and one Y chromosome in male cells (Fig. 2). The role of the Y chromosome in male sex determination arose from observations that XY and XYY (Klinefelter syndrome) individuals develop testes whereas XX and XO (Turner syndrome) individuals instead develop ovaries (72, 104): note that individuals with Turner syndrome have so-called streak gonads located below the fallopian tubes and generally show no evidence of germinal elements (89). Thus while the presence of a single Y chromosome is necessary and sufficient to generate a male gonadal phenotype, the presence of a single X chromosome, while necessary, is not sufficient to generate a full female gonadal phenotype. In 1990, the gene responsible for testicular determination, named *SRY* (sex-determining region Y) was identified (100, 111, 237, 245) (Fig. 2) and comprises a single exon encoding a 204 amino acid protein containing a DNA-binding domain (HMG-box: High Mobility Group), arguing for this protein as a regulator of gene expression. For many decades it was believed that the only role of the Y chromosome was the development of the male gonadal phenotype and the initiation of male fertility (251). This opinion was reinforced by the dearth of other obvious phenotypes that segregated with the Y chromosome; "hairy ears" being one of the few well-documented exceptions (59). However, this concept of the Y chromosome as a genetic wasteland is now being challenged (180, 210). Indeed, since the sex chromosomes account for 5% of the total human genome (1,000–2,000 genes on the X chromosome and <50 genes on the Y chromosome; 129, 224, 246), there is at least the

mathematical possibility that 1:20 proteins (and related biochemical reactions and pathways or cell biological processes) would differ between males and females. Given such odds, it is hard to imagine that cells from males and females would not differ in at least some aspects of cellular biochemistry and physiology. The Y chromosome has two genetically distinct aspects (Fig. 2). The distal part of the short arm of the Y chromosome is shared with the most distal part of the short arm of the X chromosome and can recombine with its X chromosome counterpart during meiosis in males, a region termed the "pseudoautosomal region" because loci in this region undergo recombination during spermatogenesis, akin to homologous recombination between autosomes (98, 215). A second pseudoautosomal region is also present on the distal portion of the long arms of the sex chromosomes (74). The remainder of the Y chromosome does not undergo recombination with the X chromosome and strictly comprises Y chromosome-specific DNA. Compared with other chromosomes, the Y chromosome has a limited number of genes. The roughly two dozen different genes encoded on the Y chromosome can be divided into two categories. One cohort of Y chromosome-specific genes is expressed exclusively in the testes and is likely involved in gonadal development and spermatogenesis; mutation or deletion of some of these genes leads to male infertility (113, 129, 154). A second group of Y chromosome genes consists of genes that do have homologous counterparts on the X chromosome but may yield slightly different final protein products (275). For example, the gene on the Y chromosome encoding the ribosomal protein S4 (*RPS4Y*; Fig. 2), a component of the 40S subunit, results in a slightly different protein than that expressed on the X chromosome (*RPS4X*) with a 19 amino acid difference between the two sex-distinct proteins (66, 275). While functionally equivalent isoforms of ribosomal proteins exist in yeast, these differ by no more than a few amino acids (279), the 19 amino acid difference between "male" and "female" ribosomes suggesting the possibility of differences in ribosomal function between "male" and "female" cells. Similarly, nucleotide sequence analysis of the *ZFY* (zinc finger protein) shows it to be similar but distinct (383 amino acids of 393 are identical) from its X chromosome (*ZFX*) counterpart (233) (Fig. 2). The differences or similarities between other homologous proteins remains to be determined. However, since *RP4SY* and *ZFY* are present only in males, it is possible that such "male"-specific expression can result in potentially extensive biochemical differences between "male" cells and "female" cells. Regardless of whether or not genes on the Y chromosome, other than *SRY*, are important in determining cellular function, the *SRY* genes certainly are. In the 45-day-old 46XY human fetus, these genes cause the gonadal ridge to develop into the testes (89). The fetal testes secrete Mullerian inhibiting hormone, which causes the regression of primordial Mullerian ducts; thus the fallopian tubes and uterus do not develop. The fetal testes also secrete testosterone, causing the differentiation of the primordial Wolffian duct system into the epididymis and vas deferens.

Females Have Two X Chromosomes and No Y

In contrast to male genomes that have only one X chromosome, female genomes have twice the amount of X chromosome genetic material compared with males. Thus, whereas females can be either homozygous or heterozygous with respect to X chromosome-linked traits, males (due to the presence of only one X chromosome) are hemizygous. Products of the X chromosome genes, like those of autosomes, are involved in many aspects of cellular function, metabolism development, and growth (224). Indeed, the X chromosome contains the largest number of immune-related genes within the entire genome (19). In contrast to males where genes from only one X chromosome are present, the occurrence of two X chromosomes in females gives rise to the potential expression of twice the amount of X chromosome DNA in females compared with males. This double dosage of X chromosome genes in females is, however, annulled at many loci by the process of X chromosome inactivation (39, 167, 196, 283). This fundamentally female process is never found in normal XY males (89) and only occurs in female cells outside of the germline. The process of X inactivation profoundly alters the cell's transcriptional landscape, engendering epigenic changes and differential nuclear compartmentation of chromosomes in a highly regulated fashion (97). The inactive chromosome changes conformation to yield a darkly staining mass called the sex chromatin or Barr body (143). Because of the random nature in the choice of which of the two X chromosomes are inactivated (206), females have two epigenetically distinct populations of cells, in which either the maternally or the paternally derived X chromosome is expressed (196). Males, by contrast, only express an active maternally derived X chromosome in all cells; of course the "maternally derived" X chromosome could itself be paternally derived. The random feature of X chromosome inactivation leads to a mosaic of expression of the two X chromosomes in female tissues, and this has been invoked as the basis for lack of a tight genotype-phenotype correlation in the severities of recessive X chromosome-linked diseases (156). A classic example of random X-inactivation is presented by the calico, or tortoiseshell, cat. Each X chromosome expresses either an orange or a black coat coloring, yet the calico cat coat pattern is extremely common. This illustrates the fact that both X chromosomes contribute to the cat's color and explains why almost all calico cats are female (181). Since males only have a single X chromosome, "variants" in genes on one X chromosome cannot be obviated by a second X chromosome. Thus, males demonstrate a clearer, more common or more extreme version of any variant phenotype than females do. This is exemplified at its extreme by "X-lined" genetic diseases, including Duchenne and Becker muscular dystrophies (168), hemophilia (24), porphyria (3) X-linked cone-and rod-dystrophies (160), and color blindness (174). A dramatic example of male hemizygosity for X chromosome-linked traits is seen in X chromosome-linked dominant mutations. Mutations in these genes are embryonically lethal to males in utero and are therefore only seen in females. For example, X chromosome-linked incontinentia pigmenti is a relatively benign dermatological condition in females, but it is lethal to males who inherit a mutant allele (249).

Intriguingly, in females there are reports of a strong somatic selection against cells that bear mutations on the active X chromosome (17, 277). For example, the B-cell lineage in heterozygous females carrying mutations at the X chromosome-linked agammaglobulinemia show selective inactivation of the mutant chromosome and expression of genes from the non-mutant X chromosome (45). Despite the process of X-chromosome inactivation, not all genes on the X chromosome are subject to inactivation (55, 277, 289). As much as 15% of X chromosome-linked genes have been identified as being expressed from the "inactive" X chromosome in at least some cells in culture (34, 35). A notable example of this is seen in the *ZFX* gene (185, 233) (Fig. 2), a zinc finger protein expressed only on the X chromosome and therefore completely absent from males. Moreover, some genes are transcribed with equal efficiency from both the "active" and "inactive" chromosome. For example, gastrin-releasing peptide (GRP) is known to be expressed by both the active and inactive X chromosomes. More than a curiosity, this double expression of GRP may have important clinical consequences, as elevated levels of GRP are proposed to be associated with an elevated risk of lung cancer in women who smoke (241).

Male and Female Cells Are Not the Same

Nearly all biochemical, signaling, and trafficking pathways elucidated for mammalian cells have been obtained from studies on cell lines. Some of these cell lines have been cultured for over 50 years and were considered for their functional and morphological features without regard to their sex origin. A notable exception is the HeLa cell, which is the oldest and probably most widely used of all cell lines. Obtained from a patient with cervical cancer, the cells were taken without consent from Henrietta Lacks, a female patient at Johns Hopkins hospital, who eventually died of her cancer on October 4, 1951 (78, 106, 231). Indeed, the sex of the HeLa cell is fairly well known even to the general public thanks to a recent best seller in the popular science press (247). HeLa cells have been central to many biomedical breakthroughs of the last half century, from their initial use in the development of a polio vaccine (231) to their key role in studies leading to the awarding of two Nobel Prizes in Physiology or Medicine: Harald zur Hausen, in 2008, for his discovery of human papilloma viruses causing cervical cancer (26), and Elizabeth Blackburn, Carol Greider, and Jack Szostak, in 2009, for their discovery of how chromosomes are protected by telomeres and the enzyme telomerase (231, 239, 265). More recently, HeLa cells have again gained prominence as drivers of National Institutes of Health (NIH) policy. In April 2013, a group working at the European Molecular Biology Laboratory in Heidelberg, Germany, published the genome of the HeLa cell line (130). At the same time, an NIH-funded group working at the University of Washington was preparing to publish their version of the HeLa genome (1). Given that immediate descendants of Henrietta Lacks are still alive, concern was raised by other researchers and by the Lacks family that the genome sequence could reveal heritable aspects of Lacks' germline DNA. Such sequence data could be used to draw inferences concerning the Lacks family's medical status, engendering a quagmire of legal and ethical issues. NIH has now implemented a new policy regarding the distribution and use of genome sequence data from HeLa cells (grants.nih.gov/grants/guide/notice-files/NOT-OD-13–099.html). Under the new guidelines, the DNA sequence data from HeLa cells will be subject to controlled use; applications to access the sequence data are being reviewed by a newly formed HeLa Genome Data Access working group at NIH, on which two members of the Lacks family will serve. The hardiness of the HeLa cell has, unfortunately, also proven to be one of its greatest concerns. HeLa cells have been noted to contaminate and indeed overgrow other cell cultures grown within the same laboratory, interfering with, and invalidating, many publications. The degree to which HeLa cell contamination is a problem remains unknown, as few researchers have the time, money, or knowledge for determining the purity of cell lines within their laboratories. However, contamination by HeLa cells have been estimated to range between 10% and 20% of all cell lines in use (150), and cross-contamination remains a major ongoing problem in modern cell cultures (32, 173). Despite these concerns, cell lines are vital to much of current biomedical research. The advances in basic biomedical sciences, and in the development of pharmacological treatments for numerous diseases, would not be possible without the use of cell lines obtained from human and non-human sources. As scientists, we owe a great debt to those patients who have wittingly and unwittingly provided the tissue samples upon which so many of us rely for our research.

Differences between the male and female brain have been a subject of study by philosophers, poets, and scientists alike. It has long been held that sex differences in the brain are caused by differential exposure to gonadal secretions during fetal and neonatal development (5), with distinct sexual dimorphism particularly in sex steroid-concentrating regions (145). However, there is accumulating evidence that supports the notion of sexual dimorphism in the brain in the absence of gonadal secretions (202, 218). For example, morphological and functional sex differences in dopaminergic (and probably noradrenergic) neurons are seen in cultures of rat brain tissues removed at day E14 (day of insemination = E0), whereas the male rat gonad does not start to secrete testosterone until day E15 (217). In fact, no measurable differences in whole body androgens are seen in rats until after E18. Many of the differences in brain-derived cells are retained even following growth of excised tissues, from male and female brains, in identical culture media. Studies by Dewing et al. (53) have described over 50 different sex-dimorphic genes, i.e., genes that display intrinsic differences between male and female cells that are not dependent on hormone exposure and persist in cell culture. Dopaminergic neurons, although accounting for less than 1% of brain neurons, are nonetheless critical for such diverse brain functions as voluntary movement (134, 263), stress response (135), and addictive behavior/reward (159, 220). Dopaminergic neurons from female rat fetuses, in dissociated cell cultures, are morphologically

distinct from those obtained from male rat fetuses, differences that are present even when gonadal hormones are absent (36). Moreover, cultured female neurons display a dopamine uptake rate twice that of their male counterparts (217). Gene array studies using nigral dopaminergic neurons from male and female patients with Parkinson's disease (obtained post mortem by laser capture dissection) have shown considerable sex-specific transcriptional profiles (244). Sex dissimilarities were not confined to a specific pathway but displayed differential transcription patterns in signal transduction, neuronal maturation, protein kinases, proteolysis, and WNT signaling (31, 244). Results from such studies support the notion that being male is a risk factor for Parkinson's disease. Indeed, epidemiological studies have shown that both the incidence and the prevalence of Parkinson's disease are 1.5 to 2 times greater in males that females (144, 280). Furthermore, the age of onset of Parkinson's disease is slightly earlier (mean 2.2 years) in men than women (92).

The hippocampus plays a key role in both short- and long-term memory (133, 184), as well as spatial navigation (22). Cultured male hippocampal neurons survive longer under normoxic conditions than female-derived hippocampal neurons but are more sensitive to ischemia than their female counterparts (99). In Alzheimer's disease, the hippocampus is one of the first regions of the brain to be affected; women are disproportionately affected by Alzheimer's disease, with two thirds of all sufferers being female (37). It is interesting to speculate that the sex disparity observed between male and female Alzheimer's patients may have an underlying basis in genes differentially expressed from the X and/or Y chromosomes in hippocampal neurons. In addition to differences in sensitivity to oxygen tension between male and female cells from the hippocampus, differential sensitivity to a wide range of cytotoxic agents has been shown for several neurons of the central nervous system (CNS) (60). For example, neurons from male rats are more sensitive to nitrosative ($ONOO^-$) stress than those neurons obtained from female rats. In contrast, neurons from female rats are more sensitive to apoptosis-inducing agents (staurosporine and etoposide) than neurons from their male counterparts (60). These observations are relevant to many CNS pathologies, where nitrosative stress is thought to play an important role in cerebral ischemia and traumatic brain injury. At a biochemical level, this may be related to the observation that male neurons are unable to maintain high levels of the reductant glutathione (60), a key protector from oxidative insult (73, 109). Mitochondria from female rats contain higher glutathione peroxidase (a key enzyme in maintaining cellular glutathione levels) activity than those from males (25). Such differences between the ability of male and female neurons to respond to oxidative stress and ischemia may provide an underlying mechanism for the observation that boys have a worse outcome following traumatic brain injury compared with girls (58).

Sex diversity of gene expression is not reserved for the CNS alone, however. For example, kidney cells obtained from female embryonic rats are significantly more sensitive to ethanol- and camptothecin-induced apoptosis than their male counterparts (197). While male and female splenocytes display similar responses to nitrosative stress and staurosporine-induced apoptosis, female splenocytic cells are more sensitive than their male counterparts and react to significantly lower doses of staurosporine than male cells (60). Cyp1A1 is a member of the cytochrome P-450 family, a family of proteins responsible for the metabolism and inactivation of many drugs and toxins (211, 272). Cyp1A1 plays a particularly prominent role in the metabolism of polycyclic aromatic hydrocarbons present in cigarette smoke. Female smokers have a higher level of aromatic/hydrophobic DNA adducts in lung tissue than males, due to a more responsive Cyp1A1 enzyme (166). High levels of lung DNA adducts have been related to an early onset of lung cancer (228), and several, though not all, epidemiological studies have suggested that with similar exposure to cigarette smoke, females may be at greater risk of developing lung cancer than males (295). Differences in drug metabolism are also seen in male and female livers. Female liver cells have more cytochrome CYP3A compared with male liver cells (186). Again, more than just a biochemical curiosity, such differences in CYP3A expression between male and female hepatocytes have important clinical consequences, as the actions of CYP3A account for the metabolism of half the drugs in the pharmacopeia (261, 296). Thus, for 50% of prescription drugs, the effectiveness of a particular drug dosage of 50% may be quite different in females compared with males (90).

A recent attempt to catalog differential gene expression between male and female cells examined 233 lymphoblastoid cell lines: 115 female and 118 male lines. Utilizing 4,799 probes, 10 autosomal genes were identified as having a sex-specific expression pattern (298). These genes encoded a wide variety of proteins involved in multiple cellular processes, including cell adhesion, apoptosis, zinc ion binding, transcription factors, and structural molecules. When such studies are extended to more tissues, it appears that thousands of genes may show sexually dimorphic gene expression (290). Microarray analysis of 23,574 transcripts from murine liver, adipose, muscle, and brain tissues showed highly tissue-specific patterns of sexually dimorphic gene expression (290). The degree of sexual dimorphism ranged from ~14% in brain to ~70% in liver, likely (at least in part) accounting for the differential drug metabolism observed between males and females (2). Given such differences in gene expression, the question arises as to whether such differences result in a physiological phenotype. Stem-cell mediated muscle regeneration in mouse models of muscular dystrophy has raised some interesting data related to this point (51). Female muscle-derived stem cells are less sensitive to oxidative stress and regenerate skeletal muscle much more efficiently than muscle-derived stem cells from their male counterparts when transplanted into *mdx* or *mdx/SCID* mice, a dystrophin-deficient animal model of muscular dystrophy (51). Precisely how these differences arise is not immediately apparent, although differences in handling of oxidative stress appear to be a key feature between male and female cells. The finding that male and female mus-

cle-derived stem cells have different properties is likely to have a big impact on other stem cell-mediated therapies should the findings be replicated for other diseases. Although the molecular mechanisms and genes involved in the sex disparity observed across various cell types await a fuller elucidation, what does seem to be a recurrent theme is the observation that female cells are better able to survive stress than male cells. Given the broad range of stress responses, this could arise from multiple genes present on one or both of the X chromosomes in females.

What Sex Are My Cells?

The notion that there are sex differences between cells has gained prominence through the increased use of primary cells obtained from both animals and humans. For animal studies, the sex of the primary cells can be known without difficulty, though IRB restrictions usually preclude immediate knowledge of the sex of the patient when cells are derived from human tissues. In contrast, the sex of cultured cells has been rarely considered (161, 260; Fig. 1). Indeed, while cultured cell lines have provided a plethora of data on biochemical pathways, cell biological processes, and gene expression, they have essentially been considered asexual objects of study. To facilitate the inclusion of the cells' sex in future manuscript submissions to the *AJP-Cell Physiology*, Table 1 was generated by examining the last decade of papers published in *AJP-Cell Physiology*. Although clearly not an exhaustive list, it does represent the majority of cell lines used by submitting authors. Any omission of cell lines is the responsibility of the authors of this review and was completely unintentional. While the sexes presented in Table 1 are based on published data, it is also possible that some researchers may have worked with contaminated or misidentified cell lines (41), including incorrect sex assignment (63).

Yet, how can sex be determined for a cell line? In the modern era of molecular genetics, determination of the sex of a cell line utilizes an identical approach to that taken by forensic pathologists in determining the sex of human remains. Sexing cells by polymerase chain reaction (PCR)-based methodology is accomplished by amplification of homologous genes found on the X and Y chromosomes. The amelogenin gene is one such gene, and it codes for an extracellular matrix protein found in the developing tooth (64, 230). In humans, it has been determined that there are two amelogenin genes, one on the X chromosome (in the p22 region of the short arm; 156) and the other in the pericentric region of the Y chromosomes (132, 172, 230). Nakahori et al. (172) demonstrated the presence of a 6-bp insertion in intron 1 of the amelogenin-Y sequence (Y chromosome) that was absent from the amelogenin-X (X chromosome) gene. This 6-bp insertion results in a size difference between PCR products covering the intron-1 region, and it has been used to differentiate males from females (91, 147). Since both males and females have at least one X chromosome, the PCR product derived from the X chromosome is, automatically, a positive control. Separation of the “male” and “female” PCR products can be achieved by gel electrophoresis or denaturing high-performance liquid chromatography (240). Thus, females will show a single band for the amelogenin-X gene, whereas males will have two bands, one corresponding to the amelogenin-X gene and one from the male amelogenin-Y gene (Fig. 3). Furthermore, since the area under the curve can be used to quantitate the amount of PCR product, it is also possible to identify XXY (Klinefelter syndrome) and XYY DNA. Amelogenin-based sex tests are part of various PCR multiplex reaction kits from different manufacturers and are widely used for DNA typing for samples in the forensic field (29).

While the determination of amelogenin gene expression should be relatively straightforward, there are a few cases where sex assignment based on this assay has not aligned with classic cytogenetic analysis of metaphase chromosomes. For example, cell line ATTC CRL-5873 (NCI-H1514), established in October 1986 from a 56-yr-old female with small cell lung carcinoma, shows positive for the Y chromosome-specific amelogenin sequence (63). Exactly how Y chromosome-specific PCR products end up in a female cell line is not entirely clear, although the possibility exists that there has been a misidentification of the cell line (http://www.atcc.org/Products/Cells%20and%20Microorganisms/Cell%20Lines/Misidentified%20Cell%20Lines.aspx). As can also be seen in Table 1 (noted by asterisks), some cell lines display an amelogenin test result consistent with a female genotype, yet the tissue of origin is from a documented male donor. For example, the PC-3 cell line is derived from human prostate epithelial cells (arguably an exclusively male tissue type), yet this cell line lacks the amelogenin-Y gene consistent with a male genotype. Indeed, over 100 reportedly “male” cell lines in the ATCC collection appear to have lost all trace of their Y chromosome and yield only X chromosome amelogenin during analysis (190) (Table 1). For example, in 1990, the cell line CRL-2234 was isolated from a hepatocellular carcinoma from a 52-yr-old Asian male (188). According to ATCC and “Short Tandem Repeat” (STR) analysis, CRL-2234 cells characteristically have a low amelogenin-Y peak, which decreases with passage. By *passage 17*, the Y chromosome can no longer be detected by routine amelogenin analysis. Whether other Y chromosome-specific genes are also lost with passage of CRL-2234, or indeed other “male” cell lines, is not known. Such loss of the Y chromosome, of course, severely impedes assessment of sex chromosome function on cellular functions. Several intestinal epithelial cell lines, including T84 (derived from a colonic carcinoma, isolated by H. Masui from a metastatic site in the lung from a 72-yr-old male patient; 54) and HT29 (isolated in 1964 by J. Fogh from a colonic cancer in a 44-yr-old female Caucasian; 271), have been extensively employed both by electrophysiologists and by cell biologists studying transepithelial ion/solute transport and polarized protein trafficking. Moreover, several studies have compared the biology of these cell lines (33, 40, 44). Since T84 colonic epithelial cells are derived from a male patient and HT29 colonic epithelial cells are derived from a female

patient, one would think that the T84 and HT29 cell lines would be an ideal pair to discern any male/female differences in epithelial biology. However, when direct comparisons are made within the same study, little difference has been noted between T84 and HT29 cells. However, data from ATCC reveals that T84 cells have lost their Y chromosome (as detected by amelogenin analysis). Thus, whether similarities between T84 and HT29 cells are due to a biology exclusively related to autosomal gene expression, or whether differences would exist had T84 cells retained their Y chromosome, is difficult to evaluate.

Sex Disparity

Many sex disparities in disease severity and prognosis have been ascribed to hormonal differences. It will be interesting to see how many of these differences really are hormonally mediated and which arise from intrinsic differences in male and female cells that are unrelated to hormonal exposure. However, for human tissues, such experiments are technically difficult, as testes in male fetuses start to develop by weeks 6 to 8. Studies are beginning to elucidate sex differences in gene expression levels and phenotypic responses; many of these, however, are still at the descriptive level. It will be important to define the precise mechanistic underpinnings of these observations of differences between male and female cells and how these observations may impact on cells that are maintained in cell culture. In 1993, the NIH mandated enrollment of women in human clinical trials, yet no similar initiative has been put forward for research using animals. As a result, male-to-female bias in neuroscience research studies has been estimated to be around 5.5:1 for male:female animal studies (15). The reasons for this disparity are likely varied but are mostly based on the assumption that results from males apply to females, or because the presence of hormonal cycles will increase the variance in acquired data, confounding interpretation of data (177, 278). However, based on data presented in this review it is clearly inappropriate to assume that results from studies conducted on only one sex will apply wholesale to the other (182).

Future Perspectives

We are now entering an era of physiological genomics and individualized medicine. It is clear that the presence of an XY or XX chromosome pair will have an impact on how an individual will respond to, or metabolize, a particular drug regimen (2). Many pharmaceutical companies and university research labs are developing high-throughput screening assays to identify and develop drugs for the treatment of various human diseases. Not only are cell lines being utilized, but also primary cells have a growing part in drug screening. The question arises, should we screen on male cells? Female cells? Or both? Even when it is known that there is a sex disparity in disease severity, this issue is rarely raised. For example, it is known that girls with cystic fibrosis do not grow as well as boys and have poorer lung function (223); under the age of 20 there is a 60% greater chance of girls dying compared with boys (52). The development of primary human airway cells as a model for cystic fibrosis has been a huge boon for the discovery and development of the first FDA-approved drug for the treatment of a subpopulation of patients with cystic fibrosis; yet none of the published reports has provided any information on the sex of the patient from whom the airway cells were obtained (267). A similar dearth of sex data is seen in other reports on the identification and development of other drugs for the treatment of patients with cystic fibrosis (268, 269). Such omissions, however, are not solely at the discretion of the researcher. IRB protocols prohibit the disclosure of any data that may lead to patient identification. Thus for compliance, researchers are generally not given access to such data. It can be reasonably argued that it is now time to release the restriction on revealing the sex of tissues used in drug screening. Although the sex of the cells being used in drug screens could be independently determined through amelogenin determination, this is both duplicative of existing data and may be construed as attempts at patient identification. It is clear that IRB members should at least appraise themselves of the importance of researchers knowing the sex of tissues as they develop screening assays. As noted earlier, muscle-derived stem cells from female mice regenerate muscle tissue when transplanted into dystrophic (*mdx*) mice to a greater extent than muscle-derived stem cells from male mice (51). In addition, in *mdx* mice, female hosts exhibit a significantly higher regeneration than male hosts (51). Whether this effect will be seen with human muscle-derived stem cells is not known. Muscle-derived stem cells can also undergo osteogenic differentiation with BMP4 treatment in vitro (128, 281). When male muscle-derived stem cells were used to evaluate ectopic intramuscular bone formation, male hosts (unaltered or castrated males) showed significantly more bone formation that when the same male stem cells were placed in female hosts (unaltered or ovariectomized) (157). Thus clearly, not only does the sex of the stem cell matter, but also the sex of the host into which the cells are placed. Should stem cells prove useful for the treatment of patients with dystrophies or compromised bone wound healing, the sex of the donor and recipient will have an impact on patient prognosis, raising questions of survival and function of cell grafts in the same- and opposite-sexed recipients.

Can those of us who predominantly work with cultured cells escape the impact of the sex of our cells? With few exceptions, cells are cultured in media containing serum, although some manufacturers supply media to be used without serum. Certainly, unless specifically removed, such media will contain sex steroids. What is the impact of these steroids when culturing cells of unknown or indeed known sex? The matter is further complicated if one is utilizing “male” cells that have lost their Y chromosome (Table 1). While the issue of how representative is the biology of a cell line with respect to the tissue from which it was

obtained is one with which most researchers are keenly aware, the potential impact of the loss of an entire chromosome on a cell's biology is seldom considered. It is now perhaps time that such changes are contemplated. Although it is premature (and probably unrealistic) to suggest that all studies be performed on a cohort of both male and female cells prior to publication, the notion that there may be male/female differences in experimental outcomes is clearly not something that should be dismissed out of hand.

GRANTS

Work in the authors' laboratory is supported by Cystic Fibrosis Foundation Grant CFF BRADBU05G0 and National Institutes of Health National Heart, Lung, and Blood Institute Grant HL-102208 (to N. A. Bradbury).

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

AUTHOR CONTRIBUTIONS

K.S. prepared figures; N.A.B. drafted manuscript; C.E.M. and N.A.B. edited and revised manuscript; N.A.B. approved final version of manuscript.

ACKNOWLEDGMENTS

The authors are particularly indebted to the reviewers of our manuscript for many helpful suggestions.

Footnotes

[1]This article is the topic of an Editorial Focus by Catherine M. Fuller and Paul A. Insel (74a).

REFERENCES

1. Adey A, Burton JN, Kitzman JO, Hiatt JB, Lewis AP, Martin BK, Qiu R, Lee C, Shendure J. The haplotype-resolved genome and epigenome of the aneuploid HeLa cancer cell line. *Nature* 500: 207–211, 2013 [PMCID: PMC3740412] [PubMed: 23925245]

2. Anderson GD. Sex differences in drug metabolism: cytochrome P-450 and uridine diphosphate glucuronosyltransferase. *J Gend Specif Med* 5: 25–33, 2002 [PubMed: 11859684]

3. Anderson KE, Bishop DF, Desnick RJ. Disorders of heme biosynthesis: X-linked sideroblastic anemias and the porphyrias. In: *The Metabolic and Molecular Basis of Inherited Disease*, edited by Scriver CR, Beaudet AL, Sly WS. New York: McGraw-Hill, 2001, p. 2991–3062

4. Aristotle *Historia Animalium: Books VII–X*. Cambridge, MA: Harvard Univ. Press, 350 BC

5. Arnold AP. Concepts of genetic and hormonal induction of vertebrate sexual differentiation in the twentieth century, with special reference to the brain. In: *Hormones, Brain and Behavior*, edited by Pfaff D, Arnold AP, Etgen AM, Fahrbach SE, Rubin RT. New York: Academic, 2009, p. 105–135

6. Avery B, Jorgensen CB, Madison V, Greve T. Morphological development and sex of bovine in vitro-fertilized embryos. *Mol Reprod Dev* 32: 265–270, 1992 [PubMed: 1497876]

7. Aviv A, Shay J, Christensen K, Wright W. The longevity gender gap: are telomeres the explanation? *Sci Aging Knowledge Environ* 2005: pe16, 2005 [PubMed: 15944464]

8. Bancroft FC, Levine L, Tashjian AH., Jr Control of growth hormone production by a clonal strain of rat pituitary cells. Stimulation by hydrocortisone. *J Cell Biol* 43: 432–441, 1969 [PMCID: PMC2107805] [PubMed: 5389137]

9. Banks-Schlegel SP, Gazdar AF, Harris CC. Intermediate filament and cross-linked envelope expression in human lung tumor cell lines. *Cancer Res* 45: 1187–1197, 1985 [PubMed: 2578876]

10. Barranco SC, Townsend CM, Jr, Casartelli C, Macik BG, Burger NL, Boerwinkle WR, Gourley WK. Establishment and characterization of an in vitro model system for human adenocarcinoma of the stomach. *Cancer Res* 43: 1703–1709, 1983 [PubMed: 6831414]

11. Barranco SC, Townsend CM, Jr, Quraishi MA, Burger NL, Nevill HC, Howell KH, Boerwinkle WR. Heterogeneous responses of an in vitro model of human stomach cancer to anticancer drugs. *Invest New Drugs* 1: 117–127, 1983 [PubMed: 6678861]

12. Barrett KE, Huott PA, Shah SS, Dharmsathaphorn K, Wasserman SI. Differing effects of apical and basolateral adenosine on colonic epithelial cell line T84. *Am J Physiol Cell Physiol* 256: C197–C203, 1989 [PubMed: 2536228]

13. Basson MD, Beidler DR, Turowski G, Zarif A, Modlin IM, Jena BP, Madri JA. Effect of tyrosine kinase inhibition on basal and epidermal growth factor-stimulated human Caco-2 enterocyte sheet migration and proliferation. *J Cell Physiol* 160: 491–501, 1994 [PubMed: 8077287]

14. Basson MD, Modlin IM, Madri JA. Human enterocyte (Caco-2) migration is modulated in vitro by extracellular matrix composition and epidermal growth factor. *J Clin Invest* 90: 15–23, 1992 [PMCID: PMC443057] [PubMed: 1634605]

15. Beery AK, Zucker I. Sex bias in neuroscience and biomedical research. *Neurosci Biobehav Rev* 35: 565–572, 2011 [PMCID: PMC3008499] [PubMed: 20620164]

16. Beeson PB. Age and sex associations of 40 autoimmune diseases. *Am J Med* 96: 457–462, 1994 [PubMed: 8192178]

17. Belmont JW. Genetic control of X inactivation and processes leading to X-inactivation skewing. *Am J Hum Genet* 58: 1101–1108, 1996 [PMCID: PMC1915050] [PubMed: 8651285]

18. Benda P, Lightbody J, Sato G, Levine L, Sweet W. Differentiated rat glial cell strain in tissue culture. *Science* 161: 370–371, 1968 [PubMed: 4873531]

19. Bianchi I, Lleo A, Gershwin ME, Invernizzi P. The X chromosome and immune associated genes. *J Autoimmun* 38: J187–J192, 2012 [PubMed: 22178198]

20. Biedler JL, Roffler-Tarlov S, Schachner M, Freedman LS. Multiple neurotransmitter synthesis by human neuroblastoma cell lines and clones. *Cancer Res* 38: 3751–3757, 1978 [PubMed: 29704]

21. Billiau A, Edy VG, Heremans H, Van Damme J, Desmyter J, Georgiades JA, De Somer P. Human interferon: mass production in a newly established cell line, MG-63. *Antimicrob Agents Chemother* 12: 11–15, 1977 [PMCID: PMC352146] [PubMed: 883813]

22. Bird CM, Burgess N. The hippocampus and memory: insights from spatial processing. *Nat Rev Neurosci* 9: 182–194, 2008 [PubMed: 18270514]

23. Bolin SR, Ridpath JF, Black J, Macy M, Roblin R. Survey of cell lines in the American Type Culture Collection for bovine viral diarrhea virus. *J Virol Methods* 48: 211–221, 1994 [PubMed: 7989438]

24. Bolton-Maggs PH, Pasi KJ. Haemophilias A and B. *Lancet* 361: 1801–1809, 2003 [PubMed: 12781551]

25. Borras C, Sastre J, Garcia-Sala D, Lloret A, Pallardo F, Vina J. Mitochondria from females exhibit higher antioxidant gene expression and lower oxidative damage than males. *Free Radic Biol Med* 34: 546–552, 2003 [PubMed: 12614843]

26. Boshart M, Gissmann L, Ikenberg H, Kleinheinz A, Scheurlen W, zur Hausen H. A new type of papillomavirus DNA, its presence in genital cancer biopsies and in cell lines derived from cervical cancer. *EMBO J* 3: 1151–1157, 1984 [PMCID: PMC557488] [PubMed: 6329740]

27. Brandes LJ, Hermonat MW. Receptor status and subsequent sensitivity of subclones of MCF-7 human breast cancer cells surviving exposure to diethylstilbestrol. *Cancer Res* 43: 2831–2835, 1983 [PubMed: 6850594]

28. Brandt BL, Kimes BW, Klier FG. Development of a clonal myogenic cell line with unusual biochemical properties. *J Cell Physiol* 88: 255–275, 1976 [PubMed: 178672]

29. Brinkmann B. Is the amelogenin sex test valid? *Int J Legal Med* 116: 63, 2002 [PubMed: 12056521]

30. Brower M, Carney DN, Oie HK, Gazdar AF, Minna JD. Growth of cell lines and clinical specimens of human non-small cell lung cancer in a serum-free defined medium. *Cancer Res* 46: 798–806, 1986 [PubMed: 3940644]

31. Cantuti-Castelvetri I, Keller-McGandy C, Bouzou B, Asteris G, Clark TW, Frosch MP, Standaert DG. Effects of gender on nigral gene expression and parkinson disease. *Neurobiol Dis* 26: 606–614, 2007 [PMCID: PMC2435483] [PubMed: 17412603]

32. Capes-Davis A, Theodosopoulos G, Atkin I, Drexler HG, Kohara A, MacLeod RA, Masters JR, Nakamura Y, Reid YA, Reddel RR, Freshney RI. Check your cultures! A list of cross-contaminated or misidentified cell lines. *Int J Cancer* 127: 1–8, 2010 [PubMed: 20143388]

33. Cario E, Rosenberg IM, Brandwein SL, Beck PL, Reinecker HC, Podolsky DK. Lipopolysaccharide activates distinct signaling pathways in intestinal epithelial cell lines expressing Toll-like receptors. *J Immunol* 164: 966–972, 2000 [PubMed: 10623846]

34. Carrel L, Willard HF. Heterogeneous gene expression from the inactive X chromosome: an X-linked gene that escapes X inactivation in some human cell lines but is inactivated in others. *Proc Natl Acad Sci USA* 96: 7364–7369, 1999 [PMCID: PMC22091] [PubMed: 10377420]

35. Carrel L, Willard HF. X-inactivation profile reveals extensive variability in X-linked gene expression in females. *Nature* 434: 400–404, 2005 [PubMed: 15772666]

36. Carruth LL, Reisert I, Arnold AP. Sex chromosome genes directly affect brain sexual differentiation. *Nat Neurosci* 5: 933–934, 2002 [PubMed: 12244322]

37. Carter CL, Resnick EM, Mallampalli M, Kalbarczyk A. Sex and gender differences in Alzheimer's disease: recommendations for future research. *J Womens Health (Larchmt)* 21: 1018–1023, 2012 [PubMed: 22917473]

38. Ceribelli A, Pino MS, Cecere FL. Gender differences: implications for clinical trials and practice. *J Thorac Oncol* 2: S15–S18, 2007 [PubMed: 17457223]

39. Chang SC, Tucker T, Thorogood NP, Brown CJ. Mechanisms of X-chromosome inactivation. *Front Biosci* 11: 852–866, 2006 [PubMed: 16146776]

40. Chao KL, Dreyfus LA. Interaction of *Escherichia coli* heat-stable enterotoxin B with cultured human intestinal epithelial cells. *Infect Immun* 65: 3209–3217, 1997 [PMCID: PMC175454] [PubMed: 9234777]

41. Chatterjee R. Cell biology. Cases of mistaken identity. *Science* 315: 928–931, 2007 [PubMed: 17303729]

42. Chen TR. Chromosome identity of human prostate cancer cell lines, PC-3 and PPC-1. *Cytogenet Cell Genet* 62: 183–184, 1993 [PubMed: 8428522]

43. Choi HJ, Choi YH, Yee SB, Im E, Jung JH, Kim ND. Ircinin-1 induces cell cycle arrest and apoptosis in SK-MEL-2 human melanoma cells. *Mol Carcinog* 44: 162–173, 2005 [PubMed: 16163705]

44. Christensen J, El-Gebali S, Natoli M, Sengstag T, Delorenzi M, Bentz S, Bouzourene H, Rumbo M, Felsani A, Siissalo S, Hirvonen J, Vila MR, Saletti P, Aguet M, Anderle P. Defining new criteria for selection of cell-based intestinal models using publicly available databases. *BMC Genomics* 13: 274, 2012 [PMCID: PMC3412164] [PubMed: 22726358]

45. Conley ME, Brown P, Pickard AR, Buckley RH, Miller DS, Raskind WH, Singer JW, Fialkow PJ. Expression of the gene defect in X-linked agammaglobulinemia. *N Engl J Med* 315: 564–567, 1986 [PubMed: 3488506]

46. Cozens AL, Yezzi MJ, Kunzelmann K, Ohrui T, Chin L, Eng K, Finkbeiner WE, Widdicombe JH, Gruenert DC. CFTR expression and chloride secretion in polarized immortal human bronchial epithelial cells. *Am J Respir Cell Mol Biol* 10: 38–47, 1994 [PubMed: 7507342]

47. Cussenot O, Berthon P, Berger R, Mowszowicz I, Faille A, Hojman F, Teillac P, Le Duc A, Calvo F. Immortalization of human adult normal prostatic epithelial cells by liposomes containing large T-SV40 gene. *J Urol* 146: 881–886, 1991 [PubMed: 1714974]

48. Darlington GJ. Liver cell lines. *Methods Enzymol* 151: 19–38, 1987 [PubMed: 3431441]

49. Darlington GJ, Bernhard HP, Miller RA, Ruddle FH. Expression of liver phenotypes in cultured mouse hepatoma cells. *J Natl Cancer Inst* 64: 809–819, 1980 [PubMed: 6102619]

50. de Larco JE, Todaro GJ. Epithelioid and fibroblastic rat kidney cell clones: epidermal growth factor (EGF) receptors and the effect of mouse sarcoma virus transformation. *J Cell Physiol* 94: 335–342, 1978 [PubMed: 304450]

51. Deasy BM, Lu A, Tebbets JC, Feduska JM, Schugar RC, Pollett JB, Sun B, Urish KL, Gharaibeh BM, Cao B, Rubin RT, Huard J. A role for cell sex in stem cell-mediated skeletal muscle regeneration: female cells have higher muscle regeneration efficiency. *J Cell Biol* 177: 73–86, 2007 [PMCID: PMC2064113] [PubMed: 17420291]

52. Demko CA, Byard PJ, Davis PB. Gender differences in cystic fibrosis: *Pseudomonas aeruginosa* infection. *J Clin Epidemiol* 48: 1041–1049, 1995 [PubMed: 7775991]

53. Dewing P, Shi T, Horvath S, Vilain E. Sexually dimorphic gene expression in mouse brain precedes gonadal differentiation. *Brain Res Mol Brain Res* 118: 82–90, 2003 [PubMed: 14559357]

54. Dharmsathaphorn K, McRoberts JA, Mandel KG, Tisdale LD, Masui H. A human colonic tumor cell line that maintains vectorial electrolyte transport. *Am J Physiol Gastrointest Liver Physiol* 246: G204–G208, 1984 [PubMed: 6141741]

55. Disteche CM. The great escape. *Am J Hum Genet* 60: 1312–1315, 1997 [PMCID: PMC1716151] [PubMed: 9199551]

56. Doetschman T, Maeda N, Smithies O. Targeted mutation of the Hprt gene in mouse embryonic stem cells. *Proc Natl Acad Sci USA* 85: 8583–8587, 1988 [PMCID: PMC282503] [PubMed: 3186749]

57. Doetschman TC, Eistetter H, Katz M, Schmidt W, Kemler R. The in vitro development of blastocyst-derived embryonic stem cell lines: formation of visceral yolk sac, blood islands and myocardium. *J Embryol Exp Morphol* 87: 27–45, 1985 [PubMed: 3897439]

58. Donders J, Hoffman NM. Gender differences in learning and memory after pediatric traumatic brain injury. *Neuropsychology* 16: 491–499, 2002 [PubMed: 12382988]

59. Dronamraju KR. Y-linkage in man. *Nature* 201: 424–425, 1964 [PubMed: 14110028]

60. Du L, Bayir H, Lai Y, Zhang X, Kochanek PM, Watkins SC, Graham SH, Clark RS. Innate gender-based proclivity in response to cytotoxicity and programmed cell death pathway. *J Biol Chem* 279: 38563–38570, 2004 [PubMed: 15234982]

61. Dumenco L, Oguey D, Wu J, Messier N, Fausto N. Introduction of a murine p53 mutation corresponding to human codon 249 into a murine hepatocyte cell line results in growth advantage, but not in transformation. *Hepatology* 22: 1279–1288, 1995 [PubMed: 7557882]

Daubert Response App. 0683

62. Dunn KC, Aotaki-Keen AE, Putkey FR, Hjelmeland LM. ARPE-19, a human retinal pigment epithelial cell line with differentiated properties. *Exp Eye Res* 62: 155–169, 1996 [PubMed: 8698076]

63. Durkin AS, Cedrone E, Sykes G, Boles D, Reid YA. Utility of gender determination in cell line identity. *In Vitro Cell Dev Biol Anim* 36: 344–347, 2000 [PubMed: 10949990]

64. Eastoe JE. The chemical composition of bone and tooth. *Adv Fluorine Res* 21: 5–17, 1965 [PubMed: 14315616]

65. Ehler E, Babiychuk E, Draeger A. Human foetal lung (IMR-90) cells: myofibroblasts with smooth muscle-like contractile properties. *Cell Motil Cytoskeleton* 34: 288–298, 1996 [PubMed: 8871816]

66. Fisher EM, Beer-Romero P, Brown LG, Ridley A, McNeil JA, Lawrence JB, Willard HF, Bieber FR, Page DC. Homologous ribosomal protein genes on the human X and Y chromosomes: escape from X inactivation and possible implications for Turner syndrome. *Cell* 63: 1205–1218, 1990 [PubMed: 2124517]

67. Flotte TR, Afione SA, Solow R, Drumm ML, Markakis D, Guggino WB, Zeitlin PL, Carter BJ. Expression of the cystic fibrosis transmembrane conductance regulator from a novel adeno-associated virus promoter. *J Biol Chem* 268: 3781–3790, 1993 [PubMed: 7679117]

68. Fogh J. Cultivation, characterization, and identification of human tumor cells with emphasis on kidney, testis, and bladder tumors. *Natl Cancer Inst Monogr*: 5–9, 1978 [PubMed: 571047]

69. Fogh J, Fogh JM, Orfeo T. One hundred and twenty-seven cultured human tumor cell lines producing tumors in nude mice. *J Natl Cancer Inst* 59: 221–226, 1977 [PubMed: 327080]

70. Fogh J, Trempe G. New human tumor cell lines. In: *Human Tumor Cells In Vitro*, edited by Fogh J. New York: Plenum, 1975, p. 115–159

71. Fogh J, Wright WC, Loveless JD. Absence of HeLa cell contamination in 169 cell lines derived from human tumors. *J Natl Cancer Inst* 58: 209–214, 1977 [PubMed: 833871]

72. Ford CE, Jones KW, Polani PE, De Almeida JC, Briggs JH. A sex-chromosome anomaly in a case of gonadal dysgenesis (Turner's syndrome). *Lancet* 1: 711–713, 1959 [PubMed: 13642858]

73. Franco R, Cidlowski JA. Glutathione efflux and cell death. *Antioxid Redox Signal* 17: 1694–1713, 2012 [PMCID: PMC3474185] [PubMed: 22656858]

74. Freije D, Helms C, Watson MS, Donis-Keller H. Identification of a second pseudoautosomal region near the Xq and Yq telomeres. *Science* 258: 1784–1787, 1992 [PubMed: 1465614]

74a. Fuller CM, Insel PA. *I don't know the question, but sex is definitely the answer!* Focus on "In pursuit of scientific excellence: sex matters." *Am J Physiol Cell Physiol* 302: C1269–C1270, 2012; and "Do you know the sex of your cells?" *Am J Physiol Cell Physiol* (November 6, 2013), 10.1152/ajpcell.00281.2013 [PubMed: 24225885] [CrossRef: 10.1152/ajpcell.00281.2013]

75. Gallagher R, Collins S, Trujillo J, McCredie K, Ahearn M, Tsai S, Metzgar R, Aulakh G, Ting R, Ruscetti F, Gallo R. Characterization of the continuous, differentiating myeloid cell line (HL-60) from a patient with acute promyelocytic leukemia. *Blood* 54: 713–733, 1979 [PubMed: 288488]

76. Gaush CR, Hard WL, Smith TF. Characterization of an established line of canine kidney cells (MDCK). *Proc Soc Exp Biol Med* 122: 931–935, 1966 [PubMed: 5918973]

77. Gazdar AF, Linnoila RI, Kurita Y, Oie HK, Mulshine JL, Clark JC, Whitsett JA. Peripheral airway cell differentiation in human lung cancer cell lines. *Cancer Res* 50: 5481–5487, 1990 [PubMed: 2386953]

78. Gey GO, Coffman WD, Kubicek MT. Tissue culture studies on the proliferative capacity of cervical carcinoma and normal epithelium. *Cancer Res* 12: 264–265, 1952

79. Giard DJ, Aaronson SA, Todaro GJ, Arnstein P, Kersey JH, Dosik H, Parks WP. In vitro cultivation of human tumors: establishment of cell lines derived from a series of solid tumors. *J Natl Cancer Inst* 51: 1417–1423, 1973 [PubMed: 4357758]

80. Ginsberg HS, Young CHS. Genetics of adenoviruses. In: *Comprehensive Virology*, edited by Fraenkel-Conrat RR, Wagner H. New York: Plenum, 1977, p. 27–88

81. Gluzman Y. SV40-transformed simian cells support the replication of early SV40 mutants. *Cell* 23: 175–182, 1981 [PubMed: 6260373]

82. Gompel A, Piette JC. Systemic lupus erythematosus and hormone replacement therapy. *Menopause Int* 13: 65–70, 2007 [PubMed: 17540136]

83. Graham FL, Smiley J, Russell WC, Nairn R. Characteristics of a human cell line transformed by DNA from human adenovirus type 5. *J Gen Virol* 36: 59–74, 1977 [PubMed: 886304]

84. Green H. (Inventor) *Triglyceride- accumulating clonal cell line*. US patent 4,003,789, January 18, 1977

85. Green H, Meuth M. An established pre-adipose cell line and its differentiation in culture. *Cell* 3: 127–133, 1974 [PubMed: 4426090]

86. Greene LA, Tischler AS. Establishment of a noradrenergic clonal line of rat adrenal pheochromocytoma cells which respond to nerve growth factor. *Proc Natl Acad Sci USA* 73: 2424–2428, 1976 [PMCID: PMC430592] [PubMed: 1065897]

87. Grimm MO, Jurgens B, Schulz WA, Decken K, Makri D, Schmitz-Drager BJ. Inactivation of tumor suppressor genes and deregulation of the c-myc gene in urothelial cancer cell lines. *Urol Res* 23: 293–300, 1995 [PubMed: 8839385]

88. Gruenert DC, Finkbeiner WE, Widdicombe JH. Culture and transformation of human airway epithelial cells. *Am J Physiol Lung Cell Mol Physiol* 268: L347–L360, 1995 [PubMed: 7900815]

89. Grumbach MM, Hughes IA, Conte FA. Disorders of sex differentiation. In: *Textbook of Endocrinology* (10 ed.), edited by Larsen PR, Kronenberg HM, Melmed S, Polansky S. Philadelphia: Saunders (Elsevier), 2003, p. 865

90. Guengerich FP. Human cytochrome P450 enzymes. In: *Cytochrome P450: Structure, Mechanism, and Biochemistry*, edited by Ortiz de M, ontellano PR. New York: Plenum, 1995, p. 473–535

91. Haas-Rochholz H, Weiler G. Additional primer sets for an amelogenin gene PCR-based DNA-sex test. *Int J Legal Med* 110: 312–315, 1997 [PubMed: 9387013]

92. Haaxma CA, Bloem BR, Borm GF, Oyen WJ, Leenders KL, Eshuis S, Booij J, Dluzen DE, Horstink MW. Gender differences in Parkinson's disease. *J Neurol Neurosurg Psychiatry* 78: 819–824, 2007 [PMCID: PMC2117736] [PubMed: 17098842]

93. Haggerty DF, Young PL, Popjak G, Carnes WH. Phenylalanine hydroxylase in cultured hepatoxytes. I. Hormonal control of enzyme levels. *J Biol Chem* 248: 223–232, 1973 [PubMed: 4348207]

94. Hanauske U, Hanauske AR, Clark GM, Tsen D, Buchok J, von Hoff DD. A new in vitro screening system for anticancer drugs for the treatment of non-small cell lung cancer. *Sel Cancer Ther* 5: 97–111, 1989 [PubMed: 2554453]

95. Harris SA, Enger RJ, Riggs BL, Spelsberg TC. Development and characterization of a conditionally immortalized human fetal osteoblastic cell line. *J Bone Miner Res* 10: 178–186, 1995 [PubMed: 7754797]

96. Harris SA, Spelsberg TC. (Inventors) *Immortalized human fetal osteoblastic cells*. US patent 5,681,701, October 28, 1997

97. Heard E, Disteche CM. Dosage compensation in mammals: fine-tuning the expression of the X chromosome. *Genes Dev* 20: 1848–1867, 2006 [PubMed: 16847345]

98. Henke A, Fischer C, Rappold GA. Genetic map of the human pseudoautosomal region reveals a high rate of recombination in female meiosis at the Xp telomere. *Genomics* 18: 478–485, 1993 [PubMed: 8307556]

99. Heyer A, Hasselblatt M, von Ahsen N, Hafner H, Siren AL, Ehrenreich H. In vitro gender differences in neuronal survival on hypoxia and in 17beta-estradiol-mediated neuroprotection. *J Cereb Blood Flow Metab* 25: 427–430, 2005 [PubMed: 15689954]

100. Hiort O, Holterhus PM. The molecular basis of male sexual differentiation. *Eur J Endocrinol* 142: 101–110, 2000 [PubMed: 10664515]

101. Holtkamp GM, Van Rossem M, de Vos AF, Willekens B, Peek R, Kijlstra A. Polarized secretion of IL-6 and IL-8 by human retinal pigment epithelial cells. *Clin Exp Immunol* 112: 34–43, 1998 [PMCID: PMC1904947] [PubMed: 9566787]

102. Horoszewicz JS, Leong SS, Kawinski E, Karr JP, Rosenthal H, Chu TM, Mirand EA, Murphy GP. LNCaP model of human prostatic carcinoma. *Cancer Res* 43: 1809–1818, 1983 [PubMed: 6831420]

103. Hull RN, Cherry WR, Weaver GW. The origin and characteristics of a pig kidney cell strain, LLC-PK. *In Vitro* 12: 670–677, 1976 [PubMed: 828141]

104. Jacobs PA, Strong JA. A case of human intersexuality having a possible XXY sex-determining mechanism. *Nature* 183: 302–303, 1959 [PubMed: 13632697]

105. Jensen FC, Girardi AJ, Gilden RV, Koprowski H. Infection of human and simian tissue cultures with Rous sarcoma virus. *Proc Natl Acad Sci USA* 52: 53–59, 1964 [PMCID: PMC300571] [PubMed: 14192657]

106. Jones HW, Jr, McKusick VA, Harper PS, Wuu KD. George Otto Gey. (1899–1970). The HeLa cell and a reappraisal of its origin. *Obstet Gynecol* 38: 945–949, 1971 [PubMed: 4942173]

107. Jumarie C, Malo C. Caco-2 cells cultured in serum-free medium as a model for the study of enterocytic differentiation in vitro. *J Cell Physiol* 149: 24–33, 1991 [PubMed: 1939345]

108. Kaighn ME, Narayan KS, Ohnuki Y, Lechner JF, Jones LW. Establishment and characterization of a human prostatic carcinoma cell line (PC-3). *Invest Urol* 17: 16–23, 1979 [PubMed: 447482]

109. Kalyanaraman B. Teaching the basics of redox biology to medical and graduate students: oxidants, antioxidants and disease mechanisms. *Redox Biol* 1: 244–257, 2013 [PMCID: PMC3757692] [PubMed: 24024158]

110. Kao FT, Puck TT. Genetics of somatic mammalian cells. IV. Properties of Chinese hamster cell mutants with respect to the requirement for proline. *Genetics* 55: 513–524, 1967 [PMCID: PMC1211407] [PubMed: 6068403]

111. Kashimada K, Koopman P. Sry: the master switch in mammalian sex determination. *Development* 137: 3921–3930, 2010 [PubMed: 21062860]

112. Katayose Y, Kim M, Rakkar AN, Li Z, Cowan KH, Seth P. Promoting apoptosis: a novel activity associated with the cyclin-dependent kinase inhibitor p27. *Cancer Res* 57: 5441–5445, 1997 [PubMed: 9407946]

113. Kent-First M, Muallem A, Shultz J, Pryor J, Roberts K, Nolten W, Meisner L, Chandley A, Gouchy G, Jorgensen L, Havighurst T, Grosch J. Defining regions of the Y-chromosome responsible for male infertility and identification of a fourth AZF region (AZFd) by Y-chromosome microdeletion detection. *Mol Reprod Dev* 53: 27–41, 1999 [PubMed: 10230814]

114. Kim YW, Kern HF, Mullins TD, Koriwchak MJ, Metzgar RS. Characterization of clones of a human pancreatic adenocarcinoma cell line representing different stages of differentiation. *Pancreas* 4: 353–362, 1989 [PubMed: 2734279]

115. Kimes BW, Brandt BL. Characterization of two putative smooth muscle cell lines from rat thoracic aorta. *Exp Cell Res* 98: 349–366, 1976 [PubMed: 943301]

116. Kimes BW, Brandt BL. Properties of a clonal muscle cell line from rat heart. *Exp Cell Res* 98: 367–381, 1976 [PubMed: 943302]

117. Kit S, Dubbs DR, DeTorres RA, Melnick JL. Enhanced thymidine kinase activity following infection of green monkey kidney cells by simian adenoviruses, simian papovavirus SV40, and an adenovirus-SV40 "hybrid". *Virology* 27: 453–457, 1965 [PubMed: 4285109]

118. Klein E, Ben-Bassat H, Neumann H, Ralph P, Zeuthen J, Polliack A, Vanky F. Properties of the K562 cell line, derived from a patient with chronic myeloid leukemia. *Int J Cancer* 18: 421–431, 1976 [PubMed: 789258]

119. Knowles BB, Aden DP. *Human hepatoma derived cell line, process for preparation thereof, and uses therefor*. US patent 4,393,133, July 12, 1983

120. Knowles BB, Howe CC, Aden DP. Human hepatocellular carcinoma cell lines secrete the major plasma proteins and hepatitis B surface antigen. *Science* 209: 497–499, 1980 [PubMed: 6248960]

121. Koeffler HP, Golde DW. Acute myelogenous leukemia: a human cell line responsive to colony-stimulating activity. *Science* 200: 1153–1154, 1978 [PubMed: 306682]

122. Koeffler HP, Golde DW. Human myeloid leukemia cell lines: a review. *Blood* 56: 344–350, 1980 [PubMed: 6996765]

123. Kohara H, Tabata M, Kiura K, Ueoka H, Kawata K, Chikamori M, Aoe K, Chikamori K, Matsushita A, Harada M. Synergistic effects of topoisomerase I inhibitor, 7-ethyl-10-hydroxycamptothecin, and irradiation in a cisplatin-resistant human small cell lung cancer cell line. *Clin Cancer Res* 8: 287–292, 2002 [PubMed: 11801571]

124. Konkel M. *Compounds specific for the human. alpha.1d adrenergic receptor and uses thereof*. US patent 6,706,716, March 16, 2004

125. Koren HS, Anderson SJ, Larrick JW. In vitro activation of a human macrophage-like cell line. *Nature* 279: 328–331, 1979 [PubMed: 450085]

126. Koyama H, Goodpasture C, Miller MM, Teplitz RL, Riggs AD. Establishment and characterization of a cell line from the American opossum (*Didelphys virginiana*). *In Vitro* 14: 239–246, 1978 [PubMed: 566717]

127. Kruszewski FH, Walker TL, DiPasquale LC. Evaluation of a human corneal epithelial cell line as an in vitro model for assessing ocular irritation. *Fundam Appl Toxicol* 36: 130–140, 1997 [PubMed: 9143482]

128. Kuroda R, Usas A, Kubo S, Corsi K, Peng H, Rose T, Cummins J, Fu FH, Huard J. Cartilage repair using bone morphogenetic protein 4 and muscle-derived stem cells. *Arthritis Rheum* 54: 433–442, 2006 [PubMed: 16447218]

129. Lahn BT, Page DC. Functional coherence of the human Y chromosome. *Science* 278: 675–680, 1997 [PubMed: 9381176]

130. Landry JJ, Pyl PT, Rausch T, Zichner T, Tekkedil MM, Stutz AM, Jauch A, Aiyar RS, Pau G, Delhomme N, Gagneur J, Korbel JO, Huber W, Steinmetz LM. The genomic and transcriptomic landscape of a HeLa Cell line. *G3 (Bethesda)* 3: 1213–1224, 2013 [PMCID: PMC3737162] [PubMed: 23550136]

131. Lange B, Valtieri M, Santoli D, Caracciolo D, Mavilio F, Gemperlein I, Griffin C, Emanuel B, Finan J, Nowell P, Rovera G. Growth factor requirements of childhood acute leukemia: establishment of GM-CSF-dependent cell lines. *Blood* 70: 192–199, 1987 [PubMed: 3496132]

132. Lau EC, Mohandas TK, Shapiro LJ, Slavkin HC, Snead ML. Human and mouse amelogenin gene loci are on the sex chromosomes. *Genomics* 4: 162–168, 1989 [PubMed: 2737677]

133. Lavenex P, Banta Lavenex P. Building hippocampal circuits to learn and remember: insights into the development of human memory. *Behav Brain Res* 254: 8–21, 2013 [PubMed: 23428745]

134. Lees AJ, Hardy J, Revesz T. Parkinson's disease. *Lancet* 373: 2055–2066, 2009 [PubMed: 19524782]

135. Leggio GM, Salomone S, Bucolo C, Platania C, Micale V, Caraci F, Drago F. Dopamine D_3 receptor as a new pharmacological target for the treatment of depression. *Eur J Pharmacol*. 10.1016/j.ejphar.2013.07.022 [Epub ahead of print] [PubMed: 23872400] [CrossRef: 10.1016/j.ejphar.2013.07.022]

136. Lelbovitz A, Wright WC, Pathak S, Siciliano MJ, Daniels WP. Detection and analysis of a glucose 6-phosphate dehydrogenase phenotype B cell line contamination. *J Natl Cancer Inst* 63: 635–645, 1979 [PubMed: 288927]

137. Lieber M, Smith B, Szakal A, Nelson-Rees W, Todaro G. A continuous tumor-cell line from a human lung carcinoma with properties of type II alveolar epithelial cells. *Int J Cancer* 17: 62–70, 1976 [PubMed: 175022]

138. Lightbody J. Establishment of differentiated clonal strains of glial brain cells in culture (Abstract). *Fed Proc* 27: 720, 1968

139. Littlewood-Evans AJ, Bilbe G, Bowler WB, Farley D, Wlodarski B, Kokubo T, Inaoka T, Sloane J, Evans DB, Gallagher JA. The osteoclast-associated protease cathepsin K is expressed in human breast carcinoma. *Cancer Res* 57: 5386–5390, 1997 [PubMed: 9393764]

140. Lockshin MD. Invited review: sex ratio and rheumatic disease. *J Appl Physiol* 91: 2366–2373, 2001 [PubMed: 11641382]

141. Lowy DR, Rands E, Scolnick EM. Helper-independent transformation by unintegrated Harvey sarcoma virus DNA. *J Virol* 26: 291–298, 1978 [PMCID: PMC354067] [PubMed: 26810]

142. Lozzio CB, Lozzio BB. Human chronic myelogenous leukemia cell-line with positive Philadelphia chromosome. *Blood* 45: 321–334, 1975 [PubMed: 163658]

143. Lyon MF. Gene action in the X-chromosome of the mouse (*Mus musculus* L.). *Nature* 190: 372–373, 1961 [PubMed: 13764598]

144. Lyons KE, Hubble JP, Troster AI, Pahwa R, Koller WC. Gender differences in Parkinson's disease. *Clin Neuropharmacol* 21: 118–121, 1998 [PubMed: 9579298]

145. MacLusky NJ, Naftolin F. Sexual differentiation of the central nervous system. *Science* 211: 1294–1302, 1981 [PubMed: 6163211]

146. Macpherson I, Stoker M. Polyoma transformation of hamster cell clones–an investigation of genetic factors affecting cell competence. *Virology* 16: 147–151, 1962 [PubMed: 14468055]

147. Mannucci A, Sullivan KM, Ivanov PL, Gill P. Forensic application of a rapid and quantitative DNA sex test by amplification of the X-Y homologous gene amelogenin. *Int J Legal Med* 106: 190–193, 1994 [PubMed: 8038111]

148. Mapherson I. Characteristics of a hamster cell clone transformed by polyoma virus. *J Natl Cancer Inst* 30: 795–815, 1963

149. Martin TJ, Ingleton PM, Underwood JC, Michelangeli VP, Hunt NH, Melick RA. Parathyroid hormone-responsive adenylate cyclase in induced transplantable osteogenic rat sarcoma. *Nature* 260: 436–438, 1976 [PubMed: 1062678]

150. Masters JR. HeLa cells 50 years on: the good, the bad and the ugly. *Nat Rev Cancer* 2: 315–319, 2002 [PubMed: 12001993]

151. Matsuki Y, Nakashima M, Amizuka N, Warshawsky H, Goltzman D, Yamada KM, Yamada Y. A compilation of partial sequences of randomly selected cDNA clones from the rat incisor. *J Dent Res* 74: 307–312, 1995 [PubMed: 7876422]

152. Matsunaga H, Handa JT, Aotaki-Keen A, Sherwood SW, West MD, Hjelmeland LM. Beta-galactosidase histochemistry and telomere loss in senescent retinal pigment epithelial cells. *Invest Ophthalmol Vis Sci* 40: 197–202, 1999 [PubMed: 9888444]

153. McAllister RM, Gardner MB, Greene AE, Bradt C, Nichols WW, Landing BH. Cultivation in vitro of cells derived from a human osteosarcoma. *Cancer* 27: 397–402, 1971 [PubMed: 5100401]

154. McDonough PG. The Y-chromosome and reproductive disorders. *Reprod Fertil Dev* 10: 1–16, 1998 [PubMed: 9727588]

155. McIntosh JC, Schoumacher RA, Tiller RE. Pancreatic adenocarcinoma in a patient with cystic fibrosis. *Am J Med* 85: 592, 1988 [PubMed: 3177424]

156. McKusick-Nathans Institute of Genetic Medicine *Online Mendelian Inheritance in Man*. Baltimore, MD: Johns Hopkins University; http://omim.org/

157. Meszaros LB, Usas A, Cooper GM, Huard J. Effect of host sex and sex hormones on muscle-derived stem cell-mediated bone formation and defect healing. *Tissue Eng Part A* 18: 1751–1759, 2012 [PMCID: PMC3432905] [PubMed: 22712541]

158. Metzgar RS, Gaillard MT, Levine SJ, Tuck FL, Bossen EH, Borowitz MJ. Antigens of human pancreatic adenocarcinoma cells defined by murine monoclonal antibodies. *Cancer Res* 42: 601–608, 1982 [PubMed: 7034925]

159. Meyer-Lindenberg A, Miletich RS, Kohn PD, Esposito G, Carson RE, Quarantelli M, Weinberger DR, Berman KF. Reduced prefrontal activity predicts exaggerated striatal dopaminergic function in schizophrenia. *Nat Neurosci* 5: 267–271, 2002 [PubMed: 11865311]

160. Michaelides M, Hardcastle AJ, Hunt DM, Moore AT. Progressive cone and cone-rod dystrophies: phenotypes and underlying molecular genetic basis. *Surv Ophthalmol* 51: 232–258, 2006 [PubMed: 16644365]

161. Miller VM. In pursuit of scientific excellence: sex matters. *Adv Physiol Educ* 36: 83–84, 2012 [PubMed: 22665420]

162. Milo GE, Shuler CF, Stoner G, Chen JC. Conversion of premalignant human cells to tumorigenic cells by methylmethane sulfonate and methylnitronitrosoguanidine. *Cell Biol Toxicol* 8: 193–205, 1992 [PubMed: 1337307]

163. Minesita T, Yamaguchi K. An androgen-dependent mouse mammary tumor. *Cancer Res* 25: 1168–1175, 1965 [PubMed: 5843266]

164. Minesita T, Yamaguchi K. An androgen-dependent tumor derived from a hormone- independent spontaneous tumor of a female mouse. *Steroids* 4: 815–830, 1964

165. Mizell M. *Biology of Amphibian Tumors*. New York: Springer-Verlag, 1969

166. Mollerup S, Ryberg D, Hewer A, Phillips DH, Haugen A. Sex differences in lung CYP1A1 expression and DNA adduct levels among lung cancer patients. *Cancer Res* 59: 3317–3320, 1999 [PubMed: 10416585]

167. Muller HJ. Further studies on the nature and causes of gene mutation. *Proc 6th Int Congr Genet Ithaca, New York*, 1932, p. 213–255

168. Muntoni F, Torelli S, Ferlini A. Dystrophin and mutations: one gene, several proteins, multiple phenotypes. *Lancet Neurol* 2: 731–740, 2003 [PubMed: 14636778]

169. Murakami H, Masui H. Hormonal control of human colon carcinoma cell growth in serum-free medium. *Proc Natl Acad Sci USA* 77: 3464–3468, 1980 [PMCID: PMC349637] [PubMed: 6932031]

170. Naito Y, Kino I, Horiuchi K, Fujimoto D. Promotion of collagen production by human fibroblasts with gastric cancer cells in vitro. *Virchows Arch B Cell Pathol Incl Mol Pathol* 46: 145–154, 1984 [PubMed: 6147924]

171. Nakabayashi H, Taketa K, Miyano K, Yamane T, Sato J. Growth of human hepatoma cells lines with differentiated functions in chemically defined medium. *Cancer Res* 42: 3858–3863, 1982 [PubMed: 6286115]

172. Nakahori Y, Takenaka O, Nakagome Y. A human X-Y homologous region encodes “amelogenin”. *Genomics* 9: 264–269, 1991 [PubMed: 2004775]

173. Nardone RM. Eradication of cross-contaminated cell lines: a call for action. *Cell Biol Toxicol* 23: 367–372, 2007 [PubMed: 17522957]

174. Nathans J, Piantanida TP, Eddy RL, Shows TB, Hogness DS. Molecular genetics of inherited variation in human color vision. *Science* 232: 203–210, 1986 [PubMed: 3485310]

175. Nayak SK, O'Toole C, Price ZH. A cell line from an anaplastic transitional cell carcinoma of human urinary bladder. *Br J Cancer* 35: 142–151, 1977 [PMCID: PMC2025316] [PubMed: 836756]

176. Nichols WW, Murphy DG, Cristofalo VJ, Toji LH, Greene AE, Dwight SA. Characterization of a new human diploid cell strain, IMR-90. *Science* 196: 60–63, 1977 [PubMed: 841339]

177. Nieuwenhoven L, Klinge I. Scientific excellence in applying sex- and gender-sensitive methods in biomedical and health research. *J Womens Health (Larchmt)* 19: 313–321, 2010 [PubMed: 20136550]

178. Norris JS, Cornett LE, Hardin JW, Kohler PO, MacLeod SL, Srivastava A, Syms AJ, Smith RG. Autocrine regulation of growth: II. Glucocorticoids inhibit transcription of c-$_{sis}$ oncogene-specific RNA transcripts. *Biochem Biophys Res Commun* 122: 124–128, 1984 [PubMed: 6743325]

179. O'Reilly MA, Gazdar AF, Morris RE, Whitsett JA. Differential effects of glucocorticoid on expression of surfactant proteins in a human lung adenocarcinoma cell line. *Biochim Biophys Acta* 970: 194–204, 1988 [PubMed: 3382698]

180. Ohno S. *Sex Chromosomes and Sex-Linked Genes*. New York: Springer-Verlag, 1967

181. Osgood MP. X-chromosome inactivation: the case of the calico cat. *Am J Pharm Educ* 58: 204–205, 1994

182. Ostrer H. Sex-based differences in gene transmission and gene expression. *Lupus* 8: 365–369, 1999 [PubMed: 10455514]

183. Owen NE. Effect of TPA on ion fluxes and DNA synthesis in vascular smooth muscle cells. *J Cell Biol* 101: 454–459, 1985 [PMCID: PMC2113682] [PubMed: 2410432]

184. Packard MG, Knowlton BJ. Learning and memory functions of the Basal Ganglia. *Annu Rev Neurosci* 25: 563–593, 2002 [PubMed: 12052921]

185. Page DC, Disteche CM, Simpson EM, de la Chapelle A, Andersson M, Alitalo T, Brown LG, Green P, Akots G. Chromosomal localization of ZFX–a human gene that escapes X inactivation–and its murine homologs. *Genomics* 7: 37–46, 1990 [PubMed: 1970799]

186. Paine MF, Ludington SS, Chen ML, Stewart PW, Huang SM, Watkins PB. Do men and women differ in proximal small intestinal CYP3A or P-glycoprotein expression? *Drug Metab Dispos* 33: 426–433, 2005 [PubMed: 15608139]

187. Papsidero LD, Kuriyama M, Wang MC, Horoszewicz J, Leong SS, Valenzuela L, Murphy GP, Chu TM. Prostate antigen: a marker for human prostate epithelial cells. *J Natl Cancer Inst* 66: 37–42, 1981 [PubMed: 6935463]

188. Park JG, Frucht H, LaRocca RV, Bliss DP, Jr, Kurita Y, Chen TR, Henslee JG, Trepel JB, Jensen RT, Johnson BE, Bang YJ, Kim JP, Gazdar AF. Characteristics of cell lines established from human gastric carcinoma. *Cancer Res* 50: 2773–2780, 1990 [PubMed: 2158397]

189. Park JG, Lee JH, Kang MS, Park KJ, Jeon YM, Lee HJ, Kwon HS, Park HS, Yeo KS, Lee KU, Kim ST, Chung JK, Hwang YJ, Lee HS, Kim CY, Lee YI, Chen TR, Hay RJ, Song SY, Kim WH, Kim YI. Characterization of cell lines established from human hepatocellular carcinoma. *Int J Cancer* 62: 276–282, 1995 [PubMed: 7543080]

190. Park SJ, Jeong SY, Kim HJ. Y chromosome loss and other genomic alterations in hepatocellular carcinoma cell lines analyzed by CGH and CGH array. *Cancer Genet Cytogenet* 166: 56–64, 2006 [PubMed: 16616112]

191. Partridge NC, Frampton RJ, Eisman JA, Michelangeli VP, Elms E, Bradley TR, Martin TJ. Receptors for 1,25(OH)2-vitamin D3 enriched in cloned osteoblast-like rat osteogenic sarcoma cells. *FEBS Lett* 115: 139–142, 1980 [PubMed: 6248375]

192. Pathak S, Siciliano MJ, Cailleau R, Wiseman CL, Hsu TC. A human breast adenocarcinoma with chromosome and isoenzyme markers similar to those of the HeLa line. *J Natl Cancer Inst* 62: 263–271, 1979 [PubMed: 283262]

193. Pattillo RA, Gey GO. The establishment of a cell line of human hormone-synthesizing trophoblastic cells in vitro. *Cancer Res* 28: 1231–1236, 1968 [PubMed: 4299001]

194. Pattillo RA, Gey GO, Delfs E, Mattingly RF. Human hormone production in vitro. *Science* 159: 1467–1469, 1968 [PubMed: 5753554]

195. Pattillo RA, Hussa RO, Story MT, Ruckert AC, Shalaby MR, Mattingly RF. Tumor antigen and human chorionic gonadotropin in CaSki cells: a new epidermoid cervical cancer cell line. *Science* 196: 1456–1458, 1977 [PubMed: 867042]

196. Payer B, Lee JT. X chromosome dosage compensation: how mammals keep the balance. *Annu Rev Genet* 42: 733–772, 2008 [PubMed: 18729722]

197. Penaloza C, Estevez B, Orlanski S, Sikorska M, Walker R, Smith C, Smith B, Lockshin RA, Zakeri Z. Sex of the cell dictates its response: differential gene expression and sensitivity to cell death inducing stress in male and female cells. *FASEB J* 23: 1869–1879, 2009 [PMCID: PMC2698656] [PubMed: 19190082]

198. Perantoni A, Berman JJ. Properties of Wilms' tumor line (TuWi) and pig kidney line (LLC-PK1) typical of normal kidney tubular epithelium. *In Vitro* 15: 446–454, 1979 [PubMed: 225262]

199. Pergament E, Fiddler M, Cho N, Johnson D, Holmgren WJ. Sexual differentiation and preimplantation cell growth. *Hum Reprod* 9: 1730–1732, 1994 [PubMed: 7836527]

200. Peterson MD, Mooseker MS. Characterization of the enterocyte-like brush border cytoskeleton of the C2BBe clones of the human intestinal cell line, Caco-2. *J Cell Sci* 102: 581–600, 1992 [PubMed: 1506435]

201. Peterson WD, Jr, Stulberg CS, Swanborg NK, Robinson AR. Glucose-6-phosphate dehydrogenase isoenzymes in human cell cultures determined by sucrose-agar gel and cellulose acetate zymograms. *Proc Soc Exp Biol Med* 128: 772–776, 1968 [PubMed: 5668122]

202. Pilgrim C, Reisert I. Differences between male and female brains–developmental mechanisms and implications. *Horm Metab Res* 24: 353–359, 1992 [PubMed: 1526620]

203. Pitot HC, Peraino C, Morse PA, Jr, Potter VR. Hepatomas in tissue culture compared with adapting liver in vivo. *Natl Cancer Inst Monogr* 13: 229–245, 1964 [PubMed: 14143233]

204. Pollack MS, Heagney SD, Livingston PO, Fogh J. HLA-A, B, C and DR alloantigen expression on forty-six cultured human tumor cell lines. *J Natl Cancer Inst* 66: 1003–1012, 1981 [PubMed: 7017212]

205. Ponten J, Macintyre EH. Long term culture of normal and neoplastic human glia. *Acta Pathol Microbiol Scand* 74: 465–486, 1968 [PubMed: 4313504]

206. Puck JM, Willard HF. X inactivation in females with X-linked disease. *N Engl J Med* 338: 325–328, 1998 [PubMed: 9445416]

207. Puck TT, Cieciura SJ, Robinson A. Genetics of somatic mammalian cells. III. Long-term cultivation of euploid cells from human and animal subjects. *J Exp Med* 108: 945–956, 1958 [PMCID: PMC2136918] [PubMed: 13598821]

208. Quaroni A, Isselbacher KJ, Ruoslahti E. Fibronectin synthesis by epithelial crypt cells of rat small intestine. *Proc Natl Acad Sci USA* 75: 5548–5552, 1978 [PMCID: PMC393003] [PubMed: 103096]

209. Quaroni A, Wands J, Trelstad RL, Isselbacher KJ. Epithelioid cell cultures from rat small intestine. Characterization by morphologic and immunologic criteria. *J Cell Biol* 80: 248–265, 1979 [PMCID: PMC2110349] [PubMed: 88453]

210. Quintana-Murci L, Fellous M. The human Y chromosome: the biological role of a "functional wasteland". *J Biomed Biotechnol* 1: 18–24, 2001 [PMCID: PMC79676] [PubMed: 12488622]

211. Raccor BS, Kaspera R. Extra-hepatic isozymes from the CYP1 and CYP2 families as potential chemotherapeutic targets. *Curr Top Med Chem* 13: 1441–1453, 2013 [PubMed: 23688134]

212. Ralph P. Retention of lymphocyte characteristics by myelomas and θ^+-lymphomas: sensitivity to cortisol and phytohemagglutinin. *J Immunol* 110: 1470–1475, 1973 [PubMed: 4541304]

213. Ralph P, Nakoinz I. Inhibitory effects of lectins and lymphocyte mitogens on murine lymphomas and myelomas. *J Natl Cancer Inst* 51: 883–890, 1973 [PubMed: 4542714]

214. Rapp F, Hsu TC. Viruses and mammalian chromosomes. IV. Replication of herpes simplex virus in diploid Chinese hamster cells. *Virology* 25: 401–411, 1965 [PubMed: 14328610]

Daubert Response App. 0689

215. Rappold GA. The pseudoautosomal regions of the human sex chromosomes. *Hum Genet* 92: 315–324, 1993 [PubMed: 8225310]

216. Rauchman MI, Nigam SK, Delpire E, Gullans SR. An osmotically tolerant inner medullary collecting duct cell line from an SV40 transgenic mouse. *Am J Physiol Renal Fluid Electrolyte Physiol* 265: F416–F424, 1993 [PubMed: 8214101]

217. Reisert I, Engele J, Pilgrim C. Early sexual differentiation of diencephalic dopaminergic neurons of the rat in vitro. *Cell Tissue Res* 255: 411–417, 1989 [PubMed: 2924342]

218. Reisert I, Pilgrim C. Sexual differentiation of monoaminergic neurons–genetic or epigenetic? *Trends Neurosci* 14: 468–473, 1991 [PubMed: 1722367]

219. Rheinwald JG, Beckett MA. Tumorigenic keratinocyte lines requiring anchorage and fibroblast support cultured from human squamous cell carcinomas. *Cancer Res* 41: 1657–1663, 1981 [PubMed: 7214336]

220. Robinson TE, Berridge KC. The neural basis of drug craving: an incentive-sensitization theory of addiction. *Brain Res Brain Res Rev* 18: 247–291, 1993 [PubMed: 8401595]

221. Rodan SB, Imai Y, Thiede MA, Wesolowski G, Thompson D, Bar-Shavit Z, Shull S, Mann K, Rodan GA. Characterization of a human osteosarcoma cell line (Saos-2) with osteoblastic properties. *Cancer Res* 47: 4961–4966, 1987 [PubMed: 3040234]

222. Rokaw MD, Benos DJ, Palevsky PM, Cunningham SA, West ME, Johnson JP. Regulation of a sodium channel-associated G-protein by aldosterone. *J Biol Chem* 271: 4491–4496, 1996 [PubMed: 8626803]

223. Rosenfeld M, Davis R, FitzSimmons S, Pepe M, Ramsey B. Gender gap in cystic fibrosis mortality. *Am J Epidemiol* 145: 794–803, 1997 [PubMed: 9143209]

224. Ross MT, Grafham DV, Coffey AJ, Scherer S, McLay K, Muzny D, Platzer M, Howell GR, Burrows C, Bird CP, Frankish A, Lovell FL, Howe KL, Ashurst JL, Fulton RS, Sudbrak R, Wen G, Jones MC, Hurles ME, Andrews TD, Scott CE, Searle S, Ramser J, Whittaker A, Deadman R, Carter NP, Hunt SE, Chen R, Cree A, Gunaratne P, Havlak P, Hodgson A, Metzker ML, Richards S, Scott G, Steffen D, Sodergren E, Wheeler DA, Worley KC, Ainscough R, Ambrose KD, Ansari-Lari MA, Aradhya S, Ashwell RI, Babbage AK, Bagguley CL, Ballabio A, Banerjee R, Barker GE, Barlow KF, Barrett IP, Bates KN, Beare DM, Beasley H, Beasley O, Beck A, Bethel G, Blechschmidt K, Brady N, Bray-Allen S, Bridgeman AM, Brown AJ, Brown MJ, Bonnin D, Bruford EA, Buhay C, Burch P, Burford D, Burgess J, Burrill W, Burton J, Bye JM, Carder C, Carrel L, Chako J, Chapman JC, Chavez D, Chen E, Chen G, Chen Y, Chen Z, Chinault C, Ciccodicola A, Clark SY, Clarke G, Clee CM, Clegg S, Clerc-Blankenburg K, Clifford K, Cobley V, Cole CG, Conquer JS, Corby N, Connor RE, David R, Davies J, Davis C, Davis J, Delgado O, Deshazo D, Dhami P, Ding Y, Dinh H, Dodsworth S, Draper H, Dugan-Rocha S, Dunham A, Dunn M, Durbin KJ, Dutta I, Eades T, Ellwood M, Emery-Cohen A, Errington H, Evans KL, Faulkner L, Francis F, Frankland J, Fraser AE, Galgoczy P, Gilbert J, Gill R, Glöckner G, Gregory SG, Gribble S, Griffiths C, Grocock R, Gu Y, Gwilliam R, Hamilton C, Hart EA, Hawes A, Heath PD, Heitmann K, Hennig S, Hernandez J, Hinzmann B, Ho S, Hoffs M, Howden PJ, Huckle EJ, Hume J, Hunt PJ, Hunt AR, Isherwood J, Jacob L, Johnson D, Jones S, de Jong PJ, Joseph SS, Keenan S, Kelly S, Kershaw JK, Khan Z, Kioschis P, Klages S, Knights AJ, Kosiura A, Kovar-Smith C, Laird GK, Langford C, Lawlor S, Leversha M, Lewis L, Liu W, Lloyd C, Lloyd DM, Loulseged H, Loveland JE, Lovell JD, Lozado R, Lu J, Lyne R, Ma J, Maheshwari M, Matthews LH, McDowall J, McLaren S, McMurray A, Meidl P, Meitinger T, Milne S, Miner G, Mistry SL, Morgan M, Morris S, Müller I, Mullikin JC, Nguyen N, Nordsiek G, Nyakatura G, O'Dell CN, Okwuonu G, Palmer S, Pandian R, Parker D, Parrish J, Pasternak S, Patel D, Pearce AV, Pearson DM, Pelan SE, Perez L, Porter KM, Ramsey Y, Reichwald K, Rhodes S, Ridler KA, Schlessinger D, Schueler MG, Sehra HK, Shaw-Smith C, Shen H, Sheridan EM, Shownkeen R, Skuce CD, Smith ML, Sotheran EC, Steingruber HE, Steward CA, Storey R, Swann RM, Swarbreck D, Tabor PE, Taudien S, Taylor T, Teague B, Thomas K, Thorpe A, Timms K, Tracey A, Trevanion S, Tromans AC, d'Urso M, Verduzco D, Villasana D, Waldron L, Wall M, Wang Q, Warren J, Warry GL, Wei X, West A, Whitehead SL, Whiteley MN, Wilkinson JE, Willey DL, Williams G, Williams L, Williamson A, Williamson H, Wilming L, Woodmansey RL, Wray PW, Yen J, Zhang J, Zhou J, Zoghbi H, Zorilla S, Buck D, Reinhardt R, Poustka A, Rosenthal A, Lehrach H, Meindl A, Minx PJ, Hillier LW, Willard HF, Wilson RK, Waterston RH, Rice CM, Vaudin M, Coulson A, Nelson DL, Weinstock G, Sulston JE, Durbin R, Hubbard T, Gibbs RA, Beck S, Rogers J, Bentley DR. The DNA sequence of the human X chromosome. *Nature* 434: 325–337, 2005 [PMCID: PMC2665286] [PubMed: 15772651]

225. Ross RA, Spengler BA, Biedler JL. Coordinate morphological and biochemical interconversion of human neuroblastoma cells. *J Natl Cancer Inst* 71: 741–747, 1983 [PubMed: 6137586]

226. Rousell J, Haddad EB, Mak JC, Barnes PJ. Transcriptional down-regulation of m2 muscarinic receptor gene expression in human embryonic lung (HEL 299) cells by protein kinase C. *J Biol Chem* 270: 7213–7218, 1995 [PubMed: 7706260]

227. Ryan MJ, Johnson G, Kirk J, Fuerstenberg SM, Zager RA, Torok-Storb B. HK-2: an immortalized proximal tubule epithelial cell line from normal adult human kidney. *Kidney Int* 45: 48–57, 1994 [PubMed: 8127021]

228. Ryberg D, Hewer A, Phillips DH, Haugen A. Different susceptibility to smoking-induced DNA damage among male and female lung cancer patients. *Cancer Res* 54: 5801–5803, 1994 [PubMed: 7954403]

229. Santoli D, Yang YC, Clark SC, Kreider BL, Caracciolo D, Rovera G. Synergistic and antagonistic effects of recombinant human interleukin (IL) 3, IL-1 alpha, granulocyte and macrophage colony-stimulating factors (G-CSF and M-CSF) on the growth of GM-CSF-dependent leukemic cell lines. *J Immunol* 139: 3348–3354, 1987 [PubMed: 3500218]

230. Sasaki S, Shimokawa H. The amelogenin gene. *Int J Dev Biol* 39: 127–133, 1995 [PubMed: 7626398]

231. Scherer WF, Syverton JT, Gey GO. Studies on the propagation in vitro of poliomyelitis viruses. IV. Viral multiplication in a stable strain of human malignant epithelial cells (strain HeLa) derived from an epidermoid carcinoma of the cervix. *J Exp Med* 97: 695–710, 1953 [PMCID: PMC2136303] [PubMed: 13052828]

232. Schmidt M, Deschner EE, Thaler HT, Clements L, Good RA. Gastrointestinal cancer studies in the human to nude mouse heterotransplant system. *Gastroenterology* 72: 829–837, 1977 [PubMed: 321290]

233. Schneider-Gadicke A, Beer-Romero P, Brown LG, Nussbaum R, Page DC. ZFX has a gene structure similar to ZFY, the putative human sex determinant, and escapes X inactivation. *Cell* 57: 1247–1258, 1989 [PubMed: 2500252]

234. Schneider U, Schwenk HU, Bornkamm G. Characterization of EBV-genome negative "null" and "T" cell lines derived from children with acute lymphoblastic leukemia and leukemic transformed non-Hodgkin lymphoma. *Int J Cancer* 19: 621–626, 1977 [PubMed: 68013]

235. Schoumacher RA, Ram J, Iannuzzi MC, Bradbury NA, Wallace RW, Hon CT, Kelly DR, Schmid SM, Gelder FB, Rado TA, Frizzell RA. A cystic fibrosis pancreatic adenocarcinoma cell line. *Proc Natl Acad Sci USA* 87: 4012–4016, 1990 [PMCID: PMC54034] [PubMed: 1692630]

236. Seifarth JE, McGowan CL, Milne KJ. Sex and life expectancy. *Gend Med* 9: 390–401, 2012 [PubMed: 23164528]

237. Sekido R, Lovell-Badge R. Sex determination and SRY: down to a wink and a nudge? *Trends Genet* 25: 19–29, 2009 [PubMed: 19027189]

238. Sekiguchi M, Sakakibara K, Fujii G. Establishment of cultured cell lines derived from a human gastric carcinoma. *Jpn J Exp Med* 48: 61–68, 1978 [PubMed: 209229]

239. Shay JW, Zou Y, Hiyama E, Wright WE. Telomerase and cancer. *Hum Mol Genet* 10: 677–685, 2001 [PubMed: 11257099]

240. Shinka T, Naroda T, Tamura T, Sasahara K, Nakahori Y. A rapid and simple method for sex identification by heteroduplex analysis, using denaturing high-performance liquid chromatography (DHPLC). *J Hum Genet* 46: 263–266, 2001 [PubMed: 11355016]

241. Shriver SP, Bourdeau HA, Gubish CT, Tirpak DL, Davis AL, Luketich JD, Siegfried JM. Sex-specific expression of gastrin-releasing peptide receptor: relationship to smoking history and risk of lung cancer. *J Natl Cancer Inst* 92: 24–33, 2000 [PubMed: 10620630]

242. Shulman LM, Bhat V. Gender disparities in Parkinson's disease. *Expert Rev Neurother* 6: 407–416, 2006 [PubMed: 16533144]

243. Siciliano MJ, Barker PE, Cailleau R. Mutually exclusive genetic signatures of human breast tumor cell lines with a common chromosomal marker. *Cancer Res* 39: 919–922, 1979 [PubMed: 427779]

244. Simunovic F, Yi M, Wang Y, Stephens R, Sonntag KC. Evidence for gender-specific transcriptional profiles of nigral dopamine neurons in Parkinson disease. *PLos One* 5: e8856, 2010 [PMCID: PMC2810324] [PubMed: 20111594]

245. Sinclair AH, Berta P, Palmer MS, Hawkins JR, Griffiths BL, Smith MJ, Foster JW, Frischauf AM, Lovell-Badge R, Goodfellow PN. A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif. *Nature* 346: 240–244, 1990 [PubMed: 1695712]

246. Skaletsky H, Kuroda-Kawaguchi T, Minx PJ, Cordum HS, Hillier L, Brown LG, Repping S, Pyntikova T, Ali J, Bieri T, Chinwalla A, Delehaunty A, Delehaunty K, Du H, Fewell G, Fulton L, Fulton R, Graves T, Hou SF, Latrielle P, Leonard S, Mardis E, Maupin R, McPherson J, Miner T, Nash W, Nguyen C, Ozersky P, Pepin K, Rock S, Rohlfing T, Scott K, Schultz B, Strong C, Tin-Wollam A, Yang SP, Waterston RH, Wilson RK, Rozen S, Page DC. The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. *Nature* 423: 825–837, 2003 [PubMed: 12815422]

247. Skloot R. *The Immortal Life of Henrietta Lacks*. New York: Broadway Paperbacks, 2011

248. Skubitz KM, Pessano S, Bottero L, Ferrero D, Rovera G, August JT. Human granulocyte surface molecules identified by murine monoclonal antibodies. *J Immunol* 131: 1882–1888, 1983 [PubMed: 6619543]

249. Smahi A, Courtois G, Vabres P, Yamaoka S, Heuertz S, Munnich A, Israel A, Heiss NS, Klauck SM, Kioschis P, Wiemann S, Poustka A, Esposito T, Bardaro T, Gianfrancesco F, Ciccodicola A, D'Urso M, Woffendin H, Jakins T, Donnai D, Stewart H, Kenwrick SJ, Aradhya S, Yamagata T, Levy M, Lewis RA, Nelson DL. Genomic rearrangement in NEMO impairs NF-kappaB activation and is a cause of incontinentia pigmenti. The International Incontinentia Pigmenti (IP) Consortium. *Nature* 405: 466–472, 2000 [PubMed: 10839543]

250. Soule HD, Vazguez J, Long A, Albert S, Brennan M. A human cell line from a pleural effusion derived from a breast carcinoma. *J Natl Cancer Inst* 51: 1409–1416, 1973 [PubMed: 4357757]

251. Stern C. The problem of complete Y-linkage in man. *Am J Hum Genet* 9: 147–166, 1957 [PMCID: PMC1931892] [PubMed: 13469791]

252. Stich HF, Vanhoosier GL, Trentin JJ. Viruses and mammalian chromosomes; chromosome aberrations by human adenovirus type 12. *Exp Cell Res* 34: 400–403, 1964 [PubMed: 14165598]

253. Stindl R. Tying it all together: telomeres, sexual size dimorphism and the gender gap in life expectancy. *Med Hypotheses* 62: 151–154, 2004 [PubMed: 14729022]

254. Stone KR, Mickey DD, Wunderli H, Mickey GH, Paulson DF. Isolation of a human prostate carcinoma cell line (DU 145). *Int J Cancer* 21: 274–281, 1978 [PubMed: 631930]

255. Stoner GD, Kaighn ME, Reddel RR, Resau JH, Bowman D, Naito Z, Matsukura N, You M, Galati AJ, Harris CC. Establishment and characterization of SV40 T-antigen immortalized human esophageal epithelial cells. *Cancer Res* 51: 365–371, 1991 [PubMed: 1703038]

Daubert Response App. 0691

256. Sullivan KM, Mannucci A, Kimpton CP, Gill P. A rapid and quantitative DNA sex test: fluorescence-based PCR analysis of X-Y homologous gene amelogenin. *Biotechniques* 15: 636–638, 640–631, 1993 [PubMed: 8251166]

257. Syms AJ, Norris JS, Smith RG. Proliferation of a highly androgen-sensitive ductus deferens cell line (DDT1MF-2) is regulated by glucocorticoids and modulated by growth on collagen. *In Vitro* 19: 929–936, 1983 [PubMed: 6662552]

258. Tashjian AH., Jr Clonal strains of hormone-producing pituitary cells. *Methods Enzymol* 58: 527–535, 1979 [PubMed: 218079]

259. Tashjian AH, Jr, Yasumura Y, Levine L, Sato GH, Parker ML. Establishment of clonal strains of rat pituitary tumor cells that secrete growth hormone. *Endocrinology* 82: 342–352, 1968 [PubMed: 4951281]

260. Taylor KE, Vallejo-Giraldo C, Schaible NS, Zakeri R, Miller VM. Reporting of sex as a variable in cardiovascular studies using cultured cells. *Biol Sex Differ* 2: 11, 2011 [PMCID: PMC3224776] [PubMed: 22060014]

261. Thummel KE. Gut instincts: CYP3A4 and intestinal drug metabolism. *J Clin Invest* 117: 3173–3176, 2007 [PMCID: PMC2045626] [PubMed: 17975661]

262. Tom BH, Rutzky LP, Jakstys MM, Oyasu R, Kaye CI, Kahan BD. Human colonic adenocarcinoma cells. I. Establishment and description of a new line. *In Vitro* 12: 180–191, 1976 [PubMed: 1262041]

263. Toulouse A, Sullivan AM. Progress in Parkinson's disease-where do we stand? *Prog Neurobiol* 85: 376–392, 2008 [PubMed: 18582530]

264. Tsuchiya S, Kobayashi Y, Goto Y, Okumura H, Nakae S, Konno T, Tada K. Induction of maturation in cultured human monocytic leukemia cells by a phorbol diester. *Cancer Res* 42: 1530–1536, 1982 [PubMed: 6949641]

265. Turner T. Development of the polio vaccine: a historical perspective of Tuskegee University's role in mass production and distribution of HeLa cells. *J Health Care Poor Underserved* 23: 5–10, 2012 [PMCID: PMC4458465] [PubMed: 23124495]

266. Vaccarino V, Parsons L, Every NR, Barron HV, Krumholz HM. Sex-based differences in early mortality after myocardial infarction. National Registry of Myocardial Infarction 2 Participants. *N Engl J Med* 341: 217–225, 1999 [PubMed: 10413733]

267. Van Goor F, Hadida S, Grootenhuis PD, Burton B, Cao D, Neuberger T, Turnbull A, Singh A, Joubran J, Hazlewood A, Zhou J, McCartney J, Arumugam V, Decker C, Yang J, Young C, Olson ER, Wine JJ, Frizzell RA, Ashlock M, Negulescu P. Rescue of CF airway epithelial cell function in vitro by a CFTR potentiator, VX-770. *Proc Natl Acad Sci USA* 106: 18825–18830, 2009 [PMCID: PMC2773991] [PubMed: 19846789]

268. Van Goor F, Hadida S, Grootenhuis PD, Burton B, Stack JH, Straley KS, Decker CJ, Miller M, McCartney J, Olson ER, Wine JJ, Frizzell RA, Ashlock M, Negulescu PA. Correction of the F508del-CFTR protein processing defect in vitro by the investigational drug VX-809. *Proc Natl Acad Sci USA* 108: 18843–18848, 2011 [PMCID: PMC3219147] [PubMed: 21976485]

269. Van Goor F, Straley KS, Cao D, Gonzalez J, Hadida S, Hazlewood A, Joubran J, Knapp T, Makings LR, Miller M, Neuberger T, Olson E, Panchenko V, Rader J, Singh A, Stack JH, Tung R, Grootenhuis PD, Negulescu P. Rescue of ΔF508-CFTR trafficking and gating in human cystic fibrosis airway primary cultures by small molecules. *Am J Physiol Lung Cell Mol Physiol* 290: L1117–L1130, 2006 [PubMed: 16443646]

270. Vina J, Borras C, Gambini J, Sastre J, Pallardo FV. Why females live longer than males: control of longevity by sex hormones. *Sci Aging Knowledge Environ* 2005: pe17, 2005 [PubMed: 15944465]

271. von Kleist S, Chany E, Burtin P, King M, Fogh J. Immunohistology of the antigenic pattern of a continuous cell line from a human colon tumor. *J Natl Cancer Inst* 55: 555–560, 1975 [PubMed: 1159834]

272. Wang B, Zhou SF. Synthetic and natural compounds that interact with human cytochrome P450 1A2 and implications in drug development. *Curr Med Chem* 16: 4066–4218, 2009 [PubMed: 19754423]

273. Wang D, Christensen K, Chawla K, Xiao G, Krebsbach PH, Franceschi RT. Isolation and characterization of MC3T3–E1 preosteoblast subclones with distinct in vitro and in vivo differentiation/mineralization potential. *J Bone Miner Res* 14: 893–903, 1999 [PubMed: 10352097]

274. Wang H, Schumacher AE, Levitz CE, Mokdad AH, Murray CJ. Left behind: widening disparities for males and females in US county life expectancy, 1985–2010. *Popul Health Metr* 11: 8, 2013 [PMCID: PMC3717281] [PubMed: 23842281]

275. Watanabe M, Zinn AR, Page DC, Nishimoto T. Functional equivalence of human X- and Y-encoded isoforms of ribosomal protein S4 consistent with a role in Turner syndrome. *Nat Genet* 4: 268–271, 1993 [PubMed: 8358435]

276. Weiss A, Wiskocil RL, Stobo JD. The role of T3 surface molecules in the activation of human T cells: a two-stimulus requirement for IL 2 production reflects events occurring at a pre-translational level. *J Immunol* 133: 123–128, 1984 [PubMed: 6327821]

277. Willard HF. The sex chromosomes and X-chromosome inactivation. In: *The Metabolic Basis of Inherited Disease*, edited by Scriver CR, Beaudet AL, Sly D, Valle D, Childs B, Vogelstein B. New York: McGraw-Hill, 2000

278. Wizemenn TM, Pardue ML. *Exploring the Biological Contributions to Human Health: Does Sex Matter?* Washington, DC: Natl Acad Press, 2001

279. Woolford JL., Jr The structure and biogenesis of yeast ribosomes. *Adv Genet* 29: 63–118, 1991 [PubMed: 1763709]

Daubert Response App. 0692

280. Wooten GF, Currie LJ, Bovbjerg VE, Lee JK, Patrie J. Are men at greater risk for Parkinson's disease than women? *J Neurol Neurosurg Psychiatry* 75: 637–639, 2004 [PMCID: PMC1739032] [PubMed: 15026515]

281. Wright V, Peng H, Usas A, Young B, Gearhart B, Cummins J, Huard J. BMP4-expressing muscle-derived stem cells differentiate into osteogenic lineage and improve bone healing in immunocompetent mice. *Mol Ther* 6: 169–178, 2002 [PubMed: 12161183]

282. Wu JC, Merlino G, Fausto N. Establishment and characterization of differentiated, nontransformed hepatocyte cell lines derived from mice transgenic for transforming growth factor alpha. *Proc Natl Acad Sci USA* 91: 674–678, 1994 [PMCID: PMC43011] [PubMed: 7904757]

283. Wutz A. Gene silencing in X-chromosome inactivation: advances in understanding facultative heterochromatin formation. *Nat Rev Genet* 12: 542–553, 2011 [PubMed: 21765457]

284. Xu KP, Yadav BR, King WA, Betteridge KJ. Sex-related differences in developmental rates of bovine embryos produced and cultured in vitro. *Mol Reprod Dev* 31: 249–252, 1992 [PubMed: 1571158]

285. Yaffe D, Saxel O. A myogenic cell line with altered serum requirements for differentiation. *Differentiation* 7: 159–166, 1977 [PubMed: 558123]

286. Yaffe D, Saxel O. Serial passaging and differentiation of myogenic cells isolated from dystrophic mouse muscle. *Nature* 270: 725–727, 1977 [PubMed: 563524]

287. Yamane H, Kiura K, Tabata M, Bessho A, Tsuchida T, Motoda K, Hiraki A, Ueoka H, Harada M. Small cell lung cancer can express CD34 antigen. *Anticancer Res* 17: 3627–3632, 1997 [PubMed: 9413215]

288. Yamazaki T, Yokoyama T, Akatsu H, Tukiyama T, Tokiwa T. Phenotypic characterization of a human synovial sarcoma cell line, SW982, and its response to dexamethasone. *In Vitro Cell Dev Biol Anim* 39: 337–339, 2003 [PubMed: 15038780]

289. Yang F, Babak T, Shendure J, Disteche CM. Global survey of escape from X inactivation by RNA-sequencing in mouse. *Genome Res* 20: 614–622, 2010 [PMCID: PMC2860163] [PubMed: 20363980]

290. Yang X, Schadt EE, Wang S, Wang H, Arnold AP, Ingram-Drake L, Drake TA, Lusis AJ. Tissue-specific expression and regulation of sexually dimorphic genes in mice. *Genome Res* 16: 995–1004, 2006 [PMCID: PMC1524872] [PubMed: 16825664]

291. Yee C, Krishnan-Hewlett I, Baker CC, Schlegel R, Howley PM. Presence and expression of human papillomavirus sequences in human cervical carcinoma cell lines. *Am J Pathol* 119: 361–366, 1985 [PMCID: PMC1888002] [PubMed: 2990217]

292. Yufu Y, Goto T, Choi I, Uike N, Kozuru M, Ohshima K, Taniguchi T, Motokura T, Yatabe Y, Nakamura S. A new multiple myeloma cell line, MEF-1, possesses cyclin D1 overexpression and the p53 mutation. *Cancer* 85: 1750–1757, 1999 [PubMed: 10223569]

293. Yunis AA, Arimura GK, Russin DJ. Human pancreatic carcinoma (MIA PaCa-2) in continuous culture: sensitivity to asparaginase. *Int J Cancer* 19: 128–135, 1977 [PubMed: 832918]

294. Zabner J, Karp P, Seiler M, Phillips SL, Mitchell CJ, Saavedra M, Welsh M, Klingelhutz AJ. Development of cystic fibrosis and noncystic fibrosis airway cell lines. *Am J Physiol Lung Cell Mol Physiol* 284: L844–L854, 2003 [PubMed: 12676769]

295. Zang EA, Wynder EL. Differences in lung cancer risk between men and women: examination of the evidence. *J Natl Cancer Inst* 88: 183–192, 1996 [PubMed: 8632492]

296. Zanger UM, Schwab M. Cytochrome P450 enzymes in drug metabolism: regulation of gene expression, enzyme activities, and impact of genetic variation. *Pharmacol Ther* 138: 103–141, 2013 [PubMed: 23333322]

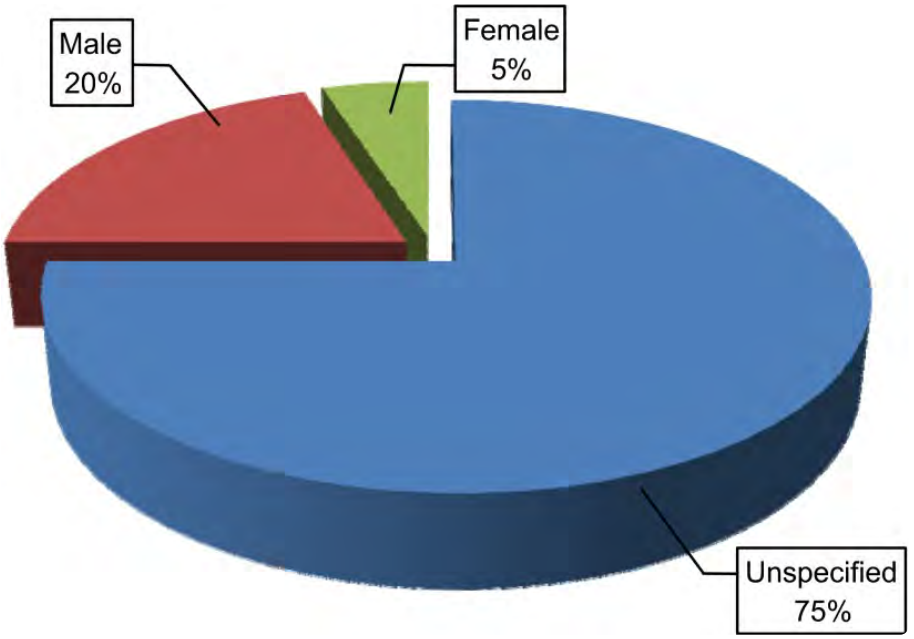
297. Zeitlin PL, Lu L, Rhim J, Cutting G, Stetten G, Kieffer KA, Craig R, Guggino WB. A cystic fibrosis bronchial epithelial cell line: immortalization by adeno-12-SV40 infection. *Am J Respir Cell Mol Biol* 4: 313–319, 1991 [PubMed: 1849726]

298. Zhang W, Bleibel WK, Roe CA, Cox NJ, Eileen Dolan M. Gender-specific differences in expression in human lymphoblastoid cell lines. *Pharmacogenet Genomics* 17: 447–450, 2007 [PMCID: PMC2716706] [PubMed: 17502836]

299. Zorgdrager A, De Keyser J. The premenstrual period and exacerbations in multiple sclerosis. *Eur Neurol* 48: 204–206, 2002 [PubMed: 12422068]

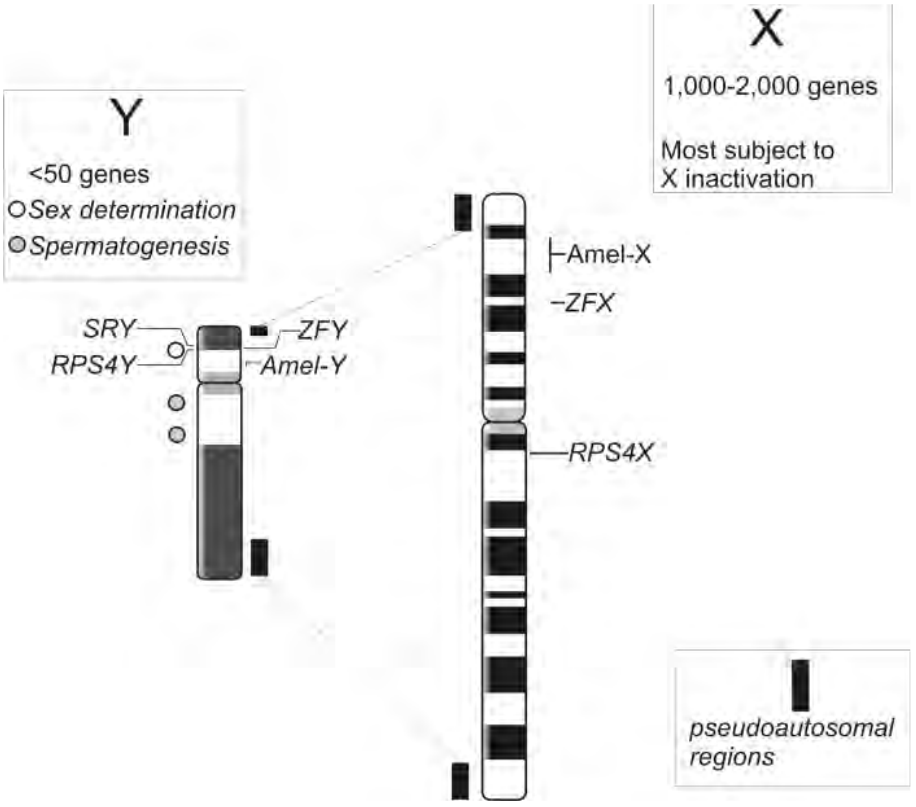
Figures and Tables

Fig. 1.



Distribution of studies by sex, published in *AJP-Cell Physiology* in 2013. Shown is the percentage of articles describing the sex of cells derived from male subjects, female subjects, or unreported (n = 100 articles randomly selected from *AJP-Cell Physiology* manuscripts published in 2013).

Fig. 2.



Comparison of size and gene organization for X and Y chromosomes. Approximate locations of chromosome-specific genes for zinc finger proteins (*ZFX* and *ZFY*) and ribosomal proteins (*RPS4X* and *RPS4Y*) are shown, as well as locations for chromosome-specific amelogenin (Amel) genes used for sex determination. See text for details.

Table 1.

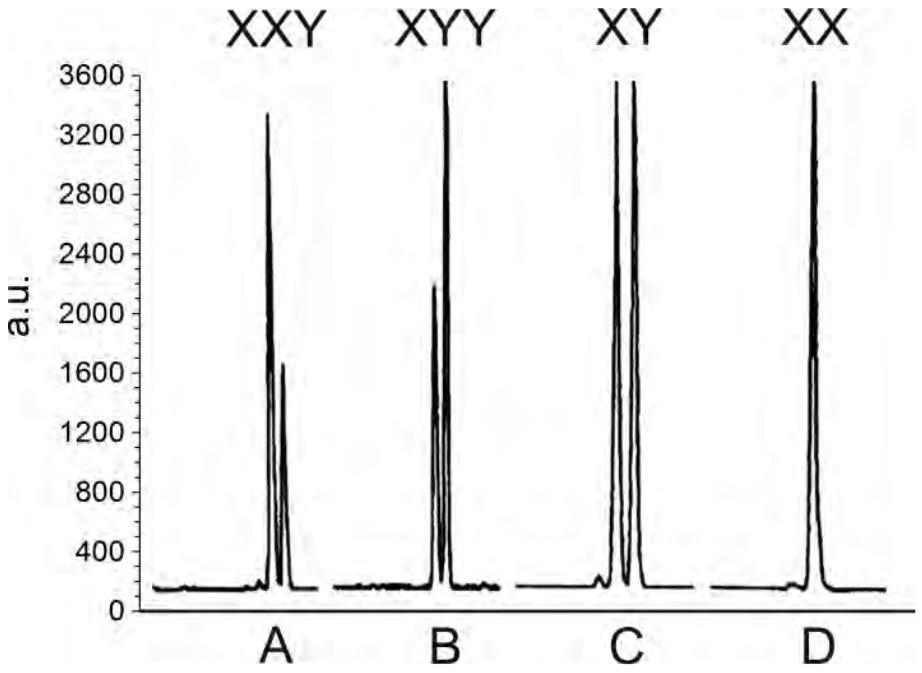
Table of the most commonly used cell lines appearing in AJP-Cell Physiology

Cell Line	Sex	Description	Species	Year	Origin	Reference
5637	Male	Urinary bladder epithelium	Human	1974	C	(68, 69)
3T3-L1	Male	Embryo fibroblast	Mouse	1962	N	(84, 85)
16HBE		Lung epithelial	Human		N	(46, 88)
A549	Male	Lung epithelial	Human	1972	C	(79, 137)
A6	Male	Kidney epithelial	*Xenopus*	1965	N	(165, 222)
A7r5		Aorta smooth muscle	Rat	1976	N	(15, 183)
AGS	Female	Stomach epithelial	Human	1979	C	(10, 11)
AML-12	Male	Liver epithelial	Mouse	1994	N	(61, 282)
AML-193	Female	Lymphoblast	Human	1987	C	(131, 229)
ARPE-19	Male	Retinal pigmented epithelial	Human	1986	N	(62, 101)
BeWo	Female	Placenta	Human	1966	N	(193, 194)
BHK		Kidney fibroblast	Syrian hamster	1961	N	(146, 148)
BT-549	Female	Breast epithelial (ductal)	Human	1978	C	(112, 139)
BW5147.3		T-lymphocyte		1973	C	(212, 213)
C2BBe1	Male	Colonic epithelial cell (a CaCo-2 subclone)	Human	1988	C	(13, 200)
C2C12		Myoblast	Mouse	1977	N	(285, 286)
C6		Glial cell	Rat	1968	N	(18, 138)
C127	Female	Mammary epithelial	Mouse			(141)
Ca SKI	Female	Cervical epithelial	Human	1977	C	(195, 291)
CaCo-2	Male*	Colonic epithelial	Human	1977	C	(14, 107)
Calu-3	Male*	Airway epithelial	Human	1975	C	(70, 94)
Capan-1	Male*	Pancreatic epithelial	Human			(70, 71)
CCL-39	Female	Fibroblast	Chinese hamster	1964	N	(214, 252)
CFPAC-1	Male	Pancreatic epithelial	Human	1990	C	(155, 235)
CHO	Female	Ovarian epithelial	Chinese hamster	1957	N	(110, 207)
COS-7		Kidney fibroblast	African green monkey	1964	N	(81)
CPAE	Female	Pulmonary endothelial	Cow	1979	N	(23)
CRL-2234	Male*	Hepatocyte	Human	1990	C	(189)
CV-1	Male	Kidney fibroblast	African green monkey	1964	N	(105, 117)
DDT1-MF-2	Male		Syrian hamster	1983	C	(178, 257)
DU 145	Male	Prostate epithelial	Human	1978	C	(187, 254)
ES-D3		Embryonic stem cell	Mouse	1985	N	(56, 57)
GH3	Female	Pituitary epithelial	Rat	1965	C	(8, 258)
GH4C1	Female	Pituitary epithelial	Rat	1968	C	(258, 259)
H441	Male	Lung epithelial	Human	1982	C	(77, 179)
H460	Male	Lung epithelial	Human	1982	C	(9, 30)
H4TG	Male	Liver epithelial	Rat	1964	C	(93, 203)
H9c2		Myocardial myoblast	Rat	1976	N	(28, 116)
H460	Male	Lung epithelial	Human	1982	C	(9, 30)

Where possible, references include the first descriptions of the cell lines. Cells were derived from "C," cancerous tissue or "N," noncancerous tissue (usually virally transformed).

*Cells derived from human "male" tissues that now express only amelogenin-X and no amelogenin-Y. NA, original deposition date could not be determined. This table is not intended to be a comprehensive data set, but rather to highlight the cell lines that are routinely used by authors in *AJP-Cell Physiology*. For a larger database, the reader is directed to such sites as American Tissue Type Collection (ATTC.org).

Fig. 3.



Comparison of electrophoretograms of sex test PCR products generated by an ABI Gene Scanner 362A. Primers for “male” and “female” amelogenin genes were employed. A, XXY cell line DNA with 1.8:1 X:Y peak area ratio; B, XYY male with 1:1.8 X:Y peak area ratio; C, normal male DNA with 1.01:1 X:Y peak area ratio; D, normal female (note absence of Y peak). a.u., arbitrary units. [From Sullivan et al. (256) with permission.]

The Secular Trend for Grip Strength in Canada and The United States

Article *in* Journal of Sports Sciences · February 2011

DOI: 10.1080/02640414.2010.547209 · Source: PubMed

CITATIONS
29

READS
289

1 author:

Irwin Silverman
Bowling Green State University
39 PUBLICATIONS **1,199** CITATIONS

SEE PROFILE

Journal of Sports Sciences
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/rjsp20

The secular trend for grip strength in Canada and the United States

Irwin W. Silverman [a]
[a] Department of Psychology , Bowling Green State University , Bowling Green, Ohio, USA
Published online: 28 Feb 2011.

To cite this article: Irwin W. Silverman (2011) The secular trend for grip strength in Canada and the United States, Journal of Sports Sciences, 29:6, 599-606, DOI: 10.1080/02640414.2010.547209

To link to this article: http://dx.doi.org/10.1080/02640414.2010.547209

Journal of Sports Sciences, March 15th 2011; 29(6): 599–606

Routledge
Taylor & Francis Group

The secular trend for grip strength in Canada and the United States

IRWIN W. SILVERMAN

Department of Psychology, Bowling Green State University, Bowling Green, Ohio, USA

(Accepted 8 December 2010)

Abstract
Worldwide, body weight has risen dramatically in recent decades, raising the question of whether there were concomitant changes in physical fitness. Past research with children and adolescents has shown that body weight and grip strength are positively correlated. Therefore, it was predicted that grip strength had increased on average in children and adolescents over the past four or five decades. To test this prediction, relevant data were extracted from 18 studies for males ($N = 5676$ in total) and 17 studies for females ($N = 5489$ in total). The studies were conducted in Canada and the United States from about 1966 on, with participants' ages ranging from 6 through 19 years. Weighted least squares regression analyses showed that grip strength was significantly predicted by age but not by country. Male grip strength decreased significantly over the period covered, but the change was very small, accounting for far less than 1% of the variance in male grip strength. Discussion focuses on potential explanations for why grip strength did not change over time as predicted.

Keywords: *Grip strength, secular trend, age differences, children, adolescents, males, females, Canada, United States*

Introduction

Recent decades have witnessed dramatic increases in the rates of overweight and obesity across much of the globe (World Health Organization, 2010). A frequently asked question is whether these changes had an effect on physical fitness. Literature reviews indicate that a single answer cannot be given to this question. Whereas aerobic fitness as measured by performance on tests of running in children and adolescents declined over time (Tomkinson, 2007b; Tomkinson, Léger, Olds, & Cazorla, 2003), anaerobic fitness as measured by performance on tests of speed and power in children and adolescents remained relatively stable over time (Tomkinson, 2007a). The present review provides evidence relative to the question of whether there has been a secular change in performance on another measure of anaerobic fitness – grip strength – in children and adolescents.

Secular changes in grip strength are of interest for several reasons. First, grip strength is correlated with a number of other measures of strength (Fleishman, 1963; Wang, Leger, & Dumas, 2005; Wind, Takken, Helders, & Engelbert, 2010). Second, grip strength is predictive of health according to a Canadian study (Payne, Gledhill, Katzmarzyk, Jamnik, & Ferguson, 2000a). Third, grip strength is predictive of mortality according to studies conducted in several countries (Gale, Martyn, Cooper, & Sayer, 2007; Ling et al., 2010; Metter, Talbot, Schrager, & Conwit, 2007; Newman et al., 2006; Saski, Kasaqi, Yamada, & Fujita, 2007; Takata et al., 2007). Thus, secular changes in grip strength may predict changes in overall strength, health, and mortality.

The prediction tested here is that grip strength in children and adolescents increased on average over the past four or five decades. This prediction was based on the well-established finding that grip strength is positively correlated with body weight in both sexes during childhood and adolescence (e.g. Bowman & Katz, 1984; Butler, 1997; Casajús, Leiva, Villarroya, Legaz, & Moreno, 2007; Häger-Ross & Rösblad, 2002; Milliken, Faigenbaum, Loud, & Westcotte, 2008; Montpetit & Montoye, 1967). Therefore, as body weight has increased in recent decades, grip strength should have also increased over this period.

Secular changes in grip strength have been examined in only one previous study (Tremblay et al., 2010). In this study, grip strength was compared in two large and representative Canadian samples, one tested in 1981 and the other in 2007 to 2009. Results show that at each of three ages (7–10, 11–14, and 15–19 years), significant decreases in grip strength occurred over time, with the declines ranging from

Correspondence: I. W. Silverman, Department of Psychology, Bowling Green State University, Bowling Green, OH 43403, USA.
E-mail: silpsych@comcast.net

ISSN 0264-0414 print/ISSN 1466-447X online © 2011 Taylor & Francis
DOI: 10.1080/02640414.2010.547209

0.185 to 0.370 kg per year in males and 0.185 to 0.222 kg per year in females. These findings are contrary to prediction, but as they involve only two samples and two time points, they may reflect a local perturbation rather than an overall trend.

As originally planned, the present review was to examine the secular trend for grip strength not only for children and adolescents but also for adults. Furthermore, the review was to include data from across the world. The literature search, however, uncovered sufficient data to provide a reasonably fine-grained picture of possible secular changes in grip strength for only a limited age range, ages 6 through 19 years, and for only two countries, Canada and the United States. Accordingly, the present data analysis is limited to this age range and these two countries. As there were no relevant Canadian studies published before 1969, the period covered by this study is from about 1966 onward. Although the year when the prevalence of overweight and obesity began to increase cannot be pinpointed, it is likely that the present review covers all or most of the period of time when body weight increased in both Canada and the United States (Shields, 2006).

Methods

Literature search

To locate articles for the present review, the following databases were searched: PubMed, Medline, CINAHL, PsycINFO, Physical Education Index, and SPORTDiscus. The search terms used were *grip strength*, *grip force*, and *neuropsychological test battery*. In addition, the reference lists of the retrieved articles were searched for relevant articles.

Selection criteria

As noted above, the present review is restricted to studies providing data on samples ages 6 through 19 years and to studies conducted in Canada and the United States from about 1966 onward. Four additional restrictions were placed on the studies reviewed herein. First, the samples had to be recruited from the general population and had to consist of participants presumed to be healthy and free of any injury or health problem that might affect grip strength. Second, the data had to be reported for males and females separately. Third, the age range within a sample could be no greater than 5 years. An exception to this was the study of Gallup and colleagues (Gallup, White, & Gallup, 2007). The age range for this study was 18–28 years, but the mean age was 19.11 years with a standard deviation of 1.68, which makes it likely that most of the participants in this study met the selection criterion for age range. Fourth, grip strength was measured in terms of force rather than pressure. (In fact, all retrieved studies met this criterion.)

Only partial use was made of data reported in a Canadian study by Tremblay and colleagues (Tremblay, Barnes, Copeland, & Esliger, 2005). In this study, there were three groups of children, one of which was composed of Old Order Mennonite children who lived a lifestyle described as ''representative of life in Canada three to four generations ago'' (p. 1188). The data for these Old Order Mennonite children were not included in the present analysis because it seems doubtful that the forces that have contributed to the recent increases in the prevalence of overweight/obesity were operative within the Old Order Mennonite community as described.

With two exceptions (Shephard, 1986; Tremblay et al., 2010), the samples used in the retrieved studies were convenience samples. The two exceptions were nationally representative Canadian samples. What effect the use of convenience samples had on the results cannot be ascertained. But as in the retrieved studies no attempt was made to screen out people except for health or physical limitations, there is no reason to believe that the data are biased for or against a particular outcome with respect to the relation between year of testing and grip strength.

With two exceptions, the participants were tested at only one age. In one study (Janz, Dawson, & Mahoney, 2000), the same participants were tested at five ages, and in the other (Trudeau, Shephard, Arsenault, & Laurencelle, 2003), the same participants were tested at three ages. Retention of the grip strength values for each of these ages would have been problematic because the data were analysed using multiple regression, which assumes that errors are independent. Therefore, for these studies, only the value obtained at the middlemost age was retained.

Data analysis

Typically, grip strength is measured for both hands. However, studies have differed with respect to how the resulting data have been reported. Specifically, grip strength has been reported for the right and left hands separately, for the dominant and non-dominant hands separately, for the preferred and non-preferred hands separately, and for the two hands summed. To make maximum use of the retrieved data, grip strength was analysed in terms of the summed values for the two hands. Hence, when grip strength was reported separately for the right/left hands, dominant/non-dominant hands or preferred/non-preferred hands, the values for the two hands were combined.

Two other cases of how grip strength is measured should be noted: in one, grip strength is tested for only one hand; in the other, grip strength is tested in both hands, but the best value regardless of hand is reported. No study had to be excluded from the data set because grip strength was so measured.

Several methodological factors could potentially affect the measurement of grip strength. These include the test instrument used, body position (standing or sitting), arm position (straight or flexed), grip position, number of practice and test trials and, given more than one test trial is used, whether the mean or maximum value is recorded. As this information was not reported in all studies, no consideration was given to these variables in the data analyses. However, it should be remarked that some of these variables may have little or no effect on grip measurement, at least according to the adult literature. For example, results have been mixed regarding the effect of body position on grip strength (Amosun, Moyo, & Matara, 1995; Balogun, Akomolafe, & Amusa, 1991; Bowman & Weaver, 2008; Swanson, Matev, & de Groot, 1970). Similarly, results have been mixed regarding whether grip strength differs depending on whether the maximum or average value is used when more than one test trial is given (Ertem et al., 2005; Haidar, Kumar, Bassi, & Deshmukh, 2004; Hamilton, Balnave, & Adams, 1994; Mathiowetz, Weber, Volland, & Kashman, 1984). No consideration was given to analysing the data according to test instrument because the test instruments used were produced by at least six different companies (three studies did not report the manufacturer of the instrument).

Thus, the only predictor variables used in the data analyses were age, country in which the study was performed (Canada or the United States), and year in which the data were collected. Year of testing was generally not reported; therefore, it was coded as the year in which the study was published minus 2 years.

The analyses were performed using SPSS® v.15.0 for Windows.

Results

The literature search yielded 18 studies for males and 17 studies for females, 11 of them for each sex having been conducted in Canada (see Table I). Year of testing (see above) ranged from about 1967 to 2009 for the studies conducted in Canada and from about 1966 to 2007 for the studies conducted in the United States. For males there were 63 means and for females there were 62 means, with the total number of males and females being 5676 and 5489, respectively.

For each sex separately, the data were subjected to a weighted least squares hierarchical multiple regression analysis with the weights being the sample sizes for each sex. Age was entered in step 1, age squared in step 2, country (Canada = 1, United States = 0) in step 3, year tested in step 4, and year tested squared in step 5. To avoid problems associated with multicollinearity, age and year tested were both centred. In addition to the prediction that grip strength increased on average over the years covered by the retrieved studies, it was expected, on the basis of past research (e.g. Mathiowetz, Wiemer, & Federman, 1986), that grip strength would increase over the age span covered here. No prediction was made regarding whether grip strength would differ between Canada and the United States.

Table II presents the results of the regression analyses for the two sexes separately. For males, the analysis yielded significant effects for age and age squared. The linear relation between age and grip strength was very strong, with age accounting for 96% of the variance in grip strength. Age squared, although significant, accounted for far less than 1% additional variance.. Inspection of Figure 1 shows grip strength increasing with age, with a slight positive acceleration around age 12. In addition, the analysis showed that year tested was significant, but it too accounted for far less than 1% of the variance in grip strength.

For females, only age and age squared were significant, with the former accounting for 87% of the variance in grip strength and the latter for an additional 6% of the variance. Inspection of Figure 2 shows grip strength increasing linearly with age until about age 17 when it begins to level off.

Given that the average age of the sample tested by Gallup et al. (2007) might have exceeded the upper age limit of 19 years, the regression analyses described above were re-run excluding Gallup and colleagues' data. These analyses yielded results very similar to those reported herein.

Discussion

The present results provide no support for the predicted secular improvement in grip strength. There was slight support for Tremblay and colleagues' (2005) finding for a decrease in grip strength over recent decades, but it was found only in males. However, the decline, which averaged 0.122 kg per year, was far less than the declines reported by Tremblay et al. A further consideration is that the decline in male grip strength may have been due to chance, as year tested accounted for far less than 1% of the variance in male grip strength.

As noted above, Tomkinson (2007a) found no evidence for a secular change in performance for two other tests of anaerobic fitness: speed and power. Tomkinson offered no explanation for why perfor-

Table I. Descriptive statistics for the studies of grip strength included in the data analyses.

Study	Year(s) tested	Country	Age (years)	Males		Females	
				n	*M* (kg)	*n*	*M* (kg)
Ager et al. (1984)	1982	USA	6.0	31	16.06	34	13.72
			7.0	33	20.08	35	15.62
			8.0	23	24.76	28	20.89
			9.0	29	27.55	29	24.44
			10.0	24	32.83	37	29.42
			11.0	30	41.01	35	33.34
			12.0	25	47.63	24	43.23
Backous et al. (1990)	1988	USA	12.0	98	47.80	–	–
Butterfield et al. (2009)	2007	USA	5.5	11	24.39	18	19.46
			7.0	15	28.33	9	24.16
			8.0	19	31.46	13	26.38
			9.0	11	33.88	15	34.13
			10.0	34	40.48	22	38.37
			11.0	37	43.63	37	41.55
			12.0	38	54.23	32	48.04
			13.0	38	64.59	39	54.85
			14.0	49	75.37	44	54.62
			15.0	57	85.91	47	59.88
			16.0	39	89.21	40	58.96
			17.0	13	102.76	23	61.77
			18.5	22	104.97	14	64.90
Finlayson & Reitan (1976)	1974	Canada	6.0	10	19.85	10	16.95
			7.0	10	23.05	10	18.10
			8.0	10	23.65	10	21.70
			12.0	10	41.50	10	42.15
			13.0	10	56.05	10	48.25
			14.0	10	69.05	10	53.95
Fromm-Auch & Yeudall (1983)	1981	Canada	16.0	17	73.80	15	54.40
Gallup et al. (2007)	2005	USA	19.0	80	96.09	61	52.11
Janz et al. (2000)	1998	USA	12.6	61	54.00	62	46.00
Katzmarzyk et al. (1997)	1966	USA	7.5	73	30.70	62	28.40
			9.5	51	41.40	80	33.60
			11.5	60	51.30	65	45.60
			7.5	57	33.50	52	31.00
			9.5	67	41.00	54	36.10
			11.5	69	45.00	50	38.90
Mathiowetz et al. (1986)	1984	USA	6.5	26	28.67	33	25.27
			8.5	30	36.70	32	30.98
			10.5	43	46.40	40	43.05
			12.5	34	51.76	36	48.85
			14.5	34	64.28	34	48.72
			16.5	31	78.25	35	56.34
			18.5	33	91.17	30	60.47
Milliken et al. (2008)	2006	USA	9.5	52	35.80	39	38.40
Morehouse et al. (2000)	1998	Canada	19.0	40	87.00	60	49.60
Payne et al. (2000b)	1998	Canada	17.0	54	94.00	59	56.00
Peters & Servos (1989)	1987	Canada	19.0	70	94.58	105	53.43
Shephard (1986)	1981	Canada	8.0	613	29.98	534	26.92
			11.0	654	44.46	655	39.16
			13.5	423	66.48	398	54.66
			17.0	904	95.65	911	59.96
Spreen & Gadden (1969)		Canada	7.0	36	19.90	24	19.10
			8.0	12	26.60	19	21.40
			9.0	13	32.70	14	25.30
			10.0	8	35.10	15	32.90
			11.0	12	42.00	21	36.20
			12.0	18	46.50	15	42.90
Tremblay et al. (2005)	2002	Canada	11.0	74	48.00	90	48.60
			11.0	52	46.30	58	40.80
Tremblay et al. (2010)	2007–2009	Canada	8.0	446	25.00	418	23.00
			12.5	316	51.00	301	42.00

(*Continued*)

Table I. (*Continued*).

Study	Year(s) tested	Country	Age (years)	Males		Females	
				n	*M* (kg)	*n*	*M* (kg)
			17.0	286	85.00	307	54.00
Trudeau et al. (2003)	1970–1977	Canada	10.0	51	37.21	55	32.94
			11.0	51	42.32	55	37.93
			12.0	51	48.13	55	45.47
Yeudell et al. (1987)	1985	Canada	17.5	32	84.04	30	58.27

Note: Grip strength measure is the sum of values for left and right hands, preferred and non-preferred hands, or dominant and non-dominant hands.

Table II. Hierarchical regression analysis of grip strength for each sex.

Step/predictor	*B*	*SEB*	β	R^2	ΔR^2
Males ($k = 63$)					
1. Age	6.70	0.23	0.95***	0.958	0.958
2. Age^2	0.12	0.06	0.06	0.961	0.003
3. Country	−0.17	1.58	−0.00	0.961	0.000
4. Year tested	−0.12	0.06	0.06*	0.965	0.004
5. Year tested^2	−0.00	0.01	0.00	0.965	0.000
Constant	50.41	1.91			
Females ($k = 62$)					
1. Age	4.03	0.16	1.21***	0.867	0.867
2. Age^2	−0.28	0.04	−0.30***	0.923	0.056
3. Country	−1.62	1.15	−0.05	0.925	0.002
4. Year tested	−0.05	0.04	−0.05	0.931	0.006
5. Year tested^2	−0.00	0.00	−0.04	0.932	0.001
Constant	45.49	1.39			

Note: *B*, *SEB*, and β are the values for the final step of the regression. k = number of means. Country: 1 = Canada, 0 = USA. *$P < 0.05$; **$P < 0.001$.

mance did not change over time, which is unremarkable given that his study was conducted for purely descriptive purposes. In contrast, the present study was undertaken to test the prediction that grip strength had increased in recent decades as a result of the increased prevalence of overweight/obesity in children and adolescents. Accordingly, it is important to attempt an explanation for the lack of support for the predicted increase in grip strength over time.

Several potential explanations for the failure to confirm the prediction tendered here regarding the secular trend for grip strength are discussed below. Each of these explanations is based on the idea that as body weight increased over recent decades, another factor, which changed more or less simultaneously, counteracted the effect of body weight on grip strength. One factor that might have an effect on grip strength is physical activity. Considerable research (see review by Dollman, Norton, & Norton, 2005) has shown that physical activity declined in Western children in recent decades. More specifically, physical activity decreased in three clearly defined contexts: walking and cycling to school, school physical education classes, and organized sports. Could the declines in these physical activities have affected grip strength? This is doubtful because these physical activities are more aerobic than anaerobic in nature. A further reason for doubting that the decline in physical activity had an effect on grip strength are the findings from the few studies (Ara, Moreno, Leiva, Gutin, & Casajús, 2007; Raudsepp & Jürimäe, 1996; Tremblay et al., 2005) that have examined the relation between physical activity and grip strength. These studies all assessed physical activity in terms of aerobic activities, but despite this similarity the findings produced by them are inconsistent. For example, Ara et al. found that grip strength was significantly greater in boys who participated in extra-curricular sports than in boys who did not participate in extra-curricular sport activities. In girls, however, there was no difference in grip strength between those who did and did not participate in extra-curricular sport activities.

A second potential explanation for why grip strength did not change over time as predicted focuses on the relation between vitamin D and grip strength. Vitamin D has long been known to play a vital role in the normal development of bone (Holick & Chen, 2008). Quite recently, research has indicated that vitamin D is also related to muscle composition and strength. Regarding muscle composition, Gilansz and colleagues (Gilansz, Kremer, Mo, Wren, & Kremer, 2010) found an inverse relation between serum vitamin D concentrations and infiltration of fat in muscle in females aged 16–22 years. Regarding muscle strength and function, Foo et al. (2009) found that grip strength was lower in adolescent girls categorized as vitamin D deficient or insufficient than in those categorized as vitamin D sufficient.

Unfortunately, there is no direct evidence to show that concentrations of vitamin D declined in recent decades, although there is evidence that vitamin D insufficiency/deficiency is common in American samples (Gilsanz et al., 2010; Mansbach, Ginde, & Carmargo, 2009). Obese children appear to be

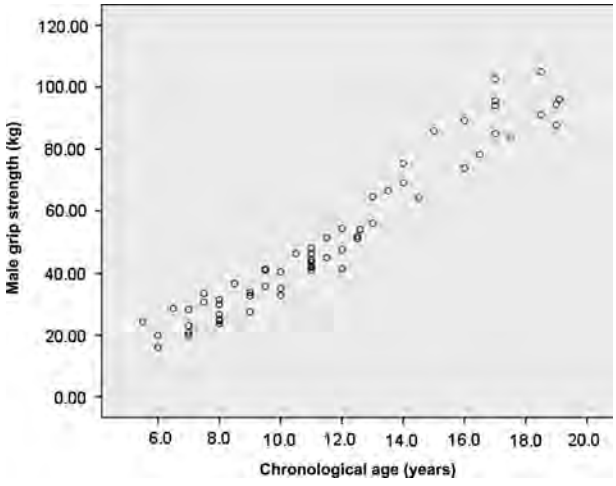


Figure 1. Scatterplot of male grip strength in relation to chronological age.

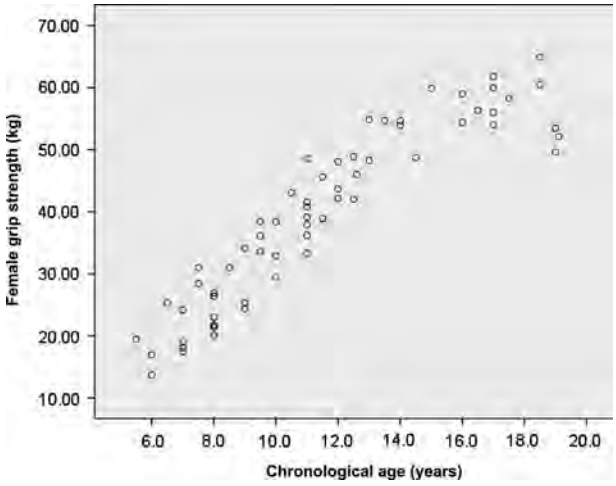


Figure 2. Scatterplot of female grip strength in relation to chronological age.

especially prone to the problem of vitamin D insufficiency/deficiency, one study finding that half of the children classified as obese had extremely low concentrations of vitamin D (Smodtkin-Tangorra et al., 2007).

Despite the absence of direct evidence showing that vitamin D declined in recent decades, there is strong indirect evidence pointing to such a trend. First, it should be pointed out that the principal source of vitamin D in humans is the action of ultraviolet B radiation in sunlight on the skin (Cannell, Hollis, Sorenson, Taft, & Anderson, 2009). Therefore, if time outdoors declined in recent decades, it is very likely that there was a concurrent decline in vitamin D in the body. Evidence for a decline in time outdoors comes from a large panel study (University of Michigan News Service, 2004) conducted with a nationally representative sample of American children and adolescents. The study found that the average amount of time spent in outdoor activities was down by a half in 2001–2002 compared with 1981–1982, 50 min per week versus 1 h and 40 min per week. Of further interest, the amount of time spent in outdoor activities for 2001–2002 averaged only a little more than 7 min per day, which suggests that some children and adolescents spent hardly any time outdoors.

A second potential explanation for why grip strength did not change over time as predicted is that fat-free mass (also called lean body mass or lean mass) declined over the same time frame as body weight increased. Underlying this explanation is the assumption that fat-free mass is predictive of grip strength. Surprisingly, a search of the literature revealed only two studies (Lebrun, van der Schouw, de Jong, Grobbee, & Lamberts, 2006; Yoon et al., 2009) providing evidence on the relation between fat-free mass and grip strength. These studies, both of which were conducted with older adults, found a strong positive correlation between fat-free mass and grip strength. However, contrary to the idea that fat-free mass has declined over time are the findings of a study of US Army recruits conducted over the period 1978 to 1998 (Sharp et al., 2002). Results show that in males there was a steady increase in fat-free mass over time, while in females, although fat-free mass differed significantly from one time period to the next, there was no overall direction in the changes over time. Furthermore, in a recent study Olds (2009) found that fat-free mass had increased by an average of 0.6 kg per decade over the period 1951 to 2004 in children and adolescents living in developed countries.

Of the potential explanations discussed above, the most likely is that vitamin D insufficiency/deficiency increased over recent decades, and this change counteracted the effect of increased body weight on grip strength. Additional research should be conducted in which vitamin D and body weight are measured in the same individuals. If grip strength depends on both body weight and vitamin D, the prediction to be tested is that grip strength will increase with body weight, but only in individuals with relatively high concentrations of vitamin D. As for changes in grip strength over time, these should depend on changes over time in body weight and concentrations of vitamin D.

Although age is not of central interest here, it is noteworthy that the present results for age are consistent with those obtained in a longitudinal study in Belgium (Taeymans, Clarys, Abidi, Hebbelinck, & Duquet, 2009). Inspection of the data reported by Taeymans et al. shows that from ages 6 to 18 years, grip strength increased according to a rising positively accelerated curve in males, whereas

grip strength increased linearly with age in girls. With one slight exception, these age changes are similar to the functions that were observed in the scatterplots of the raw data presented in Figures 1 and 2. The exception is that in the present data there was a leveling off in grip strength at around age 17 years for girls, whereas grip strength showed no leveling off at this age in the data presented by Taeymans et al.

There are three major limitations to this study. One is the lack of nationally representative samples (only two for Canada and none for the United States). In addition, the Canadian samples came from almost every part of the country, whereas the American samples came from only a handful of states. Thus, it might be questioned whether the present findings truly characterize the entire population of children and adolescents living in the United States. A second limitation is the almost total absence of data for specific racial/ethnic groups. The one exception is a study in which grip strength was compared in black and white children living in Philadelphia, Pennsylvania (Katzmaryzk, Malina, & Beunen, 1997). In Table I, the data for these two groups are listed separately under the entry for Katzmaryzk et al., with the first three rows giving the data for the black children and the next three rows giving the data for the white children. Although across age there is no consistent difference in grip strength between these two groups, it may be that differences in grip strength would be found were comparisons made with other racial/ethnic groups. Such comparisons would be especially informative regarding national differences in grip strength, because of the distinct differences in the racial/ethnic composition of Canada and the United States. A third limitation is that the estimates of the year tested might be in error because few authors provided this information. However, even with complete information on year of testing, it is likely that this variable would have accounted for, at best, only a small amount of variance in grip strength, given that age and age squared together accounted for over 90% of the variance in grip strength in both sexes.

Acknowledgement

Thanks are expressed to Mark S. Tremblay for providing unpublished data that were analysed here.

References

Ager, C. L., Olivett, B. L., & Johnson, C. L. (1984). Grasp and pinch strength in children 5 to 12 years old. *American Journal of Occupational Therapy*, *38*, 107–113.

Amosun, S. L., Moyo, A., & Matara, C. (1995). Trends in grip strength in some adult male Zimbabweans. *British Journal of Occupational Therapy*, *58*, 345–348.

Ara, I., Moreno, L. A., Leiva, M. T., Gutin, B., & Casajús, J. A. (2007). Adiposity, physical activity, and physical fitness among children from Aragón, Spain. *Obesity*, *15*, 1918–1924.

Backous, D. D., Farrow, J. A., & Friedl, K. E. (1990). Assessment of pubertal maturity in boys, using height and grip strength. *Journal of Adolescent Care*, *11*, 497–500.

Balogun, J. A., Akomolafe, C. T., & Amusa, L.O. (1991). Grip strength: Effects of testing posture and elbow position. *American Journal of Occupational Therapy*, *49*, 327–337.

Bowman, O. J., & Katz, B. (1984). Hand strength and prone extension in right-dominant, 6 to 9 year olds. *American Journal of Occupational Therapy*, *38*, 367–376.

Bowman, O. J., & Weaver, K. L. (2008). To sit or stand: Does it make a difference when one measures hand strength? *Journal of Hand Therapy*, *21*, 418.

Butler, M. (1997). Grip strength: A comparative study. *New Zealand Journal of Occupational Therapy*, *48*, 5–12.

Butterfield, S. A., Lehnhard, R. A., Loovis, E. M., Coladarci, T., & Saucier, D. (2009). Grip strength performance by 5- to 19-year-olds. *Perceptual and Motor Skills*, *109*, 362–370.

Cannell, J. J., Hollis, B. W., Sorenson, M. B., Taft, T. N., & Anderson, J. J. B. (2009). Athletic performance and vitamin D. *Medicine and Science in Sports and Exercise*, *41*, 1102–1110.

Casajús, J. A., Leiva, M. T., Villarroya, A., Legaz, A., & Moreno, L. A. (2007). Physical performance and school physical education in overweight Spanish children. *Journal of Nutrition and Metabolism*, *51*, 288–296.

Dollman, J., Norton, K., & Norton, L. (2005). Evidence for secular trends in children's physical activity behaviour. *British Journal of Sports Medicine*, *30*, 892–897.

Ertem, K., Harma, A., Cetin, A., Elmali, N., Yologlu, S., Bostan, H. et al. (2005). An investigation of hand dominance, average versus maximum grip strength, body mass index and ages as determinants for hand evaluation. *Isokinetics and Exercise Science*, *13*, 223–227.

Finlayson, M. A. J., & Reitan, R. M. (1976). Handedness in relation to measures of motor and tactile-perceptual functions in normal children. *Perceptual and Motor Skills*, *43*, 75–81.

Fleishman, E. A. (1963). Factor analysis of physical fitness tests. *Educational and Psychological Measurement*, *23*, 647–661.

Foo, L. H., Zhang, Q., Zhu, K., Guansheng, M., Hu, X., & Greenfield, H. (2009). Low vitamin D status has an adverse influence on bone mass, bone turnover, and muscle strength in Chinese adolescent girls. *Journal of Nutrition*, *139*, 1002–1007.

Fromm-Auch, D., & Yeudall, L. T. (1983). Normative data for the Halstead-Reitan Neuropsychological Tests. *Journal of Clinical Neuropsychology*, *5*, 221–238.

Gale, C. R., Martyn, C. N., Cooper, C., & Sayer, A. A. (2007). Grip strength, body composition, and mortality. *International Journal of Epidemiology*, *36*, 228–235.

Gallup, A. C., White, D. D., & Gallup, G. G. (2007). Handgrip strength predicts sexual behavior, body morphology, and aggression in male college students. *Evolution and Human Behavior*, *28*, 423–429.

Gilansz, V., Kremer, A., Mo, A. O., Wren, T. A. L., & Kremer, R. (2010). Vitamin D status and its relation to muscle mass and muscle fat in young women. *Journal of Clinical Endocrinology and Metabolism*, *95*, 1595–1601.

Häger-Ross, C., & Rösblad, B. (2002). Norms for grip strength in children aged 4–16 years. *Acta Paediatrica*, *91*, 617–625.

Haidar, S. G., Kumar, D., Bassi, R. S., & Deshmukh, S. C. (2004). Average versus maximum grip strength: Which is more consistent? *Journal of Hand Surgery (Europe)*, *29B*, 82–84.

Hamilton, A., Balnave, R., & Adams, R. (1994). Grip strength testing reliability. *Journal of Hand Therapy*, *7*, 163–170.

Holick, M. F., & Chen, T. C. (2008). Vitamin D deficiency: A worldwide problem with health consequences. *American Journal of Clinical Nutrition*, *87*, 1080S–1086S.

Janz, K., Dawson, J. D., & Mahoney, L. T. (2000). Tracking physical fitness and physical activity from childhood to adolescence: The Muscatine study. *Medicine and Science in Sports and Exercise*, *32*, 1250–1257.

Katzmarzyk, P. T., Malina, P. T., & Beunen, G. P. (1997). The contribution of biological maturation to the strength and motor fitness of children. *Annals of Human Biology*, *24*, 493–505.

Lebrun, C. E. I., van der Schouw, Y. T., de Long, F. H., Grobbee, D. E., & Lamberts, W. (2006). Fat mass rather than muscle strength is the major determinant of physical function and disability in postmenopausal women younger than 75 years of age. *Menopause*, *13*, 474–481.

Ling, C. H. Y., Taekema, D., de Craean, A. J. M., Gussekloo, J., Westendrop, R. G. J., & Mair, A. B. (2010). Handgrip strength and mortality in the oldest old population: The Leiden 85-plus study. *Canadian Medial Association Journal*, *182*, 429–435.

Mansbach, J. M., Ginde, A. A., & Camargo, C. A., Jr. (2009). Serum 25-hydroxyvitamin D levels among US children ages 1 to 11 years: Do children need more vitamin D? *Pediatrics*, *124*, 1404–1410.

Mathiowetz, V., Weber, M., Volland, K., & Kashman, N. (1984). Reliability and validity of grip and pinch strength evaluation. *Journal of Hand Surgery*, *9A*, 222–226.

Mathiowetz, V., Wiemer, D. M., & Federman, S. M. (1986). Grip and pinch strength: Norms for 6- to 19-year-olds. *American Journal of Occupational Therapy*, *40*, 705–711.

Metter, E. J., Talbot, L. A., Schrager, M., & Conwit, R. (2007). Skeletal muscle strength as a predictor of all-cause mortality in healthy men. *Journals of Gerontology A: Biological Sciences and Medical Sciences*, *57*, B359–B365.

Milliken, L. A., Faigenbaum, A. D., Loud, R. L., & Westcotte, W. L. (2008). Correlates of upper and lower body muscular strength in children. *Journal of Strength and Conditioning Research*, *22*, 1339–1346.

Montpetit, R. M., & Montoye, H. J. (1967). Grip strength of school children, Saginaw, Michigan: 1899 and 1965. *Research Quarterly*, *38*, 231–240.

Morehouse, S., Szeligo, F., & Tommaso, E. (2000). Characteristics of the bimanual deficit using grip strength. *Laterality*, *5*, 167–185.

Newman, A. B., Kupelian, V., Visser, M., Simonsick, E. M., Goodpaster, B. H., Kritchevsky, S. B. et al. (2006). Strength, but not muscle mass, is associated with mortality in the health, aging and body composition study cohort. *Journals of Gerontology A: Biological Sciences and Medical Sciences*, *61*, 72–77.

Olds, T. S. (2009). One million skinfolds: Secular trends in the fatness of young people 1951–2004. *European Journal of Clinical Nutrition*, *63*, 934–946.

Payne, N., Gledhill, N., Katzmarzyk, P. T., Jamnik, V., & Ferguson, S. (2000a). Health implications of musculoskeletal fitness. *Canadian Journal of Applied Physiology*, *25*, 114–126.

Payne, N., Gledhill, N., Katzmarzyk, P. T., Jamnik, V. K., & Kier, P. J. (2000b). Canadian musculoskeletal norms. *Canadian Journal of Applied Physiology*, *25*, 430–442.

Peters, M., & Servos, P. (1989). Performance of subgroups of left-handers and right-handers. *Canadian Journal of Psychology*, *43*, 341–358.

Raudsepp, L., & Jürimäe, T. (1996). Physical activity, fitness, and adiposity of prepubertal girls. *Pediatric Exercise Science*, *8*, 259–267.

Saski, H., Kasaqi, F., Yamada, M., & Fujita, S. (2007). Grip strength predicts cause-specific mortality in middle-aged and elderly persons. *American Journal of Medicine*, *120*, 337–342.

Sharp, M. A., Patton, J. F., Knapik, J. J., Hauret, K., Mello, R. P., Ito, M. et al. (2002). Comparison of the physical fitness of men and women entering the U.S. Army: 1978–1998. *Medicine and Science in Sports and Exercise*, *34*, 356–363.

Shephard, R. J. (1986). *Fitness of a nation: Lessons from the Canadian Fitness Survey*. Basel, Switzerland: Karger.

Shields, M. (2006). Overweight and obesity among children and youth. *Health Reports*, *17*, 27–42.

Smodtkin-Tangorra, M., Purushothaman, R., Gupta, A., Nejati, G., Anhalt, H., & Ten, S. (2007). Prevalence of vitamin D insufficiency in obese children and adolescents. *Journal of Pediatric Endocrinology and Metabolism*, *20*, 817–823.

Spreen, O., & Gadden, W. H. (1969). Developmental norms for 15 neuropsychological tests age 6 to 15. *Cortex*, *5*, 170–191.

Swanson, A., Matev, I., & de Groot, G. (1970). The strength of the hand. *Bulletin of Prosthetics Research*, *145*, 10–14.

Taeymans, J., Clarys, P., Abidi, H., Hebbelinck, M., & Duquet, W. (2009). Developmental changes and predictability of static strength in individuals of different maturity: A 30-year longitudinal study. *Journal of Sports Sciences*, *27*, 833–841.

Takata, Y., Ansai, T., Akifusa, S., Soh, I., Yoshitake, Y, Kimura, Y. et al. (2007). Physical fitness and 4-year-mortality in an 80-year-old population. *Journals of Gerontology A: Biological Sciences and Medical Sciences*, *62*, 851–858.

Tomkinson, G. (2007a). Global change in anaerobic fitness test performance of children and adolescents (1958–2003). *Scandinavian Journal of Medicine and Science in Sports*, *17*, 497–507.

Tomkinson, G. (2007b). Secular changes in pediatric aerobic fitness test performance: The global picture. In G. R. Tomkinson, & T. S. Olds (Eds.), *Pediatric fitness: Secular trends and geographic variability*. (pp. 46–66). Basel: Karger.

Tomkinson, G. R., Léger, L., Olds, T., & Cazorla, G. (2003). Secular changes in the performance of children and adolescents (1980–2000): An analysis of 55 studies of the 20 m shuttle test in 11 countries. *Sports Medicine*, *33*, 285–300.

Tremblay, M. S., Barnes, J. D., Copeland, J. L., & Esliger, D. W. (2005). Conquering childhood inactivity: Is the answer in the past? *Medicine and Science in Sports and Exercise*, *37*, 1187–1194.

Tremblay, M. S., Shields, M., Laviolette, M., Craig, C. L., Janssen, I., & Gorber, S. C. (2010). Fitness of Canadian children and youth: Results from the 2007–2009 Canadian Health Measures Survey. *Health Reports*, *21*, 7–20.

Trudeau, F., Shephard, R. J., Arsenault, F., & Laurencelle, L. (2003). Tracking of physical fitness from childhood to adulthood. *Canadian Journal of Applied Physiology*, *28*, 257–271.

University of Michigan News Service (2004). *U.S. children and teens spend more time on academics*. Retrieved July 21, 2010, from http://www.ns.umich.edu/index.html?Releases/2004/Nov04/r111704a.

Wang, M., Leger, A. B., & Dumas, G. A. (2005). Prediction of back strength using anthropometric and strength measurements in healthy females *Clinical Biomechanics*, *20*, 685–692.

Wind, A. E., Takken, T., Helders, P. J. M., & Engelbert, R. H. H. (2010). Is grip strength a predictor for total muscle strength in healthy children, adolescents, and young adults? *European Journal of Pediatrics*, *169*, 281–287.

World Health Organization (2010). *Obesity and overweight*. Retrieved July, 30, 2010 from http://www.who.int/dietphysicalactivity/publications/facts/obesity/en/.

Yeudell, L., Reddon, J. R., Gill, D. M., & Stefanyk, W. O. (1987). Normative data for the Halstead-Reitan Neuropsycholgical Test stratified by age and sex. *Journal of Clinical Psychology*, *43*, 346–376.

Yoon, B.-K., Kim, C.-H., Lim, H.-J., Kim, Y.-S., Im, J.-A., Paik, I.-Y. et al. (2009). Association of physical performance and health-related factors among elderly Korean subjects. *International SportMed Journal*, *10*, 205–215.

Published in final edited form as:
Int J Obes (Lond). 2012 October ; 36(10): 1261–1269. doi:10.1038/ijo.2012.95.

Ethnic and sex differences in body fat and visceral and subcutaneous adiposity in children and adolescents

AE Staiano and **PT Katzmarzyk**
Population Science, Pennington Biomedical Research Center, Baton Rouge, LA, USA

Abstract

Body fat and the specific depot where adipose tissue (AT) is stored can contribute to cardiometabolic health risks in children and adolescents. Imaging procedures including magnetic resonance imaging and computed tomography allow for the exploration of individual and group differences in pediatric adiposity. This review examines the variation in pediatric total body fat (TBF), visceral AT (VAT) and subcutaneous AT (SAT) due to age, sex, maturational status and ethnicity. TBF, VAT and SAT typically increase as a child ages, though different trends emerge. Girls tend to accumulate more TBF and SAT during and after puberty, depositing fat preferentially in the gynoid and extremity regions. In contrast, pubertal and postpubertal boys tend to deposit more fat in the abdominal region, particularly in the VAT depot. Sexual maturation significantly influences TBF, VAT and SAT. Ethnic differences in TBF are mixed. VAT tends to be higher in white and Hispanic youth, whereas SAT is typically higher in African American youth. Asian youth typically have less gynoid fat but more VAT than whites. Obesity per se may attenuate sex and ethnic differences. Particular health risks are associated with high amounts of TBF, VAT and SAT, including insulin resistance, hepatic steatosis, metabolic syndrome and hypertension. These risks are affected by genetic, biological and lifestyle factors including physical activity, nutrition and stress. Synthesizing evidence is difficult as there is no consistent methodology or definition to estimate and define depot-specific adiposity, and many analyses compare SAT and VAT without controlling for TBF. Future research should include longitudinal examinations of adiposity changes over time in representative samples of youth to make generalizations to the entire pediatric population and examine variation in organ-specific body fat.

Keywords

pediatric; body fat; visceral adipose tissue; subcutaneous adipose tissue; ethnic differences; sex differences

Correspondence: Dr PT Katzmarzyk, Population Science, Pennington Biomedical Research, Center, 6400 Perkins Road, Baton Rouge, LA 70808-4124, USA., Peter.Katzmarzyk@pbrc.edu.

CONFLICT OF INTEREST
The authors declare no conflict of interest.

INTRODUCTION

Pediatric obesity contributes to physical and emotional health problems during childhood[1] and leads to future co-morbidities and premature mortality in adulthood.[2] Health consequences associated with obesity differ based on where adipose tissue (AT) is stored, for instance subcutaneously between the skin and muscle versus internally, such as within the chest, abdomen and pelvis (visceral AT (VAT)) or within and between muscles (nonvisceral AT).[3] Recent advances in body composition imaging allow for the examination of specific depots of AT in the pediatric population, providing sufficient precision to evaluate individual and group differences. Although the terms body fat and AT are often used interchangeably, fat specifically refers to lipids in the form of triglycerides, located mostly in AT though also present in other tissues of the body.[3] Dual-energy X-ray absorptiometry (DXA) is the method most commonly used to quantify total body fat (TBF) *in vivo*.

AT is loose connective tissue replete with adipocytes.[3] Its chief functions are energy storage, thermal insulation and mechanical cushioning. AT is composed of 80% fat but also includes protein, minerals and water.[4] Total AT is subdivided into subcutaneous AT (SAT) and internal AT (IAT).[3] The major contributor to IAT is VAT, which is located beneath the abdominal muscles and is particularly linked to poor cardiometabolic outcomes in youth.[5,6] To measure total AT as well as AT stored in specific organs and regions, the gold standards are magnetic resonance imaging (MRI) and computed tomography (CT). MRI is favored for the measurement of AT in children because, unlike CT, it uses a magnet and does not emit radiation.[7] We primarily focus on the depots of VAT and SAT, yet because the literature contains variation in the definitions used for each internal compartment,[3] we include IAT and intra-abdominal AT (IAAT) when VAT is not reported. IAT includes all AT other than SAT,[3] and IAAT includes AT within both the abdomen and pelvic region.[8]

Recent technological advances in biomedical imaging have transformed our understanding of pediatric obesity phenotypes in recent years. Thus, the purpose of this paper is to review the current status of our understanding about the extent of human variation observed in adiposity across the pediatric age range. Our focus is on the contributions of age, sex, maturation and ethnicity to variation in TBF, VAT and SAT. Further, we provide evidence of the link between adiposity and health outcomes where possible. Finally, we also discuss potential mechanisms including biological, lifestyle and environmental influences, challenges to measuring children's adiposity, and future research directions. Studies were preferentially included if the methodology employed MRI, CT or DXA. Studies using non-imaging techniques, such as anthropometry, hydrodensitometry or bioelectrical impedance methods, were included where image-based evidence was sparse.

INFLUENCE OF AGE

Total body fat

TBF accumulates as a child ages,[9] although inter-individual variation is evident as early as the third trimester of gestation. For example, MRI examination of 27 fetuses at 38–41 weeks of gestation demonstrated a range of 11.8–25%BF.[10] Thirteen of these babies born to

mothers with poorly controlled diabetes had higher birth weight and higher %BF (average of 27.4%) compared with the normal group, revealing early maternal influences on TBF.[10]

TBF also varies considerably during infancy: for instance, %BF estimated by MRI ranged from 13.3 to 22.6% in a group of eight white infants measured within 36 h of delivery.[11] Estimates from total body electrical conductivity demonstrated a sharp increase in %BF during the first 6 months of life and a slow decline thereafter in 423 healthy white infants (aged 14–379 days).[12]

After the first year of life, absolute TBF typically declines or stabilizes until 6 years of age.[13,14] A DXA study of the first 24 months of life found that %BF increased between 15 days and 6 months then decreased thereafter, although TBF increased with age.[15] At about 6 years of age, an ‘adiposity rebound’ occurs where TBF increases throughout the rest of childhood and into adolescence.[16] This rebound was confirmed in a growth curve analysis of CT scans that documented an average 2.0±0.9 kg per year increase in TBF throughout childhood between ages 8 and 13, adjusted for ethnicity, sex and baseline age.[17] When adjusted for total lean tissue mass, the increase in TBF remained significant at 1.9±0.8 kg per year, indicating that TBF accumulated faster than the growth of lean tissue.[17]

During the peripubertal and postpubertal periods, TBF tends to fluctuate within and among individuals. Trajectories established for 678 US children between the ages of 8 and 18 years, based on 12 observations over 4 years, revealed that TBF decreased as males aged, whereas TBF increased or remained constant in females.[18] The increase in TBF co-occurred with increases in body mass index (BMI) and abdominal circumference in both males and females. In a DXA study of 112 white girls tracked from ages 11 to 18 years, TBF gained was B6 kg.[19] Increase in %BF occurred between ages 11 and 12 and after age 16, whereas %BF decreased between 13.5 and 16 years.[19]

In an underwater weighing study, girls increased from 6.4 to 16.3 kg TBF and from 20 to 26%BF between the ages of 8 and 20, increasing constantly across the age range.[9] TBF growth was less consistent for boys and was inversely related to the increase in fat-free mass.[9] The increase of TBF persisted until about age 16 in girls and age 18 in boys, at which point it typically stabilized.[20] Similarly, in a sample of 8269 5–18-year-old children from the US National Health and Nutrition Examination Survey (NHANES), %BF based on skinfold thicknesses peaked at age 11 in boys but increased throughout childhood and adolescence for girls, resulting in a 1.5 times greater %BF in girls versus boys by age 18.[21]

Visceral adipose tissue

Young children deposit less than 10% of their AT in the IAAT depot, as observed in 12–14-year-olds.[8] Within the first month of life, IAT accounted for 10% of TBF in eight infants (within 36 h of delivery), yet most was deposited in the pelvis and limbs, not in the abdomen.[11] In fact, VAT was on average 0.03±0.01 l that comprised 33% of total IAT and just 0.7% of total body weight.[11] However, IAT may act as a protective fat layer in early infancy: 10 growth-restricted infants who had lower TBF, total SAT and abdominal SAT, still had similar levels of IAT compared with 25 normal weight newborns.[22] Similarly, VAT was higher in 2–6-year-old children born small-for-gestational age compared with normal

weight babies.[23] High levels of VAT can persist for several years: by age 6 years, VAT in small-for-gestational age babies was on average 50% higher than normal birth weight children.[24]

VAT increases with age throughout childhood (ages 5–17) and into adulthood.[25] In African American prepubertal 4–10-year-old children, the rate of change of IAAT was 4.3±1.6 cm^2 per year over 2 years.[26] A longitudinal study of 138 white and African American children aged 8 years at baseline and followed for 3–5 years found that VAT grew on average 11.6±2.9 cm^2 per year, and after adjusting for abdominal SAT the growth rate remained significant at 5.2±2.2 cm^2 per year.[17] However, as age increased, the growth in VAT slowed down.[17]

Age contributed 7.3% to the variance in VAT in 7–16-year-old white and Hispanic youth.[27] Over a 2-year period in 11–13-year-olds, boys increased IAT by 69% and increased from 0.31 to 0.39 in abdominal IAT-to-SAT ratio.[28] Girls increased 48% in IAT with a reduction in abdominal IAT-to-SAT ratio from 0.39 to 0.35.[28] One cross-sectional study adjusted for total AT and ethnicity, and demonstrated a decline in VAT as females went through adolescence, in contrast to males where VAT grew larger after age 12.[25]

In summary, VAT is present at birth and increases throughout childhood and adolescence, independent of growth in TBF or SAT. VAT growth appears constant throughout prepubertal ages (4–10 years)[26] but group differences in VAT growth emerge during the peripubertal and postpubertal ages.[27]

Subcutaneous adipose tissue

SAT increases as children age,[25] and age contributed 11% of the variance in SAT in a sample of 7–16-year-old Hispanic and white children.[27] In the first year of life, SAT composes the majority of TBF, varying between 89.0 to 92.3% of TBF [29] and 15.9 to 27.8% of total body weight.[11] Abdominal SAT is slightly lower, averaging 0.11±0.06 l in eight white newborn infants but still accounting for 12.2% of total body weight.[11] Annual observations over 3–5 years in 138 children (age 8.1±1.6 years) demonstrated that abdominal SAT grew 32.6±10.7 cm^2 per year.[17] Once adjusted for TBF, however, the growth rate was no longer significant. The authors conjecture that SAT is deposited in other areas than abdominally during this period in childhood.[17]

INFLUENCE OF SEX

Total body fat

Cross-sectional and longitudinal studies indicate that girls have more TBF than boys throughout childhood and adolescence (Table 1). This has been demonstrated in a longitudinal underwater weighing study in 8–20-year-olds[9] and in DXA cross-sectional studies of 7–17-year-olds in the US[30,31] and 6–18-year-olds in China,[32] among others. A cross-sectional study of 265 4–26-year-olds revealed that DXA-measured %BF was higher in females compared with males at all ages, and %BF increased for females throughout this period but not for males.[33] A contributing factor to sex differences in TBF is the higher amount of extremity BF in girls, as demonstrated in US children and adolescents aged 5–18

years.[34,35] Girls also have higher %BF by age 5, and this sex difference continues to increase until 18 years of age.[36] One exception is a study of 194 boys and 96 girls aged 6–15 years, in which CT-measured %BF was not different among girls and boys, though this was in a sample of obese children where sex differences may be attenuated.[37]

Sex differences in body fat emerge at specific developmental periods. At infancy, there is little documented sex difference in TBF. Female infants tend to have 50 g more DXA-derived TBF at birth than male infants, yet this 1.5% difference is within the coefficient of variation, and thus may not be detected in small sample sizes.[11] For instance, there were no sex differences in TBF in a cohort of eight white infants.[11] However, male infants tend to be longer in stature and have more lean mass during the first year of life,[38] which may contribute to sex differences in %BF.

Little MRI or CT data are available on sex differences during early childhood. A multi-component study calculating TBF from water, potassium and bone content found that TBF did not differ by sex during 0 to 24 months, except at 6 and 9 months at which point girls had higher %BF than boys.[15] In prepubertal children, girls typically have more TBF than boys. A CT study of 43 boys and 58 girls in the US aged approximately 7 years demonstrated that girls had more TBF.[26] African American 4–10-year-old girls had higher TBF and %BF (measured by DXA) than boys,[39] and Italian 3–11-year-old girls had more TBF than boys based on estimated fat mass from skinfold measurements.[40] However, not all studies demonstrate sex differences in TBF before puberty. A multi-year longitudinal study of American boys and girls aged 8.1±1.6 years found similar TBF.[17] There were no significant sex differences for TBF measured by bioelectrical resistance in a study of 4 boys and 12 girls aged 6.4±1.2 years in the US.[41] A study of 129 African American and white 10–12-year-olds indicated no difference in TBF measured by DXA across sexes, though boys had a bimodal distribution of TBF whereas girls' TBF was skewed to higher values.[42] Additionally, there were no sex differences in total abdominal fat measured by CT in 31 6–7-year-olds in the Netherlands.[43] Not controlling for other influences like age, maturational status and obesity status may account for the contradictory findings on sex differences, particularly in studies with small sample sizes.

During the pubertal period, females develop more TBF and fat deposited in the arms and legs, whereas males develop more total lean and muscle mass.[38] In a group of 678 children aged 8, 11 and 14 years, BMI and waist circumference (WC) increased similarly for both sexes, yet TBF increased in females and decreased in males.[18] In contrast, one study demonstrated no sex differences in TBF in a study of 160 US girls and boys aged 12–13 years.[44] Throughout adolescence after puberty, boys continue to primarily increase lean mass with little increase in TBF, as opposed to girls who tend to gain substantial TBF but little lean mass.[38] This sex divergence was also demonstrated in skinfold measurements collected in a representative sample of 5–18-year-olds in the US: girls increased in %BF throughout childhood and adolescence, whereas boys %BF peaked at age 11 and declined thereafter.[21]

Visceral adipose tissue

Findings related to sex differences in children and adolescents VAT are mixed. Many studies indicate males have more VAT than females throughout the ages of 5–25,[5,25,44,45] while others indicate no sex differences in VAT after adjustment for abdominal SAT in 4–10-year-olds.[26] Sex explained 1.8% of the variance in VAT in a study of 497 7–16-year-old prepubertal and pubertal boys and girls, and there was no sex difference in the abdominal VAT-to-SAT ratio.[27] However, boys had more VAT than girls in an MRI study of 12–13-year-old children, and the magnitude of the difference increased as WC increased.[44]

During childhood before puberty, boys may accumulate more VAT than girls. This was demonstrated in a study of 138 US girls and boys (mean age 8.1±1.6 years and followed for 3–5 years) in which VAT was higher in boys than in girls,[17] as well as in 290 Japanese children aged 6–15 years.[37] Despite similar %BF in 64 7–11-year-old obese boys and girls, boys had more VAT.[5] A study of 138 Hispanic and African American youth aged 13–25 showed that MRI-measured VAT was higher in boys than in girls.[45] Whereas IAT increased in peripubertal boys over a 2-year period (baseline mean 13 years old), the same decreased in girls.[28] In contrast, a CT study of prepubertal boys and girls (mean 13 years old) showed that IAAT was the same.[30]

During puberty, boys develop a more android shape by depositing more fat in the abdomen, whereas girls develop more TBF in general but deposit it in the hips and limbs forming a gynoid shape.[46] Boys' abdominal fat increases independently of total AT,[38] which was demonstrated in Australian 5–35-year-old males who had more abdominal fat than girls regardless of TBF.[47] In a US study (n = 160) of 12-year-olds, pubertal boys had higher VAT, WC, BMI and waist–hip ratio, even though pubertal girls had higher %BF.[44] VAT remained similar across the ages of 5–12 years, yet males had more VAT between 12 and 17 years, which marked pubertal onset for most.[25] DXA-derived waist fat and trunk fat adjusted for extremity fat was higher in boys than girls for those in late puberty in a sample of 5–29-year-olds, even though girls had more extremity and hip fat and more %BF.[48] Although 6–16-year-olds in Japan had no sex differences in VAT, older adolescent boys aged 16–20 had more VAT than their female counterparts.[49] Importantly, obesity increases android fat distribution in both sexes, thereby decreasing sex differences in body shape.[38]

Some studies demonstrate that girls have higher VAT than boys during adolescence. In a study of 160 12–14-year-olds, despite similar WC, girls had more IAAT compared with boys;[8] girls also were more sexually mature and had higher BMI and WC. Female adolescents aged ∼13–14 years had more VAT than males,[50] and also had higher BMIs. There was no sex difference in IAT (measured at the L4 lumbar level) in 16 obese adolescents (baseline age 12.8±1.4 years) measured over a 5-year period during pubertal attainment.[51] However, none of these studies controlled for TBF, and the fact that girls tend to have higher TBF may be driving the sex difference in these studies. More research is needed to elucidate reliable sex differences in VAT, and controlling for TBF is an important consideration.

Subcutaneous adipose tissue

SAT appears to be similar across sex before puberty. A study of 4 boys and 12 girls aged 6.4±1.2 years in the US showed no significant sex difference in abdominal SAT measured by skinfold thickness.[41] Similarly, abdominal SAT was the same in ~8-year-olds boys and girls in US,[17] in 6–7-year-olds in the Netherlands[43] and in 6–15-year-olds in Japan.[37] In contrast, 4–10-year-old girls had higher CT-measured abdominal SAT than boys,[26] African American 4–10-year-old girls had higher abdominal SAT measured by CT,[39] 8-year-old prepubertal girls had more abdominal SAT than boys in a US study using CT,[30] and girls had more whole-body SAT than boys in a sample of 147 5–17-year-old Caucasian, African American, Hispanic and Asian children.[25]

After puberty girls tend to accumulate more SAT than boys, as demonstrated in a 2-year study of 11–13-year-olds where girls increased abdominal SAT by 78% versus a 19% increase in boys.[28] Additionally, girls had more SAT in a CT study of 11–20-year-old Japanese adolescents,[49] and girls had more abdominal SAT in a sample of British 12–14-year-olds.[8] Whole-body SAT was also higher in pubertal adolescent girls in an MRI study of 5–17-year-olds in the US.[25] An exception was found in boys and girls with similar SAT in a CT study of 12–13-year-olds[44] and in a study of 138 US girls and boys aged 13–25;[45] however, neither analysis controlled for TBF.

INFLUENCE OF MATURATIONAL STATUS

Total body fat

Puberty involves simultaneous hormonal, biological and behavioral changes centered on sexual maturation, including the development of primary and secondary sexual characteristics.[52] Sexual maturation influences TBF accumulation. Whereas boys decrease in gynoid body fat in late puberty compared with early prepuberty, girls accumulate more gynoid body fat.[35] For instance, girls who are more sexually mature have more TBF than those less mature, whereas it is the opposite for boys.[53] In a DXA study of 920 5–18-year-old children grouped into pre-, early- and late-puberty based on breast or genitalia and pubic hair development, gynoid BF was lower in late pubertal compared with prepubertal boys, but there were no differences across pubertal stage for girls.[35] Skeletal maturation is also related to TBF, where rapidly maturing girls had more TBF and %BF than intermediate maturing girls, and rapidly maturing boys assessed from 8 to 20 years of age had higher TBF and %BF compared with slowly maturing boys.[9]

Visceral adipose tissue

Pubertal status explained 12.4% of the variance in VAT in a study of 7–16-year-olds.[27] In fact, failing to control for children and adolescents pubertal stage may contribute to inconsistencies in VAT comparisons across studies.[25] Obesity status may also alter the effects of puberty on VAT: in normal weight children, IAT typically decreases during puberty, whereas IAT typically stabilizes in obese children.[51] A longitudinal study of 16 obese male and female adolescents aged 12.8±1.4 years indicated that over a 4-year period, during which puberty was completed, IAT did not change, nor did relative body weight.[51]

Some results indicate that pubertal status did not significantly predict VAT in 5–17-year-olds, although chronological age did.[25] Children aged 7.7±1.6 years who remained prepubertal gained a similar amount of IAT (4.6 (SD 2.1) cm^2 per year) compared with those who began puberty (5.6 (SD 2.1) cm^2 per year).[26] In 10–15-year-old normal weight and obese youth, there was no difference in abdominal SAT-to-IAT ratio by pubertal status.[54] IAT remained constant over the 4 years (under 130 cm^2) though there was a 15–100 cm^2 range in individual variation of IAT.[54] Yet an MRI study of 170 British peripubertal 12–14-year-old youth demonstrated that pubertal status explained 3.7% of the variance in IAAT and was significantly related to IAAT in boys but not girls.[8] Pubertal status did not, however, significantly relate to the abdominal IAAT-SAT ratio in girls or boys.

During pubertal onset and directly following puberty, children and adolescents typically have low amounts of VAT compared with SAT. For instance, in one MRI study of 170 British 12–14-year-olds, less than 10% of total abdominal fat was IAAT.[8] Whereas fat distribution was consistent in pre- versus late-pubertal girls aged 5–18 years, boys gained more of an android fat distribution late in puberty.[35] Differences in VAT occurring in late puberty may be predominantly due to boys accumulation of VAT during this period versus smaller VAT growth in girls.[55]

Subcutaneous adipose tissue

Pubertal status contributed to 18.6% of the variance in abdominal SAT in a study of 497 7–16-year-old white and Hispanic children.[27] SAT appears to be relatively stable during puberty, as demonstrated in 170 12–14-year-olds in which there was no effect of pubertal status on abdominal SAT.[8] However, one study demonstrated an increase in abdominal SAT during and after puberty in 16 obese male and female adolescents over a 4-year period during pubertal onset from approximately age 12 through age 16.[51] Yet age, not pubertal status, significantly predicted SAT in a full-body MRI scan of 5–17-year-olds.[25]

INFLUENCE OF ETHNICITY

Total body fat

Ethnicity is a significant correlate of %BF, independent of BMI, sex, sexual maturation and distribution of fat, as demonstrated in a study of 201 white and African American 7–17-year-olds.[31] Ethnic differences are evident in TBF and fat patterning in children (Table 2). For instance, Asian 8–10-year-olds varied in TBF depending on country of origin and a marginal trend persisted in girls once adjusted for age and BMI.[56] White youth typically have more %BF than African American youth at any given BMI as observed in a sample of 7–17-year-olds;[31] however, the population of African American 2–17-year-old children experienced a steeper rise in the prevalence of obesity measured by BMI based on a 30-year period of successive cross-sectional data.[57] In a bioelectrical impedance study of white and African American 9–19-year-old girls, white girls had higher %BF between ages 9 and 12, whereas African American girls had a higher %BF at older ages.[58] These differences are attributed to minor fluctuations of %BF in girls between the ages of 9 and 12 years, followed

by a steeper incline in %BF in African American girls after age 12 that eclipsed white girls' %BF.

In a DXA study of 920 children, Asian girls had less gynoid fat compared with white and African American girls throughout pre-, early and late puberty during the ages of 5–18, and Asian boys had less gynoid fat during early and late puberty.[35] Similarly, Asian boys and girls had less extremity and gynoid fat compared with whites, whereas gynoid fat was similar between Asian and African American boys.[35]

Some studies indicate no ethnic differences in TBF: in a sample of 36 white and 65 African American prepubertal 4–10-year-old children, there were no ethnic differences in TBF or %BF despite differences in IAAT and abdominal SAT.[26] In a sample of 40 African American and white 7–10-year-old girls, there was no difference in DXA-measured TBF or %BF calculated by bioelectric impedance, although white girls had more fat deposited in the arm and chest.[59] Similarly, a sample of 40 African American and white 8–18-year-old overweight adolescent boys demonstrated no difference in MRI-measured total AT, even though African Americans had more whole-body SAT and less VAT.[60] One study demonstrated higher TBF in African Americans than whites at approximately age 8.[17] Obesity status may attenuate racial differences in TBF, demonstrated in a study of 55 obese adolescents (mean age 14–15 years) in which TBF and %BF did not differ among white, African American or Hispanic ethnic groups.[61]

Visceral adipose tissue

Racial/ethnic differences in VAT appear as early as infancy: in a comparison of 69 healthy Asian Indian and white European infants within 2 weeks of birth, Asian Indians had more VAT, despite having lower body mass, smaller head circumference and length, and less non-abdominal superficial SAT compared with the white Europeans.[62] In fact, ethnicity is a significant predictor of IAAT and can be used in a regression equation along with skinfold thickness to predict IAAT when DXA data are absent.[63]

Similarly, white youth have more VAT than African American youth at a given BMI.[64] In a study of 55 obese adolescents aged ∼13 years, MRI-measured VAT was higher in white and Hispanic obese adolescents compared with African American obese adolescents,[61] and in a multi-ethnic sample of 118 obese adolescents aged 13–15, African Americans were less likely to be in the middle or upper tertile of VAT compared with white and Hispanic adolescents.[65] Despite similar total AT, 11–18-year-old overweight white boys had 50% more VAT than similarly aged overweight African American boys.[60] White children also accumulate IAAT relative to abdominal SAT at a 26% higher rate compared with African American children aged 4–10, demonstrated by a steeper regression line for IAAT to abdominal SAT in white compared with African American obese and non-obese children.[26] In a 3–5-year longitudinal study beginning at approximately age 8, white children had a steeper growth in VAT, growing on average 1.9±0.8 cm^2 per year in VAT more than African Americans did, with no ethnic difference in abdominal SAT or TBF growth.[17] A study of 20 African American and 20 white 7–10-year-old normal weight girls matched for BMI, bone age, chronological age, breast stage and socio-economic status, found that white girls had higher MRI-measured VAT and higher waist-to-thigh ratio compared with African

American girls.[59] For a given waist-to-height ratio, in a sample of 12-year-olds, white boys had more VAT and higher WC than African American boys, but there was no difference for girls.[44]

There were ethnic differences between 407 Hispanic and white 5–18-year-olds where Hispanics had higher VAT amounts than whites, but after correcting for abdominal SAT and BMI there was no difference in VAT.[27] Moreover, ethnicity explained just 2.1% of the variance in VAT and 5.9% of the variance in abdominal SAT.

Racial differences in VAT may be attenuated in obese children and adolescents. For instance, one study of 36 obese African American and white 6–18-year-olds found that CT-measured VAT did not differ by ethnicity after adjustment for age and pubertal stage.[66]

Subcutaneous adipose tissue

Although white youth tend to have higher WC on average, African American youth often have higher abdominal SAT and this accumulates faster at higher levels of WC.[44] African American 7–11-year-olds had more abdominal SAT and TBF than white youth, despite no ethnic difference in %BF or VAT.[5] Also, 11–18-year-old African American overweight boys had more whole-body SAT and specifically more leg and thigh SAT compared with overweight white boys, despite having similar total AT.[60] However, abdominal SAT did not differ by ethnicity in a study of 36 obese African American and white 6–18-year-olds,[66] and abdominal SAT was higher in white girls compared with African American girls in a 7–10-year-old sample matched for age, BMI, breast stage and socioeconomic status.[59] Ethnic differences are also seen at the beginning of life: in a study of 69 infants under 2 weeks old, Asian Indian infants had more deep abdominal SAT and superficial abdominal SAT than white European infants, even with lower body mass.[62]

HEALTH RISKS

The importance of where AT is stored and ensuing health risks was made evident in a comparison study of 28 obese adolescents on average 13 years old, of which half were insulin resistant and half were insulin sensitive.[67] Despite pairs being matched for age, sex, pubertal stage and body composition, obese insulin-resistant adolescents had higher VAT, indicating that VAT in particular was related to early formation of insulin resistance. One study of 118 obese adolescents demonstrated that as VAT increased, abdominal SAT decreased.[68] Interestingly, this adiposity profile of high VAT and low abdominal SAT had hepatic steatosis, insulin resistance and increased risk for metabolic syndrome. In a group of 14 obese adolescent girls aged 10–16, VAT but not BMI or waist–hip ratio highly correlated with cardiovascular risk factors including basal insulin, triglycerides and HDL cholesterol.[69]

Ultrasonography-measured VAT in 192 6–15-year-old obese children demonstrated that VAT (maximum preperitoneal fat thickness) was related to elevated systolic blood pressure, regardless of family history of hypertension.[70] MRI-measured VAT was related to adverse markers of insulin resistance syndrome in 81 obese African American and white 13–16-year-olds, independent of cardiovascular fitness.[71] In fact, VAT was a more powerful predictor than %BF (measured by DXA) for lipoproteins. An early review of the literature

determined that IAAT was related to fasting insulin, insulin secretion and sensitivity, triglyceride and cholesterol concentrations.[72] However, DXA-measured TBF was related to insulin sensitivity.

Sex influences

Sex differences in fat distribution may lead to different health outcomes. For instance, a study of 920 healthy US 5–18-year-olds revealed a relationship between trunk fat and higher fasting blood pressure in boys but not in girls.[34] This relationship remained in boys (African American, Asian, white) across all pubertal stages. Intra-abdominal obesity may only adversely influence blood pressure in males, whereas the metabolic and inflammatory responses to excess adiposity may be similar in boys and girls.[6] Obesity may attenuate sex differences in obesity-related health outcomes: in a separate study of children aged ∼11 years-old, there was no difference in insulin resistance in obese boys and girls with similar TBF, %BF, abdominal SAT and VAT.[66]

Maturational status

Pubertal status may affect the relationship between AT and health outcomes. In children aged ∼12 years whose amount of IAT remained constant across multiple measurements, IAT was related to insulin glucose metabolism after puberty only, but there was no relationship between IAT and insulin or glucose before puberty.[51] In peri- and postpubertal adolescents aged 10–15, IAT was significantly related to insulin and HDL cholesterol, demonstrating that adolescents past puberty have similar IAT-risk factor relationships as adults do.[54]

Ethnicity

The relationship between adiposity and metabolic risk may differ between African American and white youth.[64] Despite a lower VAT, African American 7–12-year-old youth tend to have higher risk for diabetes compared with white youth, as well as lower insulin sensitivity.[73] Glucose and insulin were correlated with abdominal SAT in African American girls only, and glucose/insulin were not correlated with VAT in either African American or white girls aged 7–10.[59] One study of 36 obese African American and white 6–18-year-olds found that VAT and abdominal SAT did not differ by ethnicity, nor did insulin levels or insulin resistance.[66] Further studies should investigate racial differences in how various adipose depots confer health risk.[59]

POTENTIAL MECHANISMS

It is important to discover the underlying mechanisms for the observed group differences in total and regional adiposity in children and adolescents, particularly to design clinical and public health interventions to prevent the accumulation of excess adiposity.

Biological and genetic factors

Genetic factors explain a significant proportion of the variance in total and depot-specific body fat.[74,75] However, the degree to which observed sex or ethnic differences in adiposity are explained by genetic differences is not well understood. The relative contributions of

genes versus the environment to the total phenotypic variance in BMI may differ between white and African American children;[76] however, little research exists on the differential effects of specific genetic markers for obesity in different ethnic groups in childhood.[77] The determination of genetic influences on obesity in different ethnic groups is a research priority.

Pubertal changes including insulin-sensitivity change,[53] hormonal and endocrine factors,[46] and sex steroid hormones like estrogen[38] relate to body composition changes. Body composition may differ based on growth spurts and peak height velocity.[78] The earlier that the adiposity rebound occurs in a child's life, the more likely that child is to become overweight.[16]

Current level of adiposity may determine the location of subsequent adiposity accumulation: differences between obese and non-obese children are predominantly found in the abdominal SAT compartment, although obese children also have more IAAT.[79] The difference in SAT by adiposity status may be from continued expansion of the SAT depot compared to a plateau of IAT, as demonstrated in obese 12-year-old adolescents over a 4-year longitudinal study.[51] Obesity also attenuates group differences in adiposity and fat. Ethnic differences between obese white and black 6–18-year-olds were only found for the SAT depot but not for VAT,[66] and no ethnic differences were observed in TBF or %BF in obese 13–14-year-olds.[61] One explanation may be that obesity promotes central fat distribution in an android pattern regardless of sex, ethnicity or maturational status.[38] Further research should examine how obesity status diminishes group differences that are otherwise apparent in non-obese children.

Behavior and lifestyle factors

Though the literature is sparse, lifestyle factors including physical activity and nutrition may impact TBF and depot-specific adiposity,[55] and there may be group differences that influence these daily behaviors. In a study of 42 8-year-old children, after controlling for TBF, higher amounts of physical activity measured by accelerometry was related to lower VAT, but not SAT.[80] VAT measured by MRI was inversely related to aerobic fitness measured by peak VO_2 consumption during a treadmill test in 30 male and 22 female adolescents aged ~13.[50] A sedentary lifestyle may promote excess adipose accumulation. For instance, screen-time including watching television or movies predicted an increase in %BF measured by DXA in 661 healthy African American and white 14–18-year-olds.[81]

Nutrition also affects AT accumulation, though the relationships between dietary intake and depot-specific adiposity in children is not well studied.[55] Increased energy intake from protein predicted higher %BF in a sample of white and African American 14–18-year-old adolescents, but increased fat consumption predicted higher %BF in whites only and not in African Americans.[81] Interestingly, once energy intake was controlled for in regression analyses, there was no longer a significant relationship between vigorous physical activity and %BF.[81]

Stress may also be related to abdominal fat.[82,83] Insulin and cortisol may promote lipid growth.[47] A cross-sectional study of 23 female peripubertal Hispanic girls aged 8–11-years-

old demonstrated that school-related life events were related to higher VAT and abdominal SAT for girls who had high cortisol awakening response, but not for those with lower cortisol levels.[82] This stress may derive from environmental factors, particularly socio-economic status, which is inversely related to TBF in children.[84]

CHALLENGES AND FUTURE DIRECTIONS

Childhood and adolescence is a time of rapid growth, and maturational status is difficult to quantify in children and adolescents because pubertal stage, biological development, skeletal growth and somatic growth increase at different rates for different children and vary by chronological age[78] and pubertal stage.[55] Most studies are limited to cross-sectional analyses, and longitudinal analyses of the maturation of AT across the pediatric age range and pubertal stages of development are needed.[55] AT can change quickly and dramatically, and even though the total mass of AT may remain constant, the distribution of the fat may change particularly in puberty.[85] These developmental changes demonstrate the need for whole-body MRI measured by multiple slices so that fat lost at one site may be detected as fat gained at another site.[85]

Owing to the cost and resources required, most imaging studies published to date are typically limited by small sample sizes that are not necessarily representative of populations.[55] Population-based data on depot-specific AT in children and adolescents is virtually nonexistent; thus efforts should be made to include imaging methods in large-scale studies. Further, an urgent research priority is to determine the best anthropometric measurements of total and depot-specific body fat in children, such that better measures can be incorporated into epidemiological studies.[8]

Choosing the best protocols for current imaging instruments remains an urgent need, particularly in the pediatric population. Reliability and validity across imaging measures of adiposity should be established. Errors related to slice gap and number must be overcome, and methodologies need to maximize accuracy while minimizing the burden to the participant. The measurement methodology including instrument and measurement site may alter results, particularly considering group differences in fat patterning. For example, the L4–L5 intervertebral region of the spine is predominantly used in MRI studies to measure abdominal adiposity, but it is unknown if this slice provides the most accurate estimate of abdominal adiposity in children. A single MRI slice may not adequately provide a reference standard to measure total abdominal adiposity, meaning that multiple slices are needed. Ethnic-specific MRI measurement sites of VAT have been recommended in adolescents to best predict total VAT and risk for the metabolic syndrome.[86] Understanding which specific depots and regions of fat yield the most harm is necessary to tailor physical activity and weight-reduction efforts.

Moving beyond the study of IAT and VAT to examine organ-specific fat deposits, such as liver fat and intramyocellular lipid content, warrants further research. Ectopic fat deposition in organs and other tissues may better explain ethnic differences in health risks.[87] Among obese adolescents, African Americans did not have sufficient liver fat to be detected, whereas white and Hispanic adolescents had over twice as much liver fat as considered

normal.[61] Hispanic adolescents had higher liver fat and intramyocellular lipid content than both African American and white adolescents, even at similar weight and age. Moreover, ethnicity contributed 10% of the difference in liver fat and intramyocellular lipid, independent of age, gender or %BF.[61] Interestingly, in a multi-ethnic sample of 118 obese adolescents, at higher levels of VAT adolescents had less %BF and SAT and more hepatic fat, indicating that the excess amount of AT was being deposited in the visceral area, particularly in the liver.[65]

CONCLUSIONS

TBF and depot-specific AT influence cardiometabolic health risks in children and adolescents. Understanding the age, sex, maturational status and ethnic differences in TBF, VAT and SAT can improve prevention and treatment efforts, particularly if health risks are linked to these group differences. Future research should examine longitudinal changes of adiposity over time in representative samples of youth in order to generalize to the entire pediatric population.

Acknowledgments

PTK is supported, in part, by the Louisiana Public Facilities Authority Endowed Chair in Nutrition. AES is funded, in part, by an NIH National Research Service Award.

REFERENCES

1. Krebs NF, Jacobson MS. Academy of Pediatrics Committee on Nutrition. Prevention of pediatric overweight and obesity. Pediatrics. 2003; 112:424–430. [PubMed: 12897303]
2. Reilly JJ, Kelly J. Long-term impact of overweight and obesity in childhood and adolescence on morbidity and premature mortality in adulthood: Systematic review. Int J Obes. 2011; 35:891–898.
3. Shen W, Wang Z, Punyanita M, Lei J, Sinav A, Kral JG, et al. Adipose tissue quantification by imaging methods: a proposed classification. Obesity. 2003; 11:5–16.
4. Snyder, WM.; Cook, MJ.; Nasset, ES.; Karhausen, LR.; Howells, GP.; Tipson, IH. Report of the Task Group on the Reference Man. Oxford: Paergamon Press; 1984.
5. Owens S, Gutin B, Ferguson M, Allison J, Karp W, Le NA. Visceral adipose tissue and cardiovascular risk factors in obese children. J Pediatrics. 1998; 133:41–45.
6. Syme C, Abrahamowicz M, Leonard GT, Perron M, Pitoit A, Qui X, et al. Intra-abdominal adiposity and individual components of the metabolic syndrome in adolescence: sex differences and underlying mechanisms. Arch Pediatrics Adolescent Med. 2008; 162:453–461.
7. Sopher, A.; Shen, W.; Pietrobelli, A. Pediatric body composition methods. In: Heymsfield, SB.; Lohman, TG.; Wang, Z.; Going, SB., editors. Human Body Composition. 2nd edn. Champaign, IL: Human Kinetics; 2005.
8. Benfield LL, Fox KR, Peters DM, Blake H, Rogers I, Grant C, et al. Magnetic resonance imaging of abdominal adiposity in a large cohort of British children. Int J Obes. 2008; 32:91–99.
9. Guo SS, Chumlea WC, Roche AF, Siervogel RM. Age- and maturity-related changes in body composition during adolescence into adulthood: The Fels Longitudinal Study. Int J Obes. 1997; 21:1167–1175.
10. Deans HE, Smith FW, Lloyd DJ, Law AN, Sutherland HW. Fetal fat measurement by magnetic resonance imaging. Br J Radiol. 1989; 62:603–607. [PubMed: 2758247]
11. Harrington TAM, Thomas EL, Modi N, Frost G, Coutts GA, Bell JD. Fast and reproducible method for the direct quantitation of adipose tissue in newborn infants. Lipids. 2002; 37:95–100. [PubMed: 11878317]

12. De Bruin NC, van Velthoven KAM, de Ridder M, Stijnen T, Juttmann RE, Degenhart HJ, et al. Standards for total body fat and fat-free mass in infants. Arch Dis Child. 1996; 74:386–399. [PubMed: 8669953]
13. Ellis KJ, Abrams SA, Wong WW. Body composition of a young, multiethnic female population. Am J Clin Nutr. 1997; 65:724–731. [PubMed: 9062521]
14. Ellis KJ. Body composition of a young, multiethnic male population. Am J Clin Nutr. 1997; 66:1323–1331. [PubMed: 9394682]
15. Butte NF, Hopkinson JM, Wong WW, Smith EO, Ellis KJ. Body composition during the first 2 years of life: an updated reference. Pediatric Research. 2000; 47:578–585. [PubMed: 10813580]
16. Freedman DS, Kettel Khan L, Serdula MK, Srinivasan SR, Berenson GS. BMI rebound, childhood height and obesity among adults: the Bogalusa Heart Study. Int J Obes. 2001; 25:543–549.
17. Huang TTK, Johnson MS, Figueroa-Colon R, Dwyer JH, Goran MI. Growth of visceral fat, subcutaneous abdominal fat, and total body fat in children. Obesity Res. 2001; 9:283–289.
18. Dai S, Labarthe DR, Grunbaum JA, Harrist RB, Mueller WH. Longitudinal analysis of changes in indices of obesity from age 8 years to age 18 years: Project HeartBeat! Am J Epidemiol. 2002; 156:720–729. [PubMed: 12370160]
19. Lloyd T, Chinchilli VM, Eggli DF, Rollings N, Kulin HE. Body composition development of adolescent white females: The Penn State Young Women's Health Study. Arch Pediatr Adolesc Med. 1998; 152:998–1002. [PubMed: 9790610]
20. Rolland-Cachera MF, Cole TJ, Sempe M, Tichet J, Rossignol C, Charraud A. Body mass index variations: Centiles from birth to 87 years. Eur J Clin Nutr. 1991; 45:13–21. [PubMed: 1855495]
21. Laurson KR, Eisenmann JC, Welk GJ. Body fat percentile curves for U.S. children and adolescents. Am J Preventive Med. 2011; 41(Suppl 2):S87–S92.
22. Harrington TAM, Thomas EL, Frost G, Modi N, Bell JD. Distribution of adipose tissue in the newborn. Pediatric Res. 2004; 55:437–441.
23. Ibáñez L, Lopez-Bermejo A, Suárez L, Marcos MV, Díaz M, de Zegher F. Visceral adiposity without overweight in children born small for gestational age. J Clin Endocrinol Metab. 2008; 93:2079–2083. [PubMed: 18334595]
24. Ibáñez L, Suárez L, Lopez-Bermejo A, Díaz M, Valls C, de Zegher F. Early development of visceral fat excess after spontaneou catch-up growth in children with low birth weight. J Clin Endocrinol Metab. 2008; 93:925–928. [PubMed: 18089700]
25. Shen W, Punyanitya M, Silva AM, Chen J, Gallagher D, Sardinha LB, et al. Sexual dimorphism of adipose tissue distribution across the lifespan: a cross-sectional whole-body magnetic resonance imaging study. Nutr Metab. 2009; 6:17.
26. Goran M, Nagy T, Treuth M, Trowbridge C, Dezenberg C, McGloin A, et al. Visceral fat in white and African American prepubertal children. Am J Clin Nutr. 1997; 65:1703–1708. [PubMed: 9174463]
27. Brambilla P, Bedogni G, Moreno LA, Goran MI, Gutin B, Fox KR, et al. Cross-validation of anthropometry against magnetic resonance imaging for the assessment of visceral and subcutaneous adipose tissue in children. Int J Obes. 2006; 30:23–30.
28. Fox KR, Peters DM, Sharpe P, Bell M. Assessment of abdominal fat development in young adolescents using magnetic resonance imaging. Int J Obes Relat Metab Disord. 2000; 24:1653–1659. [PubMed: 11126220]
29. Olhager E, Thomas KA, Wigstrom L, Forsum E. Description and evaluation of a method based on magnetic resonance imaging to estimate adipose tissue volume and total body fat in infants. Pediatr Res. 1998; 44:572–577. [PubMed: 9773848]
30. Herd S, Gower B, Dashti N, Goran M. Body fat, fat distribution and serum lipids, lipoproteins and apolipoproteins in African-American and Caucasian-American prepubertal children. Int J Obes. 2001; 25:198–204.
31. Daniels S, Khoury P, Morrison J. The utility of body mass index as a measure of body fatness in children and adolescents: differences by race and gender. Pediatrics. 1997; 99:804–807. [PubMed: 9164773]

32. Wang H, Story RE, Venners SA, Wang B, Yang J, Li Z, et al. Patterns and interrelationships of body-fat measures among rural Chinese children aged 6 to 18 years. Pediatrics. 2007; e120:e194–e101.
33. Ogle G, Allen J, Humphries I, Lu PW, Briody JN, Morley K, et al. Body-composition assessment by dual-energy X-ray absorptiometry in subjects aged 4–26 y. Am J Clin Nutr. 1995; 61:746–753. [PubMed: 7702015]
34. He Q, Horlick M, Thornton J, Wang J, Pierson RN Jr. Heshka S, et al. Sex and race differences in fat distribution among Asian, African-American, and Caucasian prepubertal children. J Clin Endocrinol Metab. 2002; 87:2164–2170. [PubMed: 11994359]
35. He Q, Horlick M, Thornton J, Wang J, Pierson RN, Heshka S, et al. Sex-specific fat distribution is not linear across pubertal groups in a multiethnic study. Obes Res. 2004; 12:725–733. [PubMed: 15090643]
36. Shaw NJ, Crabtree NJ, Kibirige MS, Fordham JN. Ethnic and gender differences in body fat in British schoolchildren as measured by DXA. Arch Dis Child. 2007; 92:872–875. [PubMed: 17522163]
37. Asayama K, Hayashibe H, Endo A, Okada T, Hara M, Masuda H, et al. Threshold values of visceral fat and waist girth in Japanese obese children. Pediatrics Int. 2005; 47:498–504.
38. Wells JCK. Sexual dimorphism of body composition. Best Practice Res Clin Endo-crinol Metab. 2007; 21:415–430.
39. Ku CY, Gower BA, Nagy TR, Goran MI. Relationship between dietary fat, body fat, and serum lipid profile in prepubertal children. Obes Res. 1998; 6:400–407. [PubMed: 9845229]
40. Maffeis C, Pietrobelli A, Grezzani A, Provera S, Tato L. Waist circumference and cardiovascular risk factors in prepubertal children. Obesity Res. 2001; 9:179–187.
41. Goran MI, Kaskoun MC, Shuman WP. Intra-abdominal adipose tissue in young children. Int J Obes. 1995; 19:279–283.
42. Bray GA, DeLany JP, Harsha DW, Volaufova J, Champagne CC. Evaluation of body fat in fatter and leaner 10-y-old African American and white children: the Baton Rouge Children's Study. Am J Clin Nutr. 2001; 73:687–702. [PubMed: 11273842]
43. Liem ET, Rolfe EDL, L'Abée C, Sauer PJJ, Ong KK, Stolk RP. Measuring abdominal adiposity in 6 to 7-year-old children. Eur J Clin Nutr. 2009; 63:835–841. [PubMed: 19127281]
44. Lee S, Kuk JL, Hannon TS, Arslanian SA. Race and gender differences in the relationships between anthropometrics and abdominal fat in youth. Obesity. 2008; 16:1066–1071. [PubMed: 18356853]
45. Le KA, Ventura EE, Fisher JQ, Davis JN, Weigensberg MJ, Punyanitya M, et al. Ethnic differences in pancreatic fat accumulation and its relationship withother fat depots and inflammatory markers. Diabetes Care. 2011; 34:485–490. [PubMed: 21270204]
46. Loomba-Albrecht LA, Styne DM. Effect of puberty on body composition. Curr Opin Endocrinol Diabetes Obesity. 2009; 16:10–15.
47. Cowell CT, Briody J, Lloyd-Jones S, Smith C, Moore B, Howman-Giles R. Fat distribution in children and adolescents - the influence of sex and hormones. Hormone Res Pediatrics. 1997; 48(Suppl 5):93–100.
48. Taylor RW, Grant AM, Williams SM, Goulding A. Sex differences in regional body fat distribution from pre- to postpuberty. Obesity. 2010; 18:1410–1416. [PubMed: 19893501]
49. Satake E, Nakagawa Y, Kubota A, Saegusa H, Sano S-i Ohzeki T. Age and sex differences in fat distribution in non-obese Japanese children. J Pediatric Endocrinol Metab. 2010; 23:873–878.
50. Winsley R, Armstrong N, Middlebrooke A, Ramos-Ibanez N, Williams C. Aerobic fitness and visceral adipose tissue in children. Acta Paediatrica. 2006; 95:1435–1438. [PubMed: 17062473]
51. Brambilla P, Manzoni P, Agostini G, Beccaria L, Ruotolo G, Sironi S, et al. Persisting obesity starting before puberty is associated with stable intraabdominal fat during adolescence. Int J Obes. 1999; 23:299–303.
52. Kirchengast S, Angelika G. Body composition characteristics during puberty in girls and boys from Eastern Austria. Int J of Anthropology. 2003; 18:139–151.

53. Travers SH, Jeffers BW, Bloch CA, Hill JO, Eckel RH. Gender and Tanner stage differences in body composition and insulin sensitivity in early pubertal children. J Clin Endocr Metab. 1995; 80:172–178. [PubMed: 7829608]
54. Brambilla P, Manzoni P, Sironi S, Del Maschio A, di Natale B, Chiumello G. Peripheral and abdominal adiposity in childhood obesity. Int J Obes Relat Metab Disord. 1994; 18:795–800. [PubMed: 7894517]
55. Suliga E. Visceral adipose tissue in children and adolescents: a review. Nutrition Res Rev. 2009; 22:137–147. [PubMed: 19737436]
56. Liu A, Byrne N, Kagawa M, Ma G, Kijboonchoo K, Nasreddine L, et al. Ethnic differences in body fat distribution among Asian pre-pubertal children: a cross-sectional multicenter study. BMC Public Health. 2011; 11:500. [PubMed: 21703012]
57. Freedman DS, Khan LK, Serdula MK, Ogden CL, Dietz WH. Racial and ethnic differences in secular trends for childhood BMI, weight, and height. Obesity. 2006; 14:301–308. [PubMed: 16571857]
58. Morrison JA, Barton BA, Obarzanek E, Crawford PB, Guo SS, Schreiber GB. Racial differences in the sums of skinfolds and percentage of body fat estimated from impedance in black and white girls, 9 to 19 years of age: The National Heart, Lung, and Blood Institute Growth and Health Study. Obesity Res. 2001; 9:297–305.
59. Yanovski JA, Yanovski SZ, Filmer KM, Hubbard VS, Avila N, Lewis B, et al. Differences in body composition of black and white girls. Am J Clin Nutr. 1996; 64:833–839. [PubMed: 8942404]
60. Lee S, Kim Y, Kuk JL, Boada FE, Arslanian S. Whole-body MRI and ethnic differences in adipose tissue and skeletal muscle distribution in overweight black and white adolescent boys. J Obesity. 2011; 2011:159373.
61. Liska D, Dufour S, Zern TL, Taksali S, Cali AMG, Dziura J, et al. Interethnic differences in muscle, liver and abdominal fat partitioning in obese adolescents. PLoS One. 2007; 2:e569. [PubMed: 17593968]
62. Modi N, Thomas EL, Uthaya SN, Umranikar S, Bell JD, Yajnik C. Whole body magnetic resonance imaging of healthy newborn infants demonstrates increased central adiposity in Asian Indians. Pediatr Res. 2009; 65:584–587. [PubMed: 19190541]
63. Goran M, Gower BA, Treuth MS, Nagy TR. Prediction of intra-abdominal and subcutaneous abdominal adipose tissue in healthy pre-pubertal children. Int J Obes. 1998; 22:549–558.
64. Bacha F, Saad R, Gungor N, Janosky J, Arslanian SA. Obesity, regional fat distribution, and Syndrome X in obese black versus white adolescents: race differential in diabetogenic and atherogenic risk factors. J Clin Endocrinol Metab. 2003; 88:2534–2540. [PubMed: 12788850]
65. Taksali SE, Caprio S, Dziura J, Dufour S, Cali AMG, Goodman TR, et al. High visceral and low abdominal subcutaneous fat stores in the obese adolescent: a determinant of an adverse metabolic phenotype. Diabetes. 2007; 57:367–371. [PubMed: 17977954]
66. Tershakovec AM, Kuppler KM, Zemel BS, Katz L, Weinzimer S, Harty MP, et al. Body composition and metabolic factors in obese children and adolescents. Int J Obes. 2003; 27:19–24.
67. Weiss R, Taksali SE, Dufour S, Yeckel CW, Papademetris X, Cline G, et al. The "obese insulin-sensitive" adolescent: importance of adiponectin and lipid partitioning. J Clin Endocrinol Metabolism. 2005; 90:3731–3737.
68. Cali AM, Caprio S. Ectopic fat deposition and the metabolic syndrome in obese children and adolescents. Hormone Res Pediatrics. 2009; 71(Suppl 1):2–7.
69. Caprio S, Hyman LD, McCarthy S, Lange R, Bronson M, Tamborlane WV. Fat distribution and cardiovascular risk factors in obese adolescent girls: Importance of the intraabdominal fat depot. Am J Clin Nutr. 1996; 64:12–17. [PubMed: 8669407]
70. Nishina M, Kikuchi T, Yamazaki H, Kameda K, Hiura M, Uchiyama M. Relationship among systolic blood pressure, serum insulin and leptin, and visceral fat accumulation in obese children. Hypertens Res. 2003; 26:281–288. [PubMed: 12733695]
71. Owens S, Gutin B, Barbeau P, Litaker M, Allison J, Humphries M, et al. Visceral adipose tissue and markers of the insulin resistance syndrome in obese black and white teenagers. Obesity Res. 2000; 8:287–293.

72. Goran MI, Gower BA. Relation between visceral fat and disease risk in children and adolescents. Am J Clin Nutr. 1999; 70:149S–1565.
73. Goran MI, Bergman RN, Gower BA. Influence of total vs. visceral fat on insulin action and secretion in African American and White children. Obesity Res. 2001; 9:423–431.
74. Comuzzie AG, Higgins PB, Voruganti S, Cole S. Cutting the fat: the genetic dissection of body weight. Prog Mol Biol Transl Sci. 2010; 94:197–212. [PubMed: 21036326]
75. Katzmarzyk PT, Perusse L, Bouchard C. Genetics of abdominal visceral fat levels. Am J Hum Biol. 1999; 11:225–235. [PubMed: 11533946]
76. Katzmarzyk PT, Mahaney MC, Blangero J, Quek JJ, Malina RM. Potential effects of ethnicity in genetic and environmental sources of variability in the stature, mass, and body mass index of children. Hum Biol. 1999; 71:977–987. [PubMed: 10592687]
77. Klimentidis YC, Chen G-B, Lopez-Alarcon M, Harris JJ, Duarte CW, Fernandez JR. Associations of obesity genes with obesity-related outcomes in multiethnic children. Arch Med Res. 2011; 42:509–514. [PubMed: 22051089]
78. Malina RM, Koziel S, Bielicki T. Variation in subcutaneous adipose tissue distribution associated with age, sex, maturation. Am J Hum Biol. 1999; 11:189–200. [PubMed: 11533943]
79. Fox KR, Peters DM, Armstrong N, Sharpe P, Bell M. Abdominal fat deposition in 11-year-old children. Int J Obes Relat Metab Disord. 1993; 17:11–16. [PubMed: 8383635]
80. Saelens BE, Seeley RJ, van Schaick K, Donnelly LF, O'Brien KJ. Visceral abdominal fat is correlated with whole-body fat and physical activity among 8-y-old children at risk of obesity. Am J Clin Nutr. 2007; 85:46–53. [PubMed: 17209176]
81. Stallmann-Jorgensen IS, Gutin B, Hatfield-Laube JL, Humphries MC, Johnson MH, Barbeau P. General and visceral adiposity in black and white adolescents and their relation with reported physical activity and diet. Int J Obes. 2007; 31:622–629.
82. Donoho CJ, Weigensberg MJ, Emken BA, Hsu J-W, Spruijt-Metz D. Stress and abdominal fat: preliminary evidence of moderation by the cortisol awakening response in Hispanic peripubertal girls. Obesity. 2010; 19:946–952. [PubMed: 21127479]
83. Kyrou I, Chrousos GP, Tsigos C. Stress visceral obesity, and metabolic complications. Ann NY Acad Sci. 2006; 1083:77–110. [PubMed: 17148735]
84. Rizzo NS, Ruiz JR, Hurtig-Wennlof A, Kwak L, Sjostrom M. Socioeconomic status and its association with objectively-measured total physical activity and body fat in children: The European Youth Heart Study. Int J Body Composit Res. 2010; 8:93–98.
85. Shen W, Liu H, Punyanitya M, Chen J, Heymsfield SB. Pediatric obesity pheno-typing by magnetic resonance methods. Curr Opin Clin Nutr Metab Care. 2005; 8:595–601. [PubMed: 16205458]
86. Lee S, Kuk JL, Kim Y, Arslanian S. Measurement site of visceral adipose tissue and prediction of metabolic syndrome in youth. Pediatr Diabetes. 2011; 12(pt2):250–257. [PubMed: 21129140]
87. D'Adamo E, Northrup V, Weiss R, Santoro N, Pierpont B, Savoye M, et al. Ethnic differences in lipoprotein subclasses in obese adolescents: importance of liver and intraabdominal fat accretion. Am J Clin Nutr. 2010; 92:500–508. [PubMed: 20573788]
88. Deurenberg P, Deurenberg-Yap M, Foo LF, Schmidt G, Wang J. Differences in body composition between Singapore Chinese, Beijing Chinese and Dutch children. Eur J Clin Nutr. 2003; 57:405–409. [PubMed: 12627175]
89. Gower BA, Nagy TR, Trowbridge CA, Dezenberg C, Goran MI. Fat distribution and insulin response in prepubertal African American and white children. Am J Clin Nutr. 1998; 67:821–827. [PubMed: 9583837]
90. Wong WW, Hergenroeder AC, Stuff JE, Butte NF, Smith EOB, Ellis KJ. Evaluating body fat in girls and female adolescents: advantages and disadvantages of dual-energy X-ray absorptiometry. Am J Clin Nutr. 2002; 76:384–389. [PubMed: 12145011]

Table 1

Sex differences in TBF and depot-specific adiposity in children and adolescents

Reference	*Country*	n *(boys, girls)*	*Age (y)*	*Instrument*	*TBF*	*VAT*	*Abdominal SAT*
47	Australia	169, 166	4–35	DXA	%BF: G>B		
31	US	100, 92	7–17	DXA	NS		
26	US	43, 58	7.7 ± 1.6	CT, DXA	G>B	IAAT: G>B	G>B
39	US	30, 36	4–10	CT, DXA	G>B		G> B
5	US	21, 43	7–11	MRI, DXA	NS	B>G	NS
51	US	8, 8	12.8 ± 1.4	MRI		IAT: NS	
71	US	26, 54	13–16	MRI, DXA	%BF: NS	NS	
42	US	65, 64	10–12	DXA	NS		
30	US	58, 43	AA: 8.3±1.4	CT, DXA	G>B	NS	G>B
			W: 8.6±1.2				
17	US	47, 91	8.1 ± 1.6	CT, DXA	NS	B>G	NS
			3–5 years longitudinal				
11	US	4, 4	<1.5 days	MRI	NS	NS	NS
88	Singapore, Beijing, the Netherlands	75, 75	7–12	DXA	NS		
66	US	15, 21	11.8 ± 0.5	CT, DXA	NS	NS	NS
37	Japan	194, 96	6–15	CT		B>G	NS
32	China	1328, 1165	6–18	DXA	G>B		
8	UK	74, 96	13.4 ±0.4	MRI		IAAT: G>B	G>B
			13.5 ±0.5				
44	US	84, 76	~12–13	CT, DXA	NS	B>G	NS
43	The Netherlands	14, 17	6–7	CT, DXA	NS	NS	NS
25	US	88, 59	5–17	MRI		NS (<12y)	G>B
						B>G (>12y)	
49	Japan	73, 57	6–20	CT		B>G (16–20 years)	G>B (11–20 years)
48	US	518, 491	5–29	DXA	G>B		
45	US	40, 98	13–25	MRI, DXA	%BF: G>B	B>G	NS

Abbreviations: %BF, percent body fat; B>G, boys significantly higher than girls; CT, computed tomography; DXA, dual-energy X-ray absorptiometry; G>B, girls significantly higher than boys; IAT, internal adipose tissue; IAAT, intra-abdominal adipose tissue; MRI, magnetic resonance imaging; NS, no significant difference between sexes; SAT, subcutaneous adipose tissue; TBF, total body fat; VAT, visceral adipose tissue.

Int J Obes (Lond). Author manuscript; available in PMC 2014 August 12.

Table 2

Ethnic differences in TBF and depot-specific adiposity in children and adolescents

Reference	*Country: ethnic groups*	n	*Age (y)*	*Instrument*	*TBF*	*VAT*	*Abdominal SAT*
59	US girls: AA, W	40	7–10	MRI, DXA	NS	W>AA	W>AA
31	US: AA, W	192	7–17	DXA	NS		
13	US girls: AA, Hisp, W	313	3–18	DXA	Hisp>W		
14	US boys: AA, Hisp, W	297	3–18	DXA	Hisp>W Hisp>AA W>AA		
26	US	101	7.7 ± 1.6	CT, DXA	NS	W>AA	W>AA
89	US: AA, W	73	5–10	CT, DXA	NS	NS	NS
39	US: AA, W	66	4–10	CT	NS	IAAT: W>AA	NS
5	US: AA, W	64	7–11	MRI, DXA	AA>W	NS	AA>W
71	US: AA, W	81	13–16	MRI, DXA	%BF: NS	W>AA	
42	US: AA, W	129	10–12	DXA	NS		
73	US: AA, W	119	8–11	CT, DXA	NS	NS	NS
30	US: AA, W	101	AA: 8.3±1.4 W: 8.6±1.2	CT, DXA	NS	W>AA (adj. TBF)	W>AA (adj. TBF)
17	US: AA, W	138	8.1±1.6 Longitudinal	CT, DXA	AA>W	NS	NS
90	US girls: AA, AsA, Hisp, W	141	13.0 ± 1.9	DXA	NS		
64	US: AA, W	50	AA: 13.4±0.3 W: 13.3±0.4	CT, DXA	NS	W>AA	NS
61	US: AA, Hisp, W	55 obese	AA: 14.7±2.73 Hisp: 15.2±2.4 W: 14.6±1.84	MRI, DXA	NS	W and Hisp>AA	NS
44	US: AA, W	160	~12–13	CT, DXA	NS	NS	NS
45	US: AA, Hisp	138	13–25	MRI	%BF: Hisp>AA	Hisp>AA	NS
60	US: AA, W	40	11–18	MRI		W>AA	NS whole body: AA>W

Int J Obes (Lond). Author manuscript; available in PMC 2014 August 12.

Abbreviations: %BF, percent body fat; AA, African American; adj. TBF, analyses were adjusted for total body fat; AsA, Asian American; CT, computed tomography; DXA, dual-energy X-ray absorptiometry; Hisp, Hispanic American, IAAT, intra-abdominal adipose tissue; MRI, magnetic resonance imaging; NS, no significant difference between ethnic groups; SAT, subcutaneous adipose tissue; TBF, total body fat; VAT, visceral adipose tissue; W, white.

European Journal of Sport Science, 2016
Vol. 16, No. 6, 736–746, http://dx.doi.org/10.1080/17461391.2015.1088577

ORIGINAL ARTICLE

Physical fitness normative values for 6–18-year-old Greek boys and girls, using the empirical distribution and the lambda, mu, and sigma statistical method

KONSTANTINOS D. TAMBALIS[1,2], DEMOSTHENES B. PANAGIOTAKOS[1], GLYKERIA PSARRA[1], STELIOS DASKALAKIS[3], STAVROS A. KAVOURAS[1,4], NICKOS GELADAS[2], SAVAS TOKMAKIDIS[5], & LABROS S. SIDOSSIS[1,6]

[1]*Department of Nutrition and Dietetics, Harokopio University, Athens, Greece,* [2]*Department of Physical Education and Sport Science, University of Athens, Athens, Greece,* [3]*Ministry of Education and Religions, Athens, Greece,* [4]*Department of Health, Human Performance, and Recreation, Human Performance Laboratory, University of Arkansas, Fayetteville, AR, USA,* [5]*Department of Physical Education and Sport Science, Dimokritio University, Komotini, Greece,* [6]*Department of Internal Medicine, Shriners Hospital for Children, University of Texas Medical Branch, Galveston, TX, USA*

Abstract
The aim of the this study was to establish age- and gender-specific physical fitness normative values and to compare percentiles and *Z* scores values in a large, nationwide sample of Greek children aged 6–18 years. From March 2014 to May 2014, a total of 424,328 boys and girls aged 6–18 years who attended school in Greece were enrolled. The studied sample was representative, in terms of age–sex distribution and geographical region. Physical fitness tests (i.e. 20 m shuttle run test (SRT), standing long jump, sit and reach, sit-ups, and 10 × 5 m SRT) were performed and used to calculate normative values, using the percentiles of the empirical distributions and the lambda, mu, and sigma statistical method. Normative values were presented as tabulated percentiles for five health-related fitness tests based on a large data set comprising 424,328 test performances. Boys typically scored higher than girls on cardiovascular endurance, muscular strength, muscular endurance, and speed/agility, but lower on flexibility (all p values <0.001). Older boys and girls had better performances than younger ones ($p < 0.001$). Physical fitness tests' performances tended to peak at around the age of 15 years in both sexes. The presented population-based data are the most up-to-date sex- and age-values for the health-related fitness of children and adolescents in Greece and can be used as standard values for fitness screening and surveillance systems and for comparisons among the same health-related fitness scores of children from other countries similar to Greece. Schools need to make efforts to improve the fitness level of the schoolchildren through the physical education curriculum to prevent cardiovascular risk.

Keywords: *Physical fitness, normative values, performance, children*

Introduction

Physical fitness refers to "the ability to carry out daily tasks with vigor and alertness, without undue fatigue and with ample energy to enjoy leisure-time pursuits and to meet unforeseen emergencies" (Caspersen, Powell, & Christenson, 1985, p. 128). Fitness status has long been associated with age, from youth to the middle years of lifespan and even older. However, there is a lack of data in teenagers and younger children. It is crucial to know the fitness level of children as it has been suggested that fitness level in childhood is essential to carry forward favourable behavioural and biological effects into later life (Ortega, Ruiz, Castillo, & Sjöström, 2008). Accumulating epidemiologic evidence reveals that improvement in physical fitness, mainly aerobic fitness, is related to better health in children (Andersen, Wedderkopp, Hansen, Cooper, & Froberg, 2003; Ekelund et al., 2007; Hurtig-Wennlof, Ruiz, Harro, & Sjöström, 2007; Tambalis,

Correspondence: L. S. Sidossis, Department of Internal Medicine, Shriners Hospital for Children, University of Texas Medical Branch, 301 University Blvd, Galveston, TX 77555–0177, USA, E-mail: lasidoss@utmb.edu

Panagiotakos, Psarra, & Sidossis, 2011), in a dose response manner (Anderssen et al., 2007). Moreover, subjects with high physical fitness during adolescence may have lower levels of body fatness as adults (Eisenmann, Wickel, Welk, & Blair, 2005). In contrast, low levels of physical fitness in children have been associated with a number of cardiometabolic risk factors, such as hypertension, hyperlipidaemia, and obesity (Anderssen et al., 2007). To prevent early development of cardiovascular disease risk factors, among other significant interventions (e.g. increased physical activity levels, decreased obesity levels), preventive strategies must incorporate age- and sex-specific physical fitness assessment, even from childhood.

To make useful recommendations as regards to physical fitness, percentiles and *Z* scores have been used to assess children's and adolescent's growth and fitness status (Eisenmann, Laurson, & Welk, 2011). Percentiles are easier to understand and utilize in practice. In contrast, *Z* scores are more complicated and may be of limited use in clinical settings, but are more useful in research. However, it has been suggested that the lambda, mu, and sigma (LMS) method, a technique that uses *Z* scores and is based on smoothed percentile distribution curves over ages, better fits the data than the empirical percentile method (Wang, Moreno, Caballero, & Cole, 2006). These different methodologies may affect the cut-off points for the evaluation of fitness status and may lead to different results.

Numerous data sets on physical fitness levels among children are available worldwide; moreover age-, gender-, as well as region-specific values are essential for a variety of health-risk measurements (i.e. not only for fitness), in order to develop effective public health strategies (Tambalis et al., 2011; Tambalis, Panagiotakos, Arnaoutis, & Sidossis, 2013; Tokmakidis, Kasambalis, & Christodoulos, 2006). In addition, a comparison of both methodologies may prove useful for future researchers who would also like to calculate country-specific normative curves for fitness levels, as it would allow them to select the most appropriate procedure.

Therefore, the aim of the present work was to present the Greek region-specific distribution of age and sex physical fitness test measurements for 6–18-year-old children and adolescents and to evaluate sex- and age-related differences, using both percentiles and *Z* scores values in a nationwide sample of schoolchildren. These distributions (percentiles and Z scores) may also be valid for other similar populations around the world for comparisons among the same health-related fitness scores of children similar to Greece (i.e. Caucasians), and guide preventive strategies to better prevent cardiometabolic disorders in the future.

Methods

Study design and participants

Population-based, representative data were derived from a nationwide school-based survey under the auspices of the Ministry of Education. Specifically, anthropometric, physical activity, nutrition, and physical fitness data along with information on age and sex were collected from March 2014 to May 2014. In total, 424,328 (51% boys and 49% girls) children aged 6–18 years from Greek public and private schools agreed to participate in the study (participation rate was 40% of the total population). The working sample was representative of the entire Greek population (chi-square *p* value as compared with the current sample with the age–sex distribution of all Greek areas = 0.93).

Study approval

Ethical approval for the survey was granted by the Review Board of the Ministry of Education and the Ethical Committee of Harokopio University.

Assessment of fitness status

The Euro-fit physical fitness test battery was used to evaluate children's physical fitness levels; it is a set of nine physical tests covering flexibility, speed, endurance, and strength. The standardized test battery was proposed by the Council of Europe for children of school age and has been used in many European countries schools since the 1980s. Five fitness tests of the Euro-fit Battery, representative of flexibility, explosive strength, speed/agility, and aerobic performance, were administered by trained physical education professionals during the physical education classes. Specifically, measurements were performed by one teacher of physical education in each class. All physical education professionals were instructed through a detailed and extended manual of operations and followed a standardized procedure of measurements in order to minimize the potential inter-rate variability among schools. The physical education teachers were first trained by a school advisor of physical education for accurate anatomical landmarks, subject positioning, and measurement techniques. Verbal informed consent for the child to participate in the measurements was taken from physical education teachers. As the measurements were included in an obligatory school programme, verbal informed consent was considered sufficient.

Briefly, the test assessed: (a) standing long jump (SLJ; jump as far as possible from a standing position at the start) to evaluate lower body explosive power; (b) sit and reach (SR; this test involves sitting on the floor with legs stretched out straight ahead without shoes. The soles of the feet are placed flat against the box. With the palms facing downwards, and the hands on top of each other or side by side, the participant reaches forward along the measuring line as far as possible. The score is recorded to the nearest centimetre as the distance reached by the hand, using 15 cm at the level of the feet) to measure flexibility; (c) sit-ups in 30 s (SUs; lie on the mat with the knees bent at right angles, with the feet flat on the floor and held down by a partner), to measure the endurance of the abdominal and hip-flexor muscles; (d) 10 × 5 m shuttle run test (10 × 5 m SRT; from a standing start), to evaluate speed and agility; and (e) multi-stage 20 m SRT, to estimate aerobic performance. The 20 m SRT test consists of measuring the number of laps completed by participants running up and down between two lines in groups, set 20 m apart, at an initial speed of 8.5 km/h which increases by 0.5 km/h every minute, using a pre-recorded audio tape (Leger, Lambert, Goulet, Rowan, & Dinelle, 1984; Leger, Mercier, Gadoury, & Lambert, 1988). Repeat tests (two trials) were allowed for the SLJ, SR, SU, and 10 × 5 m SRT, with the best performance of each recorded.

Statistical analysis

Descriptive statistics (mean ± standard deviation) for boys and girls were calculated. Comparisons of the physical fitness tests' performances data between boys and girls were performed using the independent samples t test, after testing for equality of variances using the Levene test. Comparisons between percentile values of the physical fitness tests' performance data of both calculation methods were performed using the paired samples t test.

Age- and sex-specific distributions and percentiles were calculated using two methods: through the empirical distribution of the data, the 3th, 10th, 25th, 50th, 75th, 90th, and 97th percentiles were calculated; also, the LMS method was used (Cole & Green, 1992). The LMS method was used in order to smooth the age-dependent skewness usually observed in fitness values. Based on this method, the data were normalized using the Box–Cox power transformation. The Box–Cox λ-power transformation of the variable y_i has the following form:

$$y_i^{(\lambda)} = \frac{y_i^{\lambda} - 1}{\lambda(\mathrm{gm}(y))^{\lambda-1}} \quad \text{if } \lambda \neq 0$$

or

$$y_i^{(\lambda)} = \mathrm{gm}(y_i)\log y_i \quad \text{if } \lambda \neq 0$$

where gm is the geometric mean of y_i. The power transformation was calculated from the raw data and skewness in the distribution where y was removed. The principle idea of the LMS method is to power transform the measurement, i.e. SLJ here, and to use the coefficient of variation (CV = standard deviation/mean) of the raw data. The optimal Box–Cox power λ is the one that gives the lowest CV (Cole & Green, 1992). Thus, the LMS method calculates the best power (L), the best mean (M), and CV (S) in each series of measurements at a specific age. The degrees of freedom used to determine the L, M, and S were chosen on the basis of that which achieved the smallest difference in the penalized deviance (–2 × log{penalised likelihood}) statistic, as well as the Schwarz Bayesian Criterion (both goodness-of-fit measures) between the estimated models. Then, centiles of physical fitness tests were calculated as follows:

$$\mathrm{Centile}(\alpha) = M_{\mathrm{age}} \times (1 + L_{\mathrm{age}} \times S_{\mathrm{age}} \times Z_{\alpha})^{1/L_{\mathrm{age}}}, \quad (1)$$

where Z_{α} is the Z score (i.e. (variable – mean)/standard deviation) corresponding to the required centile (e.g. $Z = -0.67$ gives the 25th centile, $Z = 0$ gives the median M, $Z = 0.67$ gives the 75th centile). The gender-specific cut-offs of fitness tests for each age group were calculated for various centiles by solving the previous equation. Particularly, from Equation (1) we have:

$$\left(\frac{\mathrm{Centile}(\alpha)}{M_{\mathrm{age}}}\right)^{L_{\mathrm{age}}} = 1 + L_{\mathrm{age}} \times S_{\mathrm{age}} \times Z_{\alpha},$$

$$\left(\frac{\mathrm{Centile}(\alpha)}{M_{\mathrm{age}}}\right)^{L_{\mathrm{age}}} - 1 = L_{\mathrm{age}} \times S_{\mathrm{age}} \times Z_{\alpha}.$$

Thus,

$$Z_{\alpha} = \left(\left(\frac{\mathrm{Centile}(\alpha)}{M_{\mathrm{age}}}\right)^{L_{\mathrm{age}}} - 1\right) \times \left(L_{\mathrm{age}} \times S_{\mathrm{age}}\right)^{-1}. \quad (2)$$

The LMS values can be used to calculate Z scores and therefore percentile values by looking up a Z

table, using the following formula:

$$z = \frac{(x/M) - 1}{L \times S},$$

where x is performance, L is the gender- and age-specific L value, M is the gender- and age-specific M value, and S is the gender- and age-specific S value. All statistical analyses were performed using the SPSS program (Release 18; SPSS Inc., Chicago, IL, USA). The LMSchartmaker (Pan & Cole, 2010) and the LMSgrowth (Pan & Cole, 2011) freeware packages were used to calculate L, M, and S values at ages 6–18 based on Greek reference values.

Results

In Tables I and II, normative physical fitness data for 6–18-year-old children in Greece, by gender and age as tabulated critical percentiles and LMS values from 3 to 97 (P_3, P_{10}, P_{25}, P_{50}, P_{75}, P_{90}, P_{97}), are presented. Also presented are the gender- and age-specific LMS values for all fitness tests.

For each of the fitness tests, performance was better in boys compared with girls ($p < 0.001$), except for the SR test ($p < 0.001$). Moreover, older boys and girls had better performances than younger ones ($p < 0.001$). Physical fitness tests' performances also tended to peak at about the age of 15–16 years in both sexes.

In order to investigate potential differences between percentile values from the two methods, comparisons by physical fitness test and gender were performed. Data analysis revealed no significant differences between critical percentiles and LMS percentiles in SR, SLJ, SUs, 20 m SRT, and 10 × 5 m SRT values in either sex (all p values >0.05).

Discussion

The aim of the present work was to develop up-to-date age- and sex-specific physical fitness normative values for Greek children aged 6–18 years and to compare specific percentile values from two widely applied estimation methods: the frequency percentiles and the LMS smoothed percentiles. This study provides information relating to normative values across a range of health-related fitness tests. These values could be used as approximate indicative values for comparisons among the same health-related fitness scores of children (according to National Statistical Services the vast majority of children population in Greece are Caucasians) from other countries similar to Greece: i.e. a developed country with population predominately Caucasian. Moreover, these data can be used as benchmark values for health-related fitness screening and surveillance of children 6–18 years in Greece.

The presented data could be useful in a crude classification of how well children in Greece perform on the specific health-related fitness tests relative to their age- and sex-matched peers and to identify those children with specific physical fitness characteristics that could be considered essential for sporting success. Moreover, our findings facilitate the recognition of children and adolescents with low physical fitness aiming to set appropriate goals in the future and to promote encouraging health behaviours. Previous studies in children from Greece have linked low cardiorespiratory fitness with increased metabolic score risk and inflammation (Christodoulos, Douda, & Tokmakidis, 2012; Flouris, Bouziotas, Christodoulos, & Koutedakis, 2008). According to the presented percentile classification from both methods, and for a practical use of these data, children could be classified as having a performance score: very poor ($<P_{10}$), in the poor quartile (1st), in the good quartiles (2nd–3rd), in the very good quartile (4th), and excellent ($X > P_{90}$) of the distribution. Although this classification is not criterion-referenced, Looney and Plowman (1990) suggested that test scores above the 25th percentile (the poor quartile) in fitness tests should be considered acceptable from a health perspective. On the contrary, low levels of physical fitness have been associated with many serious health problems in childhood, while sufficient levels of physical fitness may have an important cardio-protective role in children. According to Anderssen et al. (2007), there was a strong association between aerobic fitness and the clustering of cardiovascular disease risk factors. Specifically, the odds ratios for clustering in each quartile of fitness, using the quartile with the highest fitness as reference, were 13.0 [95% confidence interval (CI) 8.8–19.1] 4.8 [95% CI 3.2–7.1], and 2.5 [95% CI 1.6–3.8], respectively, after adjusting for several confounding factors (Andersen et al., 2003; Anderssen et al., 2007; Ekelund et al., 2007; Hurtig-Wennlof et al., 2007). Future studies need to examine which specific childhood and/or adolescence thresholds for aerobic fitness are significantly associated with clustered cardiovascular disease risk factors.

The presented results revealed that boys consistently scored higher than girls in almost all fitness tests (with the exception of the SR test of flexibility). Moreover, older boys and girls performed better than younger ones ($p < 0.001$). Our findings are in accordance with recent studies from Latvia (Sauka et al.,

Table I. Boys physical fitness tests percentiles and LMS values and LMS summary statistics by age in 6–18-year-old Greek children and adolescents

		Percentiles							LMS method									
Age	*N*	P_3	P_{10}	P_{25}	P_{50}	P_{75}	P_{90}	P_{97}	P_3	P_{10}	P_{25}	P_{50}	P_{75}	P_{90}	P_{97}	*L*	*M*	*S*
20 m SRT (*completed laps*)																		
6	1706	1.0	4.0	9.0	14.0	22.0	31.0	44.2	2.2	4.9	9.0	14.6	22.1	31.6	43.2	0.38	14.64	0.67
7	15,196	5.0	8.0	12.0	18.0	28.0	38.0	50.0	3.2	6.7	11.8	18.9	28.0	39.4	53.2	0.39	18.86	0.64
8	20,774	6.0	9.0	14.0	23.0	34.0	47.0	59.0	4.2	8.6	14.8	23.1	33.6	46.6	62.1	0.41	23.06	0.61
9	20,149	7.0	11.0	17.0	28.0	41.0	53.0	64.0	5.2	10.4	17.6	27.1	38.9	53.0	69.7	0.44	27.09	0.59
10	19,279	8.0	12.0	20.0	31.0	45.0	58.0	70.0	6.0	12.1	20.5	31.1	44.1	59.3	76.8	0.49	31.13	0.57
11	18,793	9.0	15.0	23.0	36.0	51.0	64.0	78.0	6.6	13.8	23.4	35.4	49.5	65.8	84.1	0.54	35.36	0.55
12	11,456	8.0	15.0	24.0	39.0	54.0	69.0	82.0	6.8	15.4	26.4	39.7	55.1	72.4	91.6	0.60	39.71	0.54
13	15,555	7.0	16.0	28.0	44.0	61.0	76.0	91.0	6.8	16.7	29.3	44.0	60.7	79.0	98.9	0.66	44.04	0.54
14	13,681	7.0	19.0	32.0	50.0	69.0	85.0	100	6.4	17.8	31.8	47.8	65.4	84.5	105	0.73	47.79	0.53
15	10,135	9.0	20.0	34.0	53.0	72.0	90.0	103	5.5	18.3	33.5	50.4	68.6	87.9	108	0.79	50.37	0.52
16	1862	7.0	20.0	36.0	54.0	75.0	90.0	103	4.4	18.4	34.5	51.9	70.2	89.3	109	0.85	51.87	0.52
17	670	7.0	18.0	33.0	50.0	68.0	86.0	100	3.9	18.3	35.1	52.8	71.0	89.7	109	0.91	52.76	0.51
18	305	7.4	13.0	28.0	48.0	66.0	87.0	115	3.2	18.0	35.6	53.5	71.6	89.9	108	0.97	53.48	0.51
SLJ (*cm*)																		
6	17,494	64.0	76.0	90.0	102	115	128	140	59.0	75.0	89.7	103	116	129	141	1.42	103.4	0.19
7	23,994	71.0	85.0	100	113	127	140	151	66.0	83.1	98.7	113	127	141	153	1.42	113.4	0.19
8	23,836	80.0	94.0	108	123	138	150	162	72.8	90.7	107	123	137	151	165	1.43	122.7	0.18
9	22,667	85.0	100	116	131	147	160	172	79.0	97.7	115	131	146	161	175	1.44	131.2	0.18
10	21,886	90.0	108	123	140	155	169	181	84.8	104	122	139	155	171	185	1.44	139.2	0.18
11	21,241	97.0	113	130	147	163	177	190	90.3	111	130	147	164	180	196	1.45	147.3	0.18
12	12,961	100.0	119	135	154	171	187	201	95.9	117	137	156	174	191	207	1.45	156.0	0.17
13	8683	105.0	125	144	164	183	200	217	101	124	145	165	184	202	219	1.46	165.0	0.17
14	7231	112.0	133	153	175	195	212	228	106	130	152	173	193	212	230	1.48	173.3	0.17
15	2633	118	140	160	183	204	220	240	110	135	159	180	201	220	239	1.51	180.3	0.17
16	7729	122	145	167	188	209	225	240	112	139	163	186	207	226	245	1.55	185.6	0.17
17	2737	120	146	167	188	210	228	245	113	142	167	190	211	231	250	1.60	189.7	0.17
18	375	110	137	160	187	222	244	250	114	144	169	193	215	235	254	1.66	193.0	0.17
SR (*cm*)																		
6	21,448	2.0	5.0	10.0	15.0	18.0	22.0	25.0	2.1	5.7	9.6	13.7	18.0	22.4	27.0	0.87	13.70	0.46
7	27,509	2.0	5.0	9.0	15.0	18.0	22.0	25.0	1.8	5.3	9.2	13.4	17.9	22.5	27.3	0.84	13.44	0.48
8	27,220	2.0	4.0	9.0	14.0	18.0	22.0	26.0	1.6	5.0	8.9	13.1	17.7	22.5	27.5	0.81	13.14	0.51
9	26,692	2.0	4.0	8.0	14.0	18.0	22.0	25.5	1.4	4.6	8.5	12.8	17.4	22.4	27.6	0.78	12.78	0.53
10	25,017	2.0	4.0	8.0	13.0	17.0	21.5	26.0	1.3	4.3	8.1	12.4	17.1	22.2	27.5	0.76	12.44	0.54
11	24,418	1.0	3.0	7.0	13.0	17.0	21.0	26.0	1.2	4.2	8.0	12.3	17.1	22.2	27.7	0.74	12.29	0.56
12	15,636	1.0	3.0	7.0	13.0	17.0	22.0	27.0	1.2	4.2	8.0	12.5	17.4	22.8	28.5	0.72	12.47	0.57
13	14,613	1.0	3.0	8.0	14.0	18.0	23.0	28.0	1.3	4.4	8.4	13.0	18.2	23.8	29.9	0.71	12.99	0.57
14	12,871	2.0	4.0	9.0	15.0	20.0	25.0	30.0	1.4	4.7	8.9	13.8	19.3	25.3	31.7	0.71	13.80	0.57
15	9537	2.0	5.0	10.0	16.0	21.0	27.0	31.0	1.6	5.1	9.5	14.8	20.6	27.0	33.8	0.71	14.77	0.56
16	7291	2.0	5.0	11.0	16.5	22.0	28.0	32.0	1.8	5.5	10.2	15.8	21.9	28.7	35.9	0.70	15.76	0.56
17	2573	2.0	5.0	11.0	17.0	22.0	28.0	33.0	2.0	5.9	10.9	16.7	23.2	30.3	37.8	0.70	16.72	0.55
18	352	1.0	5.0	10.0	16.0	22.0	29.0	33.0	2.2	6.4	11.6	17.7	24.4	31.8	39.7	0.70	17.68	0.55

SUs (no. in 30 sec)																		
6	17,381	2.0	6.0	10.0	13.0	16.0	19.0	23.0	1.0	5.7	9.8	13.5	16.8	19.9	22.8	1.34	13.46	0.38
7	23,866	4.0	8.0	12.0	15.0	18.0	21.0	25.0	2.0	7.3	11.5	15.3	18.8	22.1	25.3	1.29	15.27	0.36
8	23,782	6.0	10.0	14.0	17.0	20.0	23.0	27.0	4.0	8.9	13.1	16.9	20.6	24.1	27.5	1.25	16.94	0.33
9	22,608	7.0	12.0	15.0	18.0	22.0	25.0	29.0	5.8	10.3	14.5	18.4	22.1	25.8	29.3	1.21	18.40	0.31
10	21,770	8.0	13.0	16.0	20.0	23.0	26.0	30.0	7.3	11.7	15.7	19.6	23.4	27.0	30.6	1.17	19.62	0.29
11	21,159	10.0	14.0	17.0	21.0	24.0	27.0	31.0	8.6	12.8	16.8	20.6	24.4	28.1	31.7	1.14	20.62	0.28
12	12,941	10.0	15.0	18.0	21.0	25.0	28.0	32.0	9.7	13.7	17.6	21.5	25.2	29.0	32.7	1.09	21.47	0.27
13	15,555	12.0	15.0	19.0	22.0	26.0	29.0	33.0	10.6	14.5	18.4	22.2	26.0	29.8	33.6	1.03	22.20	0.26
14	13,681	12.0	16.0	19.0	23.0	27.0	30.0	34.0	11.3	15.1	18.9	22.8	26.6	30.5	34.4	0.97	22.76	0.25
15	10,135	13.0	17.0	20.0	23.0	27.0	30.0	35.0	11.8	15.5	19.3	23.1	27.0	31.0	35.0	0.91	23.11	0.25
16	7688	13.0	16.0	19.0	23.0	27.0	30.0	34.0	12.1	15.7	19.4	23.3	27.3	31.3	35.5	0.84	23.31	0.25
17	2753	12.0	16.0	19.0	23.0	27.0	30.0	35.0	12.3	15.8	19.5	23.4	27.4	31.6	35.9	0.77	23.40	0.25
18	367	10.0	15.0	18.0	22.0	26.0	30.0	36.0	12.4	15.8	19.5	23.4	27.5	31.7	36.2	0.70	23.40	0.26
10 × 5 m SRT (sec)																		
6	21,133	32.3	28.4	26.1	24.4	22.3	21.0	19.5	31.6	29.0	26.5	24.0	21.4	18.9	16.3	1.0	24.04	0.15
7	27,340	30.0	27.0	25.0	23.0	21.3	20.0	18.6	30.6	28.2	25.7	23.3	20.8	18.3	15.9	1.0	23.32	0.15
8	23,231	29.0	26.0	23.9	22.1	20.5	19.2	18.0	29.7	27.4	25.0	22.6	20.2	17.8	15.4	1.0	22.62	0.15
9	22,143	27.9	25.0	23.1	21.5	20.0	18.7	17.4	28.9	26.6	24.2	22.0	19.6	17.3	15.0	1.0	21.96	0.16
10	25,561	27.0	24.3	22.6	21.0	19.4	18.1	17.0	28.1	25.9	23.6	21.4	19.1	16.8	14.6	1.0	21.37	0.16
11	29,745	26.1	23.8	22.00	20.4	19.0	17.7	16.4	27.5	25.3	23.1	20.9	18.6	16.4	14.2	1.0	20.86	0.16
12	12,695	26.0	23.4	21.8	20.1	18.7	17.3	16.1	27.0	24.8	22.6	20.4	18.3	16.1	14.0	1.0	20.45	0.16
13	15,168	26.3	23.2	21.3	19.8	18.3	16.9	15.7	26.6	24.4	22.3	20.1	18.0	15.9	13.7	1.0	20.13	0.16
14	13,244	26.0	23.0	21.0	19.3	17.8	16.5	15.3	26.3	24.2	22.1	19.9	17.8	15.7	13.6	1.0	19.91	0.16
15	9528	26.5	23.0	21.0	19.2	17.8	16.4	15.0	26.1	24.0	21.9	19.8	17.7	15.6	13.5	1.0	19.77	0.17
16	7146	27.8	23.5	21.1	19.3	17.9	16.4	15.0	26.1	24.0	21.9	19.7	17.7	15.6	13.5	1.0	19.68	0.17
17	2487	29.4	23.9	21.2	19.4	17.8	16.4	5.0	26.0	23.9	21.8	19.6	17.6	15.6	13.5	1.0	19.64	0.17
18	345	31.0	25.8	22.1	19.9	18.2	16.7	15.0	26.0	23.9	21.8	19.6	17.6	15.6	13.5	1.0	19.61	0.17

Note: *P*, percentile; *L*, skew; *M*, median; *S*, coefficient of variation; age: completed age, e.g. 6 years = 6.00–6.99 years.

Table II. Girls physical fitness tests percentiles and LMS values and LMS summary statistics by age in 6–18-year-old Greek children and adolescents

		Percentiles							LMS method									
Age	*N*	P_3	P_{10}	P_{25}	P_{50}	P_{75}	P_{90}	P_{97}	P_3	P_{10}	P_{25}	P_{50}	P_{75}	P_{90}	P_{97}	*L*	*M*	*S*
20 m SRT (*completed laps*)																		
6	1640	2.0	4.0	8.0	12.0	18.0	26.0	38.0	3.5	5.7	8.8	13.0	18.6	25.9	35.1	0.25	13.02	0.56
7	14,777	5.0	8.0	11.0	15.0	21.0	29.0	39.0	4.1	6.8	10.5	15.4	21.9	30.2	40.6	0.28	15.45	0.55
8	20,084	7.0	9.0	12.0	18.0	25.0	34.0	45.0	4.9	8.0	12.4	18.1	25.6	35.0	46.5	0.30	18.14	0.54
9	19,806	7.0	10.0	14.0	20.0	29.0	40.0	51.0	5.6	9.2	14.3	20.9	29.3	39.8	52.6	0.32	20.87	0.54
10	19,143	8.0	11.0	16.0	23.0	33.0	43.0	55.0	6.1	10.3	16.0	23.4	32.8	44.3	58.2	0.35	23.42	0.54
11	18,290	9.0	12.0	18.0	26.0	36.0	48.0	61.0	6.4	11.1	17.3	25.4	35.6	48.0	62.8	0.37	25.44	0.54
12	10,693	7.0	12.0	18.0	26.0	37.0	49.0	62.0	6.4	11.3	17.9	26.5	37.2	50.1	65.4	0.40	26.49	0.54
13	6767	6.0	12.0	18.0	26.0	37.0	50.0	62.0	6.0	11.0	17.9	26.7	37.6	50.7	66.2	0.42	26.71	0.55
14	5563	5.0	12.0	18.0	26.0	37.0	50.0	63.0	5.5	10.6	17.5	26.4	37.4	50.5	65.9	0.45	26.40	0.56
15	2576	7.0	12.0	17.0	25.0	35.0	47.0	62.0	4.9	9.9	16.9	25.8	36.8	49.9	65.1	0.47	25.82	0.58
16	1462	7.0	11.0	16.0	24.0	33.0	45.0	60.0	4.2	9.2	16.2	25.2	36.2	49.2	64.3	0.49	25.21	0.60
17	490	5.0	10.0	16.0	23.0	36.0	50.0	67.0	3.5	8.5	15.6	24.6	35.7	48.7	63.6	0.52	24.63	0.61
18	240	3.0	8.0	14.0	23.0	28.0	39.5	63.7	2.8	7.8	14.8	23.9	34.9	47.8	62.5	0.54	23.92	0.63
SLJ (*cm*)																		
6	17,152	59.7	70.0	80.0	93.0	105	116	128	55.3	68.4	81.0	93.2	105	117	128	1.22	93.2	0.19
7	23,303	65.0	78.0	90.0	101	115	126	140	61.5	75.6	89.1	102	115	128	140	1.19	102.3	0.19
8	23,099	73.0	85.0	97.0	110	125	138	150	67.7	82.6	97.1	111	125	139	152	1.16	111.3	0.19
9	22,427	80.0	92.5	105	120	134	148	160	73.6	89.3	105	120	135	149	164	1.13	119.8	0.19
10	21,589	85.0	99.0	111	127	142	156	170	78.5	95.0	111	127	143	158	174	1.10	127.1	0.19
11	20,678	90.0	103	118	134	150	165	180	82.1	99.1	116	133	149	165	182	1.08	132.6	0.19
12	12,171	90.0	105	120	136	153	170	185	83.8	101	119	136	153	170	187	1.07	135.8	0.19
13	7795	90.0	104	120	137	154	170	187	84.0	102	119	137	154	172	189	1.06	137.0	0.19
14	6454	89.0	103	120	135	154	170	188	83.1	101	119	137	155	172	190	1.05	136.9	0.19
15	2058	86.0	100	116	134	151	170	186	81.6	100	118	136	154	172	190	1.05	136.3	0.20
16	7677	86.0	100	115	132	150	170	187	79.9	98.7	117	136	154	172	190	1.06	135.7	0.20
17	2637	83.0	100	116	133	152	170	190	78.1	97.3	116	135	154	172	190	1.07	135.0	0.21
18	240	81.0	99.0	112	130	150	169	197	76.2	95.8	115	134	153	172	190	1.08	134.3	0.21
SR (*cm*)																		
6	20,463	3.0	7.0	12.0	16.0	20.0	24.0	28.0	3.3	7.4	11.7	16.0	20.5	25.0	29.6	0.93	16.04	0.41
7	26,851	3.0	7.0	12.0	16.0	21.0	25.0	29.0	3.1	7.3	11.6	16.1	20.7	25.4	30.1	0.92	16.10	0.42
8	25,849	3.0	7.0	12.0	16.0	21.0	25.0	30.0	2.9	7.1	11.6	16.2	20.9	25.7	30.7	0.91	16.17	0.43
9	25,623	2.0	6.0	12.0	16.0	21.0	26.0	30.0	2.8	7.0	11.6	16.3	21.1	26.1	31.2	0.90	16.27	0.44
10	24,913	2.0	6.0	11.0	17.0	22.0	26.0	30.0	2.8	7.0	11.7	16.5	21.6	26.7	32.0	0.89	16.54	0.45
11	24,156	2.0	7.0	12.0	17.0	23.0	28.0	32.0	2.8	7.3	12.1	17.1	22.4	27.8	33.3	0.88	17.11	0.45
12	15,228	3.0	7.0	13.0	18.0	24.0	29.0	34.0	3.0	7.6	12.6	17.9	23.4	29.0	34.8	0.87	17.88	0.45
13	14,014	3.0	8.0	14.0	19.0	25.0	30.0	35.0	3.2	8.0	13.2	18.6	24.3	30.2	36.2	0.87	18.63	0.45
14	12,543	3.0	9.0	15.0	20.0	26.0	30.0	36.0	3.4	8.3	13.6	19.3	25.1	31.1	37.3	0.87	19.25	0.45
15	9135	3.0	9.0	15.0	20.0	26.0	31.0	36.0	3.6	8.6	14.0	19.7	25.6	31.6	37.8	0.88	19.68	0.44
16	7539	4.0	9.0	15.0	20.0	26.0	31.0	36.0	3.9	8.9	14.3	19.9	25.8	31.8	38.0	0.88	19.93	0.43
17	2597	4.0	10.0	15.0	20.0	26.0	30.0	35.0	4.0	9.1	14.5	20.1	25.9	31.9	38.0	0.89	20.09	0.43
18	240	3.0	8.0	12.0	18.0	25.0	29.0	31.0	4.2	9.3	14.6	20.2	26.0	31.9	37.9	0.89	20.23	0.42

SUs (no. in 30 sec)																		
6	17,044	2.0	5.0	9.0	12.0	15.0	18.0	22.0	2.0	5.0	9.5	13.2	16.7	19.9	22.9	1.36	13.24	0.41
7	23,158	4.0	8.0	11.0	14.0	18.0	21.0	24.0	2.1	6.5	10.9	14.9	18.5	21.9	25.1	1.33	14.86	0.38
8	23,030	5.0	9.0	13.0	16.0	19.0	22.0	26.0	2.5	7.9	12.4	16.4	20.1	23.6	27.0	1.29	16.35	0.35
9	22,368	6.0	10.0	14.0	17.0	21.0	24.0	28.0	4.4	9.3	13.6	17.6	21.4	25.0	28.4	1.26	17.61	0.33
10	21,511	8.0	12.0	15.0	18.0	22.0	25.0	28.0	5.9	10.5	14.7	18.6	22.3	25.9	29.4	1.22	18.58	0.31
11	20,612	9.0	13.0	16.0	19.0	22.0	25.0	30.0	7.2	11.4	15.4	19.2	22.9	26.5	30.0	1.17	19.21	0.29
12	15,324	10.0	13.0	16.0	19.0	22.0	26.0	30.0	8.0	12.0	15.8	19.5	23.1	26.7	30.2	1.12	19.52	0.28
13	14,181	10.0	13.0	16.0	19.0	22.0	26.0	30.0	8.5	12.3	16.0	19.6	23.1	26.7	30.2	1.07	19.57	0.28
14	12,637	10.0	13.0	16.0	19.0	23.0	26.0	30.0	8.8	12.3	15.9	19.4	23.0	26.5	30.0	1.01	19.42	0.27
15	9264	10.0	13.0	16.0	19.0	22.0	26.0	30.0	8.8	12.2	15.6	19.1	22.6	26.1	29.6	0.96	19.09	0.27
16	7686	9.0	12.0	15.0	18.0	21.0	25.0	29.0	8.6	11.9	15.3	18.7	22.1	25.6	29.2	0.91	18.66	0.28
17	2634	9.0	12.0	15.0	18.0	21.0	25.0	28.0	8.4	11.6	14.8	18.2	21.6	25.1	28.7	0.87	18.18	0.28
18	242	7.0	10.0	14.0	17.0	20.0	24.0	28.0	8.2	11.2	14.4	17.7	21.0	24.5	28.1	0.82	17.65	0.28
10 × 5 m SRT (sec)																		
6	20,030	32.0	29.0	26.9	24.9	23.2	21.8	20.2	32.6	29.9	27.4	25.0	22.7	20.6	18.5	0.42	24.99	0.14
7	26,271	30.3	27.8	25.8	24.0	22.2	20.9	19.2	31.4	28.9	26.4	24.1	21.9	19.8	17.9	0.42	24.10	0.14
8	25,450	29.7	26.7	24.8	23.0	21.4	20.0	18.5	30.3	27.9	25.5	23.3	21.2	19.2	17.3	0.42	23.27	0.14
9	25,394	28.3	25.8	24.0	22.3	20.7	19.4	18.0	29.4	27.0	24.7	22.6	20.5	18.6	16.7	0.42	22.56	0.14
10	24,749	27.8	25.2	23.4	21.8	20.2	18.9	17.5	28.7	26.3	24.1	22.0	20.0	18.1	16.3	0.42	22.01	0.14
11	23,862	27.1	24.6	22.9	21.2	19.7	18.4	17.0	28.2	25.9	23.7	21.6	19.7	17.8	16.1	0.42	21.64	0.14
12	14,920	27.0	24.6	22.9	21.2	19.6	18.3	17.0	27.9	25.7	23.5	21.4	19.5	17.7	15.9	0.42	21.44	0.14
13	13,569	27.8	24.9	23.0	21.3	19.7	18.2	16.9	27.9	25.6	23.4	21.4	19.4	17.6	15.9	0.42	21.39	0.14
14	12,088	28.0	25.0	23.0	21.2	19.6	18.2	17.0	28.0	25.7	23.5	21.4	19.5	17.7	15.9	0.42	21.45	0.14
15	8575	28.3	25.4	23.4	21.5	19.9	18.3	16.8	28.1	25.8	23.6	21.5	19.6	17.7	16.0	0.42	21.56	0.14
16	7004	30.0	26.1	24.0	22.0	20.0	18.4	16.8	28.3	26.0	23.8	21.7	19.7	17.9	16.1	0.42	21.70	0.14
17	2392	30.1	26.1	24.0	22.0	20.0	18.5	16.8	28.5	26.2	23.9	21.8	19.9	18.0	16.2	0.42	21.85	0.14
18	204	32.0	27.7	24.7	22.4	20.2	18.7	16.5	28.7	26.3	24.1	22.0	20.0	18.1	16.3	0.42	22.02	0.14

Note: *P*, percentile; *L*, skew; *M*, median; *S*, coefficient of variation; age: completed age, e.g. 6 years = 6.00–6.99 years.

2011), Portugal (Santos et al., 2014), and Australia (Catley & Tomkinson, 2013) that have examined similar physical fitness tests in children aged 6–18 years. In all the aforementioned studies, boys performed better than girls in cardiorespiratory endurance, speed/agility, muscular strength, and muscular endurance tests, while older ages, in both sexes, have incorporated higher percentile values in comparison with younger ones (Catley & Tomkinson, 2013; Santos et al., 2014; Sauka et al., 2011). Moreover, it seems that physical fitness test performance tends to peak from about the age of 15 years, especially in girls. This finding is in accordance with results from a previous large European epidemiological study (the HELENA study) which found stability in girls' performance in aerobic fitness, speed/agility, and flexibility tests after about the age of 15 years (Ortega et al., 2011).

A finding that deserves attention is the lack of significant differences between the two methods used to create the percentiles in 20 m SRT, SLJ, SR, SUs, and 10 × 5 m SRT, in both sexes. The LMS method uses smoothed curves to estimate the critical centiles and has been extensively been used to provide a way of obtaining normalised growth centiles standards (Cole, 1990). To the best of our knowledge, this is one of the few times that the LMS method has been used to calculate fitness centiles in children and adolescents (Bustamante, Beunen, & Maia, 2012; Catley & Tomkinson, 2013; Eisenmann et al., 2011; Gulías-González, Sánchez-López, Olivas-Bravo, Solera-Martínez, & Martínez-Vizcaíno, 2014; Santos et al., 2014; Silva, Aires, Mota, Oliveira, & Ribeiro, 2012). None of the above-mentioned studies have made comparisons between two methods. In our study, for at first time we have made comparison between the LMS method and the empirical method which showed that the results can be trusted, at least in large data analysis.

School in Greece, as in most developed countries, is the first institution providing opportunities for physical activity through the pursuit of the physical education curriculum (Cavill, Kahlmeier, & Racioppi, 2006). Given that our study was based on a yearly nationwide school-based survey programme under the auspices of the Ministry of Education, our findings may be of assistance to physical fitness educators and policy-makers when assessing schools' physical education programmes. In fact, it is an opportunity for a significant intervention focusing on elementary and secondary school children physical fitness through the relatively easy access of schools, where changes in physical fitness can be easier implemented and monitored, and affect almost all schoolchildren.

Strengths and limitations

The present study has several strengths. The sample is representative of the national gender and geographical representation, as we studied almost 40% of the total population of 6–18-year-old children in Greece. The study was performed in children aged 6–18 years. This age range is an advantageous period in the life of children and adolescents at which to apply effective prevention strategies to improve physical fitness levels. Finally, the presented data were derived using the same standardized procedures in all schools.

There are also limitations in our study design. Although a common, validated protocol was used to evaluate fitness tests in all schools, a large number of experienced, professional physical educators participated as evaluators in the study. In order to minimize the variability among the different experimenters, all educators were instructed through a detailed and extended manual of operations and followed a standardized procedure of measurements. Even so, some variability in measurement will still exist. Moreover, the 20 m SRT is an objective test of aerobic fitness which indirectly infers peak VO_2 and does not directly measure aerobic fitness. In order to fully understand the development of VO_{2max} during adolescence, longitudinal studies are required to take into consideration the tempo and timing of growth and maturation, particularly in girls (Eisenmann, Laurson, & Welk, 2011). Finally, the cross-sectional design of our study cannot provide causal relationships, but only provides hypotheses for further research.

Conclusions

In conclusion, we established sex- and age-specific normative physical fitness values for children aged 6–18 years living in Greece. Boys performed better in all measurements except flexibility than girls of the same age and older children performed better than younger ones. These findings may help policy-makers to design appropriate health-related educational and physical fitness programmes for the young, to facilitate future more detailed epidemiology research on this topic and for comparisons among the same health-related fitness scores of children from other countries similar to Greece.

Acknowledgements

The authors would like to thank the physical education professionals who organized the projects and

performed the evaluations and the students for agreeing to participate. Also, we are grateful to Mrs Sarah Toombs Smith, Ph.D., for assistance with manuscript preparation.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Graduate Program, Department of Nutrition and Dietetics of Harokopio University, the Hellenic Ministry of Education and Religions, the Hellenic Ministry of Culture and Athletics, the Secretariat General of Sports, and OPAP SA.

References

Andersen, L. B., Wedderkopp, N., Hansen, H. S., Cooper, A. R., & Froberg, K. (2003). Biological cardiovascular risk factors cluster in Danish children and adolescents: The European youth heart study. *Preventive Medicine*, *37*, 363–367. doi:10.1016/S0091-7435(03)00145-2

Anderssen, S. A., Cooper, A. R., Riddoch, C., Sardinha, L. B., Harro, M., Brage, S., & Andersen, L. B. (2007). Low cardiorespiratory fitness is a strong predictor for clustering of cardiovascular disease risk factors in children independent of country, age and sex. *European Journal of Cardiovascular Prevention & Rehabilitation*, *14*, 526–531. doi:10.1097/HJR.0b013e328011efc1

Bustamante, A., Beunen, G., & Maia, J. (2012). Evaluation of physical fitness levels in children and adolescents: Establishing percentile charts for the central region of Peru. *Revista Peruana de Medicina Experimental y Salud Publica*, *29*, 188–97.

Caspersen, C. J., Powell, K. E., & Christenson, G. M. (1985). Physical activity, exercise, and physical fitness: Definitions and distinctions for health-related research. *Public Health Representative*, *100*, 126–131.

Catley, M. J., & Tomkinson, G. R. (2013). Normative health-related fitness values for children: Analysis of 85347 test results on 9–17-year-old Australians since 1985. *British Journal of Sports Medicine*, *47*, 98–108. doi:10.1136/bjsports-2011-090218

Cavill, N., Kahlmeier, S., & Racioppi, F. (2006). What can the health sector and others do to increase physical activity? In N. Cavill, S. Kahlmeier, & F. Racioppi (Eds.), *Physical activity and health in Europe: Evidence for action* (pp. 15–23). Copenhagen: World Health Organization Regional Office for Europe.

Christodoulos, A. D., Douda, H. T., & Tokmakidis, S. P. (2012). Cardiorespiratory fitness, metabolic risk, and inflammation in children. *International Journal of Pediatrics*, *12*, Article no. 270515. doi:10.1155/2012/270515

Cole, T. J. (1990) The LMS method for constructing normalized growth standards. *European Journal of Clinical Nutrition*, *44*, 45–60. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2354692

Cole, T. J., & Green, P. J. (1992). Smoothing reference centile curves: The LMS method and penalized likelihood. *Statistics in Medicine*, *11*, 1305–1319. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1518992

Eisenmann, J. C., Laurson, K. R., & Welk, G. J. (2011). Aerobic fitness percentiles for U.S. adolescents. *American Journal of Preventive Medicine*, *41*, S106–S110. doi:10.1016/j.amepre.2011.07.005

Eisenmann, J. C., Wickel, E. E., Welk, G. J., & Blair, S. N. (2005). Relationship between adolescent fitness and fatness and cardiovascular disease risk factors in adulthood: The aerobics center longitudinal study (ACLS). *American Heart Journal*, *149*, 46–53. doi:10.1016/j.ahj.2004.07.016

Ekelund, U., Anderssen, S. A., Froberg, K., Sardinha, L. B., Andersen, L. B., Brage, S., & European Youth Heart Study Group. (2007). Independent associations of physical activity and cardiorespiratory fitness with metabolic risk factors in children: The European youth heart study. Diabetologia, *50*, 1832–1840. doi:10.1016/S0140-6736(06)69075-2

Flouris, A. D., Bouziotas, C., Christodoulos, A. D., & Koutedakis, Y. (2008). Longitudinal preventive-screening cutoffs for metabolic syndrome in adolescents. *International Journal of Obesity*, *32*, 1506–1512. doi:10.1038/ijo.2008.142

Gulías-González, R., Sánchez-López, M., Olivas-Bravo, Á., Solera-Martínez, M., & Martínez-Vizcaíno, V. (2014). Physical fitness in Spanish schoolchildren aged 6–12 years: Reference values of the battery EUROFIT and associated cardiovascular risk. *Journal of School Health*, *84*, 625–635. doi:10.1111/josh.12192

Hurtig-Wennlof, A., Ruiz, J. R., Harro, M., & Sjöström, M. (2007). Cardiorespiratory fitness relates more strongly than physical activity to cardiovascular disease risk factors in healthy children and adolescents: The European youth heart study. *European Journal of Cardiovascular Prevention & Rehabilitation*, *14*, 575–581. doi:10.1097/HJR.0b013e32808c67e3

Leger, L. A., Lambert, J., Goulet, A., Rowan, C., & Dinelle, Y. (1984). Aerobic capacity of 6 to 17-year-old Quebecois–20 meter shuttle run test with 1 minute stages. *Canadian Journal of Applied Sport Sciences*, *9*, 64–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6733834.

Leger, L. A., Mercier, D., Gadoury, C., & Lambert, J. (1988). The multistage 20 metre shuttle run test for aerobic fitness. *Journal of Sports Sciences*, *6*, 93–101. doi:10.1080/02640418808729800

Looney, M. A., & Plowman, S. A. (1990). Passing rates of American children and youth on the FITNESSGRAM criterion-referenced physical fitness standards. *Research Quarterly for Exercise and Sport*, *61*, 215–223. doi:10.1080/02701367.1990.10608682

Ortega, F. B., Artero, E. G., Ruiz, J. R., España-Romero, V., Vicente-Rodríguez, G., Molnar, D., … Gutiérrez, A., HELENA Study Group. (2011). Physical fitness levels among European adolescents: The HELENA study. *British Journal of Sports Medicine*, *45*, 20–29. doi:10.1136/bjsm.2009.062679

Ortega, F. B., Ruiz, J. R., Castillo, M. J., & Sjöström, M. (2008). Physical fitness in childhood and adolescence: A powerful marker of health. *International Journal of Obesity*, *32*(1), 1–11. doi:10.1038/sj.ijo.0803774

Pan, H., & Cole, T. J. (2010). LMS Chartmaker, a program to construct growth references using the LMS method. Version 2.43. Retrieved from http://www.healthforallchildren.co.uk/

Pan, H., & Cole, T. J. (2011). LMS growth, a Microsoft Excel add-in to access growth references based on the LMS method. Version 2.71. Retrieved from http://www.healthforallchildren.co.uk/

Santos, R., Mota, J., Santos, D. A., Silva, A. M., Baptista, F., & Sardinha, L. B. (2014). Physical fitness percentiles for Portuguese children and adolescents aged 10–18 years. *Journal of Sports Sciences*, *13*, 1–3. doi:10.1080/02640414.2014.906046

Sauka, M., Priedite, I. S., Artjuhova, L., Larins, V., Selga, G., Dahlström, O., & Timpka, T. (2011). Physical fitness in northern European youth: Reference values from the Latvian physical health in youth study. *Scandinavian Journal of Public Health*, *39*, 35–43. doi:10.1177/1403494810380298

Silva, G., Aires, L., Mota, J., Oliveira, J., & Ribeiro, J. C. (2012). Normative and criterion-related standards for shuttle run performance in youth. Pediatric *Exercise Science*, *24*, 157–69.

Tambalis, K. D., Panagiotakos, D. B., Arnaoutis, G., & Sidossis, L. A. (2013). Endurance, explosive power, and muscle strength in relation to body mass index and physical fitness in Greek children aged 7–10 years. Pediatric *Exercise Science*, *25*, 394–406. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/23877385

Tambalis, K. D., Panagiotakos, D. B., Psarra, G., & Sidossis, L. A. (2011). Inverse but independent trends in obesity and fitness levels among Greek children: A time-series analysis from 1997 to 2007. *Obesity Facts*, *4*, 165–174. doi:10.1159/000327994

Tokmakidis, S. P., Kasambalis, A., & Christodoulos, A. D. (2006). Fitness levels of Greek primary schoolchildren in relationship to overweight and obesity. *European Journal of Pediatrics*, *165*, 867–874. doi:10.1007/s00431-006-0176-2

Wang, Y., Moreno, L. A., Caballero, B., & Cole, T. J. (2006). Limitations of the current world health organization growth references for children and adolescents. *Food and Nutrition Bulletin*, *27*, S175–S188. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/17361655

Routledge
Taylor & Francis Group

Journal of Sports Sciences

ISSN: 0264-0414 (Print) 1466-447X (Online) Journal homepage: https://www.tandfonline.com/loi/rjsp20

Vertical jumping and leg power normative data for English school children aged 10–15 years

Matthew J. D. Taylor , Daniel Cohen , Christine Voss & Gavin R. H. Sandercock

To cite this article: Matthew J. D. Taylor , Daniel Cohen , Christine Voss & Gavin R. H. Sandercock (2010) Vertical jumping and leg power normative data for English school children aged 10–15 years, Journal of Sports Sciences, 28:8, 867-872, DOI: 10.1080/02640411003770212

To link to this article: https://doi.org/10.1080/02640411003770212

Journal of Sports Sciences, June 2010; 28(8): 867–872

Routledge
Taylor & Francis Group

Vertical jumping and leg power normative data for English school children aged 10–15 years

MATTHEW J. D. TAYLOR[1], DANIEL COHEN[2], CHRISTINE VOSS[1], & GAVIN R. H. SANDERCOCK[1]

[1]*Centre for Sports and Exercise Science, School of Biological Sciences, University of Essex, Colchester and* [2]*School of Health and Human Sciences, London Metropolitan University, London, UK*

(Accepted 11 March 2010)

Abstract
Although vertical jumping is often incorporated into physical activity tests for both adults and children, normative data for children and adolescents are lacking in the literature. The objectives of this study were to provide normative data of jump height and predicted peak leg power for males and females aged 10.0–15.9 years. Altogether, 1845 children from 12 state primary and secondary schools in the East of England participated in the study. Each child performed two countermovement jumps, and jump height was calculated using a NewTest jump mat. The highest jump was used for analysis and in the calculation of predicted peak power. Jump height and predicted peak leg power were significantly higher for males than females from the age of 11 years. Jump height and peak power increased significantly year on year for males. For females, jump height and predicted peak leg power reached a plateau after age 12 and 13 years respectively. This study provides normative data that can be used as a tool to classify jumping performance in children aged 10–15 years.

Keywords: *Biomechanics, jumping, children*

Introduction

The vertical jump test is a component of test batteries used to assess physical ability (Payne, Gledhill, Katzmarzyk, Jamnik, & Keir, 2000). Jumping performance has been incorporated into performance prediction equations in young rugby players (Pienaar, Spamer, & Steyn, 1998), and to distinguish between elite and non-elite performers (Bencke et al., 2002). The vertical jump test is a simple method to calculate peak leg power using prediction equations based on jump height and body mass.

To evaluate any performance variable within a population, either criterion performance levels or normative reference data are needed. Normative data for vertical jumping and leg power exist for young adults (Patterson & Peterson, 2004; Payne et al., 2000). For children, however, only normative data from the Republic of Seychelles have been published (Bovet, Auguste, & Burdette, 2007).

Raw measures of jump height may have some use in performance appraisal, but ideally an estimate of peak leg power should accompany jump height normative data, as this provides insight beyond the outcome of the jump itself. For example, for a given jump height, a heavier performer will generate more power than a light performer. Power differences due to body mass may have implications for screening for sports (helping to identify talented athletes) that require high generation of power, and provide additional information to that given by jump height alone. A common and convenient method to calculate power values is via prediction equations such as those of Sayers and colleagues (Sayers, Harackiewicz, Harman, Frykman, & Rosenstein, 1999). To date, predicted peak leg power for jumping has not been reported for children.

The aims of the present study were to generate normative vertical jump height and predicted peak power data for 10- to 15-year-olds and to investigate between-sex and age group differences in these measures. Such data would be useful in the identification of talented children and when carrying out general physical capacity tests that incorporate vertical jump height and peak power.

Correspondence: M. J. D. Taylor, Centre for Sports and Exercise Science, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, UK.
E-mail: mtaylor@essex.ac.uk

ISSN 0264-0414 print/ISSN 1466-447X online © 2010 Taylor & Francis
DOI: 10.1080/02640411003770212

Methods

The study was carried out in accordance with the Helsinki Declaration of 1975. After receiving university ethics review board approval for the study, children were recruited from 12 state primary and secondary schools in the East of England. Consent for children to be tested was obtained from parents and guardians after they were sent a letter detailing the procedures of the study. Parents were asked to advise the researchers of any known illness or condition that could result in adverse responses to heavy exertion. Such children were excluded from the study. In total, 1845 children participated in the study. Data collection was carried out over a 3-month period by the same researcher. The mass and stature of each child (without shoes and in shorts and t-shirts) was recorded before jumping, using Seca flat scales and a stadiometer, to the nearest 0.1 kg and 0.5 cm, respectively (Table I).

Data collection

A demonstration of how to jump was provided to each child and then each child was allowed to practise the jump until they met the jump criteria, which mostly took two jumps. The jump was a countermovement jump with the use of arms. The jump began from a standing position, with plantigrade foot and the leg vertically aligned (i.e. knee angle approximately 180°). On instruction, the countermovement was performed and the knees flexed to approximately 90° before rapid extension and take-off. Landing (initial contact with the jump mat) was with toes initially and the knee angle extended at approximately 180°. If these criteria were not met, the jump was performed again. A NewTest Timing mat (NewTest Ltd.) was used to measure flight time, which in turn was used to calculate jump height. Children wore sports footwear for jumping. Each child was allowed two jumps using the correct technique and the best jump height was recorded and incorporated into the power prediction equation. The prediction equation (equation 1) of Sayers et al. (1999) was used to predict peak leg power (P_{peak}):

$$P_{peak}(\mathrm{W}) = 60.7 \times (\text{jump height [cm]}) + 45.3 \times (\text{body mass [kg]}) - 2055 \quad (1)$$

To date, no prediction equation has been validated for children, thus the Sayers equation was used, which incorporates jump height and the participant's mass. This equation, developed from jumps performed on a force plate, has a reported difference in adults of 2.7% with power calculated from the force plate (Sayers et al., 1999). The Sayers equation is an improvement on the Lewis formula (Fox & Mathews, 1974), which has been reported to underestimate predicted peak power by 70% (Harman, Rosenstein, Frykman, Rosenstein, & Kraemer, 1991). It has also been recommended as a replacement for the Lewis formula for physical assessment appraisals (Payne et al., 2000).

Table I. Participant characteristics (mean ± *s*).

Age (years)	*N*	Stature (cm)	Mass (kg)
Boys			
10	110	141.9 ± 7.3	37.8 ± 8.9
11	233	147.2 ± 7.2	41.9 ± 10.2
12	200	154.0 ± 8.5	46.5 ± 10.5
13	177	160.9 ± 9.2	52.4 ± 12.1
14	160	167.4 ± 8.4	57.4 ± 14.5
15	106	173.7 ± 7.1	67.3 ± 12.9
Girls			
10	88	141.8 ± 7.4	37.9 ± 8.7
11	263	149.5 ± 7.3	43.8 ± 9.2
12	193	154.2 ± 7.8	47.5 ± 9.9
13	133	160.3 ± 7.5	53.0 ± 10.7
14	103	161.4 ± 6.2	55.9 ± 11.0
15	79	161.3 ± 6.8	56.9 ± 8.1

Statistical analysis

Two-way analysis of variance (ANOVA) was used to identify the presence of an age × sex interaction for both jump height and peak power. Data were then split by sex and a one-way ANOVA with *post hoc* tests (Tukey) was used to identify differences between age groups within sexes. Finally, independent *t*-tests were used to examine differences for jump height, peak power, and mass between the sexes within each age group. Statistical significance was set at $P < 0.05$. Percentiles for vertical jumping and peak power were calculated for both boys and girls. All statistical analyses were carried out using SPSS for Windows version 14.0.

Results

Body mass was significantly different between 10-year-old boys and girls ($P < 0.05$) and between 15-year-old boys and girls ($P < 0.001$). The 95% confidence interval for skewness and kurtosis (boys and girls) showed a normal distribution for jump height and peak power. The two-way ANOVA showed a significant age × sex interaction for both jump height ($F_{5,1832} = 22.48$, $P < 0.001$) and peak power ($F_{5,1832} = 31.11$, $P < 0.001$). One way-ANOVA revealed a significant main effect for age in jump height (jump height, $F_{5,980} = 110.5$, $P < 0.001$) and peak power ($F_{5,980} = 275.92$, $P < 0.001$) for boys. In girls, there were less marked but statistically significant main effects for jump height ($F_{5,852} = 20.74$,

$P < 0.001$) and peak power ($F_{5,852} = 100.30$, $P < 0.001$). *Post hoc* analyses within sexes showed yearly increases in jump height for boys from 10 to 14 years of age (Table II), and increases in peak power for each year measured (Table II). In girls, significant increases in jump height were observed from 10 to 12 years. Furthermore, 15-year-old girls only jumped significantly higher ($P < 0.001$) than 10- and 11-year-old girls. For girls, peak power increased yearly from 10 to 13 years before reaching a plateau.

Table II. Jump height and peak power (mean ± *s*).

Age(years)	Maximum jump height (cm)	Difference		P_{peak}, (W)
		D (cm)	%	
Boys				
10	21.9 ± 5.2			986 ± 483
11	26.9 ± 5.4**#	5.0	23	1473 ± 482#
12	29.8 ± 6.0**#	2.9	11	1862 ± 537*#
13	32.4 ± 6.5**#	2.6	9	2287 ± 612**#
14	35.4 ± 7.2**#	3.2	10	2751 ± 707**#
15	37.1 ± 6.0**	1.5	4	3242 ± 584**#
Girls				
10	21.7 ± 4.7			987 ± 421
11	25.2 ± 5.0#	3.6	17	1459 ± 429#
12	26.9 ± 4.9#	1.7	7	1727 ± 489#
13	27.1 ± 5.0	0.2	0	1985 ± 08#
14	26.9 ± 5.2	0.2	0	2122 ± 483
15	28.7 ± 6.7	1.5	6	2261 ± 517

Note: *D* = between-year differences (10–11, 11–12, 12–13, 13–14, 14–15 years) in absolute scores and % = relative change in jump height.
#Jump height and peak power significantly higher ($P < 0.05$) than for preceding age group.
Significant difference between boys and girls within the same age group: *$P < 0.05$, **$P < 0.001$.

Independent *t*-tests showed that boys jumped significantly higher than girls from 11 years of age onwards (Table II), and that peak power was greater in boys from 12 years of age (Table II). Percentiles for jump height and predicted peak power are presented in Table III.

Discussion

This study provides normative data for vertical jump height and predicted peak leg power for 10- to 15-year-old English school children that should be useful when identifying talent and carrying out general physical capacity tests that incorporate vertical jump height. The ability to generate power is important in sports performance and the inclusion of predicted peak leg power provides an extra tool with which to identify potentially talented children.

Jump height

The data reported here are similar to the normative jump data reported by Bovet et al. (2007) for a sample of Seychelles children. Unfortunately, the type of jump and method of recording jump height were not reported by Bovet et al. (2007). The results for boys in the present study are also comparable to those reported by Temfemo and colleagues (Temfemo, Hudues, Chardon, Mandengue, & Ahmaidi, 2009). Holm and colleagues (Holm, Fredriksen, Fosdahl, & Vøllestad, 2008) reported mean jump height for 11-year-old boys and girls ($n = 30$) of 39 cm and 42 cm respectively. This is approximately 13 cm greater than in the current study, a difference that may be partly due to the method of measuring jump height: flight time versus a tape attached to the child's waist. Jump height

Table III. Vertical jump height (cm) and predicted peak power (W) percentiles of males and females by age.

Percentile	Males						Females					
	10	11	12	13	14	15	10	11	12	13	14	15
Jump height												
95th	30	36	39	43	49	47	28	33	36	36	36	40
90th	29	34	37	40	44	44	27	32	33	34	34	39
75th	25	30	33	37	39	42	25	28	29	29	30	31
50th	21	27	30	32	36	37	22	25	27	26	28	28
25th	18	23	26	28	30	34	18	21	24	24	23	24
10th	16	20	23	23	26	29	15	19	21	21	21	21
Peak power												
95th	1815	2185	2914	3402	3744	4308	1834	2174	2616	2837	2903	3096
90th	1625	2046	2571	2947	3583	3918	1499	2037	2501	2537	2725	2927
75th	1205	1722	2162	2634	3247	3594	1184	1738	2055	2373	2383	2662
50th	915	1456	1787	2258	2698	3185	938	1425	1677	1954	2054	2223
25th	662	1178	1490	1910	2267	2863	698	1173	1349	1592	1815	1831
10th	474	931	1226	1496	1875	2438	525	972	1148	1372	1597	1654

using the countermovement jump was greater on average (~4 cm and ~3 cm for boys and girls respectively) than that reported by Dore and colleagues (Dore, Bedu, & Van Praagh, 2009) for children using the squat jump. The countermovement jump may allow the muscles to build a high active state and force before shortening (Bobbert, Gerritsen, Litjens, & Van Soest, 1996), thus resulting in a higher jump.

The percentile tables of the present study can be used as an aid to distinguish between or to identify talented children. For example, elite 12-year-old male gymnasts' mean jump height has been reported as 30 cm versus only 25 cm for their non-elite counterparts (Bencke et al., 2002). These values correspond to the 50th and 25th percentile of the present data, respectively.

Between-group differences in jump height by sex and age

Boys jumped significantly higher than girls from 11 years onwards. This difference between the sexes has been noted previously in children (Bovet et al., 2007; Temfemo et al., 2009) and adults (Patterson & Peterson, 2004). Jump height continued to improve year on year for boys, and past research has shown this improvement to continue up to the age of 19 years (Branta, Haubenstricker, & Seefeldt, 1984). In females, however, jump height reached a plateau after 12 years of age, which is in line with the research of both Klausen and colleagues (Klausen, Schibye, & Ramusen, 1989) and Dore et al. (2009), who reported similar findings for 13- to 15-year-olds. Strength has also been reported to increase with age from 7 to 11 years (with no difference between the sexes); after 11 years differences begin to emerge, with greater increases in strength among boys than girls (Holm et al., 2008).

The plateau in jump height is in line with the findings of Malina and Bouchard (1991), who reported an increase in standing long jump performance for girls up to the age of 12 years, followed by a plateau or even a decline in performance. Loko and colleagues (Loko, Aule, Ereline, & Viru, 2000) reported vertical jump data for 10- to 17-year-old Estonian girls, which showed a significant reduction in vertical jump height between 13 and 14 years of age. The vertical jump heights reported for Estonian females by Loko et al. (2000) are above the 95th percentile for all age groups in the present study. This is in line with the ~90th percentile performance Estonian children were found to show in the standing broad jump when compared with data from 21 other European countries. The greatest differences in height and mass in Estonian girls were observed between 12 and 13 years (Loko et al., 2000), whereas in the present study they were between 10 and 11 years of age for girls. This may be indicative of the earlier onset of puberty in English females (Whincup, Gilg, Odoki, Taylor, & Cook, 2001). Such differences in body maturation, mass, and jump performance demonstrate a weakness in using jump height alone as an index of performance. For this reason, we also analysed predicted peak leg power.

Between-group differences in predicted peak leg power by sex and age

Ferretti et al. (1994) calculated peak power for 8- to 13-year-old boys ($n = 13$) using a force plate, based on the greatest instantaneous power of four consecutive jumps. Mean peak power was reported to be 1103 ± 393 W, which is within the 10–11 age range for the current study. Mean peak power for 15-year-olds in the present study corresponds to the 11–20th and 21–30th percentiles for boys and girls respectively reported by Payne et al. (2000) for their 15–19 age group. The difference between the present findings and those of Payne et al. (2000) may be due to the inclusion of the other ages, since jump height has been shown to increase up to the age of 19 years (Branta et al., 1984). Also, the age composition of these groups was not reported and therefore it is difficult to make a direct comparison with our data.

From 12 years of age onwards, peak power was significantly greater in boys than girls. This is similar to the pattern for children during cycling (Martin et al., 2004). Peak power increased significantly year on year for boys across the age range measured, whereas it only increased in females between 11 and 13 years. Body mass and jump height are incorporated into the Sayers equation to estimate peak power, thus the significantly greater peak powers in boys than girls were solely a function of increases in jump height observed year on year in males between 12 and 14 years, as body mass was not significantly different between the sexes at these ages. Differences in peak power between boys and girls were most marked at age 15 years. The higher peak power for boys at this age was due to a combination of a significantly greater body mass and greater jump height. For girls, vertical jump height did not differ between ages 12 and 13 years. There was, however, a significant increase in peak power, showing that girls' power development continued to increase at this time due to an ability to vertically displace a greater body mass (12-year-olds: 47.5 ± 9.9 kg; 13-year-olds: 53.0 ± 10.7 kg) over a similar distance (12-year-olds: 26.9 ± 4.9 cm; 13-year-olds: 27.1 ± 5.0 cm). This highlights a weakness in solely reporting jump height as an outcome measure.

Vertical jumping *per se* is useful in only a limited number of sports and it should be noted that, in all cases, jump height itself will be mediated by the

participant's body mass. The production of muscular power, however, is likely to be a better indicator of performance potential in all sports that require rapid horizontal and vertical displacement. The peak power percentiles provide a tool to help identify talented children in sports that require movement of body mass or the manipulation of an external mass.

Limitations

Contact mats are regularly used in the ''field'' to calculate jump height, as they are relatively cheap and portable. The NewTest jump mat has been reported by its manufactures to have an error of $\pm$ 0.001 s for flight time and $\pm$ 2.0 mm for jump height. However, Enoksen and colleagues (Enoksen, Tønnessen, & Shalfawi, 2009) reported the NewTest jump mat on average to overestimate jump height by 2.8 cm compared with a force plate using the impulse–momentum relationship. Contact mats use flight time to calculate jump height, thus caution should be given to this method as it requires a number of assumptions that might not hold (Hatze, 1998), and jump heights may differ depending on how they are calculated. It is therefore recommended that when comparing jump height data the system and method of calculation used should be taken into account, and the same system should be used for pre- and post-tests. A number of jumps are used to assess performance, including countermovement jump with arms, countermovement jump without arms, squat jumps, and depth jumps. The data reported in this study are for countermovement jump with arms and so should only be used only when investigating this type of jump. The jump method used by Sayers et al. (1999), which is similar to the Sargent jump (Sargent, 1921), also used the arms to facilitate jumping. The use of arms can increase jump height by 23% in children (Holm et al., 2008). Normative data for the other jumping techniques have yet to be developed for children. Furthermore, repeatability studies for children's vertical jumps have yet to be performed.

It should be noted that the peak power results were calculated from the Sayers equation, which is based on adults aged approximately 21 years, since no such equation has been validated for children. The Sayers equation has been reported to over-predict power by 2.7% compared with that calculated using a force plate. The Keir nomogram (Keir, Jamnik, & Gledhill, 2003) uses the Sayers equation and shows that for low body mass ($<$45 kg) the prediction equation may result in increased error. This may be of particular importance in children, especially for the low body mass of the 10- to 11-year-olds in the present study. Future work should focus on developing prediction equations for children based on the work of Sayers et al. (1999). Children should be tested on a force plate to calculate power from force and velocity. Jump height and body mass would than be incorporated to develop a new set of age-dependent regression equations for children.

Conclusions

These are the first normative data for vertical jump height in 10- to 15-year-old English children using the methodology described here. The present study is the first to report predicted peak leg power in this age group. By combining body mass and vertical jump height, the predicted power development values provided here may be used to help classify and monitor vertical jump performance in children of this age. This may help to identify children who can produce large lower-body power. These data should be used in the context of the methodologies adopted here, including the NewTest jump mat and flight time calculations for jump height, and the Sayers et al. (1999) peak power prediction equations.

References

Bencke, J., Damsgaard, R., Saekmose, A., Jorgensen, P., Jorgensen, K., & Klausen, K. (2002). Anaerobic power and muscle strength characteristics of 11 year old elite and non-elite boys and girls from gymnastics, team handball, tennis and swimming. *Scandinavian Journal of Medicine and Science in Sports*, *12*, 171–178.

Bobbert, M. F., Gerritsen, K. G., Litjens, M. C., & Van Soest, A. J. (1996). Why is countermovement jump height greater than squat jump height? *Medicine and Science in Sports and Exercise*, *28*, 1402–1412.

Bovet, P., Auguste, R., & Burdette, H. (2007). Strong inverse association between physical fitness and overweight in adolescents: A large school-based survey. *International Journal of Behavioral Nutrition and Physical Activity*, *4*: 24.

Branta, C., Haubenstricker, J., & Seefeldt, V. (1984). Age changes in motor skills during childhood and adolescence. *Exercise and Sport Science Reviews*, *12*, 467–520.

Dore, E., Bedu, M., & Van Praagh, E. (2009). Squat jump performance during growth in both sexes: Comparison with cycling power. *Research Quarterly for Exercise and Sport*, *79*, 517–524.

Enoksen, E., Tønnessen, E., & Shalfawi, S. (2009). Validity and reliability of the *NewTest* Powertimer 3000-series® testing system. *Journal of Sports Sciences*, *27*, 77–84.

Ferretti, G., Nairici, M. V., Binzoni, T., Garoid, L., Le Bas, J. F., Reutenauer, H., & Cerretelli, P. (1994). Determinants of peak muscle power: Effects of age and physical conditioning. *European Journal of Applied Physiology and Occupational Physiology*, *68*, 111–115.

Fox, E. L., & Mathews, D. K. (1974). *Interval training: Conditioning for sports and general fitness* (pp. 257–258). Philadelphia, PA: W. B. Saunders.

Harman, E. A., Rosenstein, M. T., Frykman, P. N., Rosenstein, R. M., & Kraemer, W. J. (1991). Estimation of human power output from vertical jumping. *Journal of Applied Sport Science Research*, *5*, 116–120.

Hatze, H. (1998). Validity and reliability of methods for testing vertical jumping performance. *Journal of Applied Biomechanics*, *14*, 127–140.

Holm, I., Fredriksen, P. M., Fosdahl, M., & Vøllestad, N. (2008). A normative sample of isotonic and isokinetic muscle strength measurements in children 7 to 12 years of age. *Acta Paediatrica*, *97*, 602–607.

Keir, P. J., Jamnik, V. K., & Gledhill, N. (2003). Technical-methodological report: A nonogram for peak leg power output in the vertical jump. *Journal of Strength and Conditioning Research*, *17*, 701–703.

Klausen, K., Schibye, B., & Ramusen, A. (1989). A longitudinal study of changes in physical performance of 10- to 15-year-old girls and boys. In S. S. Oseid & K.-H. Carlsen (Eds.), *Children and exercise XIII* (pp. 112–123). Champaign, IL: Human Kinetics.

Loko, J., Aule, R., Ereline, J., & Viru, A. (2000). Motor performance status in 10 to 17-year-old Estonian girls. *Scandinavian Journal of Medicine and Science in Sports*, *10*, 109–113.

Malina, R., & Bouchard, C. (1991). *Growth, maturation and physical activity*. Champaign, IL: Human Kinetics.

Martin, R. J., Dore, E., Twisk, J., van Praagh, E., Hautier, C. A., & Bedu, M. (2004). Longitudinal changes of maximal short-term peak power in girls and boys during growth. *Medicine and Science in Sports Exercise*, *36*, 498–503.

Patterson, D. D., & Peterson, D. F. (2004). Vertical jump and leg power norms for young adults. *Measurement in Physical Education and Exercise Science*, *81*, 33–41.

Payne, N., Gledhill, N., Katzmarzyk, P. T., Jamnik, V. K., & Keir, P. J. (2000). Canadian Musculoskeletal fitness norms. *Canadian Journal of Applied Physiology*, *25*, 430–442.

Pienaar, A. E., Spamer, M. J., & Steyn, H. S. (1998). Identifying and developing rugby talent among 10-year-old boys: A practical model. *Journal of Sports Sciences*, *16*, 691–699.

Sargent, D. A. (1921). The physical test of a man. *American Physical Education Review*, *26*, 188–194.

Sayers, S. P., Harackiewicz, D. V., Harman, E. A., Frykman, P. N., & Rosenstein, M. T. (1999). Cross-validation of three jump power equations. *Medicine and Science in Sports and Exercise*, *31*, 572–577.

Temfemo, A., Hugues, J., Chardon, K., Mandengue, S.-H., & Ahmaidi, S. (2009). Relationship between vertical jump performance and anthropometric characteristics during growth in boys and girls. *European Journal of Paediatrics*, *168*, 457–464.

Whincup, P. H., Gilg, J. A., Odoki, K., Taylor, J., & Cook, D. G. (2001). Age of menarche in contemporary British teenagers: Survey of girls born between 1982 and 1986. *British Medical Journal*, *322*, 1095–1096.

International Journal of Obesity (1997) 21, 1082–1084

Short Communication

Gender differences in body fat content are present well before puberty

RW Taylor,[1] E Gold[2], P Manning[2] and A Goulding[2]

[1]Department of Human Nutrition and [2]Department of Medicine, Otago University, Dunedin, New Zealand.

To determine whether gender differences in body fat could be detected in prepubertal children using dual energy X-ray absorptiometry (DEXA), body composition was measured in 20 healthy boys aged 3–8 y matched for age, height and weight with 20 healthy girls. Although boys and girls did not differ in age, height, weight, body mass index (BMI) or bone mineral content, the boys had a lower percentage of body fat (13.5 ± 5.1 *vs* 20.4 ± 6.1%, $P < 0.01$), a lower fat mass (3.2 ± 2.0 *vs* 4.9 ± 3.1 kg, $P < 0.01$), and a higher bone-free lean tissue mass (18.6 ± 4.3 *vs* 17.0 ± 3.5 kg, $P < 0.01$) than the girls. Girls had approximately 50% more body fat than the boys. This is the first DEXA study to show that boys aged 3–8 y have less body fat than girls of similar age, height and weight. Thus, this technology demonstrates that significant gender differences in body composition are evident, well before the onset of puberty.

Keywords: adiposity; body composition; gender; prepubertal children

Introduction

Adult men are taller, heavier, have a significantly lower percentage of body fat, but a greater lean tissue mass and a higher bone mineral content than women. These gender differences in body fat content are usually said to arise in adolescence. Prepubertally, sexual dimorphism in body composition is considered to be slight[1] and before puberty boys and girls have similar average heights and weights.[2] However, the precise estimation of body fat content in young children has until recently been difficult. This situation has changed with the advent of dual energy X-ray absorptiometry (DEXA),[3–5] which provides precise and simple independent measurements of total body fat, lean and bone mineral content in young children.[6–8] Using this methodology, several recent papers have noted lower body fat in boys than girls aged 9–10 y.[9,10] The present study was undertaken to test the hypothesis, that DEXA measurements of total body fat content would reveal that, prior to the onset of puberty, even younger boys have a smaller proportion of body fat than girls.

Methods

Participants were healthy volunteers recruited for studies of nutrition and bone health approved by our hospital Ethics Committee. Their body mass index (BMI) values lay between the 5–95 percentiles of USA reference data.[11] All were graded as prepubertal (Tanner Stage I).[12] None were taking medication affecting bone or body composition and no child had previous fractures. Each child was weighed (electronic scales) and measured (stadiometer) in light clothing without shoes. A medical history was taken and a total body DEXA scan was performed (Lunar DPX-L Lunar Corporation, Wisconsin, USA) to measure total fat mass, bone-free lean tissue mass and bone mineral content. Our DEXA precision (coefficient of variation) *in vivo* is: fat mass 2.52%; bone-free lean tissue mass 1.11% and bone mineral content 1.52%. The total percentage fat in the body determined from DEXA measurements was calculated as fat mass (kg) divided by DEXA mass [that is bone free lean tissue mass (kg) plus bone mineral content (kg) plus fat mass (kg)] × 100/1.

The DEXA scan data of a group of healthy boys aged 3–8 y, who were enrolled consecutively in a study of bone development, were matched with scan data from a group of healthy girls collected simultaneously for our DEXA bone density data base. The scan results of each boy were compared with the scan results from the girl who most closely matched them in age, height and body weight. Gender matches obtained in this way were close (Table 1). Results are expressed as means ± s.d. Ninety-five percent confidence intervals for the differences between each pair were calculated (Mean ± s.e.m.). Gender differences were tested (Student's paired *t*-test).

Correspondence: Dr Ailsa Goulding, Department of Medicine, Otago University, PO Box 913, Dunedin, New Zealand.
Received 10 June 97; accepted 6 August 97

Table 1 Body composition characteristics of children ($n = 20$ matched pairs).

Gender		Age (y)	Height (m)	Weight (kg)	BMI (kg/m^2)	Total fat (kg)	Body fat (%)	Lean tissue mass (kg)	Total body BMC (kg)
Girls	Mean ± s.d.	6.4 ± 1.7	1.19 ± 0.12	23.5 ± 6.1	16.3 ± 1.7	4.9 ± 3.1*	20.4 ± 6.1*	17.0 ± 3.5*	0.78 ± 0.23
Boys	Mean ± s.d.	6.4 ± 1.6	1.19 ± 0.12	23.6 ± 6.2	16.4 ± 1.4	3.2 ± 2.0	13.5 ± 5.1	18.6 ± 4.3	0.80 ± 0.25
Girls	Range	3.9–8.9	1.01–1.37	15.6–39.0	13.9–20.8	2.5–14.8	12.1–37.2	11.5–23.7	0.49–1.23
Boys	Range	4.0–8.9	0.99–1.42	15.0–39.8	14.5–19.7	1.5–8.2	7.5–28.6	12.2–29.0	0.44–1.34
	Mean Difference[a]	0.04	0.0	−0.02	−0.1	1.7	6.9	−1.5	−0.02
	95% CI	−0.08–0.16	−0.01–0.02	−0.6–0.5	−0.6–0.4	0.7–2.6	4.2–9.5	−2.3–−0.8	−0.05–0.02

Analysed by paired *t*-test. *$P < 0.01$. [a]Girls relative to boys.

Results

Although the 20 pairs of prepubertal girls and boys in this study were similar in age, height, weight, BMI and bone mineral content (BMC), boys had significantly less fat, a lower % body fat and a higher bone-free lean tissue mass than the girls (Table 1 and Figure 1). Standard deviation scores[13] for height and weight respectively, were: 1.06 ± 0.65 and 0.85 ± 0.92 in girls and 0.77 ± 0.84 and 0.88 ± 1.05 in boys. The average fat mass of the girls (4.9 kg), expressed as a percentage of the average fat mass of the boys, was 152% (3.2 kg, $P < 0.05$) while the bone-free lean tissue mass was 9% lower ($P < 0.05$) and the BMC content 2% lower in girls than in boys (not statistically significant $P > 0.05$).

Discussion

This appears to be the first gender comparison of body fat content in tightly age-matched children aged 3–8 y measured by DEXA. Our results provide strong support for the view that boys have a lower fat content than girls, even before puberty, despite similarities in overall body weight, height and BMI. Interestingly, a recent paper[7] contained some information concerning percentage body fat estimated by DEXA in children aged 4–8 y; values of 19.0 ± 7.2% body fat in 22 girls and 14.1 ± 4.1% body fat in 23 boys being reported. However these children were not pair-matched for age, height and weight, and the authors[7] did not test the gender differences, nor did they report either fat mass or bone-free tissue mass. We have calculated that their young boys had significantly less fat that
the girls ($P < 0.01$), as in our study. Thus their gender difference is of the same order as that we observed, providing independent support for our present results.

Population bias could explain our findings and be a study limitation. However, the weights and heights of our subjects were within normal limits for a representative study population of 1073 local children measured at 3, 5 and 7 y (PA Silva, personal communication). Moreover, international standard deviation scores confirm that our study population is representative.[13] We acknowledge that the differences we observed, may reflect differences in physical activity, nutrition or metabolism. There is evidence that boys are more active than girls, both prepubertally[14,15] and later.[16] Gender differences in energy intake have also been noted in young children.[17] Alternatively, gender differences in leptin may be important in favouring early fat accumulation in girls.[18]

Others have observed similar total BMC values in boys and girls aged 8–9 y.[9,10] In our study of younger

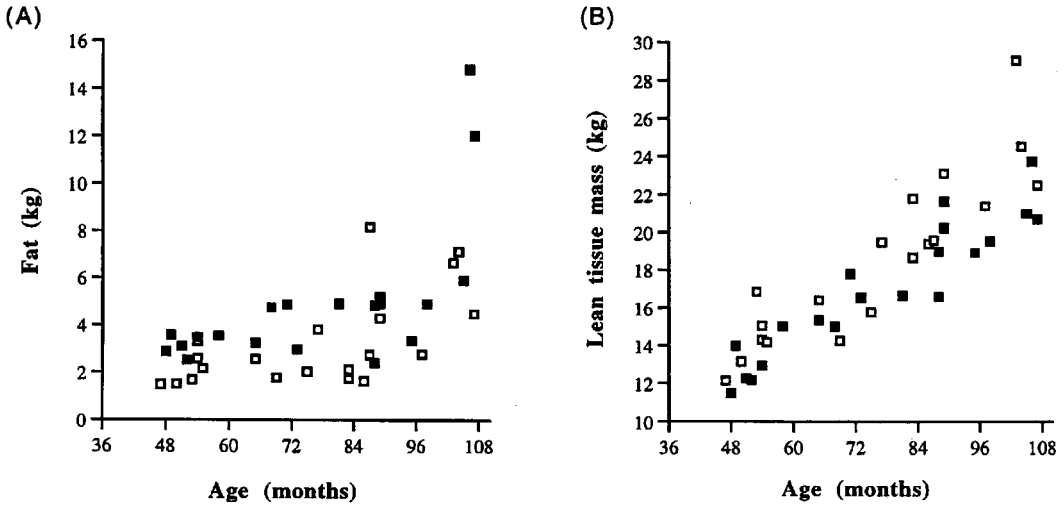


Figure 1 A, Total body fat mass (kg) and B, total body bone-free lean tissue mass (kg), in boys (open squares) and girls (closed squares) in relation to age.

children, BMC was similar in both sexes, providing support for the view that BMC is influenced by total body mass rather than lean tissue or fat tissue.[19]

The rising prevalence of obesity is currently a major concern in the western world[20] and the ability to quantify fat mass in young children may enable better definition of future health hazards associated with obesity. The common view that prepubertal children have similar body composition, may have arisen due to the lack of suitable methods for the accurate determination of body fat in this age group. With DEXA, we have now shown that significant gender differences in body fat are present in prepubertal children. Our findings need to be confirmed in a larger sample, drawn randomly from the general population.

Acknowledgements
We thank Mrs S Williams, biostatistician in the Department of Preventive and Social Medicine for statistical advice. We gratefully acknowledge funding support from The Otago Medical Research Foundation, The Health Research Council of New Zealand and the NZ Lottery Grant Board.

References

1 Forbes G. Human body composition. Growth, aging, nutrition and activity. Springer-Verlag: New York, 1987.
2 Kelly J, Stanton W, McGee R, Silva P. Tracking relative weight in subjects studied longitudinally from ages 3 to 13 years. *J Paediatr Child Health* 1992; **28**: 158–161.
3 Mazess RB, Barden HS, Bisek JP, Hanson J. Dual-energy x-ray absorptiometry for total body and regional bone-mineral and soft-tissue composition. *Am J Clin Nutr* 1990; **51**: 1106–1112.
4 Kohrt WM. Body composition by DXA: tried and true? *Med Sci Sports Med* 1995; **27**: 1349–1353.
5 Ellis K. Measuring body fatness in children and young adults: comparison of bioelectrical impedance analysis, total body electrical conductivity, and dual energy x-ray absorptiometry. *Int J Obes* 1996; **20**: 866–873.
6 Ogle GD, Allen JR, Humphries IRJ, Lu PW, Briody JN, Morley K, Howman-Giles R, Cowell CT. Body-composition assessment by dual-energy x-ray absorptiometry in subjects aged 4-26 y. *Am J Clin Nutr* 1995; **61**: 746–753.
7 Lazarus R, Baur L, Webb K, Blyth F. Body mass index in screening for adiposity in children and adolescents: systematic evaluation using receiver operating characteristic curves. *Am J Clin Nutr* 1996; **63**: 500–506.
8 Taylor RW, Cannan R, Gold E, Lewis-Barned N, Goulding A. Regional body fat distribution in New Zealand girls aged 4–16 years: a cross-sectional study by dual energy x-ray absorptiometry. *Int J Obes* 1996; **20**: 763–767.
9 Faulkner RA, Bailey DA, Drinkwater DT, Wilkinson AA, Houston CS, McKay HA. Regional and total body bone mineral content, bone mineral density and total body tissue composition in children 8–16 years of age. *Calcif Tiss Int* 1993; **53**: 7–12.
10 Nelson D, Simpson P, Johnson C, Barondess D, Kleerekoper M. The accumulation of whole body skeletal mass in third- and fourth-grade children: effects of age, gender, ethnicity and body composition. *Bone* 1997; **20**: 73–79.
11 Hammer LD, Kraemer HC, Wilson DM, Ritter PL, Dornbusch SM. Standardized percentile curves of body-mass index for children and adolescents. *Am J Dis Child* 1991; **145**: 259–263.
12 Duke PM, Litt IF, Gross RT. Adolescents' self-assessment of sexual maturation. *Pediatrics* 1980; **66**: 918–920.
13 Tanner JM, Whitehouse RH. Clinical longitudinal standards for height, weight, height velocity, weight velocity, and stages of puberty. *Arch Dis Child* 1976; **51**: 170–179.
14 Gilliam T, Freedson P, Grenen D, Shahraray B. Physical activity patterns determined by heart rate monitoring in 6–7 year-old children. *Med Sci Sports Exerc* 1981; **13**: 65–67.
15 Gutin B, Basch C, Shea S, Contento I, DeLozier M, Rips J, Irigoyen M, Zybert P. Blood pressure, fitness, and fatness in 5- and 6-year-old children. *JAMA* 1990; **264**: 1123–1127.
16 Armstrong N, Balding J, Gentle P, Kirby B. Patterns of physical activity among 11 to 16 year old British children. *BMJ* 1990; **301**: 203–205.
17 Briefel R, McDowell M, Alaimo K, Caughman C, Bischof A, Carroll M, Johnson CL. Total energy of the US population: the third National Health and Nutrition Examination Survey, 1988–1991. *Am J Clin Nutr* 1995; **62** (Suppl): 1072S–1080S.
18 Hassink SG, Sheslow DV, de Lancey E, Opentanova I, Considine RV, Caro JF. Serum leptin in children with obesity: relationship to gender and development. 1997; **98**: 201–203.
19 Moro M, Van der Meulen M, Kiratli B, Marcus R, Bachrach L, Carter D. Body mass is the primary determinant of mid-femoral bone acquisition during adolescent growth. *Bone* 1996; **19**: 519–526.
20 Prentice A, Jebb S. Obesity in Britain: gluttony or sloth? *BMJ* 1995; **311**: 437–439.

THE OBESITY SOCIETY

Professionals Collaborating to Overcome Obesity

REGISTER NOW!

WEBINARS

Personalizing Obesity Mgmt

March 2, 2022
10 am - 1:30 pm Eastern
3.25 CMEs
Price: $150

This webinar will discuss the who, when, and how of obesity management with a focus on obtaining the patient history and narrative, considering psychological drivers and barriers, staging patients according to risk and comorbidities, and tailoring medical treatments to individual needs and expectations.

Medical Mgmt of Bariatric Patients

February 15, 2022
10 am - 5:20 pm Eastern
5.25 CMEs
Price: $200

Severe obesity and its complications are best managed by an interdisciplinary team including both surgical and medical providers. Topics will include optimizing pre-operative care as well as recognizing and developing approaches to the variability in post-operative weight loss and remission rates of obesity-related complications in both adult and pediatric patients.

Pediatrics

March 17, 2022
10 am - 1:30 pm Eastern
3.25 CMEs
Price: $150

One in five adolescents in the United States have obesity and most will continue to have obesity in adulthood. Effective treatment and care of youth with obesity is urgently needed. Come join TOS Pediatric Obesity medicine experts share their experience and evidenced-based treatment approaches to caring for youth with obesity in this patient case-based webinar.

Includes OnDemand Access to Recordings throughout 2022

View program and register at Obesity.org/Webinars

ARTICLES
EPIDEMIOLOGY
nature publishing group

Sex Differences in Regional Body Fat Distribution From Pre- to Postpuberty

Rachael W. Taylor[1], Andrea M. Grant[1], Sheila M. Williams[2] and Ailsa Goulding[1]

Few large studies have evaluated the emergence of sexual dimorphism in fat distribution with appropriate adjustment for total body composition. The objective of this study was to determine the timing and magnitude of sex differences in regional adiposity from early childhood to young adulthood. Regional fat distribution was measured using dual-energy X-ray absorptiometry (trunk and extremity fat using automatic default regions and waist and hip fat using manual analysis) in 1,009 predominantly white participants aged 5–29 years. Subjects were divided into pre (Tanner stage 1), early (Tanner stages 2–3), late (Tanner stages 4–5), and post (males ≥20 years and females ≥18 years) pubertal groups. Sexual dimorphism in trunk fat (adjusted for extremity fat) was not apparent until late puberty, when females exhibited 17% less ($P < 0.001$) trunk fat than males. By contrast, sex differences in waist fat (adjusted for hip fat) were apparent at each stage of puberty, the effect being magnified with age, with prepubertal girls having 5% less ($P = 0.027$) and adult women having 48% less ($P < 0.0001$) waist fat than males. Girls had considerably more peripheral fat whether measured as extremity or hip fat at each stage. Sex differences in regional adiposity were significantly greater in young adults than in late adolescence. Exclusion of overweight participants did not materially affect the estimates. Sexual dimorphism in fat patterning is apparent even prepubertally with girls having less waist and more hip fat than boys. The magnitude of the sex difference is amplified with maturation, and particularly from late puberty to early adulthood.

Obesity (2010) **18,** 1410–1416. doi:10.1038/oby.2009.399

Introduction

Marked sexual dimorphism in total and regional body composition is apparent in adults, with men having greater lean tissue mass and a more central fat pattern compared with the more peripheral fat distribution typically observed in adult women (1). It is now well accepted that significant gender differences in total body composition are apparent even in prepubertal children (2–4). However, the emergence of sex differences in fat patterning appears to be more controversial. Although anthropometric techniques suggested that sex-specific fat patterning emerges during puberty (5,6), recent work utilizing direct measures of body fat distribution including dual-energy X-ray absorptiometry (DXA) and magnetic resonance imaging has suggested that variation in fat patterning may also exist prepubertally (3,4,7–10). However, the literature is not consistent, with several groups reporting no gender differences in regional adiposity in young children (11–13). Much of this discrepancy can probably be attributed to variations in the regional indexes of choice, the particular ages being studied, differences in relative fatness of the groups, and use of different adjustments for covariates (14).

Few previous groups have studied large samples of participants over the full span of growth from prepuberty to early adulthood to examine the timing of the appearance of sexual dimorphism in fat distribution (8,11). Elucidating such information is important because of the role of body composition during growth in predicting adult disease risk (15). The aim of this study was to examine when differences in fat distribution measured using anthropometry and DXA become apparent in a large cross-sectional cohort of children and young adults aged 5–29 years.

Methods and Procedures

Subjects

Cross-sectional data on total and regional body composition were available for 1,009 predominantly white children and young adults aged 5–29 years who had participated in various studies investigating body composition and health in our laboratory from 1996 to 2007 (16–22). All studies were approved by the Lower South Ethics Committee and written informed consent was obtained from each participant or parent/guardian. A brief medical history was obtained by questionnaire and no participant was taking medication that would affect his or her body composition and no subjects had a history of constitutional delay in growth or maturation. Tanner stage of pubertal development was assessed in all subjects <19 years of age by the parent/guardian of the younger children or by the adolescent themselves using standard descriptions and pictures (23). Participants were divided into pre

[1]Department of Medical and Surgical Sciences, University of Otago, Dunedin, New Zealand; [2]Department of Preventive and Social Medicine, University of Otago, Dunedin, New Zealand. Correspondence: Rachael W. Taylor (rachael.taylor@otago.ac.nz)

Received 19 July 2009; accepted 3 October 2009; published online 5 November 2009. doi:10.1038/oby.2009.399

(Tanner stage 1), early (Tanner stages 2 and 3), late (Tanner stages 4 and 5) (8), and post (males aged ≥20 years and females ≥18 years) pubertal groups for analysis.

anthropometric measures

Duplicate measures of height (wall-mounted stadiometer) and weight (electronic scales) were obtained with the participants wearing light clothing and no shoes. Duplicate measures of waist (minimum circumference between the rib cage and the iliac crest using a non-stretchable tape measure) and hip (maximum protuberance of the buttocks) circumferences were obtained in 908 (90%) and 846 (84%) of participants, respectively. Measurements were obtained by several trained examiners following the same protocols. In our laboratory, coefficients of variation (CVs%) for height, weight, and waist circumference in young children are <2% (17). BMI was calculated as weight in kilograms/height in meters squared. Overweight and obesity were defined using international reference norms (24) for those aged 5–18 years and BMI 25–29.9 and BMI ≥30 for overweight and obesity, respectively, for those ≥19 years of age.

dXa scanning

All DXA measurements were performed and analyzed by one experienced operator with a Lunar DPX-L scanner (software package 4.7; Lunar, Madison, WI) using standard procedures. The scanner determines total fat mass (kg) and the fat content (kg) of specific anatomical regions including trunk and extremity fat (automatic default regions) and central and peripheral fat (manual regions of interest) as shown in **Figure 1**. The trunk region consists of the area bordered by a horizontal line below the chin, vertical borders lateral to the ribs, and oblique lines passing through the femoral necks. The arm region consists of all tissues outside these lateral borders and the leg region of all tissue below the oblique lines. Extremity fat was defined as the sum of arm and leg fat (7). Central and peripheral regions of interest were also determined using manual analysis. The central region of interest (waist fat) was a box 9.6 cm high with the lower border positioned superior to the iliac crest. The peripheral region (hip fat) consists of the same sized box, positioned so that the center of the box was at the level of the greater trochanters (25). In our laboratory, the CVs for repeated *in vivo* scans on 10 adults were 2.6% for fat mass, 2.5% for fat percentage, and <3.5% for all regional measurements.

statistics

All data were analyzed using STATA (Stata Statistical Software, Release 8.0, 2003; StataCorp, College Station, TX). Analysis of covariance, which adjusted for age and included a sex × group interaction to test for a different pattern of results across the four different stages of puberty for males and females, was used to analyze the data. Logistic regression was used to analyze the data for overweight status. Regression models were used to examine the sex differences, presented as ratios of female relative to male values, for trunk fat adjusted for extremity fat, and for extremity fat adjusted for trunk fat, for the four pubertal stages. Sex differences for waist fat and hip fat, and waist girth and hip girth were also estimated. All variables not normally distributed (weight, BMI, fat masses) were log transformed before analysis.

results

Table 1 shows means (s.d.) in total body composition for each pubertal group by sex. Sex differences were observed in all indexes with the exception of the prevalence of overweight/obesity. Furthermore, the interaction between sex and stage indicated that the pattern of differences between males and females was generally higher in each successive stage of puberty. In general, higher values are apparent across puberty for most indexes and for both sexes, with the exception of body fat in males. Although absolute fat mass is approximately twofold higher in adolescent and older males compared with prepubertal males, relative fatness (fat mass index, % fat) is highest during early puberty. By contrast, females display steady gains in absolute and relative fatness with maturation. Interestingly, the significantly higher body weight of young adults compared with adolescents in late puberty, was not matched by higher stature, and was explained by a higher deposition of lean than fat mass.

Significant differences in regional fat distribution are also apparent with both sex and pubertal development (**Table 2**). The greater total body fat of females is also reflected by greater deposition of absolute regional fat at each site with most stages of maturation. In males, higher values are observed for trunk fat mass with age whereas extremity fat mass remains similar from early to postpuberty. Significant interactions between pubertal stage and sex were also observed for most measures showing that the pattern of differences between males and females depended on the level of maturation. By contrast, the nonsignificant interaction effect for waist circumference shows that the sex difference in waist circumference did not differ according to stage of puberty.

Table 3 demonstrates regional fat distribution at each stage of pubertal maturation as ratios (female values relative to males). Interestingly, sexual dimorphism in trunk fat deposition was not apparent until late puberty, and became more marked in young adults, where women had 34% less trunk fat than men, once adjusted for extremity fat. By contrast, differences in waist fat were apparent at each stage of puberty, although this effect magnified with increasing age, ranging from a difference of 5% ($P = 0.027$) in prepubertal children to 48% ($P < 0.001$) in young adults. Considerable differences in peripheral fat distribution were also apparent even in the youngest children. Thus prepubertal girls had 2% ($P = 0.283$) more extremity fat and

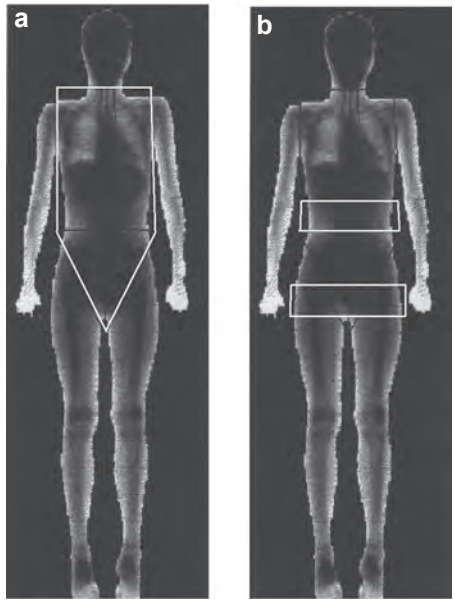


Figure 1 Automatic (**a**: trunk, arm, and leg) and manual (**b**: waist and hip) regions of interest.

table 1 total body composition by sex and puberty

	Pubertal stage				*P*		
	Pre	Early	Late	Post	Sex	Stage	Sex × stage
N							
Males	205	92	119	102			
Females	171	97	119	104			
Age (years)							
Males	7.9 (2.3)	13.6 (1.5)	17.0 (1.8)	24.0 (2.6)	<0.001	<0.001	0.002
Females	7.3 (1.7)	12.4 (1.8)	15.8 (1.5)	21.9 (3.2)			
Height (cm)							
Males	127.9 (13.5)	159.6 (10.2)	176.8 (7.1)	179.6 (7.0)	<0.001	<0.001	<0.001
Females	122.9 (11.1)	152.1 (10.0)	164.3 (5.8)	166.4 (5.7)			
Weight (kg)							
Males	29.1 (9.1)	51.8 (11.8)	69.8 (11.4)	79.4 (10.4)	<0.001	<0.001	<0.001
Females	25.9 (6.4)	48.5 (12.6)	63.0 (11.2)	67.0 (14.3)			
BMI (kg/m^2)							
Males	17.4 (2.6)	20.2 (3.4)	22.3 (2.9)	24.6 (2.5)	0.033	<0.001	0.018
Females	17.0 (2.2)	20.8 (4.0)	23.3 (4.0)	24.3 (5.2)			
Overweight/obese (%)[a]							
Males	17/5	15/5	10/3	31/4	0.987	<0.001	0.01
Females	17/5	25/10	24/8	16/13			
Lean mass (kg)							
Males	21.8 (5.5)	37.5 (7.0)	54.8 (7.4)	61.9 (7.0)	<0.001	<0.001	<0.001
Females	18.8 (3.7)	31.5 (5.6)	38.7 (3.6)	41.4 (4.6)			
Fat-free mass index[b]							
Males	13.1 (0.9)	14.6 (1.3)	17.5 (1.5)	19.2 (1.7)	<0.001	<0.001	<0.001
Females	12.3 (0.7)	13.5 (1.1)	14.3 (1.1)	15.0 (1.4)			
Fat mass (kg)							
Males	5.5 (4.6)	11.8 (8.2)	11.7 (7.2)	13.7 (6.3)	<0.001	<0.001	<0.001
Females	5.4 (3.4)	14.6 (8.7)	21.3 (8.9)	22.2 (11.5)			
Fat mass index[c]							
Males	3.2 (2.2)	4.6 (3.2)	3.7 (2.3)	4.2 (1.8)	<0.001	<0.001	<0.001
Females	3.5 (1.8)	6.2 (3.4)	7.9 (3.3)	8.1 (4.3)			
Fat (%)							
Males	17.7 (7.9)	21.5 (10.3)	16.2 (7.0)	16.9 (6.1)	<0.001	<0.001	<0.001
Females	20.4 (6.9)	28.6 (10.4)	33.0 (8.0)	31.9 (9.5)			

Analysis based on analysis of covariance adjusted for age and included terms for sex and pubertal stage and an interaction between sex and pubertal stage. Ordinal logistical regression adjusting for age was used to analyze the variable for overweight and obesity.
[a]Data are presented as mean (s.d.) except for the prevalence of overweight/obese (24) which is presented as % of group who are overweight but not obese/% of group who are obese. [b]Fat-free mass index is lean mass (kg) divided by height squared. [c]Fat mass index is fat mass (kg) divided by height squared.

6% ($P < 0.001$) more hip fat than prepubertal boys and these differences became more marked at each successive stage of puberty. Although substantial sex differences in regional adiposity were apparent by late puberty, Table 3 demonstrates that this dimorphism becomes considerably more marked in young adults. For example, adjusted trunk fat was 17% ($P < 0.001$) and 34% ($P < 0.001$) lower in late and postpubertal females, respectively, relative to males, with corresponding figures of 35% ($P < 0.001$) and 48% ($P < 0.001$) in these groups for adjusted waist fat. Similar patterns were observed for peripheral fat distribution. Waist circumference was significantly lower and hip circumference significantly higher in girls from early puberty onward.

Exclusion of overweight participants did not materially affect these estimates. The adjusted ratios at each pubertal stage were similar in the total sample ($N = 1,009$) and when analyses were restricted to the normal-weight participants ($N = 846$) for trunk or extremity fat (Figure 2) and waist or hip fat (data not shown).

table 2 r egional body composition by sex and puberty

	Pubertal stage				*P*		
	Pre	Early	Late	Post	Sex	Stage	Sex × stage
N							
Males	205	92	119	102			
Females	171	97	119	104			
Trunk fat (kg)							
Males	2.0 (2.2)	4.8 (3.9)	5.2 (3.6)	6.8 (3.3)	<0.001	<0.001	<0.001
Females	1.9 (1.6)	6.1 (4.2)	9.2 (4.4)	9.7 (5.8)			
Extremity fat (kg)							
Males	2.9 (2.2)	6.2 (4.1)	5.9 (3.5)	6.3 (3.0)	<0.001	<0.001	<0.001
Females	2.8 (1.7)	7.6 (4.4)	11.2 (4.5)	11.4 (5.4)			
Trunk fat (% total fat)							
Males	31.6 (5.6)	38.6 (5.6)	43.3 (5.3)	49.1 (4.2)	0.033	<0.001	<0.001
Females	34.1 (5.3)	39.9 (5.8)	42.4 (4.1)	42.3 (5.4)			
Trunk-to-leg fat ratio							
Males	0.73 (0.19)	0.86 (0.21)	1.04 (0.25)	1.36 (0.28)	<0.001	<0.001	<0.001
Females	0.77 (0.19)	0.91 (0.23)	0.97 (0.18)	0.99 (0.24)			
Waist fat (kg)							
Males	0.5 (0.5)	0.9 (0.8)	0.8 (0.7)	1.0 (0.6)	<0.001	<0.001	<0.001
Females	0.4 (0.4)	1.1 (0.8)	1.5 (0.8)	1.6 (1.1)			
Hip fat (kg)							
Males	0.8 (0.5)	1.4 (0.8)	1.3 (0.6)	1.4 (0.6)	<0.001	<0.001	<0.001
Females	0.9 (0.4)	1.8 (0.8)	2.5 (0.8)	2.6 (1.1)			
Waist girth (cm)							
Males	59.6 (8.2)	72.4 (9.8)	78.6 (8.1)	83.9 (7.7)	<0.001	<0.001	0.766
Females	56.6 (5.5)	69.6 (9.3)	74.0 (8.8)	79.0 (13.6)			
Hip girth (cm)							
Males	71.4 (8.9)	88.0 (8.6)	96.7 (6.6)	98.9 (6.7)	0.013	<0.001	<0.001
Females	66.9 (6.6)	87.6 (10.2)	98.3 (8.1)	101.1 (11.7)			
Waist-to-hip girth ratio							
Males	0.85 (0.05)	0.82 (0.06)	0.81 (0.05)	0.85 (0.07)	<0.001	<0.001	<0.001
Females	0.85 (0.05)	0.80 (0.06)	0.75 (0.05)	0.78 (0.08)			

Data are presented as mean (s.d.). Values for *n* for waist circumference are 202, 90, 110, and 59 and 166, 96, 113, and 72 for pre, early, late, and post stages of puberty in boys and girls, respectively. Values for *n* for hip circumference and waist-to-hip ratio are as for waist except in the prepubertal group, where *N* = 168 in boys and *N* = 138 in girls. Analysis based on analysis of covariance adjusted for age and included terms for sex and pubertal stage and an interaction between sex and pubertal stage.

dIscuss Ion

Our data demonstrate that sexual dimorphism in regional fat distribution becomes apparent at different stages of puberty, depending on the index chosen. Significant sex differences in trunk fat were not evident until late puberty, whereas even in prepubertal children, girls had significantly less waist fat than boys. Differences in peripheral fat were also evident at younger ages, whether measured using extremity fat (sum of leg and arm fat) or hip fat (fat directly over the greater trochanters). Sex differences in waist and hip circumferences were also observed from early puberty. The magnitude of the sex differences in regional adiposity increased significantly with maturation, so that striking differences were evident between adolescents in late puberty and young adults in their 20s.

The precise timing at which sex differences in regional adiposity is said to become apparent during growth is conflicting in the literature, due in part to the variety of methods used to assess centrality and to choice of statistical adjustments for other body components. In general, studies utilizing skinfold measurements have demonstrated that although absolute thicknesses are often higher in females at least from 8 years of age onward (6,26,27), demarcation in relative skinfold thickness occurs during puberty, primarily due to increased deposition of peripheral fat in females (5,6). Sexual dimorphism

table 3 sex differences (as ratios) in various measures of regional fat distribution across puberty (females relative to males)

	Adjusted ratios (95% CI)			
	Pre	Early	Late	Post
Trunk fat (adjusted for extremity fat)	1.01 (0.97, 1.06)	0.99 (0.94, 1.06)	0.83 (0.80, 0.88)	0.66 (0.62, 0.71)
Extremity fat (adjusted for trunk fat)	1.02 (0.98, 1.05)	1.06 (1.01, 1.11)	1.31 (1.25, 1.37)	1.53 (1.46, 1.61)
Waist fat (adjusted for hip fat)	0.95 (0.91, 0.99)	0.85 (0.80, 0.91)	0.65 (0.61, 0.69)	0.52 (0.49, 0.56)
Hip fat (adjusted for waist fat)	1.06 (1.03, 1.09)	1.16 (1.11, 1.21)	1.47 (1.41, 1.53)	1.66 (1.59, 1.73)
Waist girth (adjusted for hip girth)	0.99 (0.98, 1.01)	0.95 (0.94, 0.97)	0.90 (0.89, 0.92)	0.89 (0.87, 0.91)
Hip girth (adjusted for waist girth)	1.00 (0.99, 1.01)	1.04 (1.03, 1.06)	1.09 (1.08, 1.11)	1.11 (1.09, 1.12)

Data are presented as ratios (95% confidence interval (CI)) in females relative to males adjusted for height, age, sex, pubertal stage, interactions between sex and pubertal stage, and specific regional fat as shown.

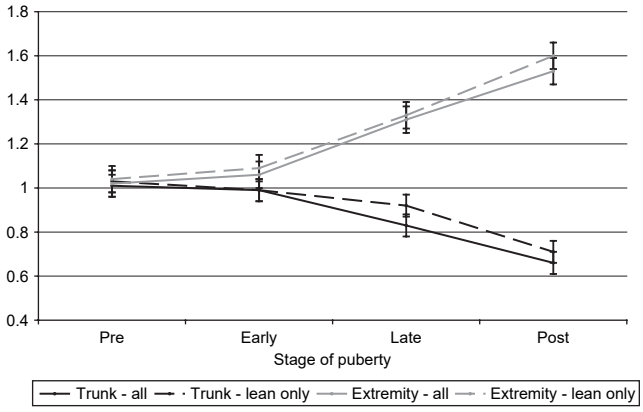


Figure 2 Ratios (females relative to males) for adjusted trunk fat (dark lines) and adjusted extremity fat (light lines) for the total group (solid lines) and lean subjects only (dashed lines).

in circumference measures is also apparent from a young age with elevated waist circumference and/or waist-to-hip ratio or lower hip circumferences being reported in boys as young as 5–7 years (4,26,28,29). Our data support this earlier work, demonstrating that females have lower waist and greater hip circumferences from early puberty. However, there are reservations regarding the use of anthropometric measures to provide indexes of regional adiposity, because of variation in body proportions with growth (15). The advent of DXA, magnetic resonance imaging, and computed tomography provide more accurate and sensitive evaluation techniques to assess changes in body composition, even in young children (15).

Central fat distribution assessed by DXA has typically been reported using trunk fat as a percentage of total fat (% trunk fat) or the trunk-to-leg fat ratio. Although two groups found that differences in trunk-to-leg fat ratio were not apparent until after puberty (11,30), others (4) reported higher trunk-to-leg fat ratio values in 7–8-year-old girls compared with boys (4). Conversely, Wang *et al.* (10) in a large sample of Chinese children, reported that 6–11-year-old boys have a higher % trunk fat than girls. Regardless of these discrepant findings, alternative methods of analyzing DXA data are warranted given the problems of using ratios in statistical analyses (31) and the difficulties in interpreting measures of adiposity which include the measure of interest (e.g., trunk fat) in both the numerator and the denominator (% trunk fat) (32). Comparing regional fat masses without making adjustments for body size is clearly inappropriate (15); girls will have more body fat in every region as they develop, given their considerably higher total body fat levels and the strong correlations which are present between regional fat and total fat. Thus as recommended recently (15), we adjusted specific central regions for their corresponding peripheral region, as well as adjusting for height, but not for weight. When we included weight as a covariate as other groups have done (7,8), the variance inflation factors indicated that multicollinearity between variables was an issue.

Our results support and extend earlier work (7,8) demonstrating that sexual dimorphism is apparent not only in peripheral fat but also in central fat during puberty. Although expected differences in trunk fat were not apparent until late puberty, we observed that even young prepubertal girls stored less fat in the waist region and more fat peripherally (over the hips) than boys. Use of the DXA manual "waist" region of interest, which focuses narrowly on fat located in the mid-abdominal regional of the body, may provide a more appropriate measure of "central" fat than is provided by the larger default trunk fat region (33), given that DXA cannot differentiate between visceral and subcutaneous fat. However, others have demonstrated that trunk fat measured by DXA is strongly correlated with visceral fat in young children (34). Furthermore, central fat (DXA) is similarly associated with adverse plasma lipid profiles in children as is visceral adipose tissue (35,36). DXA is also a more available and less invasive method of assessing body composition in children, compared with computed tomography or magnetic resonance imaging, which are more costly, take longer to complete and have a larger radiation dose (computed tomography only). Visceral fat accumulation is thought to be similar in prepubertal boys and girls (3,12), particularly after appropriate adjustment for subcutaneous fat (13), although not all studies are consistent (9,37). On the other hand, in prepubertal girls of this age, more of this central fat may be subcutaneous (3).

Interestingly, we observed greater sexual dimorphism in regional adiposity in young adults aged 18–29 years when compared with adolescents in late puberty. Much of this response was probably driven by the differences we observed between adolescent and young adult males; the latter had 2 kg more

total fat and significantly more trunk fat, but similar levels of extremity fat relative to adolescent males. Similar results have been reported recently in a longitudinal study of 17-year-old males followed for 8 years; here both active and inactive men became fatter and stored more of that fat abdominally by study end (38).

Our study is cross-sectional and therefore cannot determine actual changes in body composition within individuals. Pubertal status was self-reported rather than examined, and others (39) have reported only moderate concordance between self- and physician-based assessment. However, Duke *et al.* (23) show that 86–91% of adolescents correctly estimate their developmental age. We also do not have measures of blood hormone status, such as insulin-like growth factor I, sex steroids, and leptin which might explain variation in body composition over time and between the sexes (40,41). Others (4) have shown that variation in insulin-like growth factor I and sex steroids explain <18% of the variance in body composition in prepubertal children. However, this may underestimate the contribution of sex hormones, given that blood sampling in Garnett *et al.* (4) occurred in the early morning, and sex steroids peak during the night. Although recent research has provided greater understanding of the role of sex hormones on fat partitioning, considerable work is also required to determine the role of chromosomal sex (XX, XY) and prenatal hormones in determining fat patterning later in life (42).

Strengths of our study include measurements on a large database of children and young adults varying markedly in shape and size. Measurements were made on the same DXA scanner, and analysis of all DXA data was completed by the same person. In addition, we excluded participants with any known endocrine disorders or with a history of constitutional delay in growth or maturation from our study.

In conclusion, our study shows significant sexual dimorphism in fat patterning is apparent even prepubertally, with girls having less waist and more hip fat than boys. The magnitude of the sex difference is amplified with maturation, with young adults displaying considerably higher central fat deposition compared with those in late adolescence. Progression to young adulthood may be a time of considerable fat deposition, particularly fat that is more centrally distributed, at least in men (38). Although the predominant predictor of fat gain in men at this time was age, appropriate participation in physical activity may modulate the overall effect (38). Thus late adolescence/young adulthood could be an appropriate time for targeted lifestyle interventions to limit fat deposition.

acknowledgments

This research was supported by funding from Health Research Council, the University of Otago, the Child Health Research Foundation, Otago Medical Research Foundation, and the Dunedin School of Medicine Dean's Bequest Fund.

disclosure

The authors declared no conflict of interest.

references

1. Kissebah AH, Krakower GR. Regional adiposity and morbidity. *Physiol Rev* 1994;74:761–811.
2. Taylor RW, Gold E, Manning P, Goulding A. Gender differences in body fat content are present well before puberty. *Int J Obes Relat Metab Disord* 1997;21:1082–1084.
3. Arfai K, Pitukcheewanont PD, Goran MI *et al*. Bone, muscle, and fat: sex-related differences in prepubertal children. *Radiology* 2002;224:338–344.
4. Garnett SP, Högler W, Blades B *et al*. Relation between hormones and body composition, including bone, in prepubertal children. *Am J Clin Nutr* 2004;80:966–972.
5. Malina RM. Regional body composition: age, sex and ethnic variation. In: Roche AF, Heymsfield SB, Lohman TG (eds). *Human Body Composition*. Human Kinetics: Champaign, IL, 1996, pp 217–255.
6. Heude B, Kettaneh A, de Lauzon Guillain B *et al*.; Fleurbaix Laventie Ville Santé Group. Growth curves of anthropometric indices in a general population of French children and comparison with reference data. *Eur J Clin Nutr* 2006;60:1430–1436.
7. He Q, Horlick M, Thornton J *et al*. Sex and race differences in fat distribution among Asian, African-American, and Caucasian prepubertal children. *J Clin Endocrinol Metab* 2002;87:2164–2170.
8. He Q, Horlick M, Thornton J *et al*. Sex-specific fat distribution is not linear across pubertal groups in a multiethnic study. *Obes Res* 2004;12:725–733.
9. Dencker M, Thorsson O, Lindén C *et al*. BMI and objectively measured body fat and body fat distribution in prepubertal children. *Clin Physiol Funct Imaging* 2007;27:12–16.
10. Wang H, Story RE, Venners SA *et al*. Patterns and interrelationships of body-fat measures among rural Chinese children aged 6 to 18 years. *Pediatrics* 2007;120:e94–e101.
11. Ogle GD, Allen JR, Humphries IR *et al*. Body-composition assessment by dual-energy x-ray absorptiometry in subjects aged 4-26 y. *Am J Clin Nutr* 1995;61:746–753.
12. Huang TT, Johnson MS, Figueroa-Colon R, Dwyer JH, Goran MI. Growth of visceral fat, subcutaneous abdominal fat, and total body fat in children. *Obes Res* 2001;9:283–289.
13. Brambilla P, Bedogni G, Moreno LA *et al*. Crossvalidation of anthropometry against magnetic resonance imaging for the assessment of visceral and subcutaneous adipose tissue in children. *Int J Obes (Lond)* 2006;30:23–30.
14. Wells JC. A critique of the expression of paediatric body composition data. *Arch Dis Child* 2001;85:67–72.
15. Wells JC. Sexual dimorphism of body composition. *Best Pract Res Clin Endocrinol Metab* 2007;21:415–430.
16. Goulding A, Jones IE, Taylor RW, Manning PJ, Williams SM. More broken bones: a 4-year double cohort study of young girls with and without distal forearm fractures. *J Bone Miner Res* 2000;15:2011–2018.
17. Taylor RW, Jones IE, Williams SM, Goulding A. Evaluation of waist circumference, waist-to-hip ratio, and the conicity index as screening tools for high trunk fat mass, as measured by dual-energy X-ray absorptiometry, in children aged 3-19 y. *Am J Clin Nutr* 2000;72:490–495.
18. Goulding A, Jones IE, Taylor RW, Williams SM, Manning PJ. Bone mineral density and body composition in boys with distal forearm fractures: a dual-energy x-ray absorptiometry study. *J Pediatr* 2001;139:509–515.
19. McAuley KA, Williams SM, Mann JI *et al*. Intensive lifestyle changes are necessary to improve insulin sensitivity: a randomized controlled trial. *Diabetes Care* 2002;25:445–452.
20. Miller JC, Grant AM, Drummond BF *et al*. DXA measurements confirm that parental perceptions of elevated adiposity in young children are poor. *Obesity (Silver Spring)* 2007;15:165–171.
21. Sutherland TJ, Goulding A, Grant AM *et al*. The effect of adiposity measured by dual-energy X-ray absorptiometry on lung function. *Eur Respir J* 2008;32:85–91.
22. Taylor RW, Murdoch L, Carter P *et al*. Longitudinal study of physical activity and inactivity in preschoolers: the FLAME study. *Med Sci Sports Exerc* 2009;41:96–102.
23. Duke PM, Litt IF, Gross RT. Adolescents' self-assessment of sexual maturation. *Pediatrics* 1980;66:918–920.
24. Cole TJ, Bellizzi MC, Flegal KM, Dietz WH. Establishing a standard definition for child overweight and obesity worldwide: international survey. *BMJ* 2000;320:1240–1243.
25. Taylor RW, Keil D, Gold EJ, Williams SM, Goulding A. Body mass index, waist girth, and waist-to-hip ratio as indexes of total and regional adiposity in

women: evaluation using receiver operating characteristic curves. *Am J Clin Nutr* 1998;67:44–49.
26. Webster-Gandy J, Warren J, Henry CJ. Sexual dimorphism in fat patterning in a sample of 5 to 7-year-old children in Oxford. *Int J Food Sci Nutr* 2003;54:467–471.
27. Marini E, Cabras S, Rebato E *et al*. Sex differences in skinfold variability across human populations and during the life cycle. *Ann Hum Biol* 2007;34:377–392.
28. Mast M, Körtzinger I, König E, Müller MJ. Gender differences in fat mass of 5-7-year old children. *Int J Obes Relat Metab Disord* 1998;22:878–884.
29. Kirkby J, Metcalf BS, Jeffery AN *et al*. Sex differences in resting energy expenditure and their relation to insulin resistance in children (EarlyBird 13). *Am J Clin Nutr* 2004;80:430–435.
30. Cowell CT, Briody J, Lloyd-Jones S *et al*. Fat distribution in children and adolescents—the influence of sex and hormones. *Horm Res* 1997;48(Suppl 5):93–100.
31. Allison DB, Paultre F, Goran MI, Poehlman ET, Heymsfield SB. Statistical considerations regarding the use of ratios to adjust data. *Int J Obes Relat Metab Disord* 1995;19:644–652.
32. Cole TJ, Fewtrell MS, Prentice A. The fallacy of using percentage body fat as a measure of adiposity. *Am J Clin Nutr* 2008;87:1959; author reply 1959–1960.
33. Carey DG, Jenkins AB, Campbell LV, Freund J, Chisholm DJ. Abdominal fat and insulin resistance in normal and overweight women: direct measurements reveal a strong relationship in subjects at both low and high risk of NIDDM. *Diabetes* 1996;45:633–638.
34. Goran MI, Gower BA. Relation between visceral fat and disease risk in children and adolescents. *Am J Clin Nutr* 1999;70:149S–156S.
35. Owens S, Gutin B, Ferguson M *et al*. Visceral adipose tissue and cardiovascular risk factors in obese children. *J Pediatr* 1998;133:41–45.
36. Daniels SR, Morrison JA, Sprecher DL, Khoury P, Kimball TR. Association of body fat distribution and cardiovascular risk factors in children and adolescents. *Circulation* 1999;99:541–545.
37. Fox K, Peters D, Armstrong N, Sharpe P, Bell M. Abdominal fat deposition in 11-year-old children. *Int J Obes Relat Metab Disord* 1993;17:11–16.
38. Nordström A, Neovius MG, Rössner S, Nordström P. Postpubertal development of total and abdominal percentage body fat: an 8-year longitudinal study. *Obesity (Silver Spring)* 2008;16:2342–2347.
39. Matsudo SMM, Matsudo VKR. Self-assessment and physician assessment of sexual maturation in Brazilian boys and girls: concordance and reproducibility. *Am J Hum Biol* 1994;6:451–455.
40. Roemmich JN, Clark PA, Mai V *et al*. Alterations in growth and body composition during puberty: III. Influence of maturation, gender, body composition, fat distribution, aerobic fitness, and energy expenditure on nocturnal growth hormone release. *J Clin Endocrinol Metab* 1998;83:1440–1447.
41. Rosenbaum M, Leibel RL. Clinical review 107: role of gonadal steroids in the sexual dimorphisms in body composition and circulating concentrations of leptin. *J Clin Endocrinol Metab* 1999;84:1784–1789.
42. Lovejoy JC, Sainsbury A; Stock Conference 2008 Working Group. Sex differences in obesity and the regulation of energy homeostasis. *Obes Rev* 2009;10:154–167.

Psychological Bulletin
1985, Vol. 98, No. 2, 260–282

0033-2909/85/$00.75

Gender Differences Across Age in Motor Performance: A Meta-Analysis

Jerry R. Thomas
Departments of Physical Education and Psychology
Louisiana State University

Karen E. French
Department of Physical Education
Louisiana State University

A meta-analysis was conducted to examine gender differences in motor performance during childhood and adolescence. Data were 64 studies yielding 702 effect sizes based on 31,444 subjects. Age was regressed on effect size, and the relation was significant for 12 of 20 tasks. Several types of age-related curves were found; the curve for a throwing task was the most distinctive. Five of the tasks followed a typical curve of gender differences across age. For eight tasks, gender differences were not related to age, and effect sizes were small. Results are discussed relating the development of gender differences to biological and environmental sources.

Across the childhood and adolescent years, gender differences have been reported in performance for many motor tasks. A typical description is one in which female and male performance differences are slight but favor males in early childhood. Performance rapidly accelerates linearly across childhood, with boys maintaining a slight but increasing advantage. At puberty, female performance levels off, whereas male performance continues to improve and may even accelerate.

How large are gender differences in motor performance? Can their rate of development be estimated from literature that are basically cross-sectional in nature? If a description of one or more gender difference curves across age is developed, can the sources of the differences be inferred from descriptive data? The purpose of this article is to evaluate by meta-analysis the gender differences in motor performance across childhood and adolescence and to suggest possible sources of the differences. The obvious potential sources of explanation are biology, environment, and their interaction.

Anastasi (1981) indicated that to progress from merely describing sex differences to explaining them the function of heredity and environment (nature–nurture, biology–culture) had to be considered. For example, she suggested that the biological acceleration of girls is the mechanism to explain girls' more rapid acquisition of language skills. Anastasi noted girls' physical acceleration, but seemed to imply that boys' advantage in gross motor skills during infancy and childhood has some biological basis (p. 198). Why would girls not perform gross motor skills more effectively than boys if they are physically and psychologically accelerated? Given the small size of the prepubertal gender differences in most motor performance tasks, are environmental factors a more likely source?

Why are Gender Differences in Motor Performance Present?

Biology

The physical characteristics of boys and girls are very similar prior to puberty. In fact, gender sameness rather than difference is a more appropriate descriptor of biological characteristics such as body type, body composition, strength, and limb lengths (Malina, 1984). Thus, biology seems to offer little explanation for motor performance differences prior to puberty.

Girls have their peak growth spurt approximately 2 years earlier than boys. Ultimately, this results in the termination of long bone growth earlier, which causes girls, on the average, to be shorter than boys (Espenschade &

Karen E. French is now at the Department of Physical Education, University of South Carolina, Columbia.

Requests for reprints should be sent to Jerry R. Thomas, School of Health, Physical Education, Recreation, and Dance, Louisiana State University, Baton Rouge, Louisiana 70803.

Eckert, 1980). Also, boys during and after puberty produce increasing amounts of testosterone, which is closely associated with increased muscle tissue. The ratio of muscle to fat is similar for boys and girls prepuberty. However, after puberty, this ratio remains approximately the same for females but doubles for males (Malina & Johnson, 1967). On the average, boys become taller and heavier than girls after puberty (National Center for Health Statistics, 1977). Boys have more lean body mass and less fat (Burmeister, 1965), greater arm and calf circumference (Roche & Malina, 1983), broader shoulders and narrower hips (Roche & Malina, 1983), and greater midarm muscle circumference and smaller triceps skinfold (Frisancho, 1981). Thus, in any motor task for which size and strength are an advantage, adolescent boys will have a biological advantage in performance when compared with adolescent girls because boys are larger and have more muscle.

Environment

An important potential source of environmental influence on gender differences in motor performance is the child's perception, which evolves over time, of the appropriate gender role (Greendorfer, 1980). In particular, the child's family, peers, teachers, and coaches are potential sources for learning a gender role regarding motor skill performance. The process of gender-role identification has been attributed to three sources: imitation, socialization, and self-socialization (Maccoby & Jacklin, 1974).

Several studies have reported that, during preschool years, both parents tend to emphasize the development of gross motor behavior in boys more than girls (for a brief review, see Maccoby & Jacklin, 1974, pp. 307–311). This included the fathers engaging in more rough play with boys and treating girls as more fragile. In particular, both parents (but especially fathers) reacted more negatively when boys chose to play with dolls than when girls chose "rough and tumble" games (Fling & Manosevitz, 1972; Lansky, 1967), a fact which seems pertinent to the development of gender differences in motor skills during early childhood. This early gender difference in motor performance may also be influenced by parents' subtle messages that gross motor activities and some types of toys are more appropriate for boys than girls (Fagot, 1978). Thus, the gender differences in motor performance that are present as children enter elementary school may be largely socialized by parents, either by subtle coercion or by the children modeling what is perceived as sex-appropriate behavior.

If Sherif and Rattray (1976) are correct, physical education teachers and coaches in organized sport programs have treated the gender differences in motor performance exhibited in early childhood as naturally evolving biological factors. Even professional organizations such as the American Alliance for Health, Physical Education, Recreation and Dance (AAPHER) may contribute to gender differences by publishing separate norms for elementary school boys and girls for their physical fitness tests (AAHPER, 1976). This may lead to different expectations for boys and girls by teachers, coaches, parents, and the children themselves. As previously suggested, these motor performance differences are generally not large by the time children enter kindergarten or first grade and are most likely created by social factors. The treatment of differences as natural and biological may serve to increase them in many motor performance tasks during the elementary school years. Greendorfer (1980) and Housner (1981) provided interesting reviews and speculations about the role of physical educators and coaches in the development of gender differences in motor performance.

At puberty, gender differences in motor performance appear to be influenced by both biological and environmental factors. Although we must acknowledge the importance of biological changes as being closely associated with increasingly larger gender differences in many motor performance tasks, environmental factors may assume even greater importance at and following puberty. Boys are expected to be more masculine and girls more feminine, whereas in prepuberty some "tomboy" type behavior may have been socially acceptable in girls (Maccoby & Jacklin, 1974). Because of social pressures to conform, girls may be less inclined to participate in athletic activities and less motivated to perform well on motor tasks they do attempt. Thus, true gender differences may be overestimated.

A summary suggests that prior to puberty most gender differences in motor performance are socially induced by parents, peers, teachers, and coaches, although differences are by no means uniform and may include some type of gender-related predisposition toward certain motor tasks (a view originally suggested by Maccoby & Jacklin, 1974, pp. 363–364, as a framework to evaluate sex differences in psychological factors). Even though environmental pressures may be greater after puberty, biology plays an important role in the development of gender differences in many motor performance tasks, specifically those for which size and force production are important.

Gender Difference Inferences

We believe that quantifying the gender differences across age levels from early childhood to late adolescence provides a basis for theorizing about environmental and biological causes of any observed differences. If gender differences are small (even though reliable) across early childhood but begin to increase during the elementary school years, environment seems the more likely cause. This is particularly true if a remedial program eliminates most of the gender differences. The basis for this explanation is the previously presented evidence that parents and peers influence the early development of motor skills around sex-role models. These differences may be viewed as natural by teachers and coaches, who continue to contribute to their gradual increase during the elementary school years.

Biological factors seem to be implicated if large gender differences are noted during early childhood, particularly if these differences can be corroborated by cross-cultural findings and are difficult to reduce by training. However, even differences in biology seem to be subject to environmental reinforcement and enhancement. This is evident when a large difference noted in early childhood continues to increase during the elementary school years.

The influence of biology may also be evident in many motor tasks at puberty when male performance accelerates while female performance levels off, especially if task performance is enhanced by increased strength and size. But more likely a biology–environment interaction is involved, as social pressures are intense to conform to the feminine or masculine sex role.

"For obvious ethical and practical reasons, research on human subjects cannot expose individuals to drastic and long-lasting variations in living conditions" (Anastasi, 1981, p. 188). Thus, we are forced to use relations found in large bodies of developmental data to infer cause–effect. This methodology is essentially the same as cross-cultural research; the source of the variation is simply vertical (across time) rather than horizontal (over cultures). However, the reader should be aware that these data are only correlational, making cause–effect conclusions tenuous.

Use of Meta-Analysis

Meta-analysis (Glass, McGaw, & Smith, 1981; Hedges & Olkin, 1985) offers an objective way to evaluate the large body of literature in which gender comparisons of motor performance are reported during childhood and adolescence. In many studies, motor performance may be measured in slightly different ways. For example, running speed may be tested over varying distances. If we express male–female differences in a standard metric (standard deviation units) effect size, these studies may be combined quantitatively in a meaningful way. Thus, running speed and its development can be analyzed as a concept rather than having to consider the individual measures of running speed.

Effect size provides a way of judging the size and meaningfulness of gender differences in motor performance. Because effect size is in standard deviation units, the size of the differences and degree of overlap between the distributions can be estimated. Cohen (1969) provided a means by which to judge effect sizes: 0.20 = small, 0.50 = medium, and 0.80 = large.

The use of meta-analysis eliminates two methodological problems. First, meta-analysis overcomes the problem (Jacklin, 1981) of doing regressions within gender and then inferring between-gender differences when predictors are entered in one gender's prediction equation but not the other. The differences between genders are directly calculated as effect sizes prior to regressing age and study char-

acteristics on them. Second, meta-analysis allows the quantitative integration of a large number of cross-sectional and longitudinal studies of gender differences in motor performance. We believe that having a large number of effect sizes of gender differences in motor performance at a specific age provides the best available estimate for any true differences that may exist.

Thus, we believe that this meta-analysis of age-related changes in gender differences in motor performance provides the means to integrate and describe a large data set. In addition, it has considerable potential to further define theory about the development of sex differences, as well as to overcome some of the related methodological issues. In the following section, we have provided the details the reader needs for evaluating our work.

Method

Selection of the Data

Data were from studies that reported gender differences on motor performance during childhood and adolescence. Motor performance is defined as the outcome of movement (e.g., how fast a subject runs, how far a subject throws). In this study, motor performance was delimited to include (a) fundamental skills such as throwing, catching, running, striking, and jumping; (b) basic abilities such as balance and fine eye-motor coordination; (c) motor fitness items such as agility, arm strength, grip strength, flexibility, shuttle run, and sit-ups; and (d) information-processing responses such as reaction time, pursuit rotor tracking, and anticipation timing. The variables selected for study are those generally reported in studies and summaries in the motor-development literature (e.g., Corbin, 1980; Ridenour, 1978; Thomas, 1984; Wickstrom, 1983). Specifically excluded from the motor performance definition were cardiovascular fitness measures such as distance runs, step tests, and laboratory tests (treadmill walking and cycle ergometer riding).

Within the constraints of this definition of motor performance, a literature search was conducted using two computer bases—Educational Resources Information Center (ERIC) and PsycINFO. Health, Physical Education, and Recreation (HPER) Microform Publications (dissertations and theses) were searched by hand, as were eight journals identified as being likely to contain appropriate research reports: *Child Development, Developmental Psychology, Journal of Experimental Child Psychology, Journal of Human Movement Studies, Journal of Motor Behavior, Journal of Sport Psychology, Perceptual and Motor Skills,* and *Research Quarterly for Exercise and Sport.* The first study located was conducted in 1899, and the literature was searched through 1983.

From this search 176 studies were identified for initial consideration. Of these, 40 (23%) were from unpublished sources, mostly theses and dissertations. Only 64 (36%) of the studies could be included in the meta-analysis. (The Appendix provides citations for each of these articles.) Of the 112 (64%) studies excluded (a list can be obtained from the first author), 32 (18%) had collapsed the design and data across age levels of more than 3 years, 45 (26%) had used both girls and boys but had not provided specific information about gender ratios, 31 (18%) did not provide the minimal information necessary to calculate effect size, and 4 (2%) were eliminated because the tasks were dropped. To be included a study had to provide male–female comparisons within the childhood and adolescent years.

Data Coded From Each Study

Effect sizes (ESs) were calculated for all male–female comparisons at each age level for motor performance tasks reported. An ES was obtained for a given task at a specific age by subtracting the mean for girls from the mean for boys and dividing by the standard deviation (Glass et al., 1981; Hedges & Olkin, 1985). When the mean for boys represented better performance, the ES was positive. When the girls averaged a better performance, the ES was negative.

To calculate ES, a pooled estimate of the standard deviation (weighted for sample size) was used (Hedges & Olkin, 1985). When means and standard deviations were provided, these techniques were used to estimate ES. When complete data are not available, ES can be estimated given some minimal statistical information. These procedures were applied to estimate ES when possible (Glass et al., 1981; Hedges & Olkin, 1985).

A common practice in gender difference research is to collapse the design across gender if the test of gender is not significant. Frequently, when this is done neither the means and standard deviations nor F ratios are provided. The author could be contacted and asked to provide the information, but this technique has proved unsuccessful in the past (Hyde, 1981). The ES of interest could be discarded, but this tends to bias the data to studies reporting significant gender differences. Finally, an ES as 0 could be used because by definition the statistical test is evaluating the null hypothesis of no reliable difference between the means. The data reported are with 0 ES included because the difference with and without was trivial.

Some studies included boys and girls in the sample but did not test the gender effect or report means and standard deviations by gender. These articles were excluded because there is no way to estimate an ES.

Effect size is still a biased estimator when sample sizes are small. Hedges (1981) and Hedges and Olkin (1985) provided a correction factor for small ESs. This factor—$c = 1 - 3/\{4(n^m + n^f - 2) - 1\}$, where n^m = number of boys and n^f = number of girls—is multiplied by the ES, thereby correcting overestimation of ES with a small sample size. All ESs (now labeled ES') in this study were corrected for bias using this formula.

In addition to ES', omega-squared (ω^2) was calculated for each boy–girl comparison using the formula provided by Tolson (1980). The use of ω^2 allows an estimate of the percent of variance accounted for by gender in the dependent measure. Or ω^2 estimates how well a specific motor performance task can be predicted if the gender of the subject is known. Thomas and Nelson (1985) discussed the need for considering the magnitude of the difference

(ω^2) as well as the reliability (significance) of the difference.

Our major interest was to relate the ES′ for gender differences in 20 motor performance tasks to age. However, five additional characteristics of each study were coded, and their relation to ES′ was calculated.

1. *Internal validity* was coded as high, average, or low, using qualities of the study such as representativeness of the sample, quality of the motor performance test selected, and appropriateness of the measurement schedule.
2. *Gender of the first author* was used, based on literature suggesting that the gender of the experimenter is sometimes related to gender differences in motor performance (Rikli, 1976). This assumes that the first author is the experimenter, an assumption that may be incorrect in some instances. Thus, any relations found should be treated with caution.
3. *Type of manuscript* refers to published versus unpublished articles. Theses/dissertations were considered unpublished. If a thesis/dissertation was later published, it was coded as published and eliminated from the unpublished list.
4. *Year published* was coded as studies published before 1970 or 1970 and later. This date was selected because considerable change in sex roles occurred in the 1960s and 1970s (Anastasi, 1981, p. 202). More specifically, Title IX was passed in 1972. Thus, we wanted to see if changes in sex roles and opportunities were reflected in gender differences in motor performance.
5. *Number of male and female subjects in each study* was coded.

Characteristics of the Data

Sixty-four studies were included in the analysis. Several of the studies tested the subjects on more than one motor performance task: 16 studies had data on each subject for more than three motor tasks; 8 had three tasks; 14 had two tasks; and 26 had only one task. To some extent, when more than one measure is taken on each subject, task performance is correlated. However, these correlations are generally low for the tasks included, so we treated the tasks as if they were independent. We believe this is a valid assumption. The correlation of tests of the same characteristic (e.g., balance) are generally low (Johnson & Nelson, 1979). The specificity hypothesis (Henry, 1968) indicates that motor abilities are task specific, and that two similar tasks (e.g., throwing a football and a baseball) tend to have a correlation of zero. This is because the abilities underlying these tasks are different. Thus, correlations among different motor performance tasks (e.g., throwing and running) are expected to be low, and evidence generally supports this hypothesis (Schmidt, 1982, p. 401).

The 64 studies included 31,444 subjects of which 15,926 were boys and 15,518 were girls. Of the 64 studies included, 24 used fewer than 100 subjects, 33 used between 101 and 600 subjects, 6 used between 601 and 1,000 subjects, and 1 used more than 9,000 subjects. Several studies measured subjects on more than one motor performance task. Table 1 shows the number of measurements made for boys and girls by age level for the 64 studies.

Combining Effect Sizes

Hedges (1981) and Hedges and Olkin (1985) showed that, because of the way ES′ is distributed, the variance (s^2) of each ES′ can be calculated directly, and they provided the formula. From this formula, the observation can be made that the accuracy of ES′ is a function of sample size and ES′. Because the accuracy varies with this function, each ES′ should be weighted by the reciprocal of its variance prior to combining ES′ (Hedges, 1982; Hedges & Olkin, 1983, 1985). Thus, studies based on larger samples should be more precise estimators of true gender differences and, therefore, are given greater weight in the averaging of ES′. An overall or weighted mean estimate can be obtained by the formula provided by Hedges and Olkin (1985). An estimate of the variance for the weighted mean ES′ is obtained from the bottom half of the same formula. These formulas were used in combining and estimating the variance of the $\overline{ES}$′.

Results

We calculated the weighted mean of ES′ for each gender comparison for all tasks. Then we examined the number of ES′s for each task for all studies. When the number of studies was less than three, the task was eliminated, with one exception: Several tasks—hole punching, peg shifting, manual dexterity, tracing, and turning a screw (called speed prehension)—were combined and called fine eye–hand coordination. Table 2 lists those tasks that met the criteria for retention and those that were eliminated. Table 3 contains the number of measurements made on male and female subjects for each of the 20 tasks that were retained for the analysis.

Table 1
Number of Measurements for Boys and Girls by Age Level Included in the Meta-Analysis

Age (years)	No. for boys	No. for girls	Total
3	316	259	575
4	863	805	1,668
5	1,855	1,816	3,671
6	1,980	1,733	3,713
7	6,325	5,808	12,133
8	3,378	3,578	6,956
9	2,374	2,332	4,706
10	4,256	4,054	8,310
11	6,371	6,012	12,383
12	4,696	4,892	9,588
13	5,221	5,447	10,668
14	3,021	3,108	6,129
15	3,369	3,112	6,481
16	3,256	3,162	6,418
17	3,647	2,865	6,512
20	142	142	284
Totals	51,070	49,125	100,195

Table 2
Motor Performance Tasks Included and Excluded

Task	No. studies	No. ES's
Included		
Agility	9	19
Anticipation timing	3	23
Arm hang	3	16
Balance	14	71
Catching	4	25
Dash	19	82
Fine eye–hand	5	30
Flexibility	5	13
Grip strength	4	42
Long jump	19	85
Pursuit rotor	5	14
Reaction time	6	42
Shuttle run	7	33
Sit-ups	7	36
Tapping	4	34
Throw accuracy	5	14
Throw distance	11	58
Included (*continued*)		
Throw velocity	5	13
Vertical jump	5	20
Wall volley	4	32
Excluded		
Choice RT	2	10
Dribble	1	2
8 choice RT	2	21
4 choice RT	2	21
Free throw	1	4
Hurdle jump	2	7
Leg torque	1	5
Movement time	2	11
Pull-ups	2	13
Push-ups	2	7
Squat thrust	1	8
Striking	2	15
2 choice RT	2	21

Note. RT = reaction time. Total number of studies included = 144; this is more than the total number of studies because some contained more than one task. Number of ES's for studies included = 702. Total number of studies excluded = 22. Total number of ES's for studies excluded = 145.

The first issue is to test if ES' has homogeneity within each task (Hedges & Olkin, 1985). This test (*H* statistic) is the total sum of squares tested as a χ^2 using the number of ES's minus 1 as the degree of freedom. This test was significant ($p < .05$) for 9 of the 20 tasks (balance, grip strength, shuttle run, throw for velocity, vertical jump, dash, long jump, sit-ups, and throw for distance). However, we hypothesized a priori that a major source of this lack of homogeneity would be the large age range in the data. Other possible sources of a lack of homogeneity are outliers and study characteristics. Therefore, we used a weighted regression as outlined by Hedges and Olkin (1983, 1985, chap. 8) to examine the correlation of ES' to age within each task. We allowed the fit of the regression line to be linear (age), quadratic (age^2), or cubic (age^3).

Tasks in Which Gender Differences in Motor Performance and Age Were Related

The first step was to identify which tasks had ES's for gender that were age related. Table 4 provides a summary of the regression statistics for the 12 tasks (out of 20) that correlated significantly with age. (The *H* statistic for these 12 tasks is the sum of the χ^2 for the regression plus the χ^2 for within tested against the sum of the degrees of freedom for each.)

Table 3
Number of Measurements by Gender for the 20 Motor Performance Tasks

Task	No. for boys	No. for girls	Total
Agility	1,869	1,643	3,512
Anticipation timing	303	433	736
Arm hang	268	234	502
Balance	2,269	2,122	4,391
Catching	742	698	1,440
Dash	8,191	7,830	16,021
Fine eye–hand	1,035	982	2,017
Flexibility	1,074	995	2,069
Grip strength	1,446	1,338	2,784
Long jump	9,406	8,953	18,359
Pursuit rotor	245	237	482
Reaction time	1,478	1,597	3,075
Shuttle run	5,953	5,714	11,667
Sit-ups	5,805	5,545	11,350
Tapping	759	706	1,465
Throw accuracy	688	624	1,312
Throw distance	7,754	7,558	15,312
Throw velocity	480	465	945
Vertical jump	384	423	807
Wall volley	921	1,028	1,949

Hedges and Olkin (1983, 1985) argued that the typical way of interpreting this regression is inappropriate. They indicated that the first step is to determine if the regression is significant by treating the sum of squares for the regression as a χ^2 with the degrees of freedom associated with the regression (model). Note in Table 4 that the column labeled χ^2_{reg} is this test with its associated degrees of freedom.

Once the regression is declared significant, a test should be made to determine if the model is correctly specified. "The test for model specification provides a basis for deciding whether the variation in effect magnitude is accounted for by the explanatory variables in the model" (Hedges & Olkin, 1983, p. 137). Hedges and Olkin (1983, 1985) indicated that the sum of squares for within (error) is the appropriate test of model specification.

This sum of squares is treated as a χ^2, with the degrees of freedom associated with the within factor (see Table 4, column labeled χ^2_{wi}). Note in Table 4 that we have placed the eight tasks (balance, catching, grip strength, pursuit rotor, shuttle run, tapping, throwing velocity, and vertical jump) where the model is correctly specified in the top part of the table and the four tasks (dash, long jump, sit-ups, and distance throw) where the model is incorrectly specified at the bottom of the table. The fact that the data do not fit the model means that the results should be viewed with caution.

As indicated by the data in Table 4, the line of best fit is linear (see column labeled Age where the top value is the beta and the bottom value is the z score of the beta, all corrected according to procedures indicated by Hedges & Olkin, 1985) for balance, pursuit rotor, tapping, throwing velocity, and throwing distance. The best fit was quadratic (column labeled Age^2) for catching, grip strength, shuttle run, vertical jump, dash, long jump, and sit-ups.

For the figures that follow, the weighted $\overline{ES}'$ is bounded by a 95% confidence interval for the 12 tasks in which gender differences are related to age. Clearly, throwing for either velocity or distance (Figure 1) is very different from the other 10 motor performance tasks. Boys exceed girls in throwing velocity by 1.5 standard deviation units as early as 4 to 7 years of age. This difference rapidly increases so that by 12 years of age (last age with sufficient data for a comparison) the boys' performance exceeds the girls' performance by 3.5 standard deviation units. Throwing for distance, which to a large extent is dependent on the velocity of the throw, follows a similar pattern. The boys exceed the girls by 1.5 standard deviation units as early as 2 to 4 years of age. The increase is linear through puberty to 17 years of age when male performance is 3 standard deviation units better than female performance. The weighted $\overline{ES}'$ collapsed over age is 1.98 standard deviation units for throwing velocity and 2.18 for throwing distance. These data indicate that there is very little overlap in the distribution of throwing performance for boys and girls.

Balance, pursuit rotor tracking, and tapping (Figure 2) seem to have similar patterns. Basically, the gender difference is zero until puberty (10 to 13 years of age) when the male performance becomes better than female performance: Weighted $\overline{ES}'$ for balance is about 1 standard deviation unit, about 0.75 for pursuit rotor tracking, and about 0.50 for tapping. However, when averaged across age, the weighted $\overline{ES}$'s are relatively small: 0.09 for balance, 0.11 for pursuit rotor tracking, and 0.13 for tapping.

Five of the remaining tasks follow what has become regarded as the typical pattern for a motor performance task across childhood and adolescence. The weighted $\overline{ES}$'s are slightly in favor of the boys in early childhood by 0.25 to 0.50 standard deviation units. This difference increases somewhat during middle childhood to 0.50 to 1.00 standard deviation units. The difference increases to between 1.00 and 2.00 standard deviation units after puberty. The tasks following this pattern are the dash and situps (Figure 3) and the long jump, grip strength, and shuttle run (Figure 4).

For the vertical jump (Figure 5) the weighted $\overline{ES}'$ is essentially zero until puberty, when it increases to more than 1 standard deviation unit in favor of boys. Catching (Figure 5) follows a U-shaped pattern; boys perform about 0.75 standard deviation units better than girls during preschool. Then the differences become smaller, dropping to 0.25 units until postpuberty, when they again go up to 0.75 units. We have no data points after 13 years of age.

The model was not correctly specified for the shuttle run (Table 4) until three additional variables were included to reduce the within-

Table 4
Age-Related Studies: Regression Summary for ES′

Task	No. studies	No. ES′	$\overline{ES'}$	χ^2_{reg}	χ^2_{wi}	Age	Age^2	Validity of study	Year pub.	Gender of 1st author	Type article
Studies for which model is correctly specified											
Balance	14	67	0.09	42.9**	76.0	0.067**					
				1 *df*	65 *df*	6.55 *z*					
Catching	4	23	0.43	13.4**	15.6	−0.226*	0.019**				
				2 *df*	20 *df*	1.96 *z*	2.62 *z*				
Grip strength	4	37	0.66	78.4**	45.2	−0.419**	0.024**				
				2 *df*	34 *df*	5.24 *z*	6.43 *z*				
Pursuit rotor	5	14	0.11	5.0*	4.5	0.125*					
				1 *df*	12 *df*	2.23 *z*					
Shuttle run	7	28	0.32	673.4**	29.6	−0.002	0.006**	1.40**		−1.96**	1.91**
				5 *df*	22 *df*	0.034 *z*	2.39 *z*	14.22 *z*		9.40 *z*	11.44 *z*
Tapping	4	34	0.13	11.0*	21.5	0.080**					
				1 *df*	32 *df*	3.31 *z*					
Throw velocity	5	12	2.18	42.2**	11.3	0.454**					
				1 *df*	10 *df*	6.49 *z*					
Vertical jump	5	20	0.18	62.3**	16.6	−0.607	0.042**				
				1 *df*	17 *df*	1.95	2.62				
Studies for which model is incorrectly specified											
Dash	19	66	0.63	1061**	103**	−0.298**	0.021**				0.383**
				3 *df*	62 *df*	8.42 *z*	2.91 *z*				7.42 *z*
Long jump	19	68	0.54	1212**	123**	−0.337**	0.023**				
				2 *df*	65 *df*	4.95 *z*	16.45 *z*				
Sit-ups	7	29	0.64	697**	79**	0.799**	−0.022**		−0.975**	−0.127**	
				4 *df*	24 *df*	2.87 *z*	4.89 *z*		10.0 *z*	5.80 *z*	
Throw distance	11	47	1.98	433**	87**	0.153**					
				1 *df*	45 *df*	20.8 *z*					

Note. ES = effect size. Year pub. = published before 1970 or not. Type article = published/unpublished article. *z* = *z* score.
* $p < .05$. ** $p < .01$.

sum of squares: the validity of the study (column labeled Validity in Table 4), gender of the first author, and type of article (published or unpublished). An interpretation of these characteristics is that larger weighted $\overline{ES}$'s for gender (better male performance) were associated

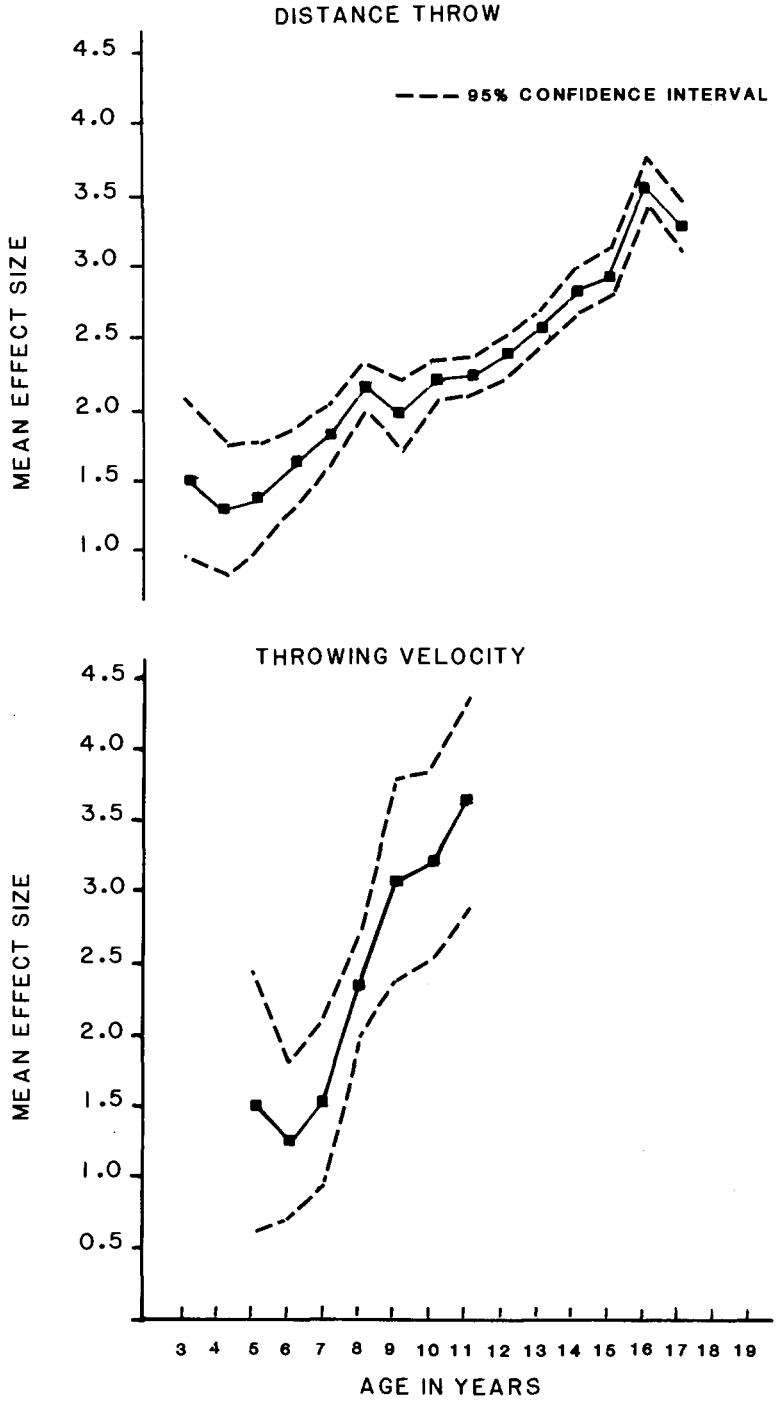


Figure 1. $\overline{ES}$' by age and gender for throwing for velocity and distance.

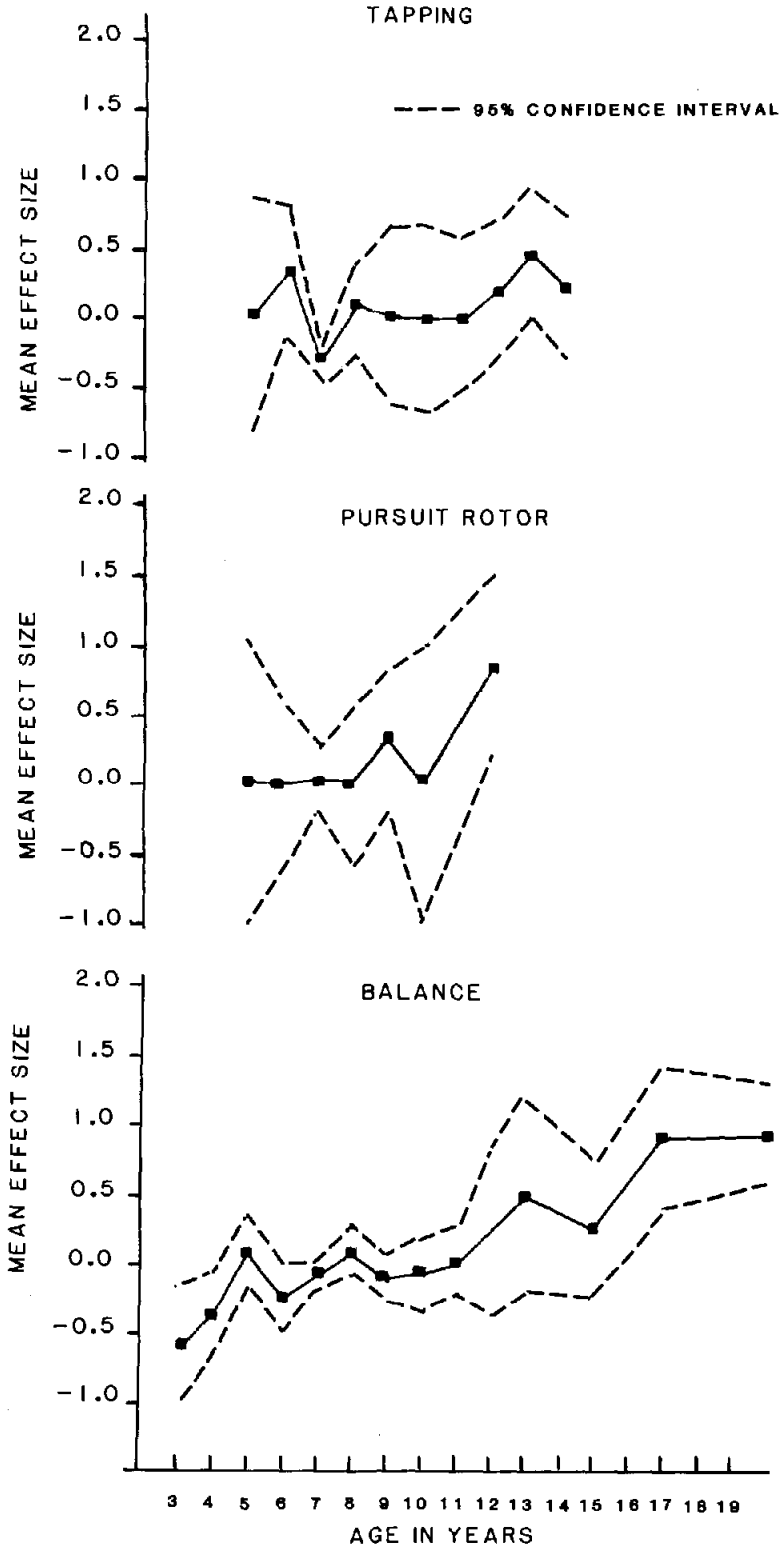


Figure 2. $\overline{ES}$ by age and gender for balance, pursuit rotor tracking, and tapping.

with studies published since 1970, having higher internal validity, and women as first author.

For the tasks in which the model was not correctly specified, a significant amount of variance was accounted for by some of the coded characteristics in addition to age. Sit-ups had larger weighted $\overline{ES}$'s associated with studies published before 1970 and studies in which women were first author. The dash had

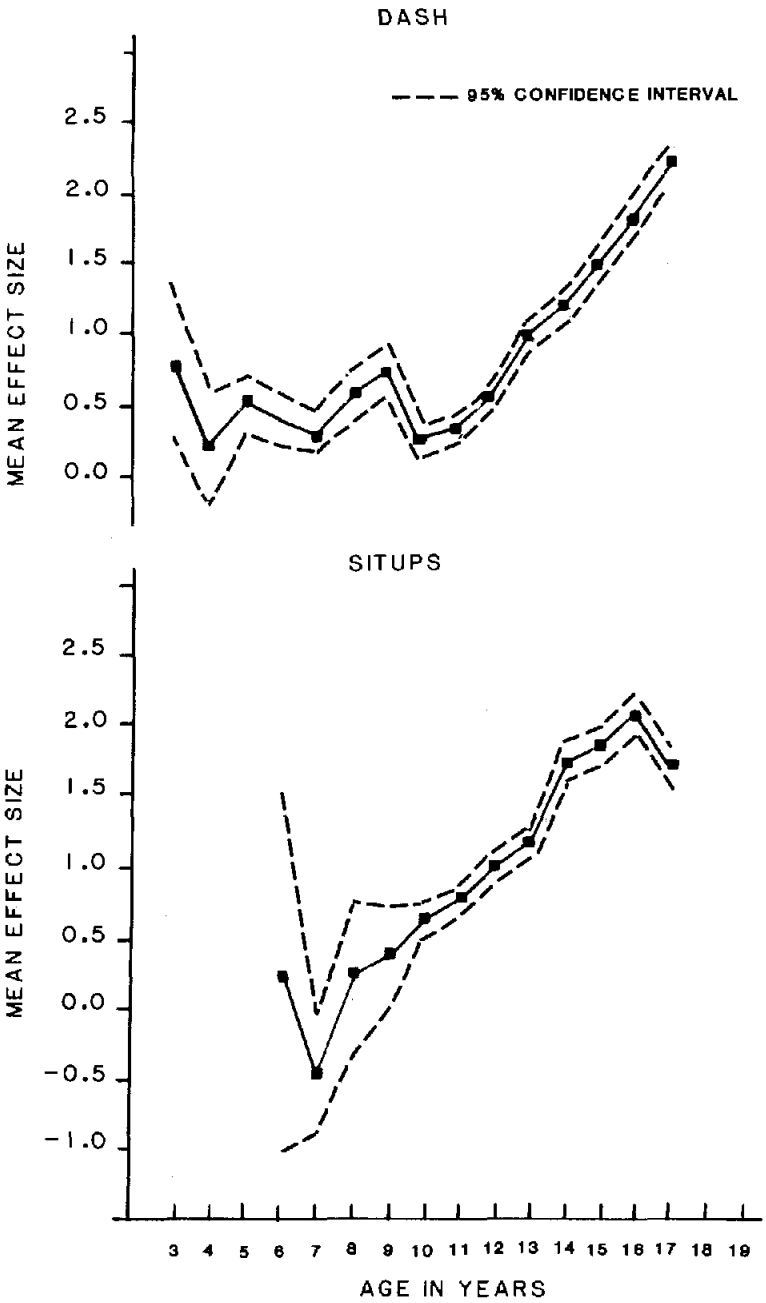


Figure 3. $\overline{ES}'$ by age and gender for dash and sit-ups.

larger weighted $\overline{ES}$'s associated with published as opposed to unpublished articles.

In any data set, outliers (extreme and unrepresentative values) are likely to exist, even when data points are ES's rather than individual performances. Outliers are particularly likely in data sets including young children (Thomas, 1984). An appropriate way to test for outliers in a meta-analysis using regression is to output the residuals from the regression, take the absolute value of the residual, calculate the mean and standard deviation of the residuals, and change the residuals to z scores (Hedges & Olkin, 1985). By definition (because the mean of a z score is 0 with a standard deviation of 1) any standardized residual larger than 2 is an outlier because it falls outside 95% of the score distribution. This procedure was used to identify outliers in the age-related data set. The data in Table 4 have the outliers eliminated. Table 5 includes the number and percentages of outliers eliminated for each task. An upper limit of eliminating 20% of the data was set and the criterion for declaring a data point an outlier was established as a z score of 2.

Although allowing 20% of the data to be declared as outliers may seem high, Table 5 shows that half of the tasks (6 of 12) actually had less than 10% of the ES's eliminated as outliers, only 4 tasks had more than 15% of the ES's eliminated, and only 1 task (long jump) actually reached the 20% maximum. Note that the four tasks (dash, long jump, sit-ups, and throw for distance) that had the highest percentage of outliers (19%–20%) were the same four for which the model was incorrectly specified (Table 4). This suggests that ES's for gender for these four tasks are not derived from the same population of ES's. Procedural differences among studies using the same general tasks may be the cause (e.g., dashes of varying lengths and sit-ups with time limits as compared with sit-ups with no time limits).

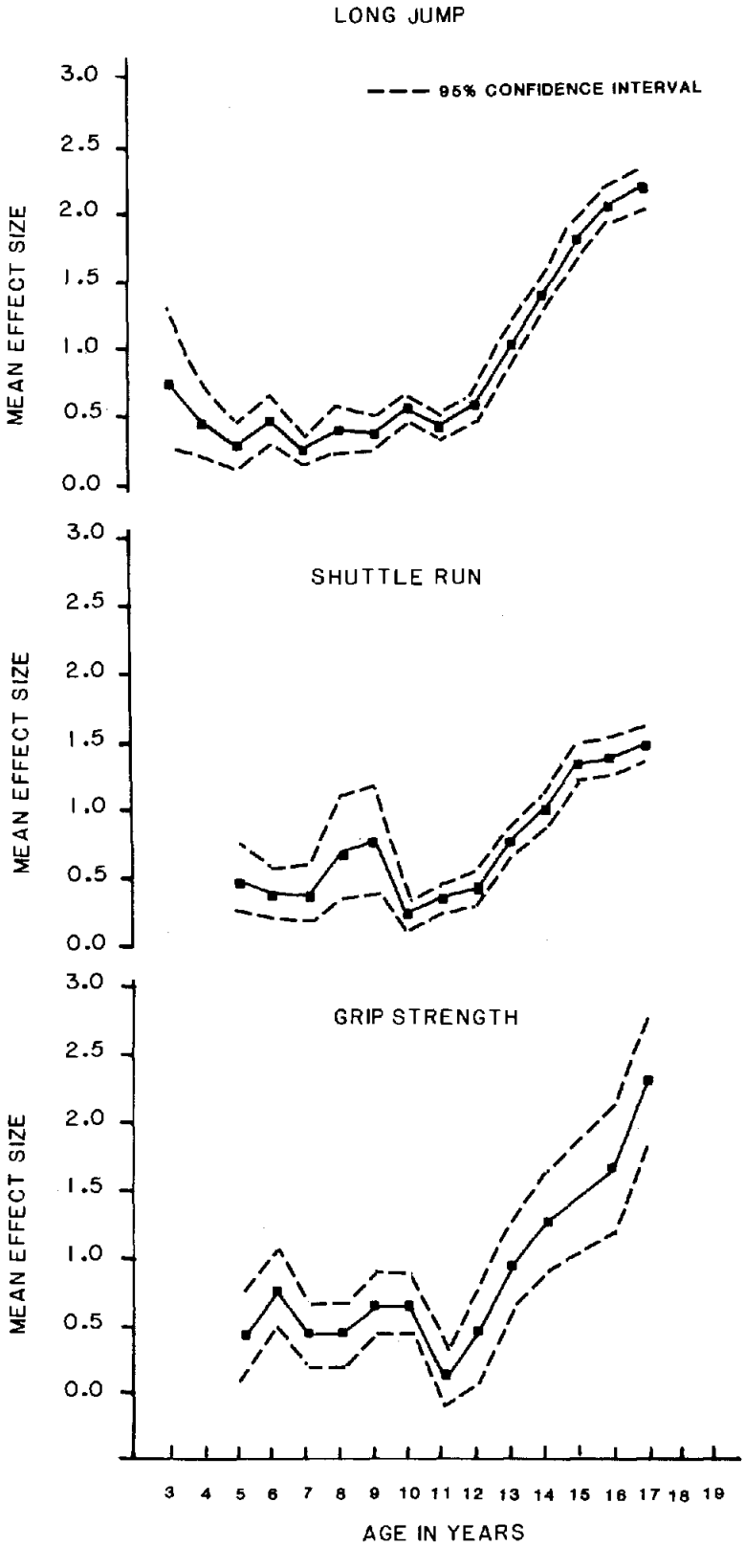
Hedges and Olkin (1985) believed the use of omega-squared (ω^2) as the dependent variable in a meta-analysis was inappropriate because the direction of the gender difference was not specified. However, the procedure is not without precedent (Hyde, 1981), and we believe analyzing ω^2 is useful because it represents the strength of the association between gender and the various motor performance tasks. Table 6 has a summary of the results of a standard regression analysis of age on ω^2. We have included the 12 tasks that were related to age in the earlier analysis of the weighted ES'. This table provides data for each task in which ω^2s were significantly related to age in a linear (shuttle run, sit-ups, throwing distance, and throwing velocity) or quadratic (balance, catching, dash, grip strength, long jump, pursuit rotor, and vertical jump) manner, or in one instance (tapping) in which the task was not related to age. The ω^2 for gender accounted for the most variance (averaged across age levels) in the two throwing tasks: 51% in throwing velocity and 47% in throwing distance. In other words, at any age about one half of the variance in throwing performance between boys and girls can be estimated by knowing the performer's gender.

The developmental patterns for ω^2 are similar to those for ES'. Thus, we have not included them. For example, in balance, catching, dash, grip strength, long jump, sit-ups, and vertical jump, little variance ($\omega^2 < .10$) is accounted for until about puberty. At puberty, there is a rapid rise in variance accounted for with ω^2 between .20 and .65. Thus, knowing the gender of the subject postpuberty allows a reasonable prediction of performance, especially in speed (dash and shuttle run), strength and endurance (grip strength and sit-ups), and power (long jump and vertical jump) motor performance tasks. The shuttle run is slightly different from the other tasks, accounting for a great amount of variance prepuberty (10%–20%).

As we showed with ES', throwing for either distance or velocity is clearly very different from the other types of motor performance tasks. Gender accounts for 35% to 60% of the variance before puberty and 60% to 70% after puberty. Given that gender is a dichotomous variable, the postpuberty prediction is approaching the upper limit of this relation. For tapping and pursuit rotor tracking, gender accounts for none of the variance prior to puberty and very little ($\omega^2 < .15$) after puberty.

Tasks for Which Gender Differences in Motor Performance Were Not Age Related

In Table 7 the summary data are reported for the motor performance tasks (8 of 20) for which gender differences were not related to



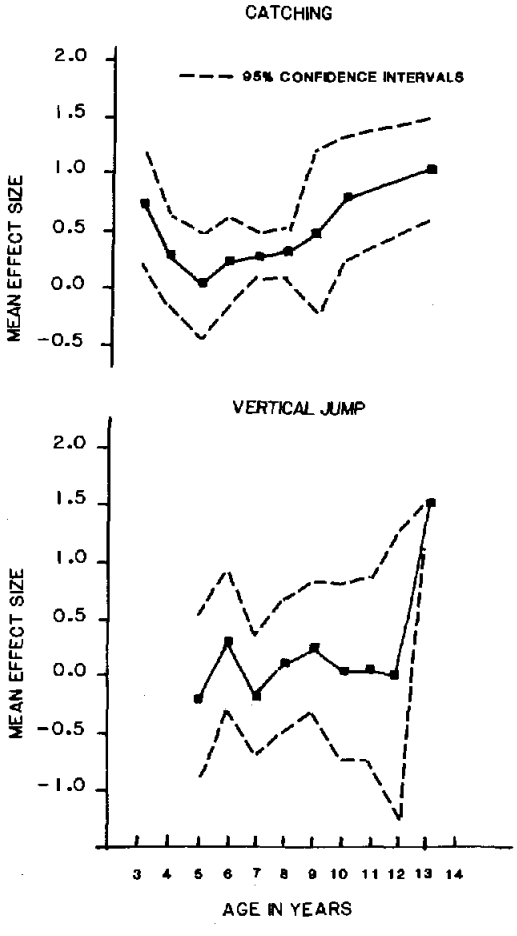


Figure 5. $\overline{ES}'$ by age and gender for vertical jump and catching.

age. The homogeneity statistic (H) is tested as a χ^2 with degrees of freedom as the number of ES's minus 1. The H was not significant ($p > .05$) for any of these tasks. The χ^2 test of the sum of squares for the regression was not significant ($p > .05$) for any of the 8 tasks.

The weighted $\overline{ES}'$ was large for throwing accuracy (0.96) and the wall volley (0.83), both being close to 1 standard deviation unit. The percent of variance accounted for (ω^2) by these two tasks was also the largest of this group. The weighted $\overline{ES}$'s were small (0.18 to 0.38) for agility, anticipation timing, flexibility, reaction time, and fine eye–hand coordination. Note that flexibility ($\overline{ES}' = -0.29$) and fine eye–motor coordination ($\overline{ES}' = -0.21$) were two tasks at which female performance was consistently better than male performance. The variance accounted for by gender in these five tasks was small ($\omega^2 < .08$). The arm hang showed no meaningful difference in $\overline{ES}'$ (0.01) or in variance accounted for ($\omega^2 = .01$).

Discussion

These data do not support the notion of uniform development of gender differences in motor performance across childhood and adolescence. Effect sizes were related to age in only 12 of the 20 (60%) motor performance tasks. These relations can be placed into several groups based on the shapes of the curves in Figures 1 to 5.

Tasks for Which Age and Gender Are Related

The major issue is how biology and environment influence the development of gender differences across childhood and adolescence. Can the nature of the developmental curves be used to make inferences about the causes of gender differences in motor performance?

Typical motor performance curve. If a typical motor performance curve exists (as described earlier) that reflects the development of gender differences, it is depicted in Figures 3 and 4 for the tasks of dash, sit-ups, long jump, grip strength, and shuttle run. In general these five tasks show small to moderate effect sizes (0.20 to 0.50) favoring boys in early childhood. The effect sizes remain moderate across the elementary school years for four of the tasks, but increase during elementary school for sit-ups. The girls, upon reaching puberty, appear slightly to reduce the difference on the dash, grip strength, and the shuttle run, but not in the long jump and sit-ups. When the boys reach puberty, all five tasks are influenced in the same way—they increase their advantage over girls rapidly until it is between 1.5 and 2 standard deviation units at 17 years of age. The nature of these five curves can be viewed with considerable trust because the 95% confidence intervals are tight.

Although girls on reaching puberty may close the performance difference on some of these five tasks, when boys reach puberty their

Figure 4. $\overline{ES}'$ by age and gender for long jump, grip strength, and shuttle run.

Table 5
Number and Percentage of Outliers by Task

Task	No. studies	No. ES's	No. ES's lost as outliners	No. ES's kept	% of ES's lost
Balance	14	71	4	67	5.6
Catching	4	25	2	23	8.0
Dash	19	82	16	66	19.5
Grip strength	4	42	5	37	13.5
Long jump	19	85	17	68	20.0
Pursuit rotor	5	14	0	14	0.0
Shuttle run	7	33	5	28	15.0
Sit-ups	7	36	7	29	19.0
Tapping	4	34	0	34	0.0
Throw distance	11	58	11	47	19.0
Throw velocity	5	13	1	12	8.0
Vertical jump	5	20	0	20	0.0

increase in size and muscle tissue is dramatically reflected in better performance across all five tasks. The average performance of boys and girls of 2 standard deviation units apart at age 17 shows that there is little overlap between the two distributions. This is an expected finding for tasks in which speed, power, muscular strength, and endurance are important.

The differences prior to puberty are more moderate and, we believe, more likely to reflect environmental influences. When differences between the mean performances of boys and girls are less than 0.50 standard deviation units, many of the girls are performing better than many of the boys. If equal expectations, encouragement, and practice opportunities were provided by parents, teachers, and coaches, differences of this size could probably be eliminated. Opportunities and encouragement to practice may be the key issue; Halverson, Roberton, and Langendorfer (1982) reported that seventh-grade boys remembered practicing throwing more over the years than did seventh-grade girls.

Several authors (Greendorfer, 1980; Housner, 1981; Sherif & Rattray, 1976) suggested that teachers and coaches take as biological the gender differences evident in motor performance when children begin elementary school. The implication is that these differences are maintained or increased because teachers and coaches have different expectations for boys than for girls. Although these data cannot speak directly to expectations, the differences

Table 6
Summary Data for Regression Using ω^2

Task	Age	Age^2	$\bar{\omega}^2$	R^2
Balance	$F(1, 64) = 22.24^{**}$	$F(1, 64) = 44.41^{**}$	.03	.63
Catching	$F(1, 20) = 12.37^{**}$	$F(1, 20) = 25.79^{**}$	.05	.81
Dash	$F(1, 63) = 38.01^{**}$	$F(1, 63) = 72.89^{**}$	.10	.79
Grip strength	$F(1, 34) = 26.48^{**}$	$F(1, 34) = 41.28^{**}$	.10	.74
Long jump	$F(1, 65) = 73.63^{**}$	$F(1, 65) = 127.19^{**}$	.08	.81
Pursuit rotor	*ns*	$F(1, 11) = 7.38^{*}$	.02	.70
Shuttle run	$F(1, 26) = 46.36^{**}$	*ns*	.12	.64
Sit-ups	$F(1, 27) = 22.93^{**}$	*ns*	.14	.46
Tapping	*ns*	*ns*	.01	.04
Throw distance	$F(1, 45) = 31.72^{**}$	*ns*	.47	.41
Throw velocity	$F(1, 10) = 30.89^{**}$	*ns*	.51	.75
Vertical jump	$F(1, 17) = 13.61^{**}$	$F(1, 17) = 17.47^{**}$	.03	.64

$^{*}p < .05$. $^{**}p < .01$.

Table 7
Summary Data for Motor Performance Tasks Where Gender Differences Were Not Related to Age

Task	Age range (years)	No. ES's	No. studies	H^a	$\overline{ES}'$	s^2	$\bar{\omega}^2$
Agility	3–17	19	9	27.15	0.21	.0019	.04
Anticipation timing	7–20	23	3	18.43	0.38	.0066	.07
Arm hang	3–12	16	3	7.86	0.01	.0100	.01
Fine eye–motor	3–10	30	5	35.82	−0.21	.0022	.04
Flexibility	5–10	13	5	19.01	−0.29	.0023	.02
Reaction time	5–20	42	6	38.90	0.18	.0013	.02
Throw accuracy	6–11	14	5	12.81	0.96	.0041	.24
Wall volley	7–13	32	4	21.31	0.83	.0023	.15

[a] H statistic is treated as a χ^2 with df equal to the number of ES's − 1. None of these are significant at $p < .05$.

present in these five tasks when children begin elementary school are maintained or increased.

Other motor performance curves. A second group of tasks (Figure 2)—balance, pursuit rotor tracking, and tapping—is similar in nature. Gender differences are not present during the elementary school years but increase to a moderate level ($\overline{ES}' = 0.50$) favoring boys at about 11 to 12 years of age. In our opinion, this small an increase is more likely to reflect increasing environmental pressures rather than any biological factor related to puberty. These three tasks do not seem to be related to the strength, endurance, and size increases noted in boys at puberty. The finding is particularly interesting for balance because balance is a task in which girls have been reported as performing better than boys (Roberton, 1984). Although girls do have a slight advantage in early childhood (3 to 4 years of age), the gender difference is essentially zero until puberty, when male performance is moderately better.

The vertical jump (Figure 5) is similar to balance, pursuit rotor tracking, and tapping in that there are no gender differences until puberty. At puberty boys show a large increase in performance so that they are more than 1 standard deviation unit better than girls. The lack of prepuberty gender differences favoring boys is unexpected. We expected that vertical jump performance would be similar to performance in the long jump, dash, and the shuttle run as reported in previous reviews (DeOreo & Keogh, 1980; Espenschade & Eckert, 1974; Wickstrom, 1983). We are hesitant to offer an explanation for this unexpected result because the 95% confidence intervals are very large prior to puberty, indicating that effect sizes show large variations about the mean.

The curve for catching (Figure 5) is different from all the others. Boys' performance exceeds girls' performance in early childhood (3 years of age), but the differences are reduced to zero by 5 years of age. Boys' performance increasingly exceeds the girls' during elementary school until the difference is 1 standard deviation unit by age 13. These differences are likely environmental due to boys having more practice. The effect sizes are also likely to be underestimated because the ball is generally projected with little velocity in the tests reported.

In our opinion, any gender differences in performance prior to puberty on these 10 tasks (balance, catching, dash, grip strength, long jump, pursuit rotor tracking, shuttle run, sit-ups, tapping, and vertical jump) are mostly environmentally induced. The attitudes, expectations, and actions of parents, teachers, peers, and age-group coaches either produce or reinforce the differences. This is reflected by the increased opportunity and encouragement to practice these tasks afforded to boys. We believe that differences this small ($\overline{ES}' < 0.50$) could easily be eliminated if girls and boys were treated similarly.

Six tasks (dash, long jump, sit-ups, grip strength, shuttle run, and vertical jump) show postpuberty changes that are probably related to biological development, even though environmental factors continue to be important. Although equal encouragement and opportunity to practice would probably reduce gender differences after puberty, boys would still, on the average, perform better than girls. For balance, pursuit rotor tracking, and tapping, the

biological changes associated with puberty do not seem to be the cause of the slight increase observed. We suspect that environmentally related variables such as competitive motivation and sex-role expectations that become increasingly significant for boys and girls create these differences. The differences in catching skills after puberty are probably not a good estimate of the true difference in sports in which the ball is projected with velocity (e.g., baseball, softball).

Throwing is different. Even a casual glance at Figure 1 indicates that throwing for velocity and distance are different from the other 10 tasks. The differences are 1.5 standard deviation units at 3 years of age. These differences increase substantially during the elementary school years. We only have data points through 12 years of age for throwing velocity, but the boys' throwing velocity is already between 3.5 and 4 standard deviation units higher than the girls'. The acceleration of gender differences is not so rapid for distance thrown but is above 2 standard deviation units at age 12 and above 3 standard deviation units after age 16. The 95% confidence intervals are moderately tight on throwing velocity and very tight on throwing for distance. Thus, we have considerable confidence in these developmental curves.

Although Maccoby and Jacklin (1974) discussed some differences in the treatment of boys and girls in early childhood, differences as large as 1.5 standard deviation units at age 3 are unlikely to be completely environmentally caused. A substantial amount of throwing practice for boys would be needed to produce a difference this large. However, the fact that this relatively large gender difference continues to increase in childhood and adolescence is probably a combination of biology and environment.

Malina (1984) reported that prepubescent boys have slightly more total lean body mass and less fat in both an absolute and relative sense than girls. In addition, sex differences in somatotypes during early childhood have been reported (Walker, 1962). Over 50% of girls have a larger endomorphic component in body type (as compared with ectomorphic and mesomorphic), whereas over 50% of boys have a larger mesomorphic component. Malina (1975) believed these sex differences were of genetic origin. These total body differences are also present specifically in the arm. Boys have a greater midarm circumference and a smaller triceps skinfold (i.e., more muscle tissue) than girls. Haubenstricker and Sapp (1980) reported that boys' forearms are 6 mm (on the average) longer than girls by 5.5 years, and that the differences continue throughout childhood. In addition, boys show an increasing advantage in late childhood in the biacromial/bicristal (shoulder/hip) ratio (Malina, 1984). Although each of these biological differences is small, taken together they may account for a portion of the gender difference in throwing performance. That biology is a factor in the differences is further confirmed by the fact that training in throwing has had little effect in reducing male–female performance differences in young children (Dusenberry, 1952; Halverson, Roberton, Safrit, & Roberts, 1977).

The increasing advantage boys gain over girls in throwing performance during elementary school is also partially attributable to environment. Both data (Halverson et al., 1982) and observation at playing fields indicate that elementary-age boys practice throwing skills much more than girls.

Thus, although gender differences in throwing velocity and throwing for distance can probably be reduced by providing equal encouragement and practice for girls, we believe that biological factors will not allow, on the average at least, their elimination. Therefore, prior to puberty, careful groupings by skill level should be made if boys and girls are to participate jointly in sports and skill drills in which throwing is important. This suggestion is reinforced by looking at the catching difference between boys and girls. The boys catch better during the elementary school years. The catching difference is most likely minimized by the low ball velocity reported in most tests of catching.

Of the 12 tasks that were age related, sports and skill drills involving throwing (or catching throws of high velocity) are the only ones for which biology appears to play an important role in the development of gender differences prior to puberty. Given the documented differences in expectations and treatment of boys and girls, the differences before puberty in the other 10 tasks are probably environmentally induced. After puberty, the importance of biology and environment are confounded in

many motor tasks, but especially in tasks requiring size, strength and endurance, and power.

Tasks Where Differences Were Not Age Related

Effect sizes for gender in 8 of the 20 (40%) motor performance tasks were not related to age. Thus, performance on these tasks did not change in any systematic way across childhood and adolescence. The gender differences favored boys and were low (0.01 to 0.38) for four of the tasks: agility, anticipation timing, arm hang, and reaction time. The differences were low and favored girls for fine eye–motor coordination (−0.21) and flexibility (−0.29). We see little reason to suspect that any of these small differences have a biological basis. The only difference of any substantial size—anticipation timing ($\overline{ES}' = 0.38$)—is likely to be a result of boys, more often than girls, practicing sports (e.g., baseball, football) that require this skill.

The effect size of throwing for accuracy ($\overline{ES}' = 0.96$), even though large, probably reflects more practice by boys rather than any biological differences. The difference in male and female performance is not developmental. The nature of the tests for throwing accuracy (i.e., the performer is a short distance from the target) is such that throwing velocity is not a factor.

The wall volley (batting an inflated ball repeatedly against a wall with the hands) also has a large effect size favoring boys. We suspect this reflects greater amounts of practice for the boys as well as their advantage in anticipation timing skills.

We see little reason to suspect that gender differences in performance on any of these eight tasks reflect anything but environmental factors. If biology were involved in the differences, we would expect them to be larger, especially after puberty.

Tasks in Which Characteristics Coded From Studies Were Important

In 3 of the 12 tasks in which gender differences were related to age, effect sizes were also related to other characteristics coded about the study. Both the shuttle run and sit-ups were negatively related to the gender of the first author. Thus, effect sizes were larger (favored boys) when the first author was female. If, in fact, the first author and the experimenter are the same person (not always true), the finding is consistent with Rikli's (1976) report that experimenters of the opposite gender cause larger gender differences in motor performance for tasks in which more effort results in better performance.

Effect sizes in the shuttle run and the dash were both positively related to whether the article was published or not, being larger in those published. This finding may reflect journals' bias toward publishing articles in which differences are significant.

Effect sizes in the shuttle run were also positively related to the validity of the study; more valid studies showed larger differences favoring boys. Sit-ups were negatively related to the year published, indicating that differences were larger in studies published before 1970. This may reflect the increasing standardization of sit-up tests. Earlier tests involved the number of sit-ups the performer could do; more recent tests typically have either a time limit (usually 30 s to 60 s) or a maximum number (50 to 100). Thus, differences between boys and girls would be reduced in more recent tests.

Issues That Remain Unaddressed

There are a number of important points that we have not considered here. First, motor performance is only the outcome of the movement. Although the outcome reflects the movement process, it does not do so perfectly and does not describe this process. Even a casual observer notes that individuals may run or throw in different ways, yet obtain similar outcomes or, conversely, run or throw in similar ways and obtain different outcomes. Thus, whether the development of gender differences across age exists in the quality of the movement cannot be determined by this meta-analysis, although work has been done in this area (see Roberton, 1984, for a summary). Roberton (1982) suggested that the form of the movement was not different between girls and boys, just that girls lag behind boys in the development of "good" form. This suggests that the nature of the underlying motor control mechanisms do not differ by gender. Although this view is consistent with most of the motor per-

formance tasks evaluated in this study (at least prior to puberty), it could be inconsistent with the large gender differences noted in throwing and in some of the other tasks after puberty.

A second point that remains unaddressed is the ethnicity and socioeconomic status of the subjects. We estimate that most subjects in the 64 studies included are white and of middle socioeconomic status. However, that information is not generally provided in the studies, particularly in the earlier ones. These issues may be of considerable importance because ethnic variation in rearing has been suggested as a factor in motor performance variation. Yet, ethnicity and socioeconomic status are confounded with gender in this report.

Third, the role of genetics (and its interaction with environment) cannot be evaluated from this data base. However, studies of twins and siblings suggest sex differences in the heritability of motor performance.

Last, an important issue is whether longitudinal and cross-sectional data points estimate gender differences in motor performance equally well at any given age. These data do not include enough longitudinal data points for any specific motor performance task to provide a fair test of this question. However, meta-analysis offers an interesting approach to this issue, if a task could be found in which a sufficient number of longitudinal and cross-sectional data points were available.

Conclusions

We believe the gender differences prior to puberty in 15 of the 20 tasks (agility, anticipation timing, arm hang, balance, dash, grip strength, finc eye–motor coordination, flexibility, long jump, pursuit rotor tracking, reaction time, shuttle run, sit-ups, tapping, vertical jump) studied are environmentally induced. This conclusion is based on the small effect sizes, usually less than 0.50 of a standard deviation unit, as well as documented observations that treatment, expectations, and practice opportunities differ by gender. If one single factor is of importance, it is that boys are involved in more competitive games than girls and generally participate in games of longer duration (Lever, 1976).

The effect sizes for 6 of the previously mentioned 15 tasks (dash, grip strength, long jump, shuttle run, sit-ups, vertical jump) show rapid increases at puberty that are probably associated with the increase in boys' size and strength. However, the same environmental variables discussed previously inflate these gender differences beyond biological explanations.

The differences in effect sizes for throwing (velocity or distance) seem to begin with biological differences (1.5 standard deviation units at 3 years of age), but are increased by more practice opportunities for boys. The effect size for throwing accuracy is also large and may have some biological basis. However, given that the throwing-accuracy tests usually require minimal force, we suspect this gender difference in performance is also the result of boys practicing more than girls.

These findings of gender differences in motor performance are generally consistent with recent theories about sex-role development. Robinson and Green (1981) suggested that both cognitive-developmental (Kohlberg & Ullian, 1974) and transcendence (Hefner, Rebecca, & Oleshansky, 1975) views of sex-role development indicate that early development is strongly influenced by parents and peers. As children enter elementary school, teachers, peers, and parents influence sex stereotyping. Finally, as adults, men and women choose the type of sex role that fits their natures. However, biology does appear to play a greater role in gender differences for throwing at all ages and postpuberty in tasks for which size, strength, and power are important. This at least limits complete choice postpuberty in cross-sex, high-level competition for sports involving size, strength, and power. We are not suggesting that women cannot become skillful performers in these types of sports, just that some levels of joint sport participation remain constrained by the biology of gender differences.

References

American Alliance for Health, Physical Education, Recreation and Dance. (1976). *Youth fitness test manual.* Washington, DC: Author.

Anastasi, A. (1981). Sex differences: Historical perspectives and methodological implications. *Developmental Review, 1,* 187–206.

Burmeister, W. (1965). Body cell mass as the basis of allometric growth functions. *Annales Paediatrici, 204,* 65–72.

Cohen, J. (1969). *Statistical power analysis for the behavioral sciences.* New York: Academic Press.

Corbin, C. B. (Ed.). (1980). *A textbook of motor development* (2nd ed.). Dubuque, IA: Brown.

DeOreo, K., & Keogh, J. (1980). Performance of fundamental motor tasks. In C. B. Corbin (Ed.), *A textbook of motor development* (2nd ed., pp. 76–91). Dubuque, IA: Brown.

Dusenberry, L. M. (1952). A study of the effects of training in ball throwing by children ages three to seven. *Research Quarterly, 23,* 9–14.

Espenschade, A., & Eckert, H. (1974). Motor development. In W. R. Johnson & E. R. Buskirk (Eds.), *Science and medicine in exercise and sport* (2nd ed., pp. 322–333). New York: Harper & Row.

Espenschade, A. S., & Eckert, H. M. (1980). *Motor development* (2nd ed.). Columbus, OH: Merrill.

Fagot, B. I. (1978). The influence of sex of child on parental reaction to toddler children. *Child Development, 49,* 459–465.

Fling, S., & Manosevitz, M. (1972). Sex typing in nursery school children's play interest. *Developmental Psychology, 7,* 146–152.

Frisancho, A. R. (1981). New norms of upper limb fat and muscle areas for assessment of nutritional status and weight. *American Journal of Clinical Nutrition, 34,* 2540–2545.

Glass, G. V., McGaw, B., & Smith, M. L. (1981). *Meta-analysis in social research.* Beverly Hills, CA: Sage.

Greendorfer, S. L. (1980). Gender differences in physical activity. *Motor Skills: Theory into Practice, 4,* 83–90.

Halverson, L. E., Roberton, M. A., & Langendorfer, S. (1982). Development of the overarm throw: Movement and ball velocity changes by seventh grade. *Research Quarterly for Exercise and Sport, 53,* 198–205.

Halverson, L. E., Roberton, M. A., Safrit, M. J., & Roberts, T. W. (1977). Effect of guided practice on overhand-throw ball velocities of kindergarten children. *Research Quarterly, 48,* 311–318.

Haubenstricker, J., & Sapp, M. (1980, April). *A longitudinal look at physical growth and motor performance: Implications for elementary and middle school activity programs.* Paper presented at the meeting of the American Alliance for Health, Physical Education, Recreation and Dance, Detroit, MI.

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics, 6,* 107–128.

Hedges, L. V. (1982). Fitting categorical models to effect sizes from a series of experiments. *Journal of Educational Statistics, 7,* 119–137.

Hedges, L. V., & Olkin, I. (1983). Regression models in research synthesis. *American Statistician, 37,* 137–140.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis.* New York: Academic Press.

Hefner, R., Rebecca, M., & Oleshansky, B. (1975). Development of sex-role transcendence. *Human Development, 18,* 143–158.

Henry, F. M. (1968). Specificity vs. generality in learning motor skill. In R. C. Brown & G. S. Kenyon (Eds.), *Classical studies on physical activity* (pp. 328–331). Englewood Cliffs, NJ: Prentice-Hall.

Housner, L. D. (1981). Sex-role stereotyping: Implications for teaching elementary physical education. *Motor Skills: Theory into Practice, 5,* 107–116.

Hyde, J. S. (1981). How large are cognitive gender differences? A meta-analysis using ω^2 and *d*. *American Psychologist, 36,* 892–901.

Jacklin, C. N. (1981). Methodological issues in the study of sex-related differences. *Developmental Review, 1,* 266–273.

Johnson, B. L., & Nelson, J. K. (1979). *Practical measurement for evaluation in physical education* (3rd ed.). Minneapolis, MN: Burgess.

Kohlberg, L., & Ullian, D. Z. (1974). Stages in the development of psychosexual concepts and attitudes. In R. C. Friedman, R. M. Richart, & R. L. Vande Wiele (Eds.), *Sex differences in behavior* (pp. 209–222). New York: Wiley.

Lansky, L. M. (1967). The family structure also affects the model: Sex-role attitudes in parents of preschool children. *Merrill-Palmer Quarterly, 13,* 139–150.

Lever, J. (1976). Sex differences in the games children play. *Social Problems, 23,* 478–487.

Maccoby, E. E., & Jacklin, C. N. (1974). *The psychology of sex differences.* Stanford, CA: Stanford University Press.

Malina, R. M. (1975). *Growth and development: The first twenty years in man.* Minneapolis, MN: Burgess.

Malina, R. M. (1984). Physical growth and maturation. In J. R. Thomas (Ed.), *Motor development during childhood and adolescence* (pp. 2–26). Minneapolis, MN: Burgess.

Malina, R. M., & Johnson, F. E. (1967). Significance of age, sex, and maturity differences in upper arm composition. *Research Quarterly, 38,* 219–230.

National Center for Health Statistics. (1977). NCHS growth curves for children, birth–18 years, United States. *Vital and Health Statistics,* Series 11, No. 165.

Ridenour, M. V. (Ed.). (1978). *Motor development: Issues and applications.* Princeton, NJ: Princeton Book.

Rikli, R. (1976). Physical performance scores as a function of experimenter sex and experimenter bias. *Research Quarterly, 47,* 776–782.

Roberton, M. A. (1982). Describing "stages" within and across motor tasks. In J. A. S. Kelso & J. E. Clark (Eds.), *The development of movement control and coordination* (pp. 293–307). New York: Wiley.

Roberton, M. A. (1984). Changing motor patterns during childhood. In J. R. Thomas (Ed.), *Motor development during childhood and adolescence* (pp. 48–90). Minneapolis, MN: Burgess.

Robinson, B. E., & Green, M. G. (1981). Beyond androgyny: The emergence of sex-role transcendence as a theoretical construct. *Developmental Review, 1,* 247–265.

Roche, A. F., & Malina, R. M. (Eds.). (1983). *Manual of physical status and performance in childhood* (Vol. 1). New York: Plenum.

Schmidt, R. A. (1982). *Motor control and learning.* Champaign, IL: Human Kinetics.

Sherif, C. W., & Rattray, G. D. (1976). Psychological development and activity in middle childhood. In J. G. Albinson & G. M. Andrew (Eds.), *Child in sport and physical activity* (pp. 97–132). Baltimore: University Park.

Thomas, J. R. (Ed.). (1984). *Motor development during childhood and adolescence.* Minneapolis, MN: Burgess.

Thomas, J. R., & Nelson, J. K. (1985). *Introduction to research in HPERD.* Champaign, IL: Human Kinetics.

Tolson, H. (1980). An adjunct to statistical significance: ω^2. *Research Quarterly for Exercise and Sport, 51,* 580–584.

Walker, R. M. (1962). Body build and behavior in young children: I. Body build and nursery school teachers' ratings. *Monographs of the Society for Research in Child Development, 27* (3, Serial No. 34).

Wickstrom, R. L. (1983). *Fundamental motor patterns* (3rd ed.). Philadelphia: Lea & Febiger.

Appendix

Studies Included in the Meta-Analysis (N = 64)

Bachman, J. C. (1961). Motor learning and performance as related to age and sex in two measures of balance coordination. *Research Quarterly, 32,* 123–137.

Boley, E. H. (1975). *Generality and specificity of motor performance of children as related to age and sex.* Unpublished master's thesis, Louisiana State University, Baton Rouge, LA.

Broekhoff, J. (1978). Longitudinal comparison of the growth, physical fitness, and motor performance of suburban and inner city elementary school children. In F. Landry & W. A. R. Orban (Eds.) *Motor learning, sport psychology, pedagogy, and didactics of physical activity* (pp. 203–210). Miami, FL: Symposia Specialists.

Caskey, S. R. (1968). Effects of motivation on standing broad jump performance of children. *Research Quarterly, 39,* 54–59.

Cowgill, S. (1978). *Nebraska physical fitness norms: Grades 1–12.* Unpublished master's thesis, Colorado State University, Fort Collins, CO.

Davol, S. H., & Breakell, S. L. (1968). Sex differences in rotary pursuit performance of young children: A follow-up. *Perceptual and Motor Skills, 26,* 1199–1202.

Davol, S. H., Hastings, M. L., & Klein, D. A. (1965). Effect of age, sex, and speed of rotation on rotary pursuit performance by young children. *Perceptual and Motor Skills, 21,* 351–357.

Dohrman, P. (1964). Throwing and kicking ability of 8-year-old boys and girls. *Research Quarterly, 35,* 464–471.

Dunham, P., Jr. (1977). Age, sex, speed, and practice in coincidence-anticipation performance of children. *Perceptual and Motor Skills, 45,* 187–193.

Eckert, H. M. (1970). Visual-motor tasks at 3 and 4 years of age. *Perceptual and Motor Skills, 31,* 560.

Eckert, H. M. (1974). Variability in skill acquisition. *Child Development, 45,* 487–489.

Eckert, H. M., & Eichorn, D. H. (1974). Construct standard in skilled action. *Child Development, 45,* 439–445.

Eckert, H. M., & Eichorn, D. H. (1977). Developmental variability in reaction time. *Child Development, 48,* 452–458.

Eckert, H. M., & Rarick, G. L. (1976). Stabilometer performance of educable mentally retarded and normal children. *Research Quarterly, 47,* 619–623.

Espenschade, A. S., & Meleney, H. E. (1961). Motor performance of adolescent boys and girls of today in comparison with those of 24 years ago. *Research Quarterly, 32,* 186–189.

Finlayson, M. A. J., & Reitan, R. M. (1976). Handedness in relation to measures of motor and tactile-perceptual functions in normal children. *Perceptual and Motor Skills, 43,* 475–481.

Fitch, J. H. (1980). *An analysis of the factors on the North Carolina motor fitness battery test among students aged ten to seventeen.* Unpublished master's thesis, North Carolina Central, Durham, NC.

Gabbard, C., Kirby, T., & Patterson, R. (1979). Reliability of the straight-arm hang for testing muscular endurance among children 2 to 5. *Research Quarterly, 50,* 735–738.

Glassow, R. B., Halverson, L. E., & Rarick, G. L. (1965). *Improvement of motor development and physical fitness in elementary school children.* Unpublished manuscript, University of Wisconsin, Cooperative Research Project #696.

Glover, E. G. (1962). *Physical fitness test items for boys and girls in the first, second, and third grades.* Unpublished master's thesis, University of North Carolina, Chapel Hill, NC.

Govatos, L. A. (1959). Relationships and age differences in growth measures and motor skills. *Child Development, 30,* 333–340.

Green, J. M. (1973). *The relative effectiveness of a perceptual motor program, a movement education program, and a traditional program in the enhancement of motor performance of kindergarten children.* Unpublished master's thesis, University of Washington, Seattle, WA.

Greene, J. L. (1973). *Effects of a prescribed physical education program upon movement characteristics of 4-year-old boys and girls.* Unpublished

doctoral dissertation, University of Utah, Salt Lake City, UT.

Halverson, L. E., Roberton, M. A., & Langendorfer, S. (1982). Development of the overarm throw: Movement and ball velocity changes by seventh grade. *Research Quarterly, 53,* 198–205.

Halverson, L. E., Roberton, M. A., Safrit, M. J., & Roberts, T. W. (1977). Effect of guided practice on overhand-throw ball velocities of kindergarten children. *Research Quarterly, 48,* 311–318.

Harper, C. J. (1975). *Movement responses of kindergarten children to a change of direction task—an analysis of selected measures.* Unpublished master's thesis, University of Wisconsin, Madison, WI.

Hartman, D. (1943). The hurdle jump as a measure of the motor proficiency of young children. *Child Development, 14,* 201–211.

Harvey, D. A. (1970). *The effects of level of aspiration and team competition as motivational techniques upon children's performances on selected sports skill tests.* Unpublished doctoral dissertation, Indiana University, Bloomington, IN.

Haywood, K. M., Greenwald, G., & Lewis, C. (1981). Contextual factors and age group differences in coincidence-anticipation performance. *Research Quarterly, 52,* 458–464.

Hunsicker, P. A. (1965). *A survey and comparison of youth fitness, 1958–1965.* Unpublished manuscript, University of Michigan, Cooperative Research Project #2418.

Ikeda, N. (1961). *A comparison of physical fitness of children in Iowa, USA and Tokyo, Japan.* Unpublished doctoral dissertation, State University of Iowa, Iowa City, IA.

Ingersoll, M. T. (1976). *The motor fitness of primary boys and girls.* Unpublished master's thesis, Ithaca College, Ithaca, NY.

Isaacs, L. D. (1980). Effects of ball size, ball color, and preferred color on catching by young children. *Perceptual and Motor Skills, 51,* 583–586.

Kane, R. J., & Meredith, H. V. (1952). Ability in the standing broad jump of elementary school children 7, 9, and 11 years of age. *Research Quarterly, 23,* 198–208.

Keating, D. P., & Bobbitt, B. L. (1978). Individual and developmental differences in cognitive-processing components of mental ability. *Child Development, 49,* 155–167.

Keogh, J. (1965). *Motor performance of elementary school children.* Unpublished manuscript, Department of Physical Education, University of California, Los Angeles. (Available through ERIC)

Knights, R. M., & Mouls, A. D. (1967). Normative and reliability data on finger and foot tapping in children. *Perceptual and Motor Skills, 25,* 717–720.

Latchaw, M. (1954). Measuring selected motor skills in fourth, fifth, and sixth grades. *Research Quarterly, 25,* 439–449.

Lee, A. M., Fant, H., Life, M. L., Lipe, L., & Carter, J. A. (1978). Field independence and performance on ball handling tasks. *Perceptual and Motor Skills, 46,* 439–442.

Maples, M. G. (1977). *Second grade children's performance on the overhand throw in relation to maternal and self preference for play activities.* Unpublished master's thesis, Purdue University, West Lafayette, IN.

McCaskill, C. L., & Wellman, B. L. (1938). A study of common motor achievements of the preschool ages. *Child Development, 9,* 141–149.

Miller, J. L. (1957). Effect of instruction on development of throwing accuracy of first grade children. *Research Quarterly, 28,* 132–137.

Milne, C., Seefeldt, V., & Reuschlein, P. (1976). Relationship between grade, sex, race, and motor performance in young children. *Research Quarterly, 47,* 726–730.

Miyashita, M., & Kanehisa, H. (1979). Dynamic peak torque related to age, sex, and performance. *Research Quarterly, 50,* 249–255.

Montpetit, R. R., Montoye, H. J., & Laeding, L. (1967). Grip strength of school children, Saginaw, Michigan: 1899 and 1964. *Research Quarterly, 38,* 231–240.

Morris, A. M., Williams, J. L., Atwater, A. E., & Wilmore, J. H. (1982). Age and sex differences in motor performance of 3 through 6-year-old children. *Research Quarterly, 53,* 214–221.

Nestroy, J. A. (1978). *Fitness levels of children taught by the physical education specialists and classroom teachers.* Unpublished master's thesis, Texas Woman's University, Denton, TX.

Pissanos, B. W., Moore, J. B., & Reeve, T. G. (1983). Age, sex, and body composition as predictors of children's performance on basic motor abilities and health-related fitness items. *Perceptual and Motor Skills, 56,* 71–77.

Pomeroy, J. E. (1938). The relation of reaction time of five-year-old children to various factors. *Child Development, 9,* 281–283.

Roberton, M. A., Halverson, L. E., Langendorfer, S., & Williams, K. (1979). Longitudinal changes in children's overarm throw ball velocities. *Research Quarterly, 50,* 256–264.

Ross, B. M. (1960). A study of the performance of boys and girls taught by the specialist and nonspecialist. *Research Quarterly, 31,* 199–207.

Seils, L. G. (1951). The relationship between measures of physical growth and gross motor performance of primary-grade school children. *Research Quarterly, 22,* 244–258.

Singer, R. N. (1969). Physical characteristic, perceptual-motor, and intelligence differences between third- and sixth-grade children. *Research Quarterly, 40,* 803–811.

Smith, J. (1956). Relation of certain physical traits and ability of motor learning in elementary children. *Research Quarterly, 27,* 221–228.

Smith, T. L. (1982). Self-concepts and movement skills of third grade children after physical education programs. *Perceptual and Motor skills, 54,* 1145–1146.

Smoll, F. L. (1966). *The influence of physical growth and muscular strength upon motor performances within and between year observations.* Unpublished master's thesis, University of Wisconsin, Madison, WI.

Stachnik, T. J. (1964). Cross-validation of psychomotor test for children. *Perceptual and Motor Skills, 18,* 913–916.

Taddonio, D. A. (1966). Effect of daily fifteen-minute periods of calisthenics upon the physical fitness of fifth grade boys and girls. *Research Quarterly, 37,* 276–281.

Thomas, J. R., Gallagher, J. D., & Purvis, G. J. (1981). Reaction time and anticipation time: Effects of development. *Research Quarterly, 52,* 359–367.

Torres, J. A. (1966). *The relationship between figure-ground perceptual ability and ball catching ability in ten and thirteen year old boys and girls.* Unpublished master's thesis, Purdue University, West Lafayette, IN.

Trussell, E. M. (1969). Relation of performance of selected physical skills to perceptual aspects of reading readiness in elementary school children. *Research Quarterly, 40,* 383–390.

Wilson, J. G., Silva, P. A., & Williams, S. M. (1981). An assessment of motor ability in seven year olds. *Journal of Human Movement Studies, 7,* 221–231.

Workman, D. J. (1979). Comparison of performance of children taught by the physical education specialist and by the classroom teacher. *Research Quarterly, 49,* 389–394.

Wright, E. J. (1967). Effects of light and heavy equipment on acquisition of sports-type skills by young children. *Research Quarterly, 38,* 705–714.

Received October 17, 1984
Revision received February 7, 1985 ■

5-20-2016

International normative 20 m shuttle run values from 1 142 026 children and youth representing 50 countries

Grant Tomkinson
University of North Dakota, grant.tomkinson@und.edu

Justin J. Lang

Mark S. Tremblay

Michael Dale

Allana G. LeBlanc

See next page for additional authors

Authors
Grant Tomkinson, Justin J. Lang, Mark S. Tremblay, Michael Dale, Allana G. LeBlanc, Kevin Belanger, Francisco B. Ortega, and Luc Léger

International normative 20 m shuttle run values from 1 142 026 children and youth representing 50 countries

Grant R Tomkinson, Justin J Lang, Mark S Tremblay, Michael Dale, Allana G LeBlanc, Kevin Belanger, Francisco B Ortega, Luc Léger

Abstract

Objective To develop sex-specific and age-specific international norms for the 20 m shuttle run test (20mSRT) in children and youth (aged 9–17 years), and to estimate the prevalence meeting the FITNESSGRAM criterion-referenced standards for healthy cardiorespiratory endurance (CRE).

Methods A systematic review was undertaken to identify papers explicitly reporting descriptive 20mSRT (with 1 min stages) data on children and youth since 1981. Data were included on apparently healthy (free from known disease/injury) 9–17 years old. Following standardisation to a common metric and for protocol differences, pseudo data were generated using Monte Carlo simulation, with population-weighted sex-specific and age-specific normative centiles generated using the Lambda Mu and Sigma (LMS) method. Sex-related and age-related differences were expressed as per cent and standardised differences in means. The prevalence with healthy CRE was estimated using the sex-specific and age-specific FITNESSGRAM criterion-referenced standards for $\dot{V}O_{2peak}$.

Results Norms were displayed as tabulated centiles and as smoothed centile curves for the 20mSRT using 4 common metrics (speed at the last completed stage, completed stages/minutes, laps and relative $\dot{V}O_{2peak}$). The final data set included 1 142 026 children and youth from 50 countries, extracted from 177 studies. Boys consistently outperformed girls at each age group (mean difference±95% CI: 0.86±0.28 km/h or 0.79±0.20 standardised units), with the magnitude of age-related increase larger for boys than for girls. A higher proportion of boys (mean±95% CI: 67±14%) had healthy CRE than girls (mean±95% CI: 54±17%), with the prevalence of healthy CRE decreasing systematically with age.

Conclusions This study provides the most comprehensive and up-to-date set of international sex-specific and age-specific 20mSRT norms for children and youth, which have utility for health and fitness screening, profiling, monitoring and surveillance.

Background

Cardiorespiratory endurance (CRE) is the ability to deliver oxygen to the muscles and to utilise it to generate energy to support muscle activity during exercise.[1, 2] In adults, low CRE is strongly associated with cardiovascular disease and all-cause mortality and morbidity (independent of adiposity),[3, 4] stroke,[5] diabetes,[6] mental health,[7] health-related quality of life[8] and cardiometabolic disease risk.[6, 9] In children and youth, CRE is a weak-to-strong predictor of cardiovascular disease risk, cancer and mental health.[10, 11] There is also evidence of an interaction between adiposity and CRE, suggesting that high levels of CRE may attenuate the deleterious effects of being overweight or obese in children and youth, the so-called 'fat but fit' phenotype.[12] Thus, CRE provides insight into the synergistic capabilities of several bodily systems and organs that are involved in the performance of physical activity and exercise, providing a strong and summative measure of health in children and youth.[10] CRE also tracks moderately well from childhood to adulthood,[13, 14] indicating that the measurement and surveillance of CRE in children provides insight into current and future population health statuses. Low fitness in childhood and adolescence is substantially linked with increased cardiometabolic disease risk,[15, 16] obesity,[17] reduced quality of life,[18] poorer skeletal[19] and mental health[11] in adulthood. In addition to health implications, CRE is an important

determinant of sporting success for many popular youth sports (eg, hockey, basketball, football (soccer), distance running, swimming, rugby).[1]

The 20 m shuttle run test (20mSRT) is arguably the most popular field-based assessment and estimate of CRE in children and youth worldwide.[20–23] It is an excellent tool for population-based surveillance and monitoring because it demonstrates strong-to-very strong test–retest reliability, and moderate-to-strong validity.[20] The 20mSRT also has excellent field-based utility due to its low cost, flexibility with testing locations (indoors, outdoors, smaller spaces) and its ability to test multiple individuals simultaneously with minimal equipment and personnel.[20, 24, 25] In order to extend the utility of the 20mSRT as a surveillance instrument, there is a need for international norms, which, to date, have only been published for a single 20mSRT metric or for selected geographical regions,[20] including Europe,[26] North America[27] and Oceania.[28] Olds *et al*[21] cumulated 20mSRT data on children and youth from 37 countries, representing six continents, suggesting that there is an international appetite for assessing CRE in children and youth. Harmonising reference values by creating international normative centiles for the 20mSRT in children and youth would provide opportunity for international surveillance and a means to compare CRE across geographic areas and time.

This study provides a 10-year update of the comprehensive 20mSRT reviews by Tomkinson *et al*[22] and Olds *et al*,[21] which were the first studies to describe a method to harmonise 20mSRT data by standardising for differences in test protocols and performance metrics. The primary aim was to develop the most comprehensive set of sex-specific and age-specific international normative centiles for CRE. The secondary aim was to estimate the proportion of children and youth meeting the FITNESSGRAM criterion-referenced standards for healthy CRE. These data will facilitate the identification of children and youth with very low CRE in order to set appropriate goals and promote positive health-related fitness behaviours, and conversely those with very high CRE, which may be important for sporting or athletic success.

Methods

Data sources

A systematic review of the scientific literature was registered (PROSPERO 2013:CRD42013003622) and completed to locate studies that reported descriptive 20mSRT data on 9–17 years old (see online supplement 1). Studies were identified up until October 2015 using the following bibliographic databases: MEDLINE (1946–2015), PsycINFO (1806–2014), EMBASE (1947–2014), SPORTDiscus (1949–2014) and Cochrane Central Register of Controlled Trials (2005–2014). The search strategy included the following terms: shuttle run*, OR beep test, OR multi-stage, OR aerobic, OR cardio*, OR endurance; with child*, adolescen*, pubescen*, boy, girl, young and youth as search term modifiers. All studies were extracted as text files, imported into Reference Manager (Thompson Reuters, San Francisco, California, USA), and assigned a unique reference identification number. Duplicate studies were first removed using Reference Manager with the remaining duplicates removed manually. Two independent reviewers screened all titles and abstracts for eligibility, with full-text copies obtained for all studies meeting initial screening criteria according to at least one reviewer. These two independent reviewers then examined all full-text articles and discrepancies were resolved by discussion and consensus. A third reviewer examined an article when the two reviewers were unable to reach consensus, with consensus reached for all included articles. Email contact with the corresponding authors of studies occurred when necessary, in order to provide clarification, to avoid ‘double counting’ previously reported data, and/or to request additional descriptive or raw data. The reference lists of all included studies were manually

reviewed to identify new studies. Reviewers contacted content experts to obtain grey literature. In addition, the extensive personal libraries of the study authors were examined for relevant studies.

Inclusion/exclusion criteria

Studies were included if they explicitly reported descriptive 20mSRT data (using the 1 min stage protocol; see Tomkinson *et al*[22]for protocol variants) at the sex by age by country level. Study participants must have been apparently healthy (free from known disease or injury) 9–17 years old who were tested from 1981 onwards—the inception year of Léger's 20mSRT with 1 min stages. Studies were excluded if they reported descriptive data on: (1) other versions of the shuttle run (eg, the 15 m test or estimates of 20mSRT performance based on the 15 m test); (2) duplicate data published in another included study; or (3) on only special interest groups of children who were atypical of their source population (eg, elite athletes, physically or mentally retarded children). Figure 1 shows a flow chart of the included studies.

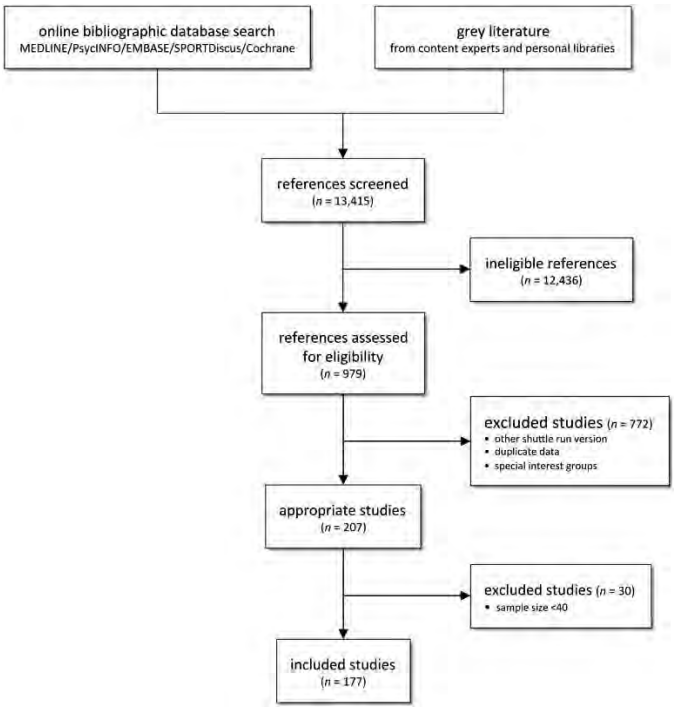


Figure 1
PRISMA flow chart outlining the identification of the included studies.

Data treatment and statistical analysis

All descriptive data were extracted into Excel (Microsoft Office 2010, USA) using a standardised data extraction table. The following descriptive data were extracted by one author and checked by another for accuracy: authors, country of testing, year of testing, sex, age, 20 m shuttle run protocol, 20 m shuttle run metric, sample size, mean, SD, median, sampling method and the sampling base. All data were examined for anomalies by running range checks and examining sex-specific and age-specific scattergrams. While only data on children and youth aged 9–17 years inclusive were retained for further analysis, individual study by sex by age by country by year groups were excluded when the sample size fell below 40 as the means and SDs for smaller samples were too labile. The final data set included 1896 study by sex by age by country by year groups.

The general procedure used to generate the sex-specific and age-specific normative centiles from extracted data is shown in figure 2. In most studies (76% or 135/177 studies), age was reported as age at last birthday; however, age was also reported in 7% (13/177) of studies as a span of years (maximum range: 3 years) and in 17% (29/177) as mean and SD years. In studies reporting age as a span of years, a new sample size was assigned to each study by sex by age by country group by dividing the reported sample size by the number of age groups (eg, a sample size of 162 was assigned to boys aged 10, 11 and 12 years in the study by Vandongen *et al*[29] which reported testing 486 boys aged 10–12 years). In studies that reported age as mean±SD years, Monte Carlo simulation was used to produce pseudo age data (using a random normal generator) based on reported means and SDs to estimate the sample size in each study by sex by age by country group.

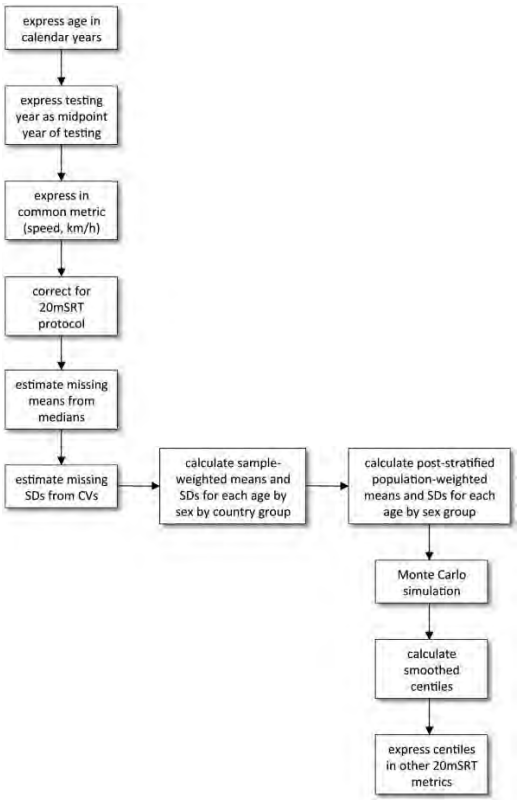


Figure 2
Flow chart showing the methodological procedure used in this study. Results from studies were first expressed in the common metric of running speed (km/h) at the final completed stage and then corrected to Léger's original 1 min protocol according to which protocol they used (Léger, Eurofit or Queen's University of Belfast). Following the estimation of missing means (from reported medians) and SDs (from calculated CVs), poststratified population-weighted means and SDs were estimated for each sex by age group, with pseudo data then generated using Monte Carlo simulation. Smoothed centiles were then generated using the LMS method, with international normative 20mSRT values expressed in several different metrics (speed at the last completed stage, the number of completed full stages/minutes and relative $\dot{V}O_{2peak}$ (ml/kg/min)). 20mSRT, 20 m shuttle run test; CV, coefficient of variation; LMS, Lambda Mu and Sigma; and $\dot{V}O_{2max}$ peak oxygen uptake.

Testing year was recorded as the midpoint year of testing (eg, 2009.5 was recorded as the measurement year for a study that reported testing children in 2009) in 55% (98/177) of studies, with 34% (60/177) reporting a span of testing years and 11% (19/177) not reporting it at all. The midpoint year was recorded for studies reporting a span of testing years (eg, 2010.0 was recorded for a study reporting testing over the period 2009–2010), with 2 years prior to the publication year assumed for studies when it was not reported, which was the median difference for those studies in which the testing year was known.

To combine data from different studies, all 20mSRT data were standardised to a common metric and for protocol differences. To do this, we used an updated version of the standardisation procedure described in detail by Tomkinson *et al*.[22] Figure 3 shows this updated standardisation procedure and summarises the steps used to express 20mSRT performances in the common metric of speed (km/h) at the last completed stage. All 20mSRT data were then standardised for protocol differences to Léger's original 1 min protocol,[30] which starts at a speed of 8.5 km/h and increases in speed by 0.5 km/h each minute.[22]

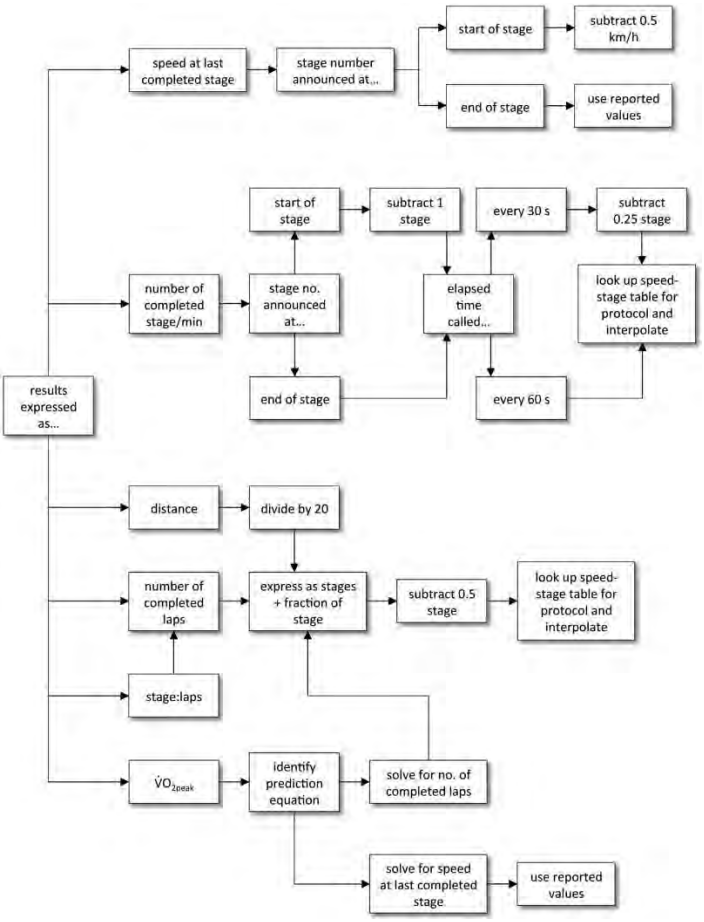


Figure 3
Performance metric conversion flow chart adapted from Tomkinson et al.[22] $\dot{V}O_{2peak}$, peak oxygen uptake.

As part of the modelling procedure used to generate sex-specific and age-specific normative centiles, means and SDs were required at the study by sex by age by country by year level. If no mean was available (1% or 2/177 studies), then mean values were estimated from the reported median values. This was done by first locating all studies reporting both median and mean values at the study by sex by age by country level, and second, by determining the best-fitting and most parsimonious linear and curvilinear (second and third order polynomials) regression models between median (predictor variable) and mean (response variable) speed values. Median and mean speed values were available for 418 study by sex by age by country groups, with the relationship nearly perfectly described by the following linear regression model: mean=0.9408×median+0.6566 (where r (95% CI)=0.988 (0.985 to 0.990) and SE=0.128). Furthermore, 5% (9/177) of studies did not report SD values. Missing SD values were estimated by first locating all studies reporting both means and SDs at the study by sex by age by country by year level; second, by calculating the corresponding coefficient of variation (CV) values; and third, by calculating the sample-weighted mean CVs for boys and girls separately. Mean and SD speed values were available

for 1585 study by sex by age by country groups, with sample-weighted mean CVs (±95% CI) of 10.8±0.1% and 9.2±0.1% for boys and girls, respectively.

Sample-weighted means and SDs (the latter calculated from sample-weighted mean CVs) were then calculated at the sex by age by country level. While these data represent the best available 20mSRT data, in order to best generate internationally representative sex-specific and age-specific normative centiles and to correct for systematic bias associated with oversampling and undersampling, means and SDs were corrected using the poststratification population-weighting procedure described by Levy and Lemeshow.[31] This procedure ensures that our population estimates reflect the underlying international age-specific and sex-specific country demographics. Thus, population estimates standardised to the mean measurement year of 2000 were extracted from the United Nations World Population Prospects report.[32] Monte Carlo simulation was then used to create pseudo data using a random normal generator based on population-weighted means and SDs at the sex by age level. Monte Carlo simulation assumes that the distributions are approximately normal, which is true of the available raw 20mSRT speed data. Pseudo data sets were repeatedly generated until the calculated mean differed from the reported mean by <0.5%, and the calculated SD differed from the reported SD by <2.5%. These pseudo data sets were then used to generate sex-specific and age-specific normative centiles in LMSchartmaker Pro (V.2.43, The Institute of Child Health, London), which analyses data using the Lambda Mu and Sigma (LMS).[33] The LMS method fits smooth centile curves to reference data by summarising the changing distribution of three sex-specific and age-specific curves representing the skewness (L; expressed as a Box-Cox power), the median (M) and the CV (S). Using penalised likelihood, the curves can be fitted as cubic splines using non-linear regression, and the extent of smoothing required can be expressed in terms of smoothing parameters or equivalent degrees of freedom.[34] The effective degrees of freedom for 20mSRT speed were 1 (L curve), 4 (M curve) and 3 (S curve) for boys and 1 (L curve), 3 (M curve) and 3 (S curve) for girls. Normative centiles were also expressed in other common 20mSRT metrics, including the number of completed stages/minutes, the number of completed laps and relative peak oxygen uptake ($\dot{V}O_{2peak}$, mL/kg/min) values using the Léger *et al*[35] prediction equation:

$$\dot{V}O_{2peak}(mL/kg/min) = 31.025 + 3.238\,speed - 3.248\,age + 0.1536\,speed \times age$$

where speed is the running speed of the last completed stage (km/h) and age is age at last birthday. In a sample of Québec children and youth, this equation had a SE of estimate of 5.9 mL/kg/min or 12.1%.[35]

The prevalence of children and youth (10–17 years old) with 'healthy' CRE was estimated using the new sex-specific and age-specific FITNESSGRAM criterion-referenced standards for $\dot{V}O_{2peak}$.[36, 37] Differences in mean 20mSRT performance (km/h) between (1) age-matched boys and girls (eg, boys aged 10 years vs girls aged 10 years), and (2) sex-matched children of different ages (eg, boys aged 10 vs 11 years), were expressed as absolute and standardised differences. Positive differences indicated that 20mSRT performances for boys or older children were higher than those for girls or younger children. Standardised differences of 0.2, 0.5 and 0.8 were used as thresholds for small, moderate and large, respectively.[38]

Results

The final data set included 1 142 026 children and youth aged 9–17 years (1896 study by sex by age by country by year groups extracted from 177 studies) from 50 countries tested over the period 1981–2014 (figure 4, see online supplement 2[w39–w190]). These 50 countries represented six major geographical regions (Africa, Asia, Europe, Latin America and the Caribbean, Northern America and Oceania),[32] including 33 high-income, 9 upper middle-income, 5 lower middle-income and 3 low-income economies.[191] Norms are

presented as tabulated centiles from 5% to 95% for four common 20mSRT metrics in tables 1⇓⇓–4, with the smoothed centile curves for 20mSRT speed presented in figure 5 and the sex-specific and age-specific LMS values shown in online supplement 3.

Figure 4
World map representing the 50 countries (shown in black) for which 20mSRT data on 9–17 years old were available. 20mSRT, 20 m shuttle run test.

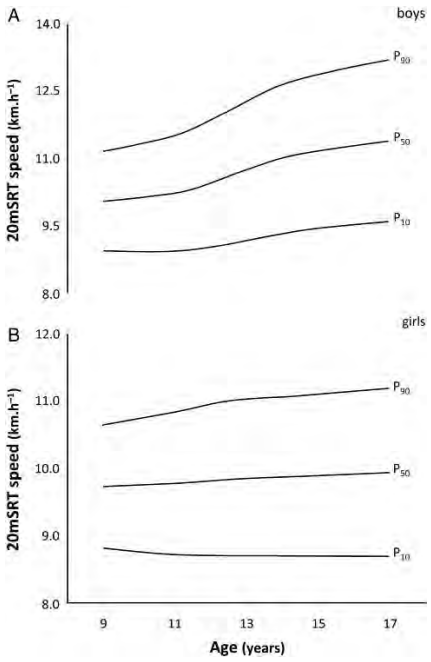


Figure 5
Smoothed centile curves (P10, P50 and P90) for the 20mSRT (speed (km/h) at the last completed stage) performance of (A) boys and (B) girls. 20mSRT, 20 m shuttle run test.

On average, approximately two-thirds of boys (mean±95% CI: 67±14%) and half of the girls (mean±95% CI: 54±17%) had healthy CRE, with the prevalence of healthy CRE decreasing by about 8% (boys) and 10% (girls) with every increasing year from age 10 to 17 years (figure 6).

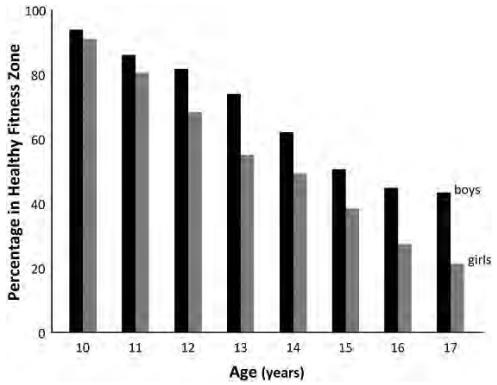


Figure 6
Prevalence of 10–17 years old from 50 countries meeting the FITNESSGRAM Healthy Fitness Zone (V.10) thresholds for $\dot{V}O_{2peak}$ (mL/kg/min). $\dot{V}O_{2peak}$, peak oxygen uptake.

Boys consistently outperformed girls at each age group (mean difference±95% CI: 0.86±0.28 km/h or 0.79±0.20 standardised units), with the sex-related differences increasing with age from a small difference at age 9 years (difference in means±95% CI: 0.32±0.01 km/h or 0.40±0.01 standardised units) to a large difference at age 17 years (difference in means±95% CI: 1.46±0.02 km/h or 1.20±0.02 standardised units; figure 7A). From age 9 to 17 years, boys' performance improved at the rate of 0.17 km/h (or 0.15 standardised units) per year, with the largest rate of increase occurring at age 12 years (0.27 km/h or 0.23 standardised units). Girls' performance steadily improved at the rate of 0.03 km/h (or 0.03 standardised units) per year (figure 7B). These age-related changes were cumulatively large for boys and small for girls.

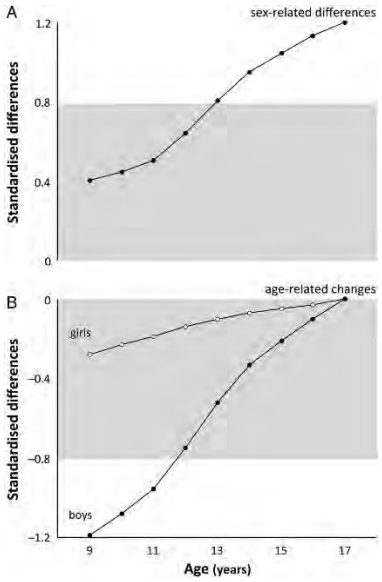


Figure 7
Standardised differences in mean 20mSRT performance (speed (km/h) at the last completed stage) between (A) age-matched boys and girls and (B) sex-matched children of different ages (anchored to age 17 years=0). Positive differences indicate that 20mSRT performances were higher for boys than for girls (top panel) or for older children than for younger children (bottom panel). The limits of the grey zone represent the thresholds for a large standardised difference (0.8 or −0.8). 20mSRT, 20 m shuttle run test.

Table 1
Twenty-metre shuttle run (speed (km/h) at the last complete stage) centiles by age and sex in 1 142 026 children and youth aged 9–17 years from 50 countries since 1981

Age (years)	P_5	P_{10}	P_{20}	P_{30}	P_{40}	P_{50}	P_{60}	P_{70}	P_{80}	P_{90}	P_{95}
Boys											
9	8.63	8.94	9.31	9.58	9.82	10.03	10.25	10.48	10.75	11.13	11.44
10	8.61	8.95	9.35	9.65	9.90	10.13	10.36	10.61	10.90	11.31	11.64
11	8.60	8.97	9.41	9.72	10.00	10.25	10.50	10.77	11.09	11.53	11.89
12	8.64	9.05	9.53	9.89	10.19	10.47	10.75	11.05	11.40	11.89	12.29
13	8.74	9.18	9.72	10.10	10.43	10.73	11.04	11.37	11.75	12.29	12.72
14	8.85	9.32	9.88	10.29	10.64	10.96	11.29	11.64	12.04	12.61	13.08
15	8.94	9.42	10.01	10.43	10.79	11.13	11.47	11.83	12.25	12.84	13.32
16	9.01	9.51	10.11	10.55	10.92	11.27	11.62	11.99	12.42	13.03	13.53
17	9.08	9.60	10.22	10.67	11.05	11.41	11.77	12.16	12.60	13.23	13.74
Girls											
9	8.56	8.82	9.13	9.35	9.54	9.72	9.90	10.08	10.31	10.61	10.87
10	8.48	8.76	9.10	9.35	9.56	9.75	9.95	10.15	10.40	10.74	11.01
11	8.42	8.72	9.09	9.35	9.57	9.78	10.00	10.22	10.48	10.85	11.15
12	8.37	8.69	9.08	9.36	9.60	9.83	10.05	10.29	10.57	10.95	11.27
13	8.36	8.69	9.09	9.38	9.63	9.86	10.09	10.34	10.63	11.03	11.36
14	8.36	8.70	9.11	9.40	9.65	9.89	10.12	10.37	10.67	11.07	11.41
15	8.36	8.70	9.12	9.42	9.67	9.91	10.15	10.40	10.70	11.11	11.44
16	8.36	8.71	9.13	9.43	9.69	9.93	10.17	10.43	10.73	11.14	11.49
17	8.36	8.72	9.14	9.45	9.71	9.96	10.20	10.46	10.77	11.19	11.54

- The ages shown represent age at last birthday (eg, 9=9.00–9.99).

Table 2
Twenty-metre shuttle run (number of completed stages/minutes) centiles by age and sex in 1 142 026 children and youth aged 9–17 years from 50 countries since 1981

Age (years)	P_5	P_{10}	P_{20}	P_{30}	P_{40}	P_{50}	P_{60}	P_{70}	P_{80}	P_{90}	P_{95}
Boys											
9	1.25	1.87	2.63	3.17	3.63	4.07	4.50	4.96	5.50	6.26	6.88
10	1.23	1.90	2.71	3.29	3.79	4.26	4.72	5.22	5.80	6.61	7.28
11	1.21	1.94	2.82	3.45	3.99	4.50	5.00	5.55	6.18	7.06	7.78
12	1.29	2.09	3.07	3.77	4.37	4.93	5.50	6.10	6.80	7.77	8.58
13	1.48	2.36	3.43	4.20	4.85	5.47	6.08	6.74	7.51	8.57	9.45
14	1.70	2.64	3.77	4.58	5.28	5.93	6.58	7.28	8.09	9.22	10.15
15	1.87	2.84	4.01	4.86	5.58	6.26	6.93	7.66	8.50	9.67	10.64
16	2.02	3.01	4.22	5.10	5.84	6.54	7.23	7.98	8.85	10.06	11.05
17	2.17	3.20	4.44	5.34	6.11	6.83	7.54	8.31	9.21	10.45	11.48
Girls											
9	1.12	1.63	2.25	2.70	3.08	3.43	3.79	4.17	4.61	5.23	5.74
10	0.97	1.53	2.21	2.70	3.11	3.50	3.89	4.31	4.80	5.47	6.03
11	0.83	1.44	2.17	2.70	3.15	3.57	3.99	4.44	4.96	5.69	6.29
12	0.74	1.39	2.17	2.73	3.21	3.65	4.10	4.57	5.13	5.90	6.54
13	0.72	1.38	2.19	2.77	3.26	3.72	4.19	4.68	5.25	6.05	6.71
14	0.72	1.40	2.21	2.80	3.31	3.78	4.24	4.75	5.33	6.14	6.81
15	0.73	1.41	2.24	2.83	3.34	3.82	4.29	4.80	5.39	6.21	6.89
16	0.73	1.42	2.26	2.87	3.38	3.86	4.34	4.86	5.46	6.29	6.97
17	0.73	1.44	2.29	2.90	3.43	3.92	4.40	4.93	5.54	6.38	7.08

Table 3
Twenty-metre shuttle run (number of laps) centiles by age and sex in 1 142 026 children and youth aged 9–17 years from 50 countries since 1981

Age (years)	P_5	P_{10}	P_{20}	P_{30}	P_{40}	P_{50}	P_{60}	P_{70}	P_{80}	P_{90}	P_{95}
Boys											
9	9	14	20	24	28	32	36	40	45	52	58
10	9	14	21	25	29	33	37	42	47	55	62
11	9	15	22	27	31	36	40	45	51	60	67
12	9	16	24	29	34	39	45	50	57	67	75
13	11	18	26	33	39	44	50	56	64	75	84
14	13	20	29	36	43	48	55	62	70	81	92
15	14	22	31	39	45	52	58	66	74	86	97
16	15	23	33	41	48	54	61	69	78	91	102
17	16	25	35	43	50	57	64	72	81	95	107
Girls											
9	8	12	17	21	24	26	29	33	36	42	47
10	7	11	17	21	24	27	30	34	38	44	49
11	6	11	16	21	24	28	31	35	40	46	52
12	5	10	16	21	25	28	32	36	41	48	54
13	5	10	17	21	25	29	33	37	42	50	56
14	5	10	17	21	25	29	33	38	43	50	57
15	5	10	17	22	26	30	34	38	44	51	58
16	5	10	17	22	26	30	34	39	44	52	59
17	5	11	17	22	26	30	35	39	45	53	60

Table 4
Relative peak oxygen uptake ($\dot{V}O_{2peak}$, mL/kg/min) centiles by age and sex in 1 142 026 children and youth aged 9–17 years from 50 countries since 1981

Age (years)	P_5	P_{10}	P_{20}	P_{30}	P_{40}	P_{50}	P_{60}	P_{70}	P_{80}	P_{90}	P_{95}
Boys											
9	41.7	43.1	44.8	46.1	47.1	48.1	49.1	50.2	51.5	53.2	54.6
10	39.7	41.3	43.2	44.6	45.8	46.9	48.0	49.2	50.6	52.5	54.1
11	37.7	39.5	41.7	43.2	44.6	45.8	47.0	48.4	49.9	52.1	53.9
12	36.0	38.0	40.5	42.3	43.8	45.2	46.7	48.2	50.0	52.4	54.5
13	34.6	36.9	39.7	41.7	43.4	45.0	46.6	48.3	50.3	53.1	55.4
14	33.3	35.8	38.8	41.0	42.9	44.6	46.4	48.3	50.5	53.5	56.0
15	31.8	34.5	37.8	40.1	42.1	44.0	45.9	47.9	50.2	53.4	56.1
16	30.4	33.2	36.6	39.1	41.3	43.2	45.2	47.3	49.8	53.3	56.1
17	28.9	32.0	35.6	38.2	40.5	42.6	44.7	46.9	49.5	53.2	56.2
Girls											
9	41.3	42.5	44.0	45.0	45.9	46.7	47.5	48.4	49.4	50.8	52.0
10	39.0	40.4	42.0	43.2	44.2	45.1	46.0	47.0	48.2	49.8	51.1
11	36.8	38.3	40.1	41.4	42.5	43.5	44.5	45.7	47.0	48.7	50.2
12	34.6	36.2	38.2	39.6	40.8	42.0	43.1	44.3	45.7	47.7	49.3
13	32.6	34.3	36.4	37.9	39.2	40.4	41.6	42.9	44.4	46.5	48.2
14	30.6	32.4	34.6	36.2	37.6	38.8	40.1	41.4	43.0	45.2	47.0
15	28.7	30.5	32.8	34.5	35.9	37.2	38.5	39.9	41.6	43.9	45.7
16	26.7	28.7	31.1	32.8	34.2	35.6	37.0	38.4	40.2	42.5	44.5
17	24.7	26.8	29.3	31.1	32.6	34.1	35.5	37.0	38.8	41.3	43.3

Discussion

This study systematically analysed 20mSRT data on 1 142 026 children and youth aged 9–17 years to generate the most comprehensive and up-to-date set of international sex-specific and age-specific norms for CRE. These international norms have utility for health and fitness screening, athlete profiling, and monitoring and surveillance in health, clinical, educational or sporting settings. They complement a

growing body of literature reporting national, regional and international growth centiles across a range of cardiometabolic disease risk factors, including adiposity (eg, body mass index[192, 193] and waist circumference[194–198]), blood pressure,[199, 200] cholesterol,[200] triglycerides,[200] glucose[200] and health-related fitness.[28, 50, 126, 201–203]

Using a quintile framework, children in the bottom 20% can be classified as having 'very low' CRE; between the 20th and 40th centiles as having 'low' CRE; between the 40th and 60th centiles as having 'moderate' CRE; between the 60th and 80th centiles as having 'high' CRE; and above the 80th centile as having 'very high' CRE. Single measures of 20mSRT performance taken in health, clinical, educational or sporting settings can then be qualitatively interpreted using these quintile-based thresholds, with longitudinal changes (eg, due to growth or exercise training interventions) monitored by tracking changes against centile bands. For example, Armstrong *et al*[1] estimated that in children an appropriate 12-week CRE training programme will induce, on average, an 8–9% increase in $\dot{V}O_{2peak}$ independent of sex, age and maturation, equivalent to an increase of ~0.5 standardised units or ~20 centile points, which should be enough for a child to shift upwards from one quintile band to the next or above the relevant criterion-referenced standard for low cardiometabolic risk.[204]

While these norms are not criterion-referenced in that they do not indicate whether children have healthy CRE or low cardiometabolic risk, this study does provide an estimate of the prevalence with healthy CRE according to the new FITNESSGRAM (V.10) standards, which have been shown to discriminate with moderate accuracy between youth with and without metabolic syndrome.[37] There are currently no agreed on international criterion-referenced standards for CRE, and while we estimated prevalence using the US-based FITNESSGRAM standards, other national and regional standards have been published elsewhere.[26, 205–211] It is important to note that the differences between published standards are (in some cases) substantial, with the new FITNESSGRAM (V.10) standards generally higher for girls and lower for boys relative to other standards,[26, 205–211] meaning that our prevalence estimates would be substantially different if calculated using other standards. In contrast, $\dot{V}O_{2peak}$ can be estimated using different test protocols and prediction equations, and special care must be taken when comparing $\dot{V}O_{2peak}$ values with standards that were estimated using different test protocols and prediction equations.[212] For example, we predicted $\dot{V}O_{2peak}$ from 20mSRT performance using the Léger *et al*[35] equation, whereas the new FITNESSGRAM standards were developed using predicted $\dot{V}O_{2peak}$ from a laboratory-based treadmill test.[37] To our knowledge, at least 17 prediction equations (from 10 studies of apparently healthy children and youth[37, 213–221]) are currently available to estimate $\dot{V}O_{2peak}$ from 20mSRT performance, equations that differ in validity and can result in substantially different estimates of $\dot{V}O_{2peak}$ and hence the prevalence of healthy CRE. Future studies need to build on multinational efforts (eg, the HELENA study in Europe) in order to develop the most valid international normative-referenced and criterion-referenced standards for CRE.[222]

Using a cross-sectional approach, this study quantified the sex-related and age-related differences in 20mSRT speed, showing that boys consistently outperformed girls and experienced larger age-related changes. Given that longitudinal data are required to determine the true developmental patterns of CRE (because they control for within-participant changes in timing and tempo), and that larger age-related increases have been observed in children followed longitudinally than in those examined cross-sectionally,[223] it is possible that our data underestimate the true developmental patterns. While the developmental patterns of children's $\dot{V}O_{2peak}$ has been well studied in non-representative longitudinal samples,[224–226] other aspects of CRE (eg, mechanical efficiency, fractional utilisation, $\dot{V}O_2$ kinetics, etc) are less well understood, making it difficult to describe the mechanistic causes underlying the developmental patterns in 20mSRT performance. Given that relative $\dot{V}O_2$ and $\dot{V}O_{2peak}$ vary linearly with

speed and peak speed, changes in 20mSRT speed should therefore reflect changes in underlying $\dot{V}O_2$ (ie, the oxygen cost of the activity). Interestingly, however, this study observed decreases in boys' relative $\dot{V}O_{2peak}$ throughout childhood and adolescence, rather than the expected plateau.[227] This unexpected developmental pattern may be due in part to $\dot{V}O_{2peak}$ prediction error or the fact that the developmental patterns were calculated using pooled cross-sectional data. In contrast, the premise of a plateau in boys' relative $\dot{V}O_{2peak}$ throughout childhood and adolescence is largely based on progressive treadmill or cycle testing of volunteer recruits who were athletically inclined, non-obese and motivated to exercise.[226, 228] This unexpected finding in the developmental pattern in boys' $\dot{V}O_{2peak}$ might reflect current maturational development, or our use of a large, internationally representative sample of children and youth.

Strengths and limitations

This study brings together 20mSRT data from 177 studies in what is to date the largest international CRE database in the world. We conducted a systematic review, using a strict set of inclusion and exclusion criteria, applied rigorous data treatment procedures to systematically control for bias (eg, differences in test protocols and performance metrics), a poststratification population weighting procedure, a novel pseudo data method and the LMS method, to generate international sex-specific and age-specific smoothed centiles (across four performance metrics) for CRE. While it is not the first comprehensive study of children's 20mSRT performance, it updates the comprehensive reviews of Olds *et al*[21] and Tomkinson *et al*[22] by: (1) extending the data coverage from 2003 to 2014 through a rigorous systematic review process, (2) producing sex-specific and age-specific international normative centiles and (3) estimating the international prevalence of children and youth with healthy CRE.

Unfortunately, there are several sampling-related and methodology-related limitations to this study. First, it pooled data from studies that used different sampling methods which raises the issue of representativeness, with some studies using probability sampling (eg, stratified proportional, stratified random or cluster sampling) and others non-probability sampling (eg, convenience sampling; see online supplement 2). Second, differences in sampling base also exist across national, state/provincial and local/city/school levels. Third, differences in testing conditions (eg, environmental conditions such as climate or altitude, practice and running surfaces) and measurement errors (eg, calibration and type of equipment, methodological drift and diurnal variation) are inherent to any large data synthesis, but the very large number of data points captured should minimise these issues. Fourth, the vigorous nature of the 20mSRT may result in difficulties in test administration in, or exclusion of individuals with physical disabilities and diseases. The absence of data from these populations is likely to have inflated our normative values within the lower centile range.

The 20mSRT data were also collected at different times over the period 1981–2014, and given convincing evidence of international declines in children's CRE in recent decades,[1, 23, 229] our norms may be biased. With a mean measurement year of 2000, and assuming an international decline of ~5% per decade since 1975,[1, 23] our normative data may overestimate values in 2014 by ~7.5% (equivalent to ~0.65 standardised units or ~24 centile points) and represent a better health-related picture than what would be true today. However, without time trend data for all included countries and evidence of recent improvements in children's CRE in some included countries (eg, Australia,[28] Japan[229] and Spain[230]), time-based corrections of our normative data are not recommended. Despite relying predominantly on data from high-income economies, there is no good evidence that 20mSRT performance is meaningfully related to a country's affluence or distribution of wealth,[21, 231] so our norms are unlikely to be biased. Future studies need to examine CRE in low-income and middle-income economies and at multiple stages of the CRE transition. It must also be remembered that the 20mSRT is affected by factors other than

underlying construct CRE.[20, 232] Validity data show that a moderate-to-large (35–70%) amount of the variance in 20mSRT performance is explained by the variance in underlying $\dot{V}O_{2peak}$,[20, 232, 233] indicating that other physiological (eg, mechanical efficiency),[234] fractional utilisation,[235] $\dot{V}O_{2peak}$ kinetics,[236] lactate threshold,[237] anaerobic capacity,[238] physical (eg, fat mass)[239] and psychosocial factors (eg, motivation, effort and self-efficacy[20]) also contribute.

Recommendations

Over the past few decades, the 20mSRT has been widely used to assess the CRE of children and youth, and yet data pooling is nearly impossible due to the difficulty with standardising performances (eg, because of differences in protocols, performance metrics, the way in which age is expressed, etc). To facilitate data pooling in the future, and to assist with the eventual update of 20mSRT norms, we make the following recommendations:

1. The test protocol used should be thoroughly and accurately reported;
2. Care should be taken to minimise and report factors that affect 20mSRT performance (eg, testing conditions and measurement errors);
3. Best practice should include 20mSRT results that are reported as the running speed (km/h) at the last completed stage;
4. Descriptive statistics (sample sizes, means and SDs) should be reported at the sex by age (at last birthday) level; and
5. The year(s) of testing should be reported;

Furthermore, because the 20mSRT is a maximal effort test, in order to ensure that a child has performed with 'good effort' (ie, they have tried very hard), perceptual (eg, ratings of perceived exertion) and/or physiological (eg, heart rate) effort should be measured in addition to performance effort.[240] Any adverse events (or lack thereof) associated with maximal effort tests such as the 20mSRT should also be reported.[241]

Conclusion

CRE is considered to be an excellent marker of current and future health. The 20mSRT is arguably the most popular measure of CRE because it is suitable for mass testing, is simple, cheap, easy, reliable, reasonably valid and is part of widely used health-related fitness test batteries (eg, Assessing Levels of PHysical Activity and fitness (ALPHA),[242] Canadian Assessment of Physical Literacy (CAPL),[243] Eurofit,[244] FITNESSGRAM[245] and even the PREFIT battery (Assessing FITness in PREscholers)).[246, 247] Using a systematic review and analytical approach, this study used the best available 20mSRT data to: (1) provide the most comprehensive and up-to-date set of international sex-specific and age-specific norms for children and youth; and (2) estimate the prevalence with healthy CRE according to the FITNESSGRAM standards. These data have utility for health and sport promotion given that they help to identify children and youth with: (1) very low CRE in order to set appropriate fitness goals, monitor longitudinal changes and promote positive health-related fitness behaviours; and (2) very high CRE in the hope of recruiting them into elite sporting or athletic development programmes.

References

1 Armstrong N, Tomkinson GR, Ekelund U. Aerobic fitness and its relationship to sport, exercise training and habitual physical activity during youth. *Br J Sports Med* 2011;45:849–58.

2 Institute of Medicine. Health-related fitness measures for youth: cardiorespiratory endurance. In: Institute of Medicine, ed. *Fitness measures and health outcomes in youth.* Washington DC: The National Academy Press, 2012:111–51.

3 Blair SN, Kohl HW III, Paffenbarger RS Jr, *et al*. Physical fitness and all-cause mortality. A prospective study of healthy men and women. *JAMA* 1989;262:2395–401.

4 Kodama S, Saito K, Tanaka S, *et al*. Cardiorespiratory fitness as a quantitative predictor of all-cause mortality and cardiovascular events in healthy men and women: a meta-analysis. *JAMA* 2009;301:2024–35.

5 Lee CD, Blair SN. Cardiorespiratory fitness and stroke mortality in men. *Med Sci Sports Exerc* 2002;34:592–5.

6 Katzmarzyk P, Church TS, Janssen I, *et al*. Metabolic syndrome, obesity, and mortality: impact of cardiorespiratory fitness. *Diabetes Care* 2005;28:391–7.

7 Dishman RK, Washburn RA, Heath GW. *Physical activity epidemiology.* Champaign, IL: Human Kinetics, 2004:358–9.

8 Gillespie LD, Robertson MC, Gillespie WJ, *et al*. Interventions for preventing falls in older people living in the community. *Cochrane Database Syst Rev* 2009;(2): CD007146.

9 Katzmarzyk PT, Church TS, Blair SN. Cardiorespiratory fitness attenuates the effects of the metabolic syndrome on all-cause and cardiovascular disease mortality in men. *Arch Intern Med* 2004;164:1092–7.

10 Ortega FB, Ruiz JR, Castillo MJ, *et al*. Physical fitness in children and adolescence: a powerful marker of health. *Int J Obes* 2008;32:1–11.

11 Ruiz JR, Castro-Piñero J, Artero EG, *et al*. Predictive validity of health-related fitness in youth: a systematic review. *Br J Sports Med* 2009;43:909–23.

12 Eisenmann JC, Katzmarzyk PT, Perusse L. Aerobic fitness, body mass index, and CVD risk factors among adolescents: the Québec family study. *Int J Obes* 2005;29:1077–83.

13 Malina RM. Physical activity and fitness: pathways from childhood to adulthood. *Am J Hum Biol* 2001;13:162–72.

14 Ortega FB, Ruiz JR, Labayen I, *et al*. Role of socio-cultural factors on changes in fitness and adiposity in youth: a 6-year follow-up study. *Nutr Metab Cardiovasc Dis* 2013;23:883–90.

15 Dwyer T, Magnussen CG, Schmidt MD, et al. Decline in physical fitness from childhood to adulthood associated with increased obesity and insulin resistance in adults. *Diabetes Care* 2009;32:683–7.

16 Ortega FB, Silventoinen K, Tynelius P, *et al*. Muscular strength in male adolescents and premature death: cohort study of one million participants. *BMJ* 2012;345:e7279.

17 Ortega FB, Labayen I, Ruiz JR, *et al*. Improvements in fitness reduce the risk of becoming overweight across puberty. *Med Sci Sports Exerc* 2011;43:1891–7.

18 Morales PF, Sánchez-López M, Moya-Martinez M, *et al*. Health-related quality of life, obesity, and fitness in schoolchildren: the Cuenca study. *Qual Life Res* 2013;22:1515–23.

19 Moliner-Urdiales D, Ruiz JR, Ortega FB, *et al*. Association of physical activity with muscular strength and fat-free mass in adolescents: the HELENA study. *Eur J Appl Physiol* 2010;109:1119–27.

20 Tomkinson GR, Olds TS. Field tests of fitness. In: Armstrong N, van Mechelen W, eds. *Paediatric exercise science and medicine.* New York, NY: Oxford University Press, 2008:109–28.

21 Olds TS, Tomkinson GR, Léger L, *et al*. Worldwide variation in the performance of children and adolescents: an analysis of 109 studies of the 20-m shuttle run test in 37 countries. *J Sports Sci* 2006;24:1025–38.

22 Tomkinson GR, Léger LA, Olds TS, *et al*. Secular trends in the performance of children and adolescents (1980–2000): an analysis of 55 studies of the 20m shuttle run test in 11 countries. *Sports Med* 2003;33:285–300.

23 Tomkinson G, Olds T. Secular changes in pediatric aerobic fitness test performance: the global picture. *Med Sport Sci* 2007;50:46–66.

24 Ruiz JR, Silva G, Oliveira N, *et al*. Criterion related validity of the 20m shuttle run test in adolescents aged 13–19 years. *J Sports Sci* 2009;27:899–06.

25 Melo X, Santa-Clara H, Almeida JP, *et al*. Comparing several equations that predict peak VO_ using the 20-m multistage-shuttle run-test in 8-10-year-old children. *Eur J Appl Physiol* 2011;111:839–49.

26 Silva G, Aires L, Mota J, *et al*. Normative and criterion-related standards for shuttle run performance in youth. *Pediatr Exerc Sci* 2012;24:157–69.

27 Carrel AL, Bowser J, White D, *et al*. Standardized childhood fitness percentiles derived from school-based testing. *J Pediatr* 2012;161:120–4.

28 Catley MJ, Tomkinson GR. Normative health-related fitness values for children: analysis of 85347 test results on 9-17-year-old Australians since 1985. *Br J Sports Med* 2013;47:98–108.

29 Vandongen R, Jenner DA, Thompson C, *et al*. A controlled evaluation of a fitness and nutrition intervention program on cardiovascular health in 10- to 12-year-old children. *Prev Med* 1995;24:9–22.

30 Léger L, Lambert J, Goulet A, *et al*. Capacité aérobie des Québécois de 6 à 17 ans —Test navette de 20 mètres avec paliers de 1 minute. *Can J Appl Sport Sci* 1984;9:64–9.

31 Levy PS, Lemeshow S. Stratification random sampling: further issues. In: Levy PS, Lemeshow S, eds. *Sampling of populations: methods and application*. Hoboken, NJ: John Wiley & Sons, Inc, 2008:143–88.

32 United Nations, Department of Economic and Social Affairs, Population Division (2015). *World Population Prospects: The 2015 Revision, Key Findings and Advance Tables* Working Paper No. ESA/P/WP.241.

33 Cole TJ, Green PJ. Smoothing reference centile curves: the LMS method and penalized likelihood. *Stats Med* 1992;11:1305–19.

34 Pan H, Cole T. *User's guide to LMSchartmaker*. UK: Medical Research Council, 2010:1–42.

35 Léger LA, Mercier D, Gadoury C, *et al*. The multistage 20 metre shuttle run test for aerobic fitness. *J Sports Sci* 1988;6:93–101.

36 Plowman SA, Meredith MD. *Fitnessgram/activitygram reference guide*. 4th edn. Dallas, TX: The Cooper Institute, 2013.

37 Welk GJ, Laurson KR, Eisenmann JC, *et al*. Development of youth aerobic-capacity standards using receiver operating characteristic curves. *Am J Prev Med* 2011;41: S111–16.

38 Cohen J. *Statistical power analysis for the behavioral sciences*. 2nd edn. New Jersey: Lawrence Erlbaum, 1988.

39 Secchi JD, García GC, Espana-Romero V, *et al.* Physical fitness and future cardiovascular risk in Argentine children and adolescents: an introduction to the ALPHA test battery. *Arch Argent Pediatr* 2014;112:132–40.

40 Hardy LL, King L, Espinel P, *et al. NSW Schools Physical Activity and Nutrition Survey (SPANS) 2010: full report*. Sydney: NSW Ministry of Health, 2010.

41 McIntyre S. *Trends in aerobic fitness from 1994 to 2009 in 10- and 11-year-old Australian children [thesis].* Notre Dame, IN: University of Notre Dame, 2009.

42 Booth M, Okely AD, Denney-Wilson E, *et al. NSW Schools Physical Activity and Nutrition Survey (SPANS) 2004: full report.* Sydney: NSW Department of Health, 2006.

43 Lloyd KC, Antonas KN. Nutritional habits and fitness levels of schoolchildren. In: *Proceedings of the Nutrition Society of Australia—twenty-fourth annual scientific meeting, Fremantle, WA, 3-6 December*. Adelaide, SA: Nutrition Society of Australia, 2000:138.

44 Cooley D, McNaughton L. Aerobic fitness of Tasmanian secondary school children using the 20-m shuttle run test. *Percept Mot Skills* 1999;88:188–98.

45 Booth M, Macaskill P, McLellan L, *et al*. *NSW schools fitness and physical activity survey 1997*. Sydney: NSW Department of Education and Training, 1997.

46 Okely AD, Gray T, Cotton WG. Effect of an extended stay outdoor education program on aerobic fitness. In: Gray T, Hayllar B, eds. *Catalysts for change: proceedings from the 10th National Outdoor Education Conference*. Sydney, Australia, 1997:206–10.

47 Australian Council for Health, Physical Education and Recreation. *Australian fitness education award: user's manual and curriculum ideas*. Adelaide, SA: Australian Council for Health, Physical Education and Recreation, 1996.

48 Australian Sports Commission. *Norms for sport related fitness tests in Australian students aged 12-17 years.* Belconnen, ACT: Australian Sports Commission, 1994.

49 Jenner DA, Vandongen R, Beilin L. Relationships between blood pressure and measures of dietary energy intake, physical fitness, and physical activity in Australian children aged 11-12 years. *J Epidemiol Community Health* 1992;46:108–13.

50 Ortega FB, Artero EG, Ruiz JR, *et al.* Physical fitness levels among European adolescents: the HELENA study. *Br J Sports Med* 2011;45:20–9.

51 Vandendriessche JB, Vandorpe BFR, Vaeyens R, *et al*. Variation in sport participation, fitness and motor coordination with socioeconomic status among Flemish children. *Pediatr Exerc Sci* 2012;24:113–28.

52 Seghers J, Rutten C. Clustering of multiple lifestyle behaviours and its relationship with weight status and cardiorespiratory fitness in a sample of Flemish 11- to 12-year olds. *Public Health Nutr* 2010;13:1838–46.

53 Verstraete SJ, Cardon GM, De Clercq DL, *et al*. A comprehensive physical activity promotion programme at elementary school: the effects on physical activity, physical fitness and psychosocial correlates of physical activity. *Public Health Nutr* 2007;10:477–84.

54 Cardon G, De Bourdeaudhuij I, De Clercq D, *et al*. Physical fitness, physical activity, and self-reported back and neck pain in elementary schoolchildren. *Pediatr Exerc Sci* 2004;16:147–57.

55 Heyters C, Marique T. *Le baromètre de la condition physique*. Bruxelles: Ministère de la Communauté française, Direction générale du sport, 2004.

56 Telama R, Naul R, Nupponen H, *et al*. *Physical fitness, sporting lifestyles and Olympic ideals: cross-cultural studies on youth sport in Europe*. Schorndorf, Germany: Verlag Karl Hofmann, 2002.

57 Baquet G, Berthoin S, Padovano C, *et al*. Effects d'un cycle de course de duree de type intermittent (court-court) sur la condition physique des adolescents. *Rev Educ Physique* 2000;40:51–60.

58 Lefèvre J, Bouckaert J, Duquet W, *et al*. De barometer van de fysieke fitheid van de Vlaamse jeugd 1997. *De resultaten. Sport (Bloso Brussel)* 1998;4:16–22.

59 Beunen G, Borms J, Vrijens J, *et al. Fysieke fitheid en sportbeoefening van de Vlaamse jeugd*. Volumen 1: Fysieke fitheid van de jeugd van 6 tot 18 jaar. Brussels: Bloso, 1991.

60 Poortmans J, Vlaeminck M, Collin M, *et al*. Estimation indirecte de la puissance aérobie maximale d'une population Bruxelloise masculine et féminine âgée de 6 à 23 ans. Comparaison avec une technique directe de la mesure de la consommation maximale d'oxygène. *J Physiol (Paris)* 1986;81:195–201.

61 Cazorla G, Gouthon P, Arémou, *et al*. *État de développement des capacités physiques des jeunes béninois 8-18 ans et plus. Rapport pour le Ministère Béninois de la Jeunesse et Sports et pour le Ministère Français de la Coopération*. Paris, France: Ministrère Béninois de la Jeunesse et Sports et pour le Ministrère Français de la Coopération, 1987.

62 Ribeiro RR, Santos KD, Carvalho WCG, *et al.* Aerobic fitness and biological and sociodemographic indicators in female school children. *Rev Bras Cineantropom Desempenho Hum* 2013;15:448–57.

63 Hobold E. *Indicadores de aptidão física relacionada à saúde de crianças e adolescentes do município de Marechal Cândido Rondon* [dissertation]. Paraná, Brasil: Universidade Federal de Santa Catarina, 2003.

64 Pieta S. *Estudio de la aptitud física de una muestra de la población escolar del estado de Paraná mediante la bateria Eurofit* [dissertation]. León, Brasil: Universidad de León, 2000.

65 Veldhuizen S, Cairney J, Hay J, *et al.* Relative age effects in fitness testing in a general school sample: how relative are they? *J Sports Sci* 2014;33:109–15.

66 Voss C, Sandercock G, Higgins JW, *et al*. A cross-cultural comparison of body composition, physical fitness, and physical activity between regional samples of Canadian and English children and adolescents. *Can J Public Health* 2014;104: e245–50.

67 Leone M, Kalinova E, Comtois AS. *Global motor skill assessment from the UQAC-UQAM test battery: Canadian normative values by age and gender.* Québec: Université du Québec à Chicoutimi, 2011.

68 Reed KE, Warburton DER, Whitney CL, *et al*. Secular changes in shuttle-run performance: a 23-year retrospective comparison of 9- to 11-year-old children. *Pediatr Exerc Sci* 2006;18:364–73.

69 Massicotte D. *Partial curl-ups, push-ups and multistage 20 meter shuttle run, national norms for 6 to 17 year-olds. Final report submitted to Canadian Association for Health, Physical Education and Recreation (CAHPER) and Fitness and Amateur Sport Canada*. Montréal: University of Quebec at Montréal, 1990.

70 Garber MD, Sajuria M, Lobelo F. Geographical variation in health-related physical fitness and body composition among Chilean 8th graders: a nationally representative cross-sectional study. *PLoS ONE* 2014;9:e108053.

71 Kain J, Uauy R, Albala, *et al*. School-based obesity prevention in Chilean primary school children: methodology and evaluation of a controlled study. *Int J Obes* 2004;28:483–93.

72 Wong TW, Yu TS, Wang XR, *et al*. Predicted maximal oxygen uptake in normal Hong Kong Chinese schoolchildren and those with respiratory diseases. *Pediatr Pulmonol* 2001;31:126–32.

73 Wang P-G, Gong G, Wang SQ, *et al*. Relationship of body fat and cardiorespiratory fitness with cardiovascular risk in Chinese children. *PLoS ONE* 2011;6:e27896.

74 Cubides RC, Alarcón LGA, Galvi ARG. Diferencias en la actividad física y la condición física entre los escolares de secundaria de dos programas curriculares oficiales de Bogotá, Colombia. *Nutrición Hospitalaria* 2015;32:2228–34.

75 Ortega FJA. Estudio transversal de las cualidades funcionales de los escolares bogotanos: Valores de potencia aeróbica, potencia muscular, velocidad de desplazamiento y velocidad de reacción, de los siete a los dieciocho años. *Med Deporte Trab* 2013;32:1151–70.

76 Mojica GT, Poveda JG, Pinilla MI, *et al*. Sobrepeso, inactividad física y baja condición física en un colegio de Bogotá, Colombia. *Arch Latinoam Nutr* 2008;58:265–73.

77 Cazorla G, Dudal J, Dieu JL. *État de développement des capacités physiques des jeunes Ivoiriens 7-18 ans et plus. Rapport pour le Ministère Ivoirien de la Jeunesse et Sports et pour le Ministère Français de la Coopération.* Paris, France: Ministrère Ivoirien de la Jeunesse et Sports et pour le Ministrère Français de la Coopération, 1985.

78 Shahin A. Ying and yang of body composition assessment. *Chin J Integr Med* 2011;17:675–9.

79 Tinazci C, Emíroğlu O. Assessment of physical fitness levels, gender and age differences of rural and urban elementary school children. *Turkiye Klinikleri J Med Sci* 2010;30:1–7.

80 Gajda V. The applications of the UNIFITTEST battery at basic schools (age groups 7-14): summary report. Ostrava: PdF OU, 1994.

81 Nielsen GA, Andersen LB. The association between high blood pressure, physical fitness, and body mass index in adolescents. *Prev Med* 2003;36:229–34.

82 Cazorla G, Dudal J, Lefrancq L. *État de développement des capacités physiques des jeunes Djiboutiens 11-18 ans et plus. Rapport pour le Ministère Djiboutien de la Jeunesse et Sports et pour le Ministère Français de la Coopération*. Paris, France: Ministrère Djiboutien de la Jeunesse et Sports et pour le Ministrère Français de la Coopération, 1986.

83 Jürimäe T, Volbekiene V, Jürimäe J. Changes in Eurofit test performance of Estonian and Lithuanian children and adolescents (1992–2002). In: Tomkinson GR, Olds TS, eds. *Pediatric fitness: secular trends and geographic variability*. Basel: Karger, 2007:129–42.

84 Jürimäe T, Saar M. Self-perceived and actual indicators of motor abilities in children and adolescents. *Percet Mot Skills* 2003;97:862–6.

85 Jürimäe T, Volbekiene V. Eurofit test results in Estonian and Lithuanian 11 to 17-year-old children: a comparative study. *Eur J Phys Educ* 1998;3:178–84.

86 Raudsepp L, Jürimäe T. Relationships of physical activity and somatic characteristics with physical fitness and motor skill in prepubertal girls. *Am J Hum Biol* 1997;9:513–21.

87 Raudsepp L, Jürimäe T. Relationships between somatic variables, physical activity, fitness and fundamental motor skills in prepubertal boys. *Biol Sport* 1996;13:279–89.

88 Kull M, Jürimäe T. Using the Eurofit test battery in Estonian 16-18 years old adolescents. *Acta Commentationes Univ Tartuensis* 1994;967:49–52.

89 Palomäki S, Heikinaro-Johansson P, Huotari P. Cardiorespiratory performance and physical activity in normal weight and overweight Finnish adolescents from 2003 to 2010. *J Sports Sci* 2015;33:588–96.

90 Viljanen T, Taimela S, Kujala U. Koululaisten fyysinen aktiivisuus, kestävyyskunto ja ponnistuskorkeus. *Liikunta Tiede* 2000;37:23–6.

91 Baquet G, Twisk JWR, Kemper HCG, *et al*. Longitudinal follow-up of fitness during childhood: interaction with physical activity. *Am J Hum Biol* 2006;18:51–8.

92 Baquet G, Berthoin S, Gerbeaux M, *et al*. High-intensity aerobic training during a 10 week one-hour physical education cycle: effect on physical fitness of adolescents aged 11 to 16. *Int J Sports Exerc Med* 2001;22:295–300.

93 Baquet G, Berthoin S, Gerbeaux M, *et al*. Assessment of the maximal aerobic speed with the incremental running field tests in children. *Biol Sport* 1999;16:23–30.

94 Cazorla G, Portes A, James F. *Opération Martinique-Eval. Centre d'Evaluation Sport Santé, Fort de France (Martinique). Rapport pour L'Inspection d'Académie de la Martinique*. Fort de France, Martinique: L'Education d'Académie de la Martique, 1997.

95 Cazorla G. *Batterie France-Éval: Mesures, épreuves et barêmes: Évaluation des qualités physiques des jeunes Français d'âge scolaire: 7-11 ans. Rapport pour le Secrétariat d'Etat Auprès du*

Premier Ministre Chargé de la Jeunesse et de Sports. Paris: Ministère de la Jeunesse et de Sports, 1987.

96 Brunet J, Van Praagh E. *Batterie experimentale de tests moteurs Eurofit: Rapport d'Activitie de la Région Auvergne—1984-1985*. Clermont-Ferrand, France: Universite de Clermont-Ferrand, 1985.

97 Tambalis KD, Panagiotakos DB, Psarra G, *et al*. Physical fitness normative values for 6-18-year-old Greek boys and girls, using the empirical distribution and the lambda, mu, and sigma statistical method. *Eur J Sport Sci* 2015:1–11.

98 Tambalis K, Panagiotakos D, Sidossis L. Greek children living in rural areas are heavier but fitter compared to their urban counterparts: a comparative, time-series (1997–2008) analysis. *J Rural Health* 2011;27:270–7.

99 Tokmakidis SP, Kasambalis A, Christodoulos AD. Fitness levels of Greek primary schoolchildren in relationship to overweight and obesity. *Eur J Pediatr* 2006;165:867–74.

100 Manios Y, Yiannakouris N, Papoutsakis C. Behavorial and physiological indices related to BMI in a cohort of primary schoolchildren in Greece. *Am J Hum Biol* 2004;16:639–47.

101 Georgiadis G. *Evaluation of physical fitness of Greek youth aged 6-18 years* [dissertation]. Athens, Greece: University of Athens, 1993.

102 Welk GJ, Saint-Maurice PF, Csányi T. Health-related physical fitness in Hungarian youth: age, sex, and regional profiles. *Res Q Exerc Sport* 2015;86:S45–57.

103 Barabás A. Measurement of aerobic power by field tests. In: Coudert J, Van Praagh E, eds. *Pediatric work physiology: children and exercise XVI*. Paris: Masson, 1992:39–41.

104 Gunnarsson HG, Sigríksson R. *Eurofit: physical fitness of Icelandic pupils at age of 6-15 years old*. Reykjavik, Iceland, 1999.

105 Grassi GP, Turci M, Sforza C. Aerobic fitness and somatic growth in adolescents: a cross sectional investigation in a high school context. *J Sports Med Phys Fitness* 2006;46:412–18.

106 Cilia G, Bazzano C, Bellucci M, *et al*. I risultati dei test Eurofit nella scuola Matteuccii di Roma. *Alcmeone* 1998;2:16–20.

107 Cilia G, Bellucci M, Bazzano C, *et al*. Eurofit 1997: Banche dati per la scuola. *Alcmeone* 1997;3:13–32.

108 Cilia G, Bellucci M, Riva M. *Eurofit 1995.* Roma: Istituto Superiore Statale di Educazione Fisica, 1996.

109 Cilia G, Bellucci M. *Eurofit: tests Europei di attitudine fisica.* Roma: Istituto Superiore Statale di Educazione Fisica, 1993.

110 Council of Europe. *Évaluation de l'aptitude physique: Eurofit batterie expérimentale*. Rome: Council of Europe, 1986.

111 Ministry of Education, Culture, Sports, Science and Technology. *Report book on the survey of physical fitness and athletic ability.* Tokyo: Ministry of Education, Culture, Sports, Science and Technology, 1999–2015.

112 Sauka M, Priedite IS, Artjuhova L, *et al*. Physical fitness in northern European youth: reference values from the Latvian physical health in youth study. *Scand J Public Health* 2011;39:35–43.

113 Cazorla G, Dudal J, Garinet P. État de développement des capacités physiques des jeunes mauriciens 7-18 ans et plus. Rapport pour le Ministère Mauricien de la Jeunesse et Sports et pour le Ministère Français de la Coopération, 1986.

114 Galaviz KI, Tremblay MS, Colley R, *et al*. Associations between physical activity, cardiorespiratory fitness, and obesity in Mexican children. *Salud Pública Méx* 2012;54:463–9.

115 Dadouchi MF, Boudiab K, Yahia J, *et al. De la détermination du profil athlétique marocain au contenu et à l'évaluation. Ministère de l'Éducation nationale et de la jeunesse.* École normale supérieure de Marrakech, 2003.

116 Brouwer SI, Stol RP, Liem ET*, et al*. The role of fitness in the association between fatness and cardiometabolic risk from children to adolescence. *Pediatr Diabetes* 2013;14:57–65.

117 Slinger J, van Breda E, Kuipers H. Aerobic fitness data for Dutch adolescents (2002-2005). *Pediatr Exerc Sci* 2009;21:10–18.

118 van Mechelen W, van Lier WH, Hlobil H, *et al*. *Eurofit: Handleiding met referentieschalen voor 12- tot en met 16-jarige jongens en meisjes in Nederland*. Haarlem: Uitgeverij de Vrieseborch, 1991.

119 Haugen T, Høigaard R, Seiler S. Normative data of BMI and physical fitness in a Norwegian sample of early adolescents. *Scand J Public Health* 2014;42:67–73.

120 Cossio-Bolaños MA, Arruda M. Propuesta de valores normativos para la evaluación de la aptitud física en niños de 6 a 12 años de Arequipa, Perú. *Rev Med Herediana* 2009;20:206–12.

121 Gonzalez-Suarez CB, Grimmer-Sommers K. The association of physical activity and physical fitness with pre-adolescent obesity: an observational study in Metromanila, Phillipines. *J Phy Act Health* 2011;8:804–10.

122 Bronikowski M, Bronikowska M. Salutogenesis as a framework for improving health resources of adolescent boys. *Scand J Public Health* 2009;37:525–31.

123 Pilicz S, Przeweda R, Dobosz J, *et al*. *Punktacja sprawnosci fizycznej mlodziezy Polskiej wg miedzynarodowego testu sprawnosci fizycznej: Kryteria pomiaru wydolnosci organizmu testem Coopera.* Warszawa: Akademia Wychowania Fizycznego Józefa Pilsudskiego, 2003.

124 Maciaszek J, Osinski W. Poziom sprawnosci fizycznej u chlopców i dziewczat Poznanskich w wieku 10-14 lat. *Rocz Naukowe AWF Poznaniu* 2001;50:3–17.

125 Mleczko E, Ozimek M. *Rozwój somatyczny i motoryczny mlodziezy Krakowskiej miedzy 15 a 19 rokiem zycia z uwzglednieniem czynników srodowiskowych*. Kraków: Akademia Wychowania Fizycznego, 2000.

126 Santos R, Mota J, Santos DA, *et al.* Physical fitness percentiles for Portuguese children and adolescents aged 10-18 years. *J Sports Sci* 2014;32:1510–81.

127 Coelho-Silva MJ, Ronque ERV, Cyrino ES, *et al*. Nutritional status, biological maturation and cardiorespiratory fitness in Azorean youth aged 11-15 years. *BMC Public Health* 2013;13:495.

128 Rodrigues LP, Leitao R, Lopes VP. Physical fitness predicts adiposity longitudinal changes over childhood and adolescence. *J Sci Med Sport* 2013;16:118–23.

129 Marta CC, Marinho DA, Barbosa TM. Physical fitness differences between prepubescent boys and girls. *J Strength Cond Res* 2012;26:1756–66.

130 Mota J, Guerra S, Leandro C. Association of maturation, sex, and body fat in cardiorespiratory fitness. *Am J Hum Biol* 2002;14:707–12.

131 Kim H-B, Stebbins CL, Chai J-H, *et al*. Taekwondo training and fitness in female adolescents. *J Sports Sci* 2011;29:133–8.

132 Cazorla G, Dudal J, Faye J. *État de développement des capacités physiques des jeunes sénégalais 7-18 ans et plus. Rapport pour le Ministère Sénégalais de la Jeunesse et Sports et pour le Ministère Français de la Coopération*. Paris, France: Ministrère Sénégalais de la Jeunesse et Sports et pour le Ministrère Français de la Coopération, 1988.

133 Bovet P, Auguste R, Burdette H. Strong inverse association between physical fitness and overweight in adolescents: a large school-based survey. *Int J Behav Nutr Phys Act* 2007;4:24.

134 Cazorla G, Rousseau G, Dudal J, *et al. Évaluation des capacités motrices de l'enfant, de l'adolescent et du jeune seychellois : 7-18 ans et plus. Rapport pour le Ministère Seychelle de l'Education et pour le Ministère Français de la Coopération*. Paris, France: Ministère Seychellois de l'Education et Ministrère Français de la Coopération, 1990.

135 Krska P, Sedláček J, Hubinak A, *et al.* General motor fitness and somatic parameters comparison between former population and present primary school girls in Ruzomberok. In: Proceedings of the 9th International Conference in Physical Education, Sport and Physical Therapy, Iaşi, 13–14 November. Iaşi, Romania: Faculty of Sports and Physical Education, Alexandru Ioan Cuza University, 2015.

136 Krska P, Sedláček J, Hubinak A, *et al.* Physical development and general motor performance of present primary school boys population in Ruzomberok. In: Zvonař M, Sajdlová Z, eds. Proceedings of the 10th International Conference of Kinanthropology, Brno, Czech Republic, 18–20 November. Brno: Faculty of Sport Studies, Masaryk University, 2015:289-96.

137 Kyselovicová O. Programy aerobiku z aspektu rozvoja telesnej zdatnosti dievcat na gymnáziu. *Acta Facultatis Educationis Physicae Univ Comenianae* 2000;41:50–61.

138 Kasa J, Majherová M. Physical and motor development of children by Eurofit. *Stud Psychol (Bratisl)* 1997;39:270–4.

139 Moravec R. Eurofit—physique and motor fitness of the Slovak school youth. In: Moravec R, Kampmiller T, Sedláček J, eds. *Eurofit—physique and motor fitness of the Slovak school youth.* Bratislava: Slovak Scientific Society for Physical Education and Sports, 1996:9–22.

140 Belej M, Junger J, Mikus M. Vykonnost ziakov prijatych na stredne skoly vo vychodoslovenskom regione zistovana systemom EUROFIT. *Tel Vych Sport* 1995;5:12–15.

141 Pienaar C, Coetzee B, Monyeki AM. The use of anthropometric measurements and the influence of demographic factors on the prediction of VO_ 2max in a cohort of adolescents: the PAHL study. *Ann Hum Biol* 2015;42:135–43.

142 Du Toit D, Pienaar AE, Truter L. Relationship between physical fitness and academic performance in South African children. *S Afr J Res Sport PH* 2011;33:23–35.

143 Pienaar AE, Viljoen A. Physical and motor ability, anthropometrical and growth characteristics of boys in the northwest province of South Africa: a sport talent perspective. *S Afr J Res Sport Phys Educ Recreation* 2010;32:71–93.

144 Du Preez SM. *The effect of physical activity on the body composition and health related fitness of 9 to 13-year-old boys* [thesis]. Potchefstroom, South Africa: North-West University, 2008.

145 Stadler MC. *The influence of a physical activity intervention program (PAI) on the physical fitness levels, body composition and health risk behaviour of 9 to 13 year old girls* [dissertation]. Potchefstroom, South Africa: North-West University, 2007.

146 Du Toit L, Venter RE, Potgieter JR. The relationship between cardiorespiratory fitness, body composition and physical self-perception of adolescent girls. *J Hum Mov Stud* 2005;48:353–64.

147 Van Gent M, Malan DDJ, Pienaar AE. A comparison of the anthropometric, physical and motor growth characteristics of 12-15 year old girls in the North West province with Australian girls. *Afr J Phys Health, Educ Recr Dance* 2002;8:309–20.

148 Du Randt R. *The 1996 South African sport talent identification project: report prepared for the Sports Information and Science Agency*. Pretoria: South African Sports Commission, 1996.

149 Gulías-González R, Martínez-Vizcaíno V, García-Prieto JC, *et al*. Excess of weight, but not underweight, is associated with poor physical fitness in children and adolescents from Castilla-La. *Eur J Pediatr* 2014;173:727–35.

150 Torrijos-Niño C, Martínez-Vizcaíno V, Pardo-Guijarro MJ, *et al*. Physical fitness, obesity, and academic achievement in schoolchildren. *J Pediatr* 2014;165:104–9.

151 Chillón P, Ortega FB, Ferrando JA, *et al*. Physical fitness in rural and urban children and adolescents from Spain. *J Sci Med Sport* 2011;14:417–23.

152 Castro-Pinero J, González-Montesinos JL, Mora J, *et al.* Percentile values for muscular strength field tests in children aged 6 to 17 years: influence of weight status. *J Strength Cond Res* 2009;23:2295–310.

153 Ortega FB, Ruiz JR, Castillo MJ, *et al.* Low level of physical fitness in Spanish adolescents. Relevance for future cardiovascular health (AVENA Study). *Rev Esp Cardiol* 2005;58:898–909.

154 García Baena J. *La condición fisica en la educación secundaria. Trabajo de investigación* [thesis]. Madrid, Spain: Universidad Nacional de Educacion a Distancia, 1999.

155 Prat JA, Casamort J, Balagué N*, et al*. *Eurofit: La batería Eurofit en Catalunya*. Barcelona: Secretaria General de l'Esport, 1998.

156 Tercedor P, Delgado-Fernandez M. Condicion fisica relacionada con la salud en escolares de 10 anos de edad de Granada. *Proceedings of the II congreso internacional sobre la enseñanza de la educación física y el deporte escolar (second international congress about teaching physical education and school sport)*. 1998.

157 Sainz RM. *La batería Eurofit en Euskadi*. Vitoria-Gasteiz: Instituto Vasco de Educación Fisica, 1996.

158 Ureña F. *Valoración y baremación de la aptitud fisica en el alumnado de segundo ciclo de educación secundaria obligatoria de la comunidad autonoma de Murcia. Su utilización según los postulados de la reforma* [dissertation]. Murcia, Spain: Universidad de Murcia, 1996.

159 Brito Ojeda EM, Navarro Valdivielso M, García Afonso D, *et al. La condición física en la población escolar de gran Canaria (10-19 años). Las Palmas de Gran Canaria, Spain: Excmo*. Cabildo Insular de Gran Canaria, 1995.

160 Sainz RM. *Aptitudes psiquicas y fisicas: Estudio ed la aptitud fisica de los adolescentes de la provincia de Vizcaya y su relacion con la personalidad* [dissertation]. Bilboa, Spain: Universidad de Deusto, 1992.

161 Rivas FJ. Valoracion de la evolucion anthropoetrica y de las caracteristicas motrices en la poblacion escolar de un centro de E.G.B. mediante un estudio transversal. In: *II Congreso Galego da Educacion Fisica e o Deporte*. La Coruña: Escola Galega do Deporte de la Xunta de Galicia, 1987:17–28.

162 Lieveld J, Vrijens J, Bouckaert J. Ethnic differences in body structure and physical fitness in Surinamean boys aged 14 years. In: Claessens AL, Lefevre J, Vanden Eynde B, eds. *World-wide variation in physical fitness.* Leuven: Katholieke Universiteit Leuven, 1993:87–92.

163 Schmid M, Romann M, Kriemler S, *et al*. Wie kann die Fitness von Schulkindern gemessen werden? *Sportmedizin Sporttraumatologie* 2007;55:52–61.

164 Cauderay M, Narring F, Michaud P-A. A cross-sectional survey assessing physical fitness of 9- to 19-year-old girls and boys in Switzerland. *Pediatr Exerc Sci* 2000;12:398–412.

165 Aandstad A, Berntsen S, Hageberg R, *et al*. A comparison of estimated maximal oxygen uptake in 9 and 10 year old schoolchildren in Tanzania and Norway. *Br J Sports Med* 2006;40:278–92.

166 Çalis M, Ergen E, Turnagöl H, *et al*. Beden egitimi derslerinin bir ögretim yili boyunca 15-16 yas grubu ögrenciler üzerindekifizyolojik etkilerinin Eurofit test bataryasi ile izlenmesi. In: *3 Ulusal Spor Bilinleri Kongresi*. Ankara, Turkey: Hacettepe Universitesi, 1992:367–72.

167 Richards J, Foster C, Townsend N. Physical fitness and mental health impact of a sport-for-development intervention in a post-conflict setting: randomised controlled trial nested within an observational study of adolescents in Gulu, Uganda. *BMC Public Health* 2014;14:619.

168 Sandercock G, Ogunleye A, Voss C. Six-year changes in body mass index and cardiorespiratory fitness of English schoolchildren from an affluent area. *Int J Obes* 2015;39:1504–7.

169 Sandercock G, Voss C, Cohen D, *et al*. Centile curves and normative values for the twenty metre shuttle-run test in English schoolchildren. *J Sports Sci* 2012;30:679–87.

170 Boddy L, Fairclough SJ, Atkinson G, *et al.* Changes in cardiorespiratory fitness in 9- to 10.9-year-old children: Sports Linx 1998-2010. *Med Sci Sports Exerc* 2012;44:481–6.

171 Sandercock G, Voss C, McConnell D. Ten year secular declines in the cardiorespiratory fitness of affluent English children are largely independent of changes in body mass index. *Arch Dis Child* 2010;95:46–7.

172 Liverpool City Council. *Liverpool Sports Linx Project 01-03: Report on the health and fitness of Liverpool primary and secondary school children*. Liverpool, UK: Liverpool City Council, 2003.

173 Boreham C, Twisk J, Murray L, *et al*. Fitness, fatness, and coronary heart disease risk in adolescents: the Northern Ireland Young Hearts Project. *Med Sci Sports Exerc* 2001;33:270–4.

174 Watkins DC. *Ten year trends (1990-2000) in biological and behavioural risk factors for coronary heart disease in northern Irish adolescents* [thesis]. Belfast, UK: The Queen's University of Belfast, 2001.

175 Twisk JWR, Boreham C, Cran G, *et al.* Clustering of biological risk factors for cardiovascular disease and the longitudinal relationship with lifestyle of an adolescent population: the Northern Ireland Young Hearts Project. *J Cardiovasc Risk* 1999;6:355–62.

176 Mahoney C. 20-MST and PWC170 validity in non-Caucasian children in the UK. *Br J Sports Med* 1992;26:45–7.

177 Mahoney CA, Boreham CAG. Eurofit in Belfast primary schools. *Scot J Phys Educ* 1991;19:1–4.

178 Riddoch C, Savage JM, Murphy N, *et al*. Long term health implications of fitness and physical activity patterns. *Arch Dis Child* 1991;66:1426–33.

179 Boreham CAG, Paliczka VJ Nichols AK. Fitness testing of Belfast schoolchildren. In: 5th European research seminar on testing physical fitness. *Strasbourg: Council of Europe*, 1987:52–7.

180 Lewitt MS, Baker JS, Mooney GP*, et al*. Pubertal stage and measures of adiposity in British schoolchildren. *Ann Hum Biol* 2012;39:440–7.

181 Ranson R, Stratton G, Taylor S. Digit ratio (2D:4D) and physical fitness (Eurofit test battery) in school children. *Early Hum Dev* 2015;91:327–31.

182 Liu W, Zillifro TD, Nichols RA. Tracking of health -related physical fitness for middle school boys and girls. *Pediatr Exerc Sci* 2012;24:549–62.

183 Welk GJ, Maduro PFDS-M, Laurson KR, *et al*. Field evaluation of the new FITNESSGRAM criterion-referenced standards. *Am J Prev Med* 2011;41:S131–42.

184 Beets MW, Pitetti KH, Cardinal BJ. Progressive aerobic cardiovascular endurance run and body mass index among an ethnically diverse sample of 10–15 year olds. *Res Q Exerc Sport* 2005;76:389–97.

185 Beets MW, Pitetti KH. A comparison of shuttle-run performance between mid-western youth and their national and international counterparts. *Pediatr Exerc Sci* 2004;16:94–112.

186 Welk GJ, Schaben JA, Shelley M. Physical activity and physical fitness in children schooled at home and children attending public schools. *Pediatr Exer Sci* 2004;16:310–23.

187 Lloyd LK, Bishop PA, Walker JL, *et al*. The influence of body size and composition on FITNESSGRAM® test performance and the adjustment of FITNESSGRAM® test scores for skinfold thickness in youth. *Meas Phys Educ Exerc Sci* 2003;7:205–26.

188 Chun DM, Corbin CB, Pangrazi RP. Validation of criterion-referenced standards for the mile run and progressive aerobic cardiovascular endurance tests. *Res Q Exerc Sport* 2000;71:125–34.

189 Wolford N. *The difference in physical fitness levels of fifth graders according to socioeconomic groups and genders* [dissertation]. Ann Arbor, MI: University Microforms International, 1998.

190 Mahar MT, Rowe DA, Parker CR, et al. Criterion-referenced and norm-referenced agreement between the mile run/walk and PACER. *Meas Phys Educ Exerc Sci* 1997;1:245–58.

191 http://data.worldbank.org/about/country-and-lending-groups (accessed 24 Dec 2015).

192 Cole TJ, Freeman JV, Preece MA. Body mass index reference curves for the UK, 1990. *Arch Dis Child* 1995;73:25–9.

193 Cole TJ, Bellizzi MC, Flegal KM, *et al.* Establishing a standard definition for child overweight and obesity worldwide: international survey. *BMJ* 2000;320:1240–3.

194 Eisenmann JC. Waist circumference percentiles for 7- to 15-year-old Australian children. *Acta Pediatr* 2005;94:1182–5.

195 Fernandez JR, Redden DT, Pietrobelli A, *et al*. Waist circumference percentiles in nationally representative samples of African-American, European-American, and Mexican-American children and adolescents. *J Pediatr* 2004;145:439–44.

196 McCarthy HD, Jarret KV, Crawley HF. The development of waist circumference percentiles in British children aged 5.0-16.9 y. *Eur J Clin Nutr* 2001;55:902–7.

197 Shields M. Overweight and obesity among children and youth. *Health Rep* 2006;17:27–42.

198 Tremblay MS, Shields M, Laviolette M, *et al*. Fitness of Canadian children and youth: results from the 2007–2009 Canadian Health Measures Survey. *Health Rep* 2009;21:1–14.

199 Jackson LV, Thalange NKS, Cole TJ. Blood pressure centiles for Great Britain. *Arch Dis Child* 2007;92:298–303.

200 Jolliffe CJ, Janssen I. Development of age-specific adolescent metabolic syndrome criteria that are linked to the Adult Treatment Panel III and International Diabetes Federation criteria. *J Am Coll Cardiol* 2007;49:891–8.

201 Castro-Piñero J, González-Montesinos JL, Keating XD, *et al.* Percentile values for running sprint field tests in children ages 6–17 years: influence of weight status. *Res Q Exerc Sport* 2010;81:143–51.

202 De Miguel-Etayo P, Gracie-Marco L, Ortega FB, *et al.* Physical fitness reference standards in European children: the IDEFICS study. *Int J Obes* 2014;38 2):S57–66.

203 Eisenmann JC, Laurson KR, Welk G J. Aerobic fitness percentiles for U.S. adolescents. *Am J Prev Med* 2011;41:S106–10.

204 Armstrong N, Barker AR. Endurance training and elite young athletes. *Med Sport Sci* 2011;56:59–83.

205 Ruiz JR, Ortega FB, Rizzo NS, *et al*. High cardiovascular fitness is associated with low metabolic risk score in children: the European Youth Heart Study. *Pediatr Res* 2007;61:350–5.

206 Lobelo F, Pate RR, Dowda M, *et al.* Validity of cardiorespiratory fitness criterion-referenced standards for adolescents. *Med Sci Sports Exerc* 2009;41:1222–9.

207 Adegboye ARA, Anderssen SA, Froberg K, *et al.* Recommended aerobic fitness level for metabolic health in children and adolescents: a study of diagnostic accuracy. *Br J Sports Med* 2011;45:722–8.

208 Boddy LM, Thomas NE, Fairclough SJ, *et al.* ROC generated thresholds for field-assessed aerobic fitness related to body size and cardiometabolic risk in schoolchildren. *PLoS ONE* 2012;7:e45755.

209 Mesa JL, Ruiz JR, Ortega FB, *et al*. Aerobic fitness in relation to blood lipids and fasting glycaemia in adolescents: influence of weight status. *Nutr Metab Cardiovasc Dis* 2006;16:285–93.

210 Moreira C, Santos R, Ruiz JR, *et al*. Comparison of different VO_ 2max equations in the ability to discriminate the metabolic risk in Portuguese adolescents. *J Sci Med Sport* 2011;14:79–84.

211 Ruiz JR, Huybrechts I, Cuenca-Garcia M, *et al*. Cardiorespiratory fitness and ideal cardiovascular health in European adolescents. *Heart* 2015;101:766–73.

212 Tomkinson G. Aerobic fitness thresholds for cardiometabolic health in children and adolescents. *Br J Sports Med* 2011;45:686–7.

213 Barnett A, Chan LYS, Bruce IC. A preliminary study of the 20-m multistage shuttle run as a predictor of peak VO_ 2 in Hong Kong Chinese students. *Pediatr Exerc Sci* 1993;5:42–50.

214 Stickland MK, Petersen SR, Bouffard M. Prediction of maximal aerobic power from the 20-m multi-stage shuttle run test. *Can J Appl Physiol* 2003;28: 272–82.

215 Flouris AD, Metsios GS, Koutedakis Y. Enhancing the efficacy of the 20 m multistage shuttle run test. *Br J Sports Med* 2005;39:166–70.

216 Matsuzaka A, Takahashi Y, Yamazoe M, *et al.* Validity of the multistage 20-m shuttle-run test for Japanese children, adolescents, and adults. *Pediatr Exerc Sci* 2004;16:113–25.

217 Mahar MT, Welk GJ, Rowe DA*, et al*. Development and validation of a regression model to estimate VO_ 2peak from PACER 20-m shuttle run performance. *J Phys Act Health* 2006;3:S34–46.

218 Ruiz JR, Ramirez-Lechuga J, Ortega FB*, et al.* Artificial neural network-based equation for estimating VO_ 2max from the 20 m shuttle run test in adolescents. *Artif Intell Med* 2008;44:233–45.

219 Oliveira J. *Validação directa do teste de vaivém em 20 metros, de Luc- Léger, em adolescentes portugueses*. Cruz-Quebrada: FMH-UTL, 1998.

220 Pitetti KH, Fernhall B, Figoni S. Comparing two regression formulas that predict VO_ 2peak using the 20-m shuttle run for children and adolescents. *Pediatr Exerc Sci* 2002;14:125–34.

221 McVeigh SK, Payne AC, Scott S. The reliability and validity of the 20-meter shuttle test as a predictor of peak oxygen uptake in Edinburgh school children, age 13 to 14 years. *Pediatr Exerc Sci* 1995;7:69–79.

222 Moreno LA, González-Gross M, Kersting M, *et al*. Assessing, understanding and modifying nutritional status, eating habits and physical activity in European adolescents: the HELENA (Healthy lifestyle in Europe by nutrition in adolescence) study. *Public Health Nutr* 2008;11:288–99.

223 Andersen KL, Seliger V, Rutenfranz J, *et al.* Physical performance capacity of children in Norway. Part IV. The rate of growth in maximal aerobic power and the influence of improved physical education of children in a rural community— population parameters in a rural community. *Eur J Appl Physiol Occup Physiol* 1976;35:49–58.

224 Armstrong N, Welsman JR. Assessment and interpretation of aerobic fitness in children and adolescents. *Exerc Sport Sci Rev* 1994;22:435–76.

225 Krahenbuhl GS, Skinner JS, Kohrt WM. Developmental aspects of maximal aerobic power in children. *Exerc Sport Sci Rev* 1985;13:503–38.

226 Rowland TW. Evolution of maximal oxygen uptake in children. *Med Sport Sci* 2007;50:200–9.

227 Rowland TW. *Children's exercise physiology*. 2nd edn. Champaign, IL: Human Kinetics, 2005:89–112.

228 Bar-Or O, Rowland TW. Physiologic and perceptual responses to exercise in the healthy child. In: Bar-Or O, Rowland TW, eds. *Pediatric exercise medicine: from physiologic principles to health care application*. Champaign, IL: Human Kinetics, 2004:3–59.

229 Tomkinson G, Macfarlane DJ, Noi S, *et al*. Temporal changes in long-distance running performance of Asian children between 1964 and 2009. *Sports Med* 2012;42:267–79.

230 Moliner-Urdiales D, Ruiz JR, Ortega FB, *et al.* Secular trends in health-related physical fitness in Spanish adolescents: the AVENA and HELENA Studies. *J Sci Med Sport* 2010;13:584–8.

231 Tomkinson G, Olds T, Borms J. Who are the Eurofittest? *Med Sport Sci* 2007;50:104–28.

232 Castro-Piñero J, Artero EG, España-Romero V, *et al.* Criterion-related validity of field-based fitness tests in youth: a systematic review. *Br J Sports Med* 2010;44:934–43.

233 Mayorga-Vega D, Aguilar-Soto P, Viciana J. Criterion-related validity of the 20-M Shuttle run test for estimating cardiorespiratory fitness: a meta-analysis. *J Sports Sci Med* 2015;14:536–47.

234 Lussier L, Buskirk ER. Effects of an endurance training regimen of assessment of work capacity in prepubertal children. *Ann N Y Acad Sci* 1977;301:734–47.

235 Krahenbuhl GS, Pangrazi RP, Chomokos EA. Aerobic responses of young boys to submaximal running. *Res Q* 1979;50:413–21.

236 Péronnet F, Thibault G. Mathematical analysis of running performance and world records. *J Appl Physiol* 1989;67:453–65.

237 Sjödin B. The relationship among running economy, aerobic power, and onset of blood lactate accumulation in young boys (11–15 years). In: Komi PV, ed. *Exercise and sport biology*. Champaign, IL: Human Kinetics, 1982:57–60.

238 Mayers N, Gutin B. Physiological characteristics of elite prepubertal cross-country runners. *Med Sci Sports* 1979;11:172–6.

239 Cureton KJ, Boileau RA, Lohman TG, *et al.* Determinants of distance running performances in children: analysis of a path model. *Res Q* 1977;42:270–9.

240 Voss C, Sandercock G. Does the twenty meter shuttle-run test elicit maximal effort in 11- to 16-year-olds? *Pediatr Exerc Sci* 2009;21:55–62.

241 Longmuir PE, Colley RC, Wherley VA, *et al*. Canadian Society for Exercise Physiology position stand: benefit and risk for promoting childhood physical activity. *Appl Physiol Nutr Metab* 2014;39:1271–9.

242 Meusel D, Ruiz JR, Ortega FB, *et al.* Assessing levels of physical activity in the European population—the ALPHA project. *Seleccion* 2007;16:9–12.

243 Longmuir PE, Boyer C, Lloyd M, *et al*. The Canadian Assessment of Physical Literacy: methods for children in grades 4 to 6 (8 to 12 years). *BMC Public Health* 2015;15:767.

244 Council of Europe. *Eurofit: handbook for the EUROFIT tests of physical fitness*. Rome: Secretariat of the Committee for the Development of Sport within the Council of Europe, 1998.

245 Welk GJ, Going SB, Morrow Jr JR, *et al*. Development of new criterion-references fitness standards in the FITNESSGRAM® program: rationale and conceptual overview. *Am J Prev Med* 2011;41:S63–7.

246 Ortega FB, Cadenas-Sánchez C, Sánchez-Delgado G, *et al.* Systematic review and proposal of a field-based physical fitness-test battery in preschool children: the PREFIT battery. *Sports Med* 2015;45:533–55.

247 Cadenas-Sánchez C, Alcántara-Moral F, Sánchez-Delgado G, *et al*. [Assessment of cardiorespiratory fitness in preschool children: adaptation of the 20 metre shuttle run test]. *Nutr Hosp* 2014;30:1333–43.

European normative values for physical fitness in children and adolescents aged 9–17 years: results from 2 779 165 Eurofit performances representing 30 countries

Grant R Tomkinson,[1,2] Kevin D Carver,[1] Frazer Atkinson,[1] Nathan D Daniell,[2] Lucy K Lewis,[2,3] John S Fitzgerald,[1] Justin J Lang,[4] Francisco B Ortega[5,6]

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/bjsports-2017-098253).

[1]Department of Kinesiology and Public Health Education, University of North Dakota, Grand Forks, North Dakota, USA
[2]Alliance for Research in Exercise, Nutrition and Activity (ARENA), School of Health Sciences & Sansom Institute for Health Research, University of South Australia, Adelaide, Australia
[3]Discipine of Physiotherapy, Flinders University, Adelaide, Australia
[4]Healthy Active Living and Obesity (HALO) Research Group, Children's Hospital of Eastern Ontario Research Institute, Ottawa, Canada
[5]The PROFITH Research Group, Department of Physical Education and Sports, Faculty of Sports Sciences, University of Granada, Granada, Spain
[6]Department of Biosciences and Nutrition, Karolinska Institute, Huddinge, Sweden

Correspondence to
Dr Grant R Tomkinson, Department of Kinesiology and Public Health Education, University of North Dakota, Grand Forks, ND 58202, USA; grant.tomkinson@und.edu

Accepted 30 October 2017
Published Online First 30 November 2017

To cite: Tomkinson GR, Carver KD, Atkinson F, *et al*. *Br J Sports Med* 2018;**52**:1445–1456.

ABSTRACT

Objective To develop sex-specific and age-specific normative values for the nine Eurofit tests in European children and adolescents aged 9–17 years.

Methods A systematic review was undertaken to identify papers that explicitly reported descriptive results for at least one of nine Eurofit tests (measuring balance, muscular strength, muscular endurance, muscular power, flexibility, speed, speed-agility and cardiorespiratory fitness (CRF)) on children and adolescents. Data were included on apparently healthy (free from known disease/injury) children and adolescents aged 9–17 years. Following harmonisation for methodological variation where appropriate, pseudodata were generated using Monte Carlo simulation, with population-weighted sex-specific and age-specific normative centiles generated using the Lambda Mu Sigma (LMS) method. Sex-specific and age-specific differences were expressed as standardised differences in means, with the percentage of children and adolescents with healthy CRF estimated at the sex-age level.

Results Norms were displayed as tabulated centiles and as smoothed centile curves for the nine Eurofit tests. The final dataset included 2 779 165 results on children and adolescents from 30 European countries, extracted from 98 studies. On average, 78% of boys (95% CI 72% to 85%) and 83% of girls (95% CI 71% to 96%) met the standards for healthy CRF, with the percentage meeting the standards decreasing with age. Boys performed substantially (standardised differences >0.2) better than girls on muscular strength, muscular power, muscular endurance, speed-agility and CRF tests, but worse on the flexibility test. Physical fitness generally improved at a faster rate in boys than in girls, especially during the teenage years.

Conclusion This study provides the largest and most geographically representative sex-specific and age-specific European normative values for children and adolescents, which have utility for health and fitness screening, profiling, monitoring and surveillance.

BACKGROUND

Physical fitness is a good summative measure of the body's ability to perform physical activity and exercise, and it also provides an important summative indicator of health.[1] In adults, cardiorespiratory fitness (CRF) and musculoskeletal fitness (MSF) are strongly associated with mortality and cancer, independent of obesity and physical activity levels.[2–5] Several studies have shown considerably stronger inverse relationships between CRF and mortality than between physical activity and mortality,[6 7] indicating that changes in CRF may be more important to monitor in response to intervention (eg, exercise training). In children and adolescents, favourable associations have been reported linking CRF and MSF to cardiometabolic disease risk, adiposity, mental health and cognition as well as MSF to bone health.[1 8–10] Direct evidence has also emerged indicating that low CRF and MSF in adolescence are significantly associated with all-cause mortality later in life.[11–13] In addition to the health implications, physical fitness is an important determinant of success for many popular youth sports and athletic events (eg, hockey, basketball, football (soccer), running, swimming, rugby).[14]

Since its inception in 1988, the Eurofit has become the most popular test battery used to assess the physical fitness of European children and adolescents and the effectiveness of national physical education curricula.[15 16] The Eurofit comprises numerous health-related and skill-related fitness tests, including: (1) flamingo balance (balance), plate tapping (upper body speed), sit-and-reach (extent flexibility), standing broad jump (lower body muscular power), handgrip strength (upper body muscular strength), sit-ups (abdominal muscular endurance), bent arm hang (upper body muscular endurance), 10×5 m agility shuttle run (running speed-agility) and the 20 m shuttle run (CRF) (see online supplement 1); (2) anthropometric tests measuring height, mass and skinfold (various sites) and (3) age-identification and sex-identification data.[17] The Eurofit has excellent field-based utility because it is cheap and simple to administer, is practical in the school and club settings, requires minimal equipment and personnel and is appropriate for mass testing.[16] The Eurofit tests demonstrate very good test-retest reliability and good criterion validity for tests where appropriate criterion measures have been identified (eg, the 20 m shuttle run, standing broad jump, handgrip strength),[18–21] suggesting that it is a good test battery to measure physical fitness in youth. Criterion-referenced standards have also been developed for some Eurofit tests (eg, CRF) to help identify children and adolescents with apparently healthy cardiometabolic

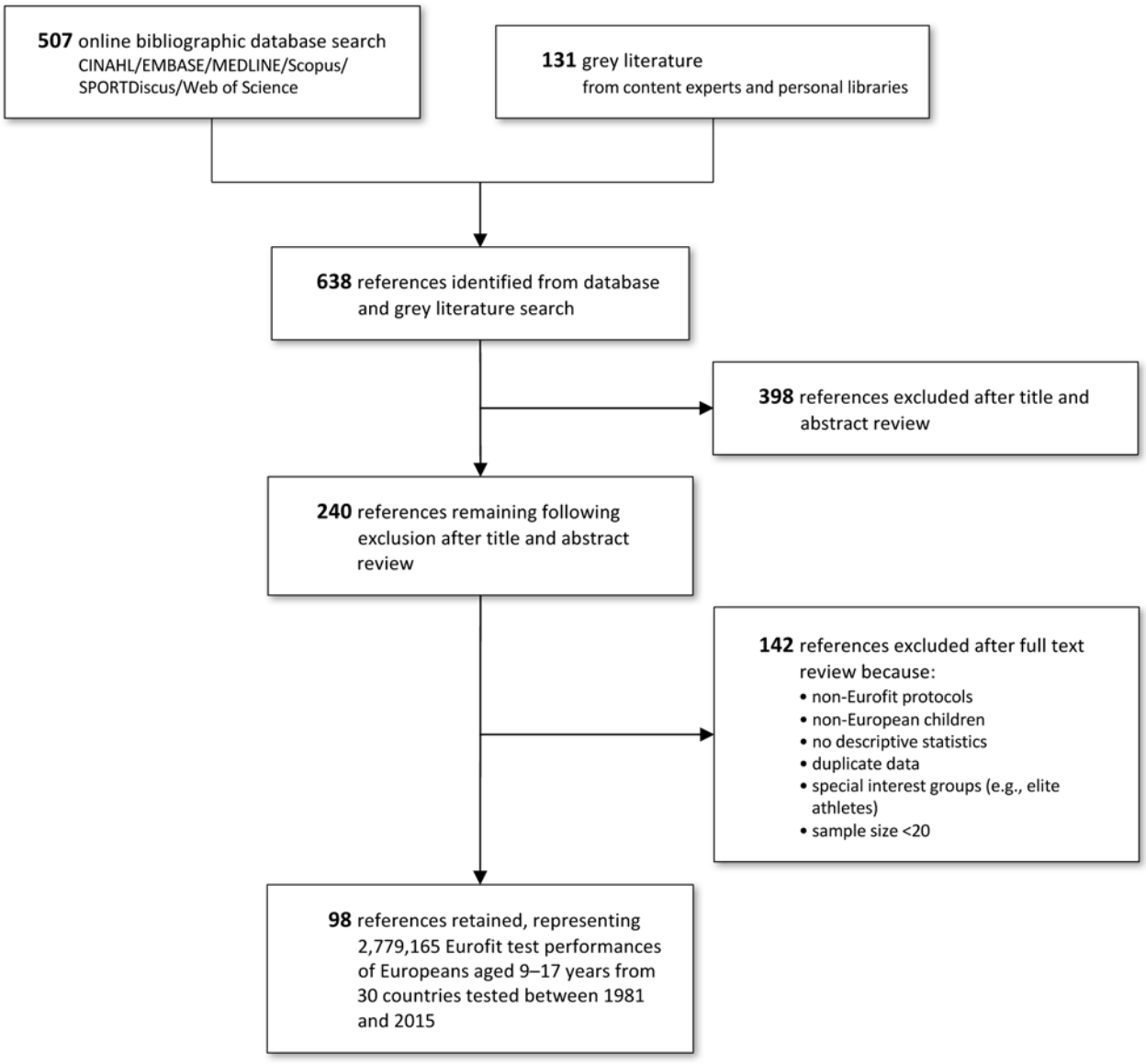


Figure 1 PRISMA flow chart outlining the flow of studies through the review.

profiles.[22 23] Several of the Eurofit tests have been supported by European experts from the ALPHA (Assessing Levels of Physical Activity) project[20] and by North American experts from the IOM (Institute of Medicine) report,[24] both of which provide strong and consistent guidelines about fitness testing in children and adolescents.

In order to extend the utility of the Eurofit as a surveillance instrument, there is a clear need for European normative-referenced standards to help interpret test scores, which are currently only available at the local, state/provincial or national level.[25–29] Previously, Tomkinson *et al*[16] used a method to match and compare Eurofit data in children and adolescents by standardising differences in test protocols and performance metrics. These data helped describe the geographical variability in the Eurofit performance of 1.2 million European children and adolescents aged 7–18 years from 23 countries,[16] and could be updated to provide European norms. Thus, the primary aim of this study was to develop sex-specific and age-specific normative values for physical fitness in European children and adolescents using the Eurofit, which implies a 10-year update to the previous Tomkinson *et al* review.[16] The secondary aim was to estimate the sex-related differences in Eurofit test performance as well as the percentage of European children and adolescents meeting the new international criterion-referenced standards for healthy CRF.[23]

METHODS

Data sources

A systematic review of the scientific literature was prospectively registered (PROSPERO 2013:CRD42013003646) and completed to locate studies that reported descriptive Eurofit data on European children and adolescents aged 9–17 years (see online supplement 2). This review was undertaken according to the Preferred Reporting Items for Systematic review and Meta-Analysis (PRISMA) guidelines for systematic reviews.[30] Studies were identified from January 1988 up until December 2016 using the following bibliographic databases: CINAHL, EMBASE, MEDLINE, Scopus, SPORTDiscus and Web of Science. This search strategy was developed by the author group

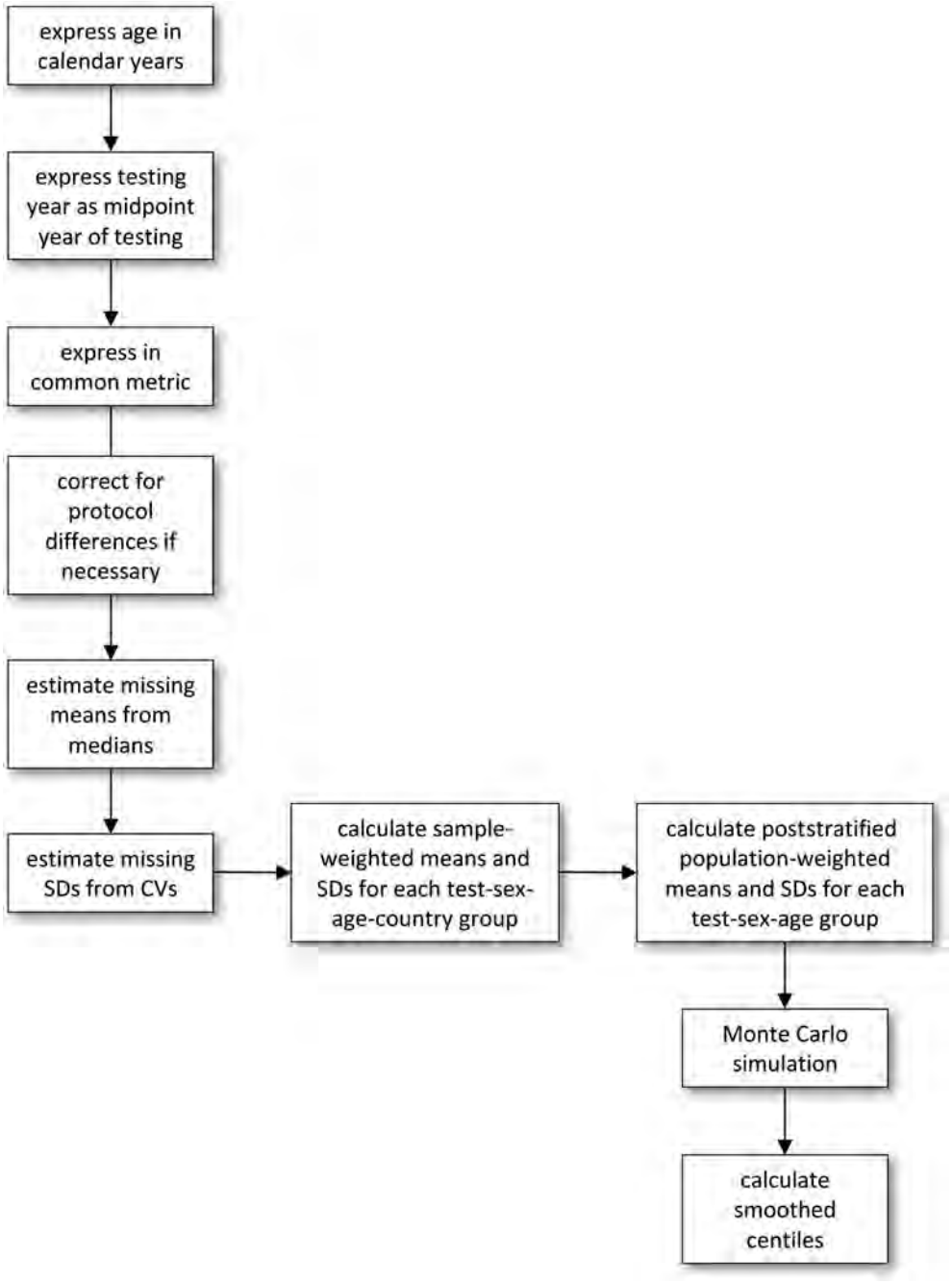


Figure 2 Flow chart showing the methodological procedure used in this study. Results from studies were first expressed in a common metric and corrected for protocol differences. Following the estimation of missing means and SDs if necessary, poststratified population-weighted means and SDs were estimated for each test-sex-age group, with pseudodata and smoothed centiles subsequently generated. CV, coefficient of variation.

in conjunction with a trained academic librarian. The search strategy included the term: Eurofit; with child*, OR adolescen*, OR youth, OR boy*, OR girl*, OR teen*, OR paediatric*, OR pediatric*, as search term modifiers. All studies were extracted as text files, imported into RefWorks (ProQuest, Ann Arbor, Michigan, USA) and assigned a unique reference identification number. Duplicate studies were first removed using RefWorks with the remaining duplicates removed manually. Two independent reviewers screened all titles and abstracts for eligibility, with full-text copies obtained for all studies meeting initial screening criteria according to at least one reviewer. These two independent reviewers then examined all full-text articles and discrepancies were resolved by discussion and consensus. A third reviewer examined an article when the two reviewers were unable to reach consensus, with consensus reached for all included articles. Email contact with the corresponding authors of studies occurred when necessary, in order to provide clarification, to avoid 'double counting' previously reported data and/or to request additional descriptive or raw data. The reference lists of all included studies were manually reviewed by two reviewers to identify new studies. Reviewers contacted content experts to obtain grey literature. In addition, the personal libraries of the authors were examined for relevant studies not identified through the search strategy.

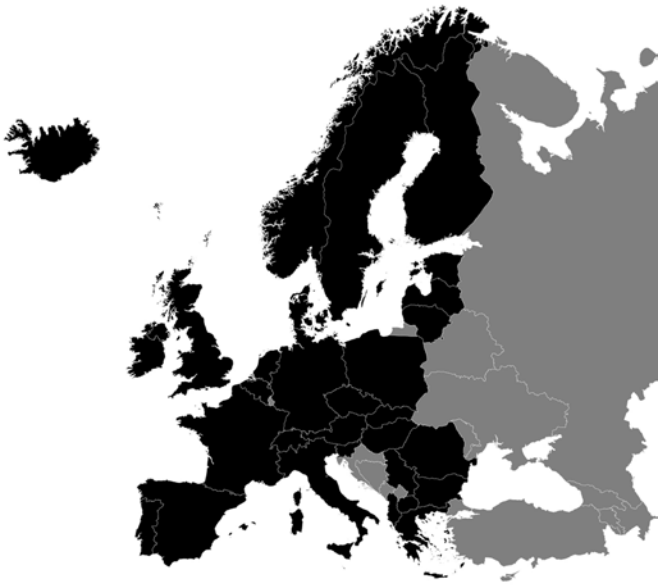


Figure 3 European map indicating the 30 countries (filled in black) for which Eurofit data on children and adolescents aged 9–17 years were available.

Inclusion/exclusion criteria

Studies were included if they explicitly reported descriptive Eurofit data at the test-sex-age-country-year level. Study participants must have been apparently healthy (free from known disease or injury) European children and adolescents aged 9–17 years who were tested from 1981 onwards—the inception year of the provisional Eurofit test battery. Studies were excluded if they reported descriptive Eurofit data on: (1) test-sex-age-country-year groups for which the sample size was less than 20 (because the means and SDs for smaller samples were too labile); (2) duplicate data published in another included study or (3) on only special interest groups that were atypical of their source population (eg, elite athletes, physically or mentally impaired children). Figure 1 shows a PRISMA flow chart of the included studies.

Data treatment and statistical analysis

All descriptive data were extracted into Excel (Microsoft Office 2010, USA) using a standardised data extraction table. The following descriptive data were extracted by one author and checked for accuracy by another: authors, country of testing, year of testing, sex, age, Eurofit test (including data on the name of test, measurement units, sample size, mean, SD and median), sampling method and the sampling base. Mean data were examined for anomalies by running range checks and examining sex-specific and age-specific scatter plots, with means±2 SEs of the mean away from the respective sex-age-test level mean identified and checked for transcription errors. Only data on children and adolescents aged 9–17 years were retained for further analysis.

The general procedure used to generate the sex-specific and age-specific normative centiles from extracted data is described elsewhere[31] and summarised in figure 2. Age was reported as age at last birthday (70% or 69/98 studies), a span of years (6% or 6/98 studies) or as mean and SD years (24% or 23/98 studies). Testing year was recorded as the midpoint year of testing (47% or 46/98 studies), a span of testing years (38% or 37/98 studies) or not reported at all (15% or 15/98 studies). Age and testing year were therefore expressed as age at last birthday and the midpoint year of testing, respectively.[31]

To combine data from different studies, all Eurofit data were standardised to a common metric and protocol. Measurement units reported in the Eurofit handbook[17] were used as the test-specific common metrics and for the presentation of normative centiles. All 20 m shuttle run data were standardised to Léger's 1-min protocol,[32] which starts at a speed of 8.5 km/hour and increases by 0.5 km/hour each minute and the speed at the last completed stage using the procedures described elsewhere.[31 33] The accuracy of the 20 m shuttle run data standardisation procedure is excellent.[33]

As part of the modelling procedure used to generate sex-specific and age-specific norms, means and SDs were required at the study-test-sex-age-country-year level. If no mean was available (1% or 1/98 studies), then mean values were estimated from the reported median values. This was done by first locating all studies reporting both median and mean values at the study-test-sex-age-country-year level and second, by determining the best-fitting and most parsimonious linear or curvilinear (second-order and third-order polynomials) regression models between median (predictor variable) and mean (response variable) values. Furthermore, 4% (4/98) of studies did not report SD values. Missing SD values were estimated by first locating all studies reporting both means and SDs at the study-test-sex-age-country-year level; second, by calculating the corresponding coefficient of variation (CV) values and third, by calculating the sample-weighted mean CVs for boys and girls separately.

Sample-weighted means and SDs (the latter calculated from sample-weighted mean CVs) were then calculated at the test-sex-age-country level. While these data represent the best available Eurofit data, in order to best generate European representative sex-specific and age-specific normative centiles and to correct for systematic bias associated with oversampling and undersampling, means and SDs were corrected using a poststratification population-weighting procedure.[34] This procedure ensures that our norms were standardised to underlying country-sex-age demographics. Thus, population estimates standardised to the mean testing year of 2000 were extracted from the United Nations World Population Prospects report.[35] Monte Carlo simulation was then used to create pseudodata using the detailed methods described elsewhere.[36] This simulation procedure attempts to 'recreate' the unavailable raw data by using a random number generator to produce data points based on population-weighted means and SDs at the sex-age level. Monte Carlo simulation assumes that the distributions are approximately normal, which was not true of all available raw Eurofit data. The simulation procedure described by Tomkinson *et al*[36] however allowed for the recreation of both normal and non-normal pseudodata, with Eurofit data considered to be either normal or non-normal following the assessment of normality by the d'Agostino-Pearson K^2 test[37] using available raw data of the same test. Pseudo-datasets were repeatedly generated until the calculated mean differed from the reported mean by <0.5%, and the calculated SD differed from the reported SD by <2.5%. These pseudo-datasets were then used to generate sex-specific and age-specific normative centiles in LMSchartmaker Pro (V.2.43, The Institute of Child Health, London, UK), which analyses data using the Lambda Mu Sigma (LMS) method.[38] The LMS method fits smooth centile curves to reference data by summarising the changing distribution of three sex-specific and age-specific curves representing the skewness (L; expressed as a Box-Cox power), the median (M) and the CV (S). Using penalised likelihood, the curves can be fitted as cubic splines using non-linear regression, and the extent

Table 1 Flamingo balance (n/60 s) centiles by age and sex based on 123 655 test performances of children and adolescents aged 9–17 years representing 19 countries

Age (years)	n	P_5	P_{10}	P_{20}	P_{30}	P_{40}	P_{50}	P_{60}	P_{70}	P_{80}	P_{90}	P_{95}
Boys												
9	3691	24	21	18	15	13	12	10	9	7	5	4
10	5140	25	22	18	16	14	12	10	8	7	5	3
11	6409	26	22	18	16	14	12	10	8	7	4	3
12	8313	26	23	18	16	14	12	10	8	7	4	3
13	8750	26	23	18	16	14	12	10	8	6	4	3
14	9466	25	21	18	15	13	11	10	8	6	4	3
15	7605	21	18	15	13	11	10	9	7	6	4	3
16	6665	21	18	15	13	11	10	8	7	6	4	3
17	5940	21	18	15	13	11	10	8	7	6	4	3
Girls												
9	3654	23	20	17	14	13	11	10	8	7	5	3
10	4935	23	20	17	15	13	11	10	8	7	5	3
11	6247	24	20	17	15	13	11	10	8	7	5	3
12	8271	24	21	17	15	13	11	10	8	7	5	3
13	8958	23	20	17	15	13	11	10	8	7	5	3
14	9279	23	20	16	14	13	11	10	8	7	5	3
15	7956	21	18	15	13	12	10	9	8	6	4	3
16	6644	19	17	14	12	11	9	8	7	6	4	3
17	5732	18	16	13	12	10	9	8	7	5	4	3

Note: the ages shown represent age at last birthday (eg, 9=9.00–9.99).

of smoothing required can be expressed in terms of smoothing parameters or equivalent df.[39]

The percentage of children and adolescents with healthy CRF (ie, healthy cardiometabolic profiles) was estimated using the new international criterion-referenced standards of 42 and 35 mL/kg/min for boys and girls, respectively.[23] Sex-specific differences in mean Eurofit performance were expressed as standardised differences. Positive differences indicated that Eurofit performances for boys were better than those for girls. Standardised differences of 0.2, 0.5 and 0.8 were used as thresholds for small, moderate and large effect sizes (ES), respectively.[40]

RESULTS

The final dataset included 2 779 165 Eurofit test performances of European children and adolescents aged 9–17 years (6458 study-sex-age-country-year groups extracted from 98 studies), representing 30 countries (figure 3). These 30 countries

Table 2 Plate tapping (s) centiles by age and sex based on 148 093 test performances of children and adolescents aged 9–17 years representing 19 countries

Age (years)	n	P_5	P_{10}	P_{20}	P_{30}	P_{40}	P_{50}	P_{60}	P_{70}	P_{80}	P_{90}	P_{95}
Boys												
9	7543	24.05	22.04	20.00	18.74	17.78	16.96	16.21	15.48	14.70	13.73	13.02
10	9090	21.55	19.90	18.19	17.13	16.31	15.61	14.97	14.33	13.65	12.80	12.17
11	8198	19.48	18.11	16.68	15.77	15.07	14.46	13.90	13.35	12.75	12.00	11.44
12	9799	17.91	16.74	15.51	14.72	14.10	13.57	13.07	12.58	12.05	11.37	10.87
13	9104	16.44	15.44	14.37	13.69	13.15	12.68	12.25	11.81	11.34	10.74	10.28
14	9964	15.12	14.26	13.34	12.74	12.27	11.86	11.48	11.09	10.67	10.13	9.72
15	7797	14.00	13.25	12.45	11.92	11.51	11.14	10.80	10.45	10.07	9.59	9.22
16	7217	13.38	12.70	11.95	11.46	11.08	10.74	10.42	10.10	9.74	9.29	8.94
17	6157	13.11	12.45	11.73	11.26	10.89	10.56	10.25	9.94	9.59	9.15	8.82
Girls												
9	7121	25.25	22.05	19.29	17.77	16.70	15.83	15.06	14.34	13.60	12.72	12.09
10	8904	22.35	19.95	17.77	16.54	15.64	14.90	14.25	13.62	12.97	12.19	11.63
11	8561	19.93	18.11	16.38	15.38	14.63	14.01	13.45	12.91	12.35	11.66	11.16
12	10 089	18.41	16.96	15.53	14.68	14.04	13.50	13.01	12.53	12.03	11.41	10.95
13	9031	16.92	15.76	14.60	13.89	13.35	12.88	12.46	12.05	11.60	11.05	10.64
14	9476	15.51	14.58	13.63	13.03	12.57	12.18	11.81	11.45	11.06	10.58	10.21
15	7690	14.95	14.12	13.25	12.70	12.28	11.91	11.57	11.24	10.87	10.41	10.07
16	6790	14.58	13.80	12.99	12.48	12.07	11.73	11.41	11.08	10.74	10.30	9.97
17	5562	14.54	13.77	12.96	12.45	12.05	11.71	11.39	11.07	10.72	10.28	9.95

Table 3 Sit-and-reach (cm) centiles by age and sex based on 464 807 test performances of children and adolescents aged 9–17 years representing 27 countries

Age (years)	n	P_5	P_{10}	P_{20}	P_{30}	P_{40}	P_{50}	P_{60}	P_{70}	P_{80}	P_{90}	P_{95}
Boys												
9	34 495	6.0	8.1	10.7	12.7	14.4	16.0	17.6	19.4	21.4	24.3	26.8
10	35 532	6.0	8.1	10.8	12.7	14.4	16.1	17.7	19.4	21.5	24.5	26.9
11	35 413	6.0	8.1	10.8	12.7	14.4	16.1	17.7	19.4	21.5	24.5	26.9
12	29 962	6.0	8.2	10.8	12.8	14.5	16.1	17.8	19.6	21.7	24.6	27.1
13	26 840	6.1	8.3	11.1	13.1	14.8	16.5	18.2	20.0	22.2	25.2	27.7
14	25 302	6.7	9.1	12.1	14.3	16.2	18.0	19.9	21.9	24.2	27.5	30.3
15	21 644	7.7	10.3	13.7	16.1	18.3	20.3	22.4	24.6	27.2	30.9	34.0
16	16 285	8.4	11.1	14.6	17.1	19.3	21.4	23.6	25.9	28.6	32.4	35.6
17	9696	9.1	11.9	15.5	18.1	20.4	22.6	24.8	27.2	30.0	33.9	37.2
Girls												
9	33 008	7.9	10.2	13.1	15.2	16.9	18.6	20.3	22.1	24.2	27.2	29.6
10	34 803	8.5	10.8	13.7	15.7	17.5	19.2	20.9	22.7	24.8	27.7	30.1
11	35 250	9.4	11.7	14.5	16.6	18.4	20.1	21.7	23.5	25.6	28.6	31.0
12	29 835	10.6	12.9	15.8	17.9	19.7	21.4	23.1	24.9	27.1	30.0	32.5
13	26 090	11.9	14.4	17.3	19.5	21.3	23.1	24.8	26.7	28.9	31.9	34.4
14	24 563	13.1	15.6	18.6	20.8	22.7	24.5	26.3	28.2	30.4	33.5	36.1
15	20 540	13.9	16.4	19.5	21.7	23.6	25.4	27.2	29.1	31.3	34.4	37.0
16	16 197	14.4	16.9	20.0	22.2	24.1	25.9	27.6	29.5	31.8	34.9	37.5
17	9352	14.7	17.2	20.3	22.5	24.4	26.1	27.9	29.8	32.1	35.2	37.8

Note: a score of 15 cm corresponds to the participant reaching their toes.

represented approximately 65% of Europe's population and 49% of Europe's land area and included 25 high-income and five upper-middle-income countries. Online supplement 3 provides a summary of the 98 included studies.

Tables 1–9 provide normative values as tabulated centiles from 5% to 95% for all nine Eurofit tests. Smoothed centile curves are presented in figure 4 with additional 20 m shuttle run norms (speed at last completed stage, number of laps and relative $\dot{V}O_{2peak}$) presented in online supplement 4.

On average, 78% of boys (95% CI 72% to 85%) and 83% of girls (95% CI 71% to 96%) had healthy CRF, with the percentage of those with healthy CRF decreasing by about 3% (boys) and 7% (girls) per year from the age of 9 years onwards (figure 5). There was considerable variability in healthy CRF levels among different European countries, which increased with age (see online supplement 5). When dividing Europe into two segments at the 45th parallel north,[41 42] a gradient existed where Northern-Central European countries had a higher percentage

Table 4 Standing broad jump (cm) centiles by age and sex based on 464 900 test performances of children and adolescents aged 9–17 years representing 29 countries

Age (years)	n	P_5	P_{10}	P_{20}	P_{30}	P_{40}	P_{50}	P_{60}	P_{70}	P_{80}	P_{90}	P_{95}
Boys												
9	35 148	100.5	107.9	116.8	123.2	128.7	133.8	138.9	144.3	150.7	159.5	166.8
10	36 069	107.6	115.3	124.6	131.3	137.0	142.4	147.7	153.4	160.1	169.3	176.9
11	35 618	115.4	123.5	133.3	140.3	146.3	151.9	157.5	163.5	170.5	180.2	188.2
12	30 631	122.5	131.0	141.2	148.5	154.8	160.7	166.5	172.8	180.1	190.3	198.6
13	24 760	129.7	138.5	149.3	157.0	163.6	169.7	175.9	182.5	190.2	200.9	209.7
14	24 061	138.7	148.1	159.6	167.8	174.8	181.4	188.0	195.0	203.2	214.6	223.9
15	20 334	147.8	157.8	169.8	178.5	186.0	192.9	199.8	207.2	215.9	227.9	237.8
16	18 967	154.2	164.5	176.9	185.9	193.6	200.8	207.9	215.6	224.6	237.0	247.2
17	12 108	158.3	168.9	181.6	190.7	198.5	205.8	213.1	221.0	230.1	242.7	253.2
Girls												
9	34 339	91.2	98.4	107.1	113.4	118.9	123.9	129.0	134.5	140.8	149.7	157.1
10	35 339	98.5	105.9	114.9	121.4	127.0	132.3	137.5	143.2	149.8	159.0	166.6
11	34 992	105.6	113.3	122.6	129.4	135.2	140.6	146.0	151.9	158.7	168.2	176.1
12	29 974	111.1	119.0	128.6	135.6	141.6	147.1	152.7	158.7	165.8	175.6	183.7
13	23 749	113.9	121.9	131.6	138.7	144.8	150.4	156.1	162.2	169.3	179.3	187.5
14	22 416	115.6	123.7	133.6	140.7	146.8	152.5	158.3	164.4	171.6	181.7	190.0
15	16 394	116.8	124.9	134.8	142.0	148.1	153.9	159.6	165.8	173.1	183.1	191.5
16	18 459	117.5	125.6	135.5	142.7	148.8	154.6	160.4	166.6	173.8	183.9	192.2
17	11 542	119.0	127.2	137.2	144.4	150.6	156.4	162.3	168.5	175.8	186.0	194.4

Table 5 Handgrip strength (kg) centiles by age and sex based on 203 295 test performances of children and adolescents aged 9–-17 years representing 24 countries

Age (years)	n	P_5	P_{10}	P_{20}	P_{30}	P_{40}	P_{50}	P_{60}	P_{70}	P_{80}	P_{90}	P_{95}
Boys												
9	10 180	8.6	10.1	11.9	13.2	14.3	15.3	16.4	17.5	18.8	20.6	22.1
10	11 965	9.5	11.1	13.0	14.5	15.7	16.8	18.0	19.2	20.6	22.6	24.2
11	11 358	10.8	12.6	14.8	16.4	17.7	19.0	20.3	21.6	23.2	25.4	27.2
12	13 107	13.1	15.2	17.7	19.6	21.2	22.6	24.1	25.7	27.6	30.1	32.3
13	13 070	16.9	19.4	22.5	24.7	26.6	28.4	30.2	32.1	34.3	37.4	39.9
14	13 843	21.6	24.5	27.9	30.4	32.6	34.6	36.6	38.7	41.2	44.7	47.6
15	10 944	25.9	28.9	32.5	35.2	37.4	39.5	41.6	43.9	46.5	50.1	53.2
16	10 062	29.1	32.1	35.8	38.5	40.7	42.9	45.0	47.2	49.9	53.6	56.7
17	8157	31.3	34.3	38.0	40.6	42.9	45.0	47.1	49.4	52.1	55.7	58.8
Girls												
9	9690	7.2	8.7	10.4	11.6	12.6	13.6	14.6	15.6	16.8	18.5	19.9
10	11 804	8.0	9.6	11.5	12.9	14.1	15.2	16.3	17.5	18.8	20.7	22.3
11	11 582	9.4	11.2	13.4	14.9	16.3	17.5	18.8	20.1	21.7	23.9	25.6
12	13 331	12.0	13.9	16.2	17.9	19.3	20.6	21.9	23.3	25.0	27.3	29.1
13	13 182	16.1	18.0	20.3	21.9	23.3	24.6	25.9	27.3	29.0	31.2	33.1
14	13 168	18.5	20.4	22.7	24.3	25.7	27.1	28.4	29.8	31.4	33.7	35.6
15	10 586	19.1	21.1	23.5	25.2	26.7	28.0	29.4	30.8	32.5	34.9	36.8
16	9672	19.3	21.2	23.6	25.4	26.9	28.2	29.6	31.1	32.8	35.2	37.2
17	7594	19.4	21.4	23.8	25.5	27.0	28.4	29.8	31.3	33.0	35.5	37.4

of children and adolescents with healthy CRF than Southern European countries (average difference in means (range): 7% (0% to 27%) at the sex-age level).

On average, boys performed substantially better than girls at each age group on muscular strength (ES: large), muscular power (ES: large), muscular endurance (ES: moderate to large), speed-agility (ES: moderate) and CRF (ES: large) tests, with the magnitude of the sex-specific differences increasing with age and accelerating from about 12 years (figure 6). Boys also developed at a faster rate than girls on these tests, especially during the teenage years. Conversely, girls performed substantially better at each age group on the flexibility test (ES: moderate), with boys and girls developing with age at similar rates. There were negligible sex-specific differences overall on the balance and upper body speed tests, although boys developed at a faster rate than girls on the upper body speed test.

DISCUSSION

This study systematically analysed 2 779 165 Eurofit performances of children and adolescents aged 9–17 years to generate

Table 6 Sit-ups (n/30 s) centiles by age and sex based on 481 032 performances of children and adolescents aged 9–17 years representing 23 countries

Age (years)	n	P_5	P_{10}	P_{20}	P_{30}	P_{40}	P_{50}	P_{60}	P_{70}	P_{80}	P_{90}	P_{95}
Boys												
9	31 757	9	11	13	15	16	17	18	20	21	23	25
10	33 748	11	13	15	17	18	19	20	22	23	25	27
11	35 559	13	14	16	18	19	20	22	23	24	26	28
12	29 338	14	15	17	19	20	21	22	24	25	27	29
13	30 805	14	16	18	20	21	22	23	24	26	28	29
14	29 024	15	17	19	20	22	23	24	25	27	29	30
15	22 541	17	18	20	22	23	24	25	26	28	30	31
16	18 751	18	19	21	22	24	25	26	27	29	30	32
17	12 059	18	20	22	23	24	25	27	28	29	31	33
Girls												
9	31 091	9	11	13	14	15	17	18	19	21	23	25
10	33 131	10	12	14	16	17	18	19	20	22	24	26
11	34 525	11	13	15	16	17	19	20	21	22	24	26
12	31 415	12	13	15	17	18	19	20	21	23	24	26
13	29 168	12	14	15	17	18	19	20	21	23	24	26
14	27 377	12	14	16	17	18	19	20	21	23	25	26
15	21 072	13	14	16	17	19	20	21	22	23	25	26
16	18 365	13	15	16	18	19	20	21	22	23	25	27
17	11 306	13	15	17	18	19	20	21	22	24	25	27

Table 7 Bent-arm hang (s) centiles by age and sex based on 189 673 test performances of children and adolescents aged 9–17 years representing 23 countries

Age (years)	n	P_5	P_{10}	P_{20}	P_{30}	P_{40}	P_{50}	P_{60}	P_{70}	P_{80}	P_{90}	P_{95}
Boys												
9	8282	1.48	2.13	3.29	4.49	5.85	7.48	9.55	12.38	16.74	25.36	35.62
10	9584	1.56	2.25	3.48	4.76	6.20	7.92	10.10	13.08	17.65	26.62	37.23
11	11 079	1.63	2.35	3.66	5.00	6.51	8.32	10.60	13.71	18.46	27.73	38.62
12	11 899	1.71	2.48	3.87	5.29	6.89	8.79	11.19	14.44	19.39	28.99	40.19
13	12 321	1.90	2.77	4.33	5.92	7.70	9.81	12.44	15.99	21.34	31.57	43.30
14	12 550	2.50	3.67	5.72	7.78	10.05	12.70	15.96	20.26	26.61	38.39	51.45
15	10 576	3.73	5.40	8.26	11.05	14.04	17.43	21.50	26.72	34.18	47.44	61.48
16	9165	5.19	7.39	10.98	14.36	17.87	21.75	26.28	31.94	39.77	53.13	66.71
17	7425	6.48	9.03	13.07	16.74	20.45	24.46	29.04	34.64	42.19	54.66	66.92
Girls												
9	7681	0.98	1.43	2.24	3.08	4.02	5.14	6.55	8.46	11.36	16.94	23.40
10	9287	0.97	1.42	2.24	3.08	4.03	5.15	6.57	8.50	11.42	17.06	23.60
11	10 942	0.96	1.42	2.23	3.08	4.03	5.16	6.59	8.53	11.48	17.18	23.79
12	13 198	0.96	1.41	2.23	3.08	4.03	5.17	6.60	8.54	11.50	17.22	23.86
13	13 613	0.96	1.41	2.23	3.08	4.03	5.18	6.62	8.58	11.56	17.33	24.04
14	13 322	0.94	1.40	2.22	3.09	4.06	5.23	6.72	8.73	11.82	17.83	24.86
15	11 324	0.92	1.38	2.23	3.11	4.13	5.35	6.91	9.05	12.34	18.80	26.41
16	9639	0.91	1.38	2.27	3.21	4.30	5.63	7.33	9.68	13.33	20.57	29.19
17	7786	0.93	1.43	2.40	3.45	4.67	6.16	8.11	10.82	15.07	23.61	33.92

the largest and most geographically representative sex-specific and age-specific European normative values for physical fitness. These norms add to existing norms across a range of other cardiometabolic risk factors, including adiposity (eg, body mass index[43 44] and waist circumference,[45–49] blood pressure,[50 51] cholesterol,[51] triglycerides[51] and glucose).[51] More importantly, they expand the normative data bank for health-related fitness, building on existing norms studies such as the recently published international CRF norms[31] and other European health-related fitness norms.[52 53]

Despite these norms not being linked to a health outcome, they nonetheless have utility for health and fitness screening, profiling, monitoring and surveillance by identifying the centile rank of children and adolescents in comparison with their peers. For instance, several authors[31 52 54] have suggested using a normative quintile-based framework to classify the fitness levels of children and adolescents, where those below the 20th centile are classified as 'very low/poor'; 20–40th centiles as 'low/poor'; 40–60th centiles as 'moderate'; 60–80th centiles as 'high/good' and those above the 80th centile as 'very high/

Table 8 10×5 m agility shuttle run (s) centiles by age and sex based on 258 618 test performances of children and adolescents aged 9–17 years representing 19 countries

Age (years)	n	P_5	P_{10}	P_{20}	P_{30}	P_{40}	P_{50}	P_{60}	P_{70}	P_{80}	P_{90}	P_{95}
Boys												
9	15 409	29.26	27.58	25.79	24.64	23.73	22.94	22.20	21.46	20.66	19.64	18.87
10	16 773	28.00	26.54	24.98	23.96	23.15	22.44	21.78	21.11	20.38	19.44	18.73
11	17 925	26.77	25.53	24.16	23.27	22.55	21.92	21.33	20.73	20.07	19.22	18.57
12	16 152	25.68	24.59	23.39	22.60	21.96	21.40	20.86	20.32	19.72	18.94	18.35
13	18 549	24.77	23.79	22.70	21.98	21.40	20.88	20.39	19.88	19.33	18.61	18.05
14	16 914	24.10	23.18	22.15	21.47	20.92	20.43	19.96	19.48	18.95	18.27	17.73
15	12 649	23.61	22.72	21.73	21.06	20.53	20.05	19.60	19.13	18.62	17.95	17.43
16	11 783	23.22	22.35	21.37	20.72	20.20	19.73	19.28	18.83	18.32	17.67	17.16
17	6423	22.89	22.03	21.07	20.43	19.91	19.45	19.01	18.56	18.06	17.42	16.91
Girls												
9	16 273	30.96	28.96	26.93	25.67	24.70	23.88	23.12	22.37	21.57	20.57	19.83
10	15 703	28.87	27.35	25.76	24.74	23.95	23.27	22.63	21.99	21.30	20.43	19.78
11	15 063	27.11	25.92	24.64	23.81	23.15	22.58	22.04	21.50	20.90	20.14	19.57
12	18 344	26.36	25.29	24.13	23.37	22.77	22.24	21.74	21.24	20.68	19.97	19.43
13	16 678	26.06	25.03	23.90	23.16	22.58	22.06	21.58	21.08	20.54	19.85	19.32
14	15 589	25.98	24.95	23.83	23.09	22.51	22.00	21.51	21.03	20.49	19.79	19.27
15	11 479	25.97	24.94	23.82	23.09	22.51	22.00	21.51	21.02	20.48	19.79	19.26
16	11 018	25.95	24.92	23.81	23.07	22.49	21.98	21.50	21.01	20.47	19.78	19.25
17	5895	25.93	24.90	23.79	23.06	22.48	21.96	21.48	20.99	20.46	19.77	19.24

Table 9 20 m shuttle run (min/stages) centiles by age and sex based on 445 092 test performances of children and adolescents aged 9–17 years representing 24 countries

Age (years)	n	P_5	P_{10}	P_{20}	P_{30}	P_{40}	P_{50}	P_{60}	P_{70}	P_{80}	P_{90}	P_{95}
Boys												
9	36 079	1.27	1.96	2.80	3.41	3.93	4.43	4.92	5.45	6.08	6.95	7.68
10	36 935	1.53	2.25	3.13	3.77	4.31	4.83	5.34	5.90	6.55	7.46	8.22
11	30 786	1.79	2.53	3.45	4.11	4.68	5.22	5.75	6.33	7.01	7.96	8.75
12	26 552	2.04	2.82	3.77	4.46	5.06	5.61	6.18	6.78	7.49	8.47	9.30
13	29 467	2.31	3.12	4.11	4.82	5.44	6.02	6.60	7.23	7.97	8.99	9.85
14	28 262	2.71	3.55	4.57	5.31	5.95	6.55	7.15	7.80	8.56	9.62	10.51
15	23 754	3.08	3.92	4.95	5.70	6.34	6.95	7.56	8.21	8.98	10.05	10.94
16	13 417	3.35	4.19	5.22	5.96	6.61	7.21	7.81	8.47	9.23	10.30	11.19
17	11 326	3.80	4.64	5.67	6.42	7.06	7.66	8.26	8.91	9.67	10.74	11.63
Girls												
9	35 027	0.87	1.41	2.08	2.56	2.98	3.38	3.77	4.20	4.70	5.40	5.98
10	36 270	1.03	1.60	2.29	2.79	3.22	3.63	4.04	4.48	5.00	5.72	6.33
11	30 751	1.31	1.91	2.64	3.18	3.64	4.07	4.51	4.98	5.53	6.30	6.94
12	26 119	1.27	1.89	2.66	3.21	3.69	4.14	4.60	5.08	5.66	6.46	7.13
13	20 066	1.25	1.87	2.64	3.20	3.68	4.13	4.58	5.07	5.65	6.46	7.13
14	19 557	1.24	1.87	2.64	3.20	3.68	4.13	4.58	5.07	5.65	6.46	7.13
15	15 682	1.24	1.87	2.63	3.19	3.67	4.13	4.58	5.07	5.65	6.46	7.13
16	13 317	1.21	1.84	2.61	3.17	3.66	4.11	4.57	5.06	5.64	6.45	7.13
17	11 725	1.20	1.83	2.60	3.17	3.65	4.11	4.56	5.06	5.64	6.45	7.13

Note: 20 m shuttle run centiles are available for other metrics in online supplement 4.

good'. Single test measures can be qualitatively interpreted using these quintile-based thresholds and longitudinal changes tracked against centile bands to identify expected, better than expected or worse than expected developmental changes. In addition, long-term intervention studies are required to determine whether changes in fitness in response to exercise training are over and above expected developmental changes illustrated by our age-related reference values. While individual fitness test scores can be benchmarked and tracked, a composite or overall fitness score could also be generated as an aggregate score summarising centiles across all fitness components or across multiple components or subdomains of interest (eg, a composite score for health-related fitness should aggregate centiles for CRF, MSF and flexibility). This scoring structure, similar to that used in the Canadian Assessment of Physical Literacy,[55 56] could help identify the fitness components/subdomains in need of attention in order to provide appropriate feedback and advice to children about how to best improve their overall physical fitness. In this context, the lowest quintile has extensively been used as a threshold for defining low fitness or unfit youth.[57] In prospective cohort studies, this group has been shown to have a disproportionately higher risk for future diseases.[58] Even more stringent cut-points (eg, 10th centile) have been proposed for individuals who should be checked for the existence of other risk factors or developmental problems. In a cohort study conducted in more than 1 million Swedish male adolescents, it was observed that those in the lowest decile of muscular strength had significantly higher risk of all-cause mortality, cardiovascular disease mortality and suicide mortality, supporting the notion that this should be considered a group at risk.[12]

To date, research examining criterion-referenced standards in children and adolescents has focused on CRF,[22 23 59] with new international standards recently published for healthy CRF recently published.[23] While not the first study to estimate the percentage of European children and adolescents with apparently healthy CRF,[52] this study provides the most current and best available estimate using the new international criterion-referenced standards. This study is consistent with previous studies showing a latitudinal gradient, where children and adolescents from Northern-Central Europe typically have better CRF than their peers from Southern Europe.[16 41 42] This study also identified considerable variability in healthy CRF levels among different European countries. Variability in CRF was previously identified as a strong unfavourable correlate of country-specific income inequality (operationalised as the Gini index); meaning, countries with a large population spread of income tend to have poor CRF levels.[42] The observed age gradient in healthy CRF levels may reflect that children are generally healthier than adolescents or it may be an artefact of the new international standards being age-independent. Unfortunately, criterion-referenced standards for fitness components other than CRF do not currently exist. In addition, CRF criterion-referenced standards do not exist for outcomes other than cardiometabolic health (ie, poor bone health, mental health, cognitive health and so on), which is a limitation and represents an area for future research.

This study systematically identified and quantified the sex-specific differences in Eurofit performance, showing that boys outperformed girls on CRF, MSF and speed-agility tests and experienced larger age-specific changes, while girls outperformed boys on the flexibility test. While the underlying causes of the sex-specific differences are clear for some fitness components (eg, differences in MSF are largely explained by physical differences such as differences in body size/composition), they are less clear for others (eg, differences in CRF may be explained by physiological differences such as differences in mechanical efficiency and/or the fractional utilisation of oxygen).[21 60 61] It is, nonetheless, beyond the scope of this paper to discuss these mechanistic causes. However, there is a need for longitudinal cohort studies to better understand what mechanisms drive sex-specific and age-specific differences in physical fitness throughout childhood and adolescence.

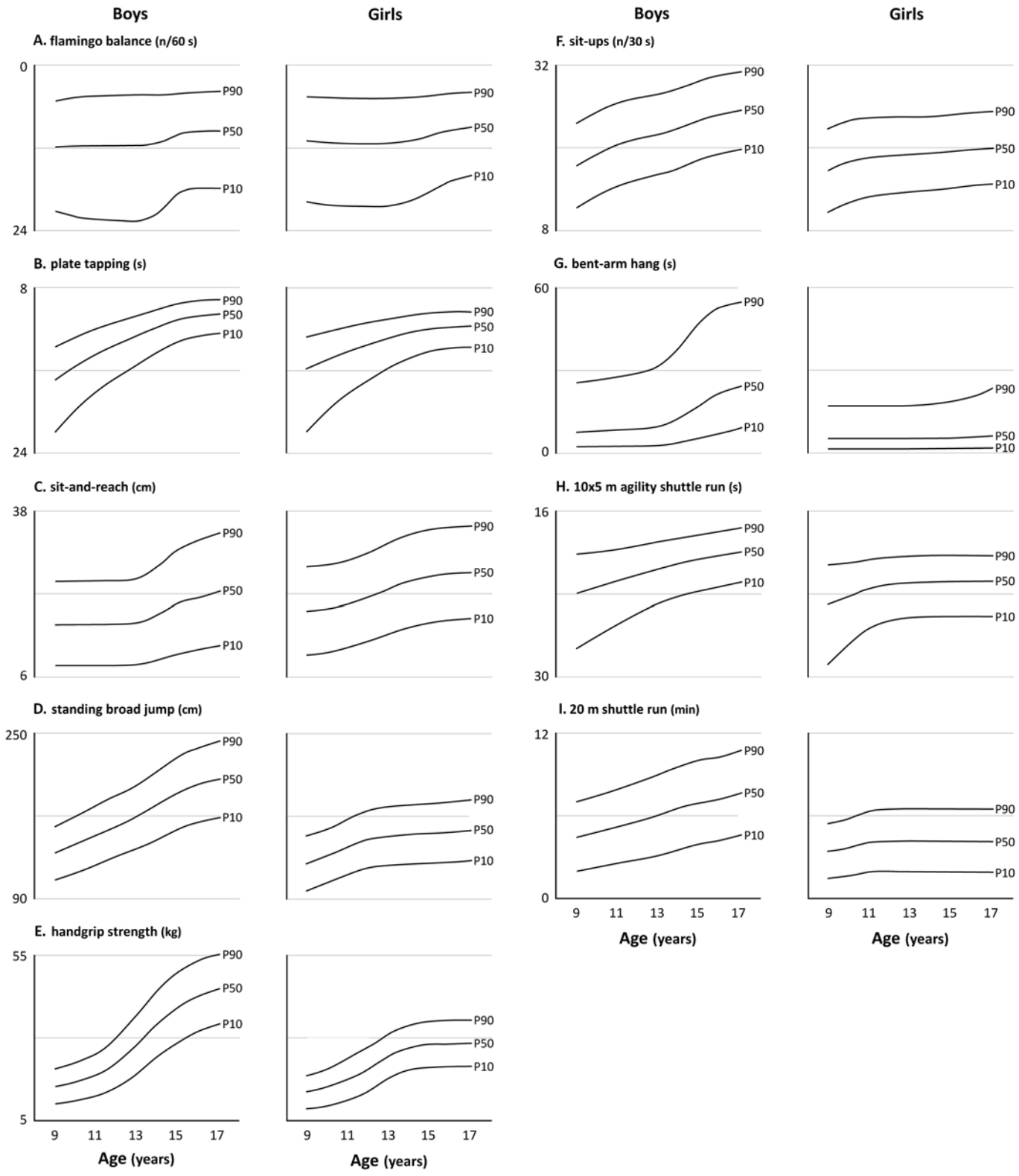


Figure 4 Smoothed centile curves (P_{10}, P_{50} and P_{90}) for (A) flamingo balance (n/60 s), (B) plate tapping (s), (C) sit-and-reach (cm), (D) standing broad jump (cm), (E) handgrip strength (kg), (F) sit-ups (n/30 s), (G) bent-arm hang (s), (H) 10×5 m agility shuttle run (s) and (I) 20 m shuttle run (min).

Strengths and limitations

This study summarised cross-sectional Eurofit data from 98 studies to generate probably Europe's largest physical fitness database for children and adolescents. Although not the first comprehensive review of children's Eurofit performance, it does provide an update to a previous review[16] by: (1) extending the data coverage from 2001 to 2015 through a rigorous systematic review process, (2) producing sex-specific and age-specific European normative values and (3) estimating the percentage of European children and adolescents with healthy CRF.

Despite the strengths of this study, it is not without limitations. First, we pooled data from studies that used different

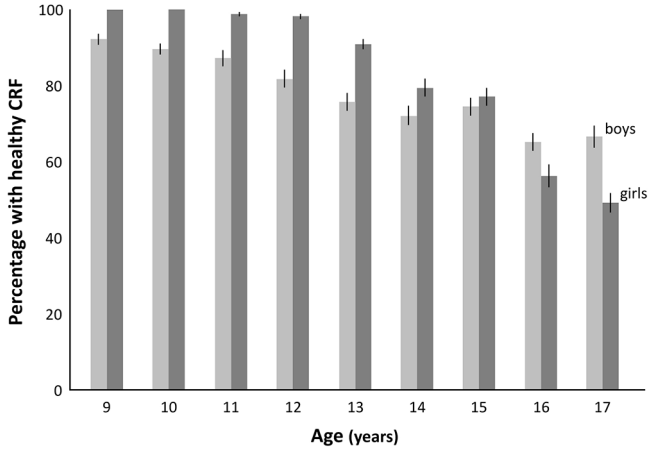


Figure 5 Percentage of European children and adolescents aged 9–17 years meeting the new international criterion-referenced standards of 42 mL/kg/min (boys, light grey bars) and 35 mL/kg/min (girls, dark grey bars) for healthy CRF. The thin black vertical lines show the 95% CIs. CRF, cardiorespiratory fitness.

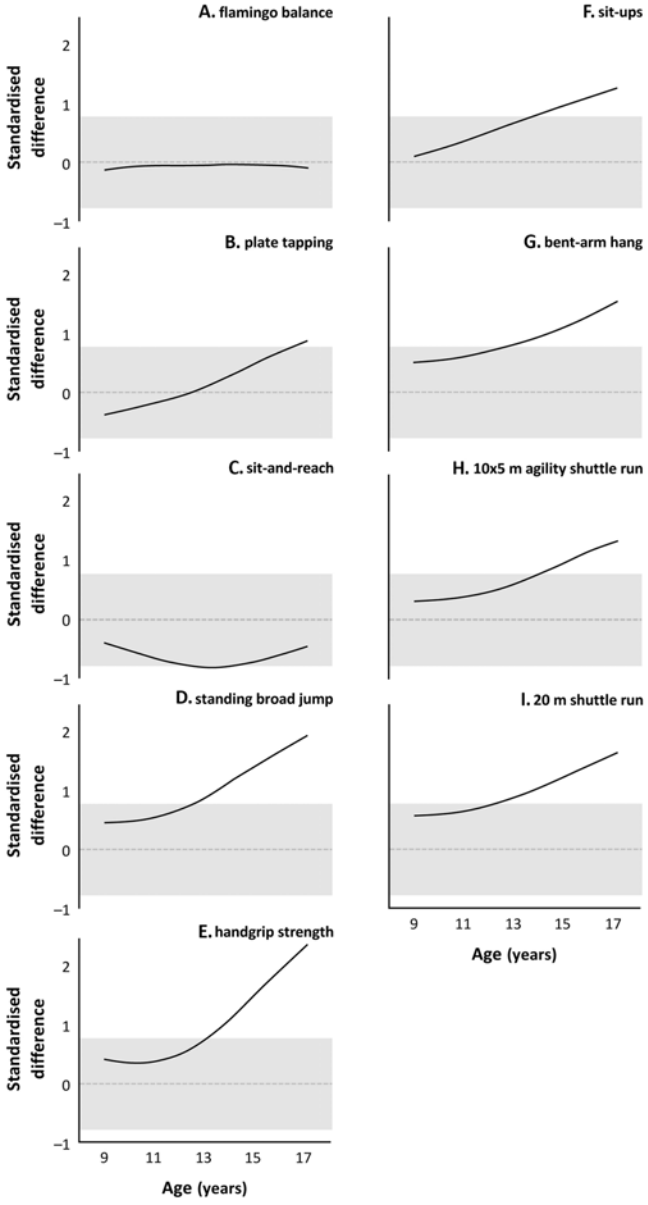


Figure 6 Standardised sex-specific differences in mean Eurofit performance for European children and adolescents aged 9–17 years. The limits of the grey zone represent the threshold for a large standardised difference (ie, 0.8 or −0.8). Positive differences indicated that Eurofit performances for boys were better than those for girls.

sampling methods (probability and non-probability sampling) and sampling frames (national-level, state/provincial-level and community-level), which raises the issue of representativeness. However, we used the best available data and a poststratification population weighted approach to control for oversampling and undersampling across studies and countries. Second, differences in testing conditions (eg, climate, altitude, practice and testing surfaces) and measurement errors (eg, methodological drift and diurnal variation) might have occurred, although the large number of included data points should have minimised these issues. Third, the vigorous nature of the Eurofit may have resulted in difficulties in testing, or exclusion of, individuals with a lower level of physical function. The absence of data from these populations may have inflated our norms within the lower centile range. Fourth, our sex-specific and age-specific norms and differences in Eurofit performance are also limited by the potential for unmeasured confounding. For example, biological maturation, which was rarely reported in the included studies and was therefore not included in our analysis, confounds sex-specific and age-specific differences in physical fitness.[62] Large-scale longitudinal studies focused on the influence of maturation on physical fitness are needed. Finally, Eurofit data were also collected at different times in the period between 1981 and 2015 and given evidence of temporal changes in some (but not all) fitness components in European children,[21 28 63–69] it is possible that our norms represent a different health-related picture than what would actually be observed today. However, without the availability of temporal trends data for all included countries, temporal corrections of our norms are not possible.

Recommendations

Given the widespread use of the Eurofit and other test batteries such as the ALPHA, there is a need for consistent reporting of results across studies to assist future data pooling and the update of normative values. In addition to recommending that the Eurofit be routinely administered (in part or in whole) in schools to improve national and regional surveillance of health and fitness, we also make the following recommendations:

1. An online multilingual operations and procedures manual, including instructional videos, should be made available (eg, the ALPHA project manual, http://profith.ugr.es/alpha-children). Researchers should make de-identified raw data available through an online data repository[42 70] in order to help improve surveillance efforts across the region. For example, scheduled for official release in 2018 is a free website (http://www.activehealthykids.org/kids-fit-guide/) that will compute a report comparing individual 20 m shuttle run performances to national, regional and international normative values and criterion-referenced standards, providing researchers with valuable analytical support.
2. Care should be taken to minimise and report factors that may impact fitness test performance (eg, climate, temperature, humidity, altitude, clothing, ground surfaces/conditions, pretest instructions and test familiarisation). Studies should be conducted to assess the effect of these factors on fitness test performance.

What are the new findings?

- This study presents the largest and most geographically representative sex-specific and age-specific European normative values for physical fitness in children and adolescents.
- This study estimated that 78% (95% CI 72% to 85%) of boys and 83% (95% CI 71% to 96%) of girls met the new international criterion-referenced standards of 42 and 35 mL/kg/min respectively for healthy cardiorespiratory fitness (CRF), with the percentage meeting the standards decreasing with age.
- This study showed that boys performed better than girls on muscular strength, muscular power, muscular endurance, speed-agility and CRF tests, but worse on the flexibility test. Boys' fitness also generally improved at a faster rate than girls' fitness, especially during the teenage years.

How might it impact on clinical practice in the future?

- Sex-specific and age-specific European normative values for physical fitness in children and adolescents are important for health and fitness screening, profiling, monitoring and surveillance.

3. Best practice should include that: (1) test protocols be followed and test results be reported as per the operations and procedures manual; (2) biological age (sexual maturation) be measured (if appropriate) in addition to chronological age; (3) descriptive statistics (sample sizes, means and SDs) be reported in 1 year age and sex groups based on age at last birthday and (4) the year(s) of testing be reported.

CONCLUSION

Physical fitness is an important indicator of good health, and the Eurofit is probably the most popular way to measure physical fitness throughout Europe. This study pooled 2 779 165 Eurofit performances, representing children and adolescents from 30 European countries. This large summary analysed the best available Eurofit data to: (1) provide the largest and most geographically representative sex-specific and age-specific European normative values for physical fitness in children and adolescents and (2) estimate the percentage of children and adolescents with healthy CRF according to the new international criterion-referenced standards. These data have utility for both health and sport promotion given that they help to identify children and adolescents with: (1) very low/poor fitness in order to set appropriate fitness goals, monitor longitudinal changes and promote positive health-related fitness behaviours (eg, physical activity and exercise promotion) and (2) very high/good fitness in the hope of recruiting them into sporting or athletic development programmes.

Acknowledgements We would like to thank the authors of the included studies for generously clarifying details of their studies and/or for providing raw data.

Contributors GRT developed the systematic review research question and objectives. GRT, ND and LL created the search strategy and provided guidance on review methodology. KC, FA, LL and ND screened and extracted the data. GRT and KC led the data analysis, data synthesis and writing of the manuscript. All authors contributed to interpretation of the results, edited, reviewed and approved the final manuscript.

Funding A College Research Council Summer Research Professorship from the College of Education and Human Development at the University of North Dakota supported this project. FBO research activity is by the Spanish Ministry of Economy and Competitiveness—MINECO (RYC-2011-09011, DEP2016-79512-R); from the University of Granada, Plan Propio de Investigación 2016, Excellence actions: Units of Excellence; Unit of Excellence on Exercise and Health (UCEES); from the EXERNET Research Network on Exercise and Health in Special Populations (DEP 2005-00046/ACTI) and from the SAMID III network, RETICS, funded by the PN I+D+I 2017-2021 (Spain), ISCIII-Sub-Directorate General for Research Assessment and Promotion, the European Regional Development Fund (ERDF) (Ref. RD16/002).

Competing interests None declared.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 Ortega FB, Ruiz JR, Castillo MJ, *et al*. Physical fitness in childhood and adolescence: a powerful marker of health. *Int J Obes* 2008;32:1–11.
2 Lee DC, Sui X, Ortega FB, *et al*. Comparisons of leisure-time physical activity and cardiorespiratory fitness as predictors of all-cause mortality in men and women. *Br J Sports Med* 2011;45:504–10.
3 Kodama S, Saito K, Tanaka S, *et al*. Cardiorespiratory fitness as a quantitative predictor of all-cause mortality and cardiovascular events in healthy men and women: a meta-analysis. *JAMA* 2009;301:2024–35.
4 Katzmarzyk PT, Craig CL. Musculoskeletal fitness and risk of mortality. *Med Sci Sports Exerc* 2002;34:740–4.
5 Sawada SS, Lee IM, Naito H, *et al*. Cardiorespiratory fitness, body mass index, and cancer mortality: a cohort study of Japanese men. *BMC Public Health* 2014;14:1012.
6 Erikssen G, Liestøl K, Bjørnholt J, *et al*. Changes in physical fitness and changes in mortality. *Lancet* 1998;352:759–62.
7 Slattery ML, Jacobs DR. Physical fitness and cardiovascular disease mortality. The US railroad study. *Am J Epidemiol* 1988;127:571–80.
8 Smith JJ, Eather N, Morgan PJ, *et al*. The health benefits of muscular fitness for children and adolescents: a systematic review and meta-analysis. *Sports Med* 2014;44:1209–23.
9 Ruiz JR, Castro-Piñero J, Artero EG, *et al*. Predictive validity of health-related fitness in youth: a systematic review. *Br J Sports Med* 2009;43:909–23.
10 Ortega FB, Labayen I, Ruiz JR, *et al*. Improvements in fitness reduce the risk of becoming overweight across puberty. *Med Sci Sports Exerc* 2011;43:1–7.
11 Högström G, Nordström AN. Aerobic fitness in late adolescence and the risk of early death: a prospective cohort study of 1.3 million Swedish men. *Int J Epi* 2016;45:1159–68.
12 Ortega FB, Silventoinen K, Tynelius P, *et al*. Muscular strength in male adolescents and premature death: cohort study of one million participants. *BMJ* 2012;345:e7279.
13 Sato M, Kodama S, Sugawara A, *et al*. Physical fitness during adolescence and adult mortality. *Epidemiology* 2009;20:463–4.
14 Armstrong N, Tomkinson G, Ekelund U. Aerobic fitness and its relationship to sport, exercise training and habitual physical activity during youth. *Br J Sports Med* 2011;45:849–58.
15 Kemper HCG, Van Mechelen W. Physical fitness testing of children: a European perspective. *Pediatr Exerc Sci* 1996;8:201–14.
16 Tomkinson GR, Olds TS, Borms J. Who are the Eurofittest? *Med Sport Sci* 2007;50:104–28.
17 Council of Europe. *Eurofit: handbook for the Eurofit tests of physical fitness*. Rome: Council of Europe, 1988.
18 Artero EG, España-Romero V, Castro-Piñero J, *et al*. Reliability of field-based fitness tests in youth. *Int J Sports Med* 2011;32:159–69.
19 Castro-Piñero J, Artero EG, España-Romero V, *et al*. Criterion-related validity of field-based fitness tests in youth: a systematic review. *Br J Sports Med* 2010;44:934–43.
20 Ruiz JR, Castro-Piñero J, España-Romero V, *et al*. Field-based fitness assessment in young people: the ALPHA health-related fitness test battery for children and adolescents. *Br J Sports Med* 2011;45:518–24.
21 Tomkinson G, Olds T. Field tests of fitness. In: Armstrong N, Van Mechelen W, eds. *Paediatric exercise science and medicine*. 2 ed. United Kingdom: Oxford, 2008:109–28.
22 Lang JJ, Tremblay MS, Ortega FB, *et al*. Review of criterion-referenced standards for cardiorespiratory fitness: what percentage of 1 142 026 international children and youth are apparently healthy? *Br J Sports Med* 2017 doi: 10.1136/bjsports-2016-096955. [Epub ahead of print 02 Mar 2017].
23 Ruiz JR, Cavero-Redondo I, Ortega FB, *et al*. Cardiorespiratory fitness cut points to avoid cardiovascular disease risk in children and adolescents; what level of fitness should raise a red flag? A systematic review and meta-analysis. *Br J Sports Med* 2016;50:1451–8.
24 IOM (Institute of Medicine). *Fitness measures and health outcomes in youth*. Washington, DC: The National Academies Press, 2012.

25 Cauderay M, Narring F, Michaud P-A. A cross-sectional survey assessing physical fitness of 9- to 19-year-old girls and boys in Switzerland. *Pediatr Exerc Sci* 2000;12:398–412.
26 Haugen T, Høigaard R, Seiler S. Normative data of BMI and physical fitness in a Norwegian sample of early adolescents. *Scand J Public Health* 2014;42:67–73.
27 Jürimäe T, Volbekiene V. Eurofit test results in Estonian and Lithuanian 11 to 17-year-old children: a comparative study. *European Journal of Physical Education* 1998;3:178–84.
28 Lefèvre J, Bouckaert J, Duquet W, *et al*. *De barometer van de fysieke fitheid van de Vlaamsejeugd: de resultaten*. *Sport* 1998;4:16–22.
29 Tambalis KD, Panagiotakos DB, Psarra G, *et al*. Physical fitness normative values for 6-18-year-old Greek boys and girls, using the empirical distribution and the lambda, mu, and sigma statistical method. *Eur J Sport Sci* 2016;16:736–46.
30 Moher D, Liberati A, Tetzlaff J, *et al*. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 2009;6:e1000097.
31 Tomkinson GR, Lang JJ, Tremblay MS, *et al*. International normative 20 m shuttle run values from 1 142 026 children and youth representing 50 countries. *Br J Sports Med* 2017;51:1545–54.
32 Léger LA, Mercier D, Gadoury C, *et al*. The multistage 20 metre shuttle run test for aerobic fitness. *J Sports Sci* 1988;6:93–101.
33 Tomkinson GR, Léger LA, Olds TS, *et al*. Secular trends in the performance of children and adolescents (1980-2000): an analysis of 55 studies of the 20m shuttle run test in 11 countries. *Sports Med* 2003;33:285–300.
34 Levy PS, Lemeshow S. Stratification random sampling: further issues. In: Levy PS, Lemeshow S, eds. *Sampling of populations: methods and application*. Hoboken, NJ: John Wiley & Sons, Inc, 2008:143–88.
35 United Nations, Department of Economic and Social Affairs, Population Division. *World population prospects: the 2015 revision, key findings and advance tables working paper*, 2015. No. ESA/P/WP.241.
36 Tomkinson GR, Hamlin MJ, Olds TS. Secular trends in anaerobic test performance in Australasian children and adolescents. *Pediatr Exerc Sci* 2006;18:314–28.
37 d'Agostino RB, Pearson ES. Tests of departure from normality: empirical results for the distribution of b2 and √b1. *Biometrika* 1973;60:613–22.
38 Cole TJ, Green PJ. Smoothing reference centile curves: the LMS method and penalized likelihood. *Stat Med* 1992;11:1305–19.
39 Pan H, Cole T. *User's guide to LMSchartmaker*. UK: Medical Research Council, 2010:1–42.
40 Cohen J. *Statistical power analysis for the behavioral sciences*. 2nd ed. New Jersey: Lawrence Erlbaum, 1988.
41 Ortega FB, Ruiz JR, Labayen I, *et al*. Health inequalities in urban adolescents: role of physical activity, diet, and genetics. *Pediatrics* 2014;133:e884–e895.
42 Lang JJ, Tremblay MS, Léger L, *et al*. International variability in 20 m shuttle run performance in children and youth: who are the fittest from a 50-country comparison? A systematic literature review with pooling of aggregate results. *Br J Sports Med* 2016 doi: 10.1136/bjsports-2016-096224. [Epub ahead of print 20 Sept 2016].
43 Cole TJ, Bellizzi MC, Flegal KM, *et al*. Establishing a standard definition for child overweight and obesity worldwide: international survey. *BMJ* 2000;320:1240–3.
44 Cole TJ, Freeman JV, Preece MA. Body mass index reference curves for the UK, 1990. *Arch Dis Child* 1995;73:25–9.
45 Eisenmann JC. Waist circumference percentiles for 7- to 15-year-old Australian children. *Acta Paediatr* 2005;94:1182–5.
46 Fernández JR, Redden DT, Pietrobelli A, *et al*. Waist circumference percentiles in nationally representative samples of African-American, European-American, and Mexican-American children and adolescents. *J Pediatr* 2004;145:439–44.
47 McCarthy HD, Jarrett KV, Crawley HF. The development of waist circumference percentiles in British children aged 5.0–16.9 y. *Eur J Clin Nutr* 2001;55:902–7.
48 Shields M. Overweight and obesity among children and youth. *Health Rep* 2006;17:27–42.
49 Tremblay MS, Shields M, Laviolette M, *et al*. Fitness of Canadian children and youth: results from the 2007–2009. *Health Rep* 2009;21:1–14.
50 Jackson LV, Thalange NK, Cole TJ. Blood pressure centiles for Great Britain. *Arch Dis Child* 2007;92:298–303.
51 Jolliffe CJ, Janssen I. Development of age-specific adolescent metabolic syndrome criteria that are linked to the Adult Treatment Panel III and International Diabetes Federation criteria. *J Am Coll Cardiol* 2007;49:891–8.
52 Ortega FB, Artero EG, Ruiz JR, *et al*. Physical fitness levels among European adolescents: the HELENA study. *Br J Sports Med* 2011;45:20–9.
53 De Miguel-Etayo P, Gracia-Marco L, Ortega FB, *et al*. Physical fitness reference standards in European children: the IDEFICS study. *Int J Obes* 2014;38 Suppl 2:S57–S66.
54 Catley MJ, Tomkinson GR. Normative health-related fitness values for children: analysis of 85347 test results on 9–17-year-old Australians since 1985. *Br J Sports Med* 2013;47:98–108.
55 Longmuir PE, Boyer C, Lloyd M, *et al*. The Canadian assessment of physical literacy: methods for children in grades 4 to 6 (8 to 12 years). *BMC Public Health* 2015;15:767.
56 Francis CE, Longmuir PE, Boyer C, *et al*. The Canadian assessment of physical literacy: development of a model of children's capacity for a healthy, active lifestyle through a delphi process. *J Phys Act Health* 2016;13:214–22.
57 Ortega FB, Ruiz JR, Labayen I, *et al*. The Fat but Fit paradox: what we know and don't know about it. *Br J Sports Med* 2017.
58 Ortega FB, Lavie CJ, Blair SN. Obesity and cardiovascular disease. *Circ Res* 2016;118:1752–70.
59 Tomkinson G. Aerobic fitness thresholds for cardio metabolic health in children and adolescents. *Br J Sports Med* 2011;45:686–7.
60 Rowland TW. Evolution of maximal oxygen uptake in children. *Med Sport Sci* 2007;50:200–9.
61 Armstrong N, Welsman JR. Aerobic fitness: what are we measuring? *Med Sport Sci* 2007;50:5–25.
62 Baxter-Jones ADG. Growth and maturation. In: Armstrong N, Van Mechelen W, eds. *Children's sport and exercise medicine*. 3 ed. Oxford, United Kingdom: Oxford University Press, 2017:13–24.
63 Costa AM, Costa MJ, Reis AA, *et al*. Tendências seculares dos níveis antropométricos e de aptidão física em crianças Portuguesas. *Acta Med Port* 2017;30:108–14.
64 Ekblom O, Oddsson K, Ekblom B. Health-related fitness in Swedish adolescents between 1987 and 2001. *Acta Paediatr* 2004;93:681–6.
65 Jürimäe T, Volbekiene V, Jürimäe J, *et al*. Changes in Eurofit test performance of Estonian and Lithuanian children and adolescents (1992–2002). *Med Sport Sci* 2007;50:129–42.
66 Mahmoud O, Mészáros J, Szabo T. Secular trend and motor performance scores in Hungarian schoolboys. *Kinesiology* 2002;2:127–33.
67 Matton L, Duvigneaud N, Wijndaele K, *et al*. Secular trends in anthropometric characteristics, physical fitness, physical activity, and biological maturation in Flemish adolescents between 1969 and 2005. *Am J Hum Biol* 2007;19:345–57.
68 Moliner-Urdiales D, Ruiz JR, Ortega FB, *et al*. Secular trends in health-related physical fitness in Spanish adolescents: the AVENA and HELENA studies. *J Sci Med Sport* 2010;13:584–8.
69 Sjolie A, Mønness E. Truncus endurance, hip and ankle mobility and aerobic fitness in 15-year-old Norwegian adolescents in 1968 and 1997. *Scand J Med Sci Sports* 2007;17:488–96.
70 Lang JJ, Tomkinson GR, Janssen I, *et al*. Making a case for cardiorespiratory fitness surveillance among children and youth. *Exerc Sport Sci Rev*. In Press.

Reaction Time Aspects of Elite Sprinters in Athletic World Championships

Espen Tønnessen,[1] Thomas Haugen,[2] and Shaher A.I. Shalfawi[3]

[1]Department of Sports, University of Nordland, Bodø, Norway; [2]Department of Physical Training, Norwegian Olympic Sport Center, Oslo, Norway; and [3]Department of Physical Education, Prince Sultan University, Riyadh, Saudi Arabia

Abstract

Tønnessen, E, Haugen, T, and Shalfawi, SAI. Reaction time aspects of elite sprinters in athletic world championships. *J Strength Cond Res* 27(4): 885–892, 2013–The aim of this study was to quantify world-class sprinters' reaction times as a function of performance level, gender, body height, finalists' heat round development, and age. A database of 100-m sprint results and corresponding reaction times from 1,319 sprinters participating in different International Association of Athletics Federations world championships during the time period 2003–9 was compiled for this investigation. Seiko was the official timekeeper of the world championships in this study. Seiko uses a silent gun system for time initiation and false start detection. Their Slit Video system captures the runners at the finish line up to 2,000 images per second with high-resolution cameras. Our results indicate that there was a significant relationship ($p < 0.01$) between reaction time and 100 m running time, with a shared variance of 8.5 and 10.8% for males ($r = 0.292$) and females ($r = 0.328$), respectively. Reaction times (0.166 ± 0.030 seconds) of males were significantly shorter ($p < 0.01$) than those for females (0.176 ± 0.034 seconds). No relationship was observed between reaction time and height. Male finalist sprinters had substantially shorter reaction times in the finals (0.142 ± 0.017 seconds) compared with round 1 (0.161 ± 0.024 seconds), round 2 (0.155 ± 0.020 seconds), and the semifinals (0.153 ± 0.022 seconds). Female finalist sprinters obtained their fastest reaction times during the semifinals (0.153 ± 0.018 seconds). The best reaction times were registered at the age of 26–29 years for males (0.150 ± 0.017 seconds) and >30 years for females (0.153 ± 0.020 seconds), but reaction times across different age categories were also positively correlated with 100 m performance ($p < 0.05$). Considering the findings of this study, the results suggest that sprinters' reacting abilities affect their sprint performance over 100 m. This study provides magnitude estimates for the influence of performance level, gender, body height, finalists' heat round, and age on reaction time among world-class sprinters, which we believe to be of great interest for coaches and athletes in sports involving reacting skills.

Key Words 100 m performance, false start, female athletes, male athletes, age

Introduction

Reaction time (response time) has been defined as the time between the detection of a sensory stimulus and subsequent behavioral response (22). Collet (5) defined total reaction time as the time from the gun signal until the athlete's production of force against the starting blocks. This includes the sound traveling time between the sound source and the athlete, the athlete's reaction to the sound, and the mechanical delay of false start equipment integrated in the start block (16,17). According to the International Association of Athletics Federations (IAAF) competition rules, a reaction time less than 100 ms is considered a false start.

Reaction times can be an important determinant of success in the 100-m sprint, where medals are often decided by hundredths or even thousandths of a second. Therefore, a poor start or long reaction time can rule an athlete out of the medal hunt in a 100-m sprint competition. Reaction time to sensory stimuli has been widely examined in the literature and on various populations (18,21,22).

Unfortunately, only a few studies have examined track and field sprinters' reaction time and their moderators; Mero et al. (16) and Smirniotou et al. (20) reported no correlation between reaction time and performance level. Delalija et al. (9) observed a significant correlation between reaction time and sprint results from the 2004 Olympic Games in Athens. Meckel et al. (14) found significant differences between fast and average sprinters, whereas no differences were observed between the groups of fast and slow sprinters. The literature also remains unclear for possible gender differences in track and field; Mero et al. (16) concluded that the average reaction times for females are longer than those for males, and Dapena (8) does not necessarily support the claim by Mero et al. (16). Collet (5) reported that sprinters' reaction

Address correspondence to: Espen Tønnessen, espen.tonnessen@olympiatoppen.no.

27(4)/885–892

Journal of Strength and Conditioning Research

times decreased from one round to another toward the final for the 8 finalists. However, no available studies have so far examined world-class sprinters' reaction times related to body height or analyzed the development of reaction time through different age categories. Additionally, most of the previously published studies have examined competitions using the loud gun starts, a system that has been criticized for not delivering the "go" signal to the different lanes at the same time (2,8).

Therefore, we have created a database of reaction times, collected under highly standardized conditions based on 100-m sprint result lists from world championships for youths and seniors between the years 2003 and 2009. This provides the potential for addressing several different questions related to reaction times among world-class sprinters. Therefore, the purpose of this study was to examine and analyze reaction times of 100-m sprints from world championships and take a deeper look at its function and relation to athletes' performance level, body height, gender, heat round, and age.

Methods

Experimental Approach to the Problem

In this study, 100 m sprinters' reaction times were defined as dependent variable, whereas 100-m performance level, body height, gender, heat round, and age were defined as independent variables. All data were collected from different IAAF world championships for youths and seniors in the time period 2003–9 (Table 1), which were presented in the IAAF official website through the competition archive (12) and biography section (11). Only the 100-m competitions were included because it has been shown that reaction time increases from short dashes to longer sprints (5). World championships before 2003 and after 2009 were excluded because of different false start rules. The 2004 and 2008 Olympic Games were also excluded from this investigation because of different reaction time monitoring systems.

The meet organizer was responsible for timing, reaction time monitoring, heat seeding, lane draws, and further qualification/advancement in accordance to the IAAF competition rule guidelines (11). Body height and date of birth were identified for each athlete by self-report through the sign up procedures administrated and controlled by each of the national athletics federations, including passport identification.

Subjects

Data were collected from 1,319 sprinters in the age range 16–47 years, representing a broad range of performance levels and with varying training background. Athletes aged 16 or 17 years on the 31st of December in the year of the competition are allowed to participate in the World Youth Championships. For the World Youth Championships, only athletes aged 16–19 years may compete. Athletes younger than 16 years were not allowed to enter in the senior World Championships (11). In total, 1,719 reaction times formed the basis of this investigation. The athletes signed up and participated voluntarily for these competitions on the basis of being timed; thus, no informed consent was obtained. This study was approved by the Ethical Committee of Nordland University.

Table 1. Description of competition with corresponding subjects.*

Year	Competition	Male (*n*)	Female (*n*)	Age (y)
2003	9th IAAF World Championships, Paris, France	72	56	16–43
2003	IAAF World Youth Championships, Sherbrooke, Canada	67	56	16–17
2005	10th IAAF World Championships, Helsinki, Finland	59	55	16–34
2005	IAAF World Youth Championships, Marrakesh, Morocco	92	85	16–17
2007	11th IAAF World Championships, Osaka, Japan	67	69	16–47
2007	IAAF World Youth Championships, Ostrava, Czech Republic	82	69	16–17
2009	12th IAAF World Championships, Berlin, Germany	92	63	16–40
2009	IAAF World Youth Championships, Bressanone, Italy	90	61	16–17

*IAAF = International Association of Athletics Federations.

Procedures

All competitions included in this investigation were arranged in the middle of the summer (July or August) and were (in an athletics context) considered the most important competition of the year for the participants. All athletes had to qualify to the championships in accordance to the entry standards set by the IAAF; thus, familiarization with the competition procedures was ensured. On the first competition day, initial heats were held in the morning or middle of the day (between 10 AM and 2 PM), whereas the round 2 races were held

in the afternoon or evening (between 6 and 9 PM). On the second day, the semifinals and finals took place in the afternoon or evening (between 4 and 10 PM), intercepted by a 1.5- to 3-hour break. The athletes followed their individual warm up routines until the obligatory call room procedures 20 minutes before the start.

Regarding nutrition, hydration, sleep, and physical activity, the athletes were assumed to have prepared themselves as they would for the most important competition of the year. All subjects were familiar with the competition procedures through national qualifications.

Statistical Analyses

All data from the database were transferred to SPSS 17 (SPSS, Inc., Chicago, IL, USA) for analyses. First and for all variables in this study, the normality distribution of the data was explored by histogram plot and tested using Shapiro-Wilk test. Then descriptive statistics were calculated and reported graphically together with the 95% confidence interval (CI).

In the percentile analyses, all athletes from 2003 to 2009 world championships were included, and the reaction times were subtracted from the 100 m times to isolate running time performance from reaction time performance. Each athlete's fastest running time and their corresponding reaction time were included in the analyses; then, the relationship between reaction time and 100 m running performance were calculated using Pearson correlation coefficient (r).

In the performance-level analyses, the athletes were divided into 4 performance level categories, namely, round 1 athletes, round 2 athletes, semifinalists, and finalists. Furthermore, each athlete was represented once with their best position from the world championships 2003–9. In the event of an athlete attaining the same position in more than one championship, the fastest running time and the corresponding reaction time were included in the analyses.

In the heat round analyses, only finalists from senior level were included. In the event of an athlete attaining the same position in more than one championship, the fastest running time and corresponding reaction time from round 1, round 2, semifinal, and final were included in the analyses.

In the age analyses, athletes were divided into 6 age groups: younger than 18, 18–19, 20–22, 23–25, 26–29, and older than 30 years. The athlete's age was calculated based on their date of birth and the competition date. In the event of an athlete falling under the same age group twice through the 2003–9 championships, the fastest running time within that age group and the corresponding reaction time were included in the analyses.

For all analyzed variables in this study, if the data were found to follow a normal distribution, the differences in the reaction time between the variables analyzed were determined using a 1-way analysis of variance followed by Tukey's post hoc test. However, if the data were not normally distributed, the nonparametric Kruskal-Wallis test was assessed, followed by Mann-Whitney Test. The level of significance was set at $p \leq 0.05$ for all analyses.

Results

With regard to reliability, all competitions included in this investigation were arranged in an athletics context, considered the most important competition of the year for the participants. All athletes had to qualify to the championships in accordance with the entry standards set by the IAAF; thus, familiarization with the competition procedures was ensured. All subjects were familiar with the competition procedures through national qualifications. Thus, we have not conducted any performance testing for the purpose of reliability; however, because all the data presented in this study were collected under highly controlled and standardized procedures, we believe they are reliable.

Figure 1 presents percentile data of the athletes' reaction time from the IAAF championships in the time period 2003–9. There was a significant ($p < 0.01$) relationship between reaction time and 100 m running time for both male ($r = 0.292$) (reaction time = 0.166 ± 0.030 seconds) and female ($r = 0.328$) (reaction time = 0.176 ± 0.034 seconds). This relationship had a shared variance of 8.5 and 10.8% for males and females, respectively. Mean 100 m running time was 10.8 ± 0.6 seconds (±*SD*) for males and 12.0 ± 0.8 seconds for females. No relationship was observed between reaction time and body height.

Figure 2 demonstrates that male finalists had a substantially ($p < 0.05$) shorter reaction time (0.142 ± 0.017 seconds) compared with semifinalists (0.153 ± 0.022 seconds), athletes from round 2 (0.155 ± 0.020 seconds), and athletes from round 1 (0.161 ± 0.024 seconds). Furthermore, the semifinalists' reaction times were significantly shorter ($p < 0.05$) than round 1 athletes.

Corresponding results for females show that reaction times for round 1 athletes (0.175 ± 0.029 seconds) were significantly longer ($p < 0.05$) than finalists (0.154 ± 0.025 seconds), semifinalists (0.153 ± 0.018 seconds), and round 2 athletes (0.161 ± 0.018 seconds). Semifinalists had a significantly shorter reaction time ($p < 0.05$) compared with round 2 athletes. No further differences were observed. Semifinalists achieved the shortest reaction times for females (Figure 2).

Figure 3 shows the development in reaction times through different heat rounds for the finalists' best performance in the time period 2003–9. For females, reaction times in the semifinals were significantly shorter ($p < 0.05$) compared with round 1 and 2. No significant heat round development was observed for males. However, the 95% CI demonstrates a slight trend toward faster reaction times from the preliminary rounds to the finals. The shortest reaction times were observed in the semifinals for females and in the finals for males.

Figure 4 demonstrates that male athletes younger than 18 years had a significantly longer reaction time (0.170 ± 0.031 seconds) ($p < 0.01$) than the other age groups. Male athletes

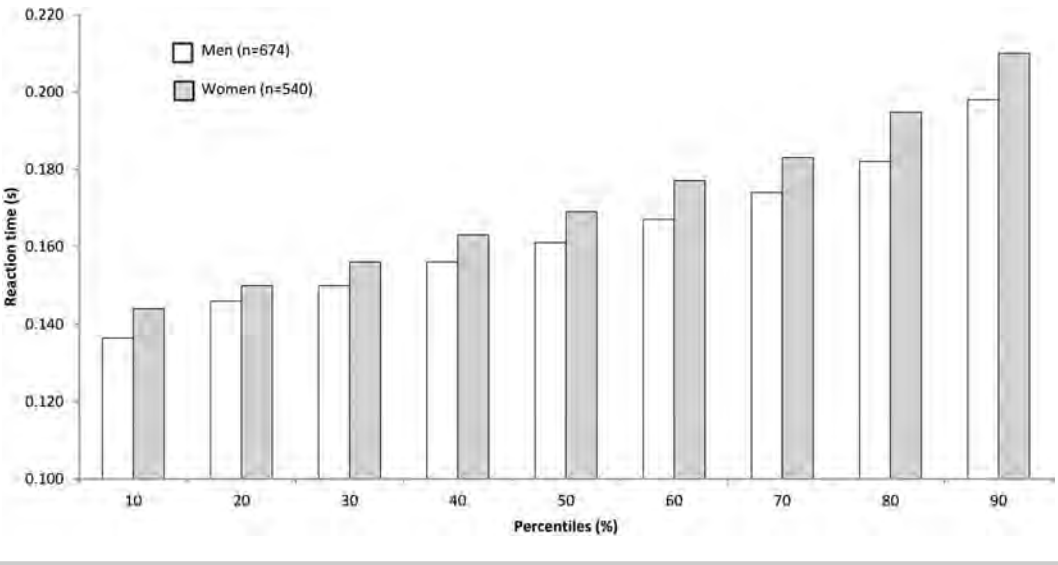


Figure 1. Reaction time for men and women in percentiles.

in the age category of 18–19 years had a significantly longer reaction time (0.164 ± 0.030 seconds) ($p < 0.05$) compared with the athletes in the age category of 26–29 years (0.150 ± 0.017 seconds) but not with the age group of 20–22 years (0.160 ± 0.024 seconds) and 23–25 years (0.158 ± 0.026 seconds). For females, a significantly longer reaction time was observed for the athletes younger than 18 years (0.180 ± 0.039 seconds) compared with other age groups ($p < 0.05$). The 18- to 19-year female athletes had a significantly longer reaction time (0.171 ± 0.029 seconds) ($p < 0.05$) compared with the female athletes older than 30 years (0.153 ± 0.020 seconds). No further significant differences were observed between age groups.

The results show that the 100-m performance level among male sprinters peaks in the age range of 26–29 years: younger than 18 years (11.16 ± 0.52 seconds), 18–19 years (10.86 ± 0.41 seconds), 20–22 years (10.73 ± 0.51 seconds), 23–25 years (10.56 ± 0.61 seconds), 26–29 years

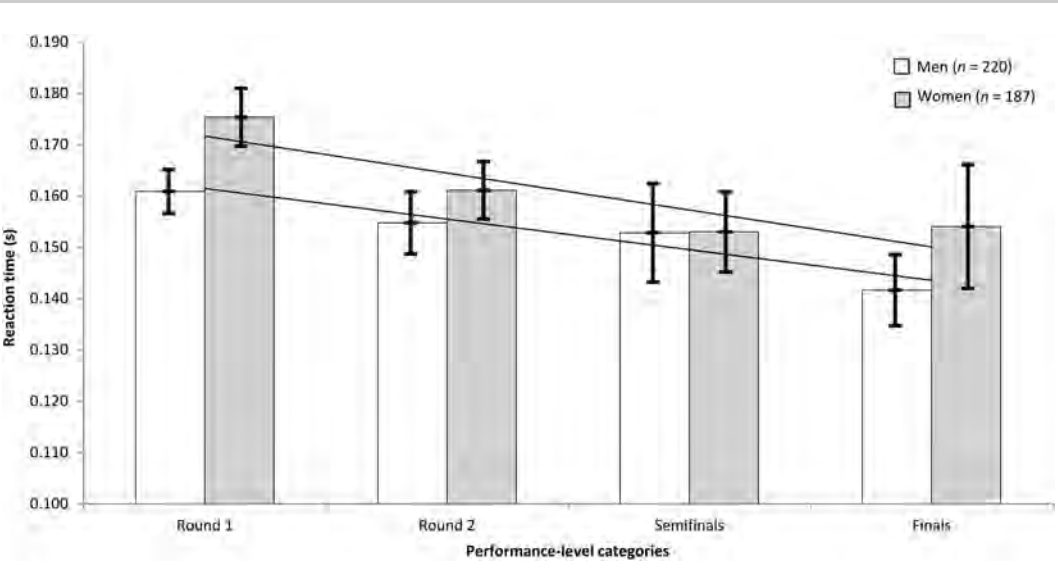


Figure 2. Mean reaction time and 95% confidence interval for athletes who went out of competition at different performance level categories.

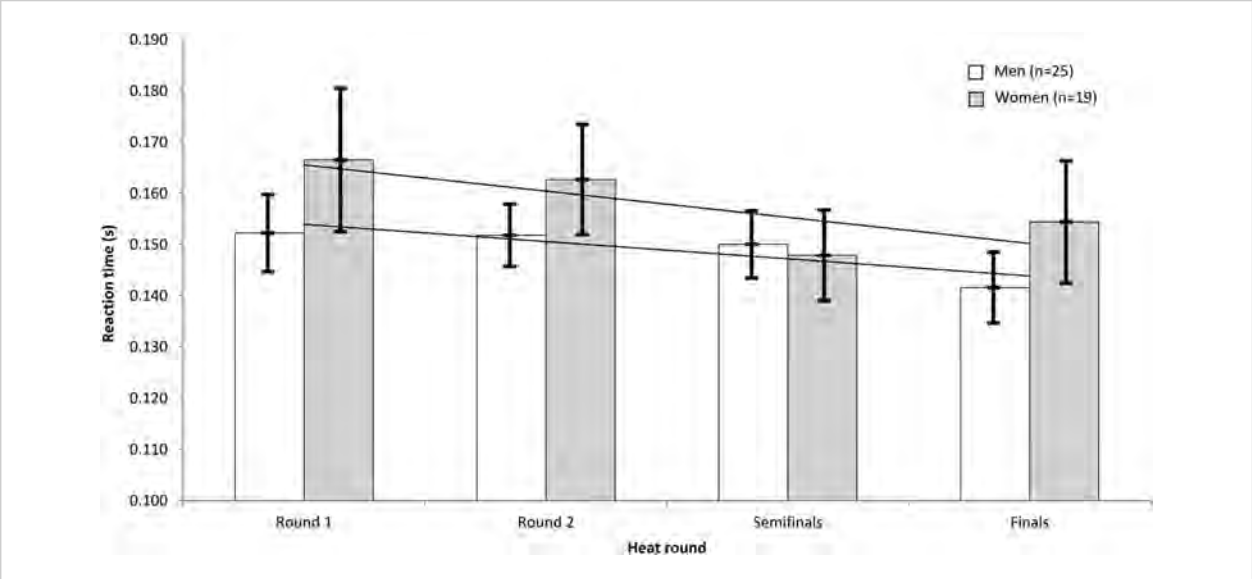


Figure 3. Mean reaction time and 95% confidence interval of the finalists' reaction time development through competition.

(10.40 ± 0.36 seconds), and older than 30 years (10.54 ± 0.49 seconds). Corresponding results for female sprinters demonstrate an improvement in 100 m performance through almost all the age categories: younger than 18 years (12.37 ± 0.71 seconds), 18–19 years (12.17 ± 0.72 seconds), 20–22 years (12.31 ± 1.08 seconds), 23–25 years (11.73 ± 0.84 seconds), 26–29 years (11.58 ± 0.55 seconds), and older than 30 years (11.46 ± 0.54 seconds).

Discussion

The percentiles in this study show that the reaction times of male and female 100-m sprinters generally vary between 0.14 and 0.20 seconds (Figure 1). The 10th percentile of reaction times was approximately 0.02 second faster than the average. These relatively modest variations can still be decisive in a sport where competitive placing is separated by mere hundredths of a second. To our knowledge, this study

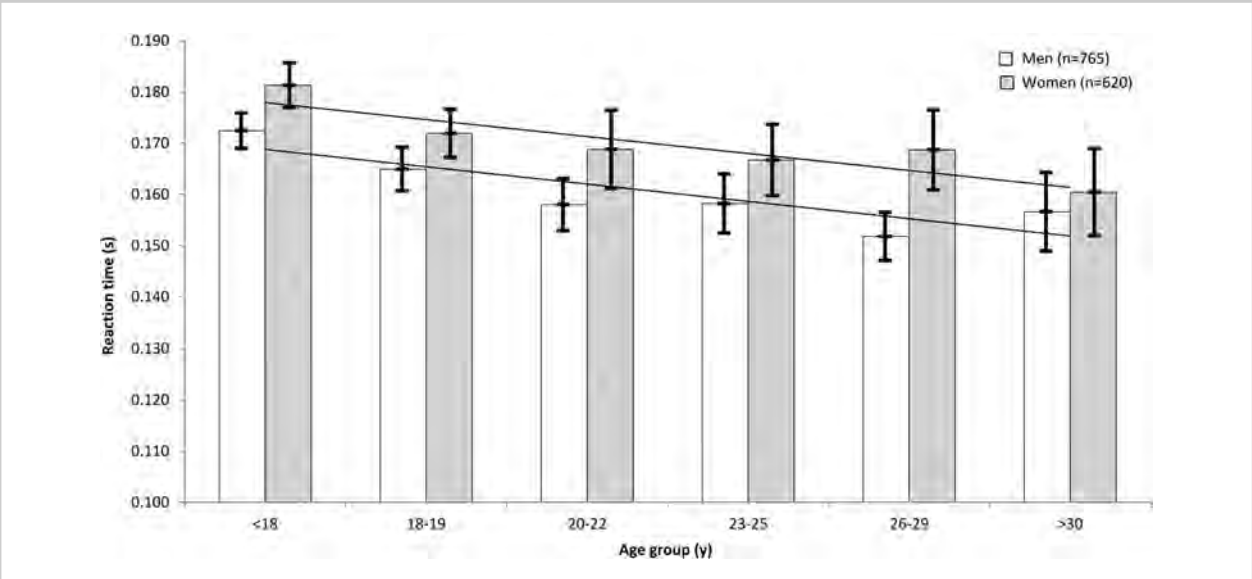


Figure 4. Mean reaction time and 95% confidence interval of the reaction time at different age categories.

is the only one to date with so many participants at a high level that uses only data in which only the silent gun system was used. Therefore, no other study is available for comparison purposes.

Our data show a weak significant correlation between reaction time and 100 m running time for both males and females. Reaction time and 100 m running time shared only approximately 10% of common variance. The detection of the relationship between reaction time and performance in 100 m running time may have been facilitated by the large sample size and the large spread in performance level observed in this study. No significant relationship has been observed in smaller groups of subjects of approximately the same performance level (14). However, a weak but significant correlation has been reported in larger heterogeneous groups (9,20). Mero et al. (16) concluded in his review article that there was no correlation between reaction time and performance level. However, his review was based on studies in which the loud gun system was used. Furthermore, a possible explanation for the weak correlation between reaction time and performance observed in this study could be that both variables are largely determined by the same physiological factors. As early as 1976, Costill et al. (6) found that sprinters had far more type II fibers than athletes from other sports. Muscle fiber type composition determines, to a large extent, anaerobic power and neuromuscular conditions such as the arrival of the stimulus at the sensory organ, conversion by the sensory organ to a neural signal, neural transmission and processing, and muscular activation (5,17). Similarly, neuromuscular conditions may explain the relationship between simple jumping tests, such as squat jump and countermovement jump, and reaction time and 100 m time in sprinters (20).

In theory (19), taller athletes could have a slower reaction time than shorter athletes with otherwise identical physiological characteristics because of the distance the nerve impulse has to travel. However, no relationship between body height and reaction time was observed in this study.

The results of this study show that males and females have an average reaction time of 0.166 ± 0.030 seconds and 0.176 ± 0.034 seconds, respectively (Figure 1). In addition, the results show that the reaction times of the male athletes were significantly shorter ($p < 0.01$) than the female athletes' reaction times. Our findings are consistent with Mero and Komi (15) who showed that male sprinters had a significantly shorter reaction time than female sprinters at the Finnish national level. Other studies of elite sprinters have also reported a significant sex difference of approximately 0.02 seconds (1,8). We are not aware of any studies that are able to explain this difference, but Spierer et al. (22) have found that women react faster to a visual stimulus compared with sound-based stimulus. Adam et al. (1) explain that differences in reaction time between gender are related to information processing speed, whereas Spierer et al., (22) indicated that the processing speed could be caused by an inherent neurological function that may differ by gender. We can also speculate that it may be because of factors such as genetics, training background, and, in particular, differences in performance level. The data for both males and females show significant differences in reaction time between athletes of different performance levels (Figure 2). However, this cannot fully account for the sex-related difference, as generally observed in this study that the reaction times of female runners at a given 100-m performance level is shorter than that of male sprinters of the same level (Figure 2). Furthermore, even though the best female group (race time = 11.46 seconds) has a longer 100-m sprint time than the slowest male group (race time = 11.16 second), these 11.46-second female athletes have better start times than the 11.16-second male athletes (Figure 3). This may be because these female athletes generally have more years of dedicated training behind them compared with the male athletes. The female runners in the above performance category are vying for international medals, whereas the men who run as fast barely qualify for the World Championships. Collet (5) found a shorter reaction time in elite sprinters with a good training background and high level of experience. It was suggested that this was because of the older and more experienced athletes being able to memorize and anticipate starting procedures to a greater degree than inexperienced sprinters. Colakoglu et al. (4) proposed that reaction time is a motor skill that can be developed through training and maturation. Whether differences in reaction time are because maturation (age) or training background cannot be determined from data in this study or other studies. Reaction time is defined as the time taken from the firing of the start signal to the athlete generating a force (pressure) against the starting blocks (5). Athletes with high maximal strength and rate of force development in the leg extensors will be able to develop a force faster than athletes with poorer physical characteristics. This could be an explanation for the difference in reaction time between males and females.

Analyses of male finalists showed that their reaction time in the final was significantly shorter ($p < 0.05$) than in the semifinals, round 2, and round 1 (Figure 3). Female finalists had shorter reaction times in the final than in round 1 and round 2 ($p < 0.05$). However, from the semifinal to the final, there were no significant differences in reaction time for females (Figure 3). Collet (5) found similar results when studying semifinalists and finalists in the Olympic Games and World Championships in the period 1987–97. Reaction times in our study were around 0.02 seconds shorter than in the investigation by Collet (5). However, the results are not fully comparable because of differences in false start rules and measuring instruments used. In the 1990s, each athlete was permitted 2 false starts before he or she was disqualified. Our data collection covered the period between 2003 and 2009 when only 1 false start per heat was permitted. The old rules therefore allowed athletes greater scope to anticipate when the starting signal occurred. It is not surprising that the

reaction time of finalists is reduced from the initial rounds to the final. The best athletes do not need to perform optimally in round 1 because they are relatively certain to go through to the next round of the competition. The focus for these athletes is therefore on saving energy and not risking a false start. The female athletes in this study had a markedly slower reaction time in the final compared with the semifinals, although this difference was not statistically significant (Figure 3). Increased external pressure and fear of disqualification may inhibit the ability to react quickly. The arousal levels of some athletes are likely to have been higher than optimal. It has been shown that if arousal exceeds a certain level, performance becomes impaired (3,24).

The results show a significant difference in reaction time ($p < 0.01$) for male athletes younger than 18 years of age compared with all other age groups (Figure 4). Furthermore, male athletes aged 18–19 years old had a slower reaction time ($p < 0.05$) than athletes in the age group 26–29 years. The data for female athletes showed that the reaction times of athletes younger than the age of 18 years were significantly slower than in all the other age groups ($p < 0.05$). The trend for both female and male athletes was that the reaction time decreased with increasing age. This is consistent with results from previous studies of elite sprinters (5). A key question is why reaction time decreases with increasing age? Neither this study nor other studies have so far been able to provide a satisfactory answer to this question, and the suggested explanations therefore remain pure conjecture. Maturation and training could be possible explanations. Most researchers who have studied reaction time propose that an individual's ability to react quickly to a stimulus is to a large extent related to the nervous system (2,5,15,17) and muscular system (2,14,20). This may explain the relationship between vertical jump height, reaction time, and performance in male sprinters (20).

The significant improvements in reaction time at the age of late 20s could also be caused by the differences in the performance level between the different age categories. As we found a significant correlation between reaction time and 100 m running time, this could affect the reaction times as a function of age. Several studies of elite athletes have shown that it takes many years of training to develop physical ability and the performance determining factors necessary to win medals at international championships (7,13,23). Another explanation could be that older athletes have a greater genetic predisposition for reacting and running quickly. A dropout study showed that those individuals who gave up athletics at the elite level were athletes who had not experienced progress or achieved their goals over the last couple of years (10). This dropout may have led to athletes in the older age categories having a better genetic predisposition for fast reaction times and sprint running performance than athletes in the younger age categories.

The data indicate that females in this study showed shorter reaction times and increased performance in 100 m with increasing age. To our knowledge, there are no obvious physiological explanations for why female sprinters reduce their reaction time in their 30s, whereas male athletes achieve the fastest reaction times in the second half of their 20s. One possible reason may be a greater dropout of female athletes with increasing age (10), so that only the very best athletes comprise the majority in the oldest age categories.

Practical Applications

The present investigation demonstrates a significant relationship between reaction time and 100 m running time of sprinters from a broad range of performance levels. This relationship indicates that reaction time effects performance in 100-m sprint. Furthermore, the variations of the reaction time found in this study can be decisive in a sport where competitive placing is separated by mere hundredths of a second. The fact that our results showed that reaction times of male athletes were significantly shorter than female athletes, the slower reaction time in females' finals compared with females' semifinals, and the trend of reaction time development through age for both males and females suggest a different training strategies approach by both male and female coaches to achieve faster reaction times. Practitioners should explore possible training methods to improve the athletes' reacting skills. Future research could focus more on the cause-effect relationships between reaction time and performance level. Mental training of sprinters might ensure an optimized arousal level at the start line to obtain fast reaction times. A poor reaction time can definitely rule an athlete out of the medal hunt.

References

1. Adam, JJ, Paas, FG, Buekers, MJ, Wuyts, IJ, Spijkers, WA, and Wallmeyer, P. Gender differences in choice reaction time: Evidence for differential strategies. *Ergonomics* 42: 327–335, 1999.
2. Brown, AM, Kenwell, ZR, Maraj, BK, and Collins, DF. "Go" signal intensity influences the sprint start. *Med Sci Sports Exerc* 40: 1142–1148, 2008.
3. Causer, J, Holmes, PS, Smith, NC, and Williams, AM. Anxiety, movement kinematics, and visual attention in elite-level performers. *Emotion* 11: 595–602, 2011.
4. Colakoglu, H, Akgun, N, Yalaz, G, and Ertat, A. The effects of speed training in acoustic and optic reaction times. *Turk J Sports Med* 22: 37–46, 1987.
5. Collet, C. Strategic aspects of reaction time in world-class sprinters. *Percept Mot Skills* 88: 65–75, 1999.
6. Costill, DL, Daniels, J, Evans, W, Fink, W, Krahenbuhl, G, and Saltin, B. Skeletal muscle enzymes and fiber composition in male and female track athletes. *J Appl Physiol* 40: 149–154, 1976.
7. Coyle, E. Improved muscular efficiency displayed as Tour de France champion matures. *J Appl Physiol* 98: 2191–2196, 2005.
8. Dapena, J. The "Loud Gun" starting system currently used at the Olympic Games does not work properly. Available at: http://www.trackandfieldnews.com/features/2005/start-problem.html Accessed November 8, 2011.
9. Delalija, A and Babić, V. Reaction time and sprint results in athletics. *Int J Perform Anal Sport* 8: 67–75, 2008.

10. Enoksen, E. Drop-out rate and drop-out reasons among promising Norwegian track and field athletes–a 25 year study. *Scand Sport Stud Forum* 2: 19–43, 2011.
11. IAAF. Athletes' Biographies Section. Available at: http://www.iaaf.org/athletes/biographies/index.html Accessed November 8, 2011.
12. IAAF. Competition web sites. Available from: http://www.iaaf.org/history/index.html# Accessed November 8, 2011.
13. Jones, AM. A five year physiological case study of an Olympic runner. *Br J Sports Med* 32: 39–43, 1998.
14. Meckel, Y, Atterbom, H, Grodjinovsky, A, Ben-Sira, D, and Rotstein, A. Physiological characteristics of female 100 metre sprinters of different performance levels. *J Sports Med Phys Fitness* 35: 169–175, 1995.
15. Mero, A and Komi, PV. Reaction time and electromyographic activity during a sprint start. *Eur J Appl Physiol Occup Physiol* 61: 73–80, 1990.
16. Mero, A, Komi, PV, and Gregor, RJ. Biomechanics of sprint running. A review. *Sports Med* 13: 376–392. 1992.
17. Pain, MT and Hibbs, A. Sprint starts and the minimum auditory reaction time. *J Sports Sci* 25: 79–86. 2007.
18. Salonikidis, K and Zafeiridis, A. The effects of plyometric, tennis-drills, and combined training on reaction, lateral and linear speed, power, and strength in novice tennis players. *J Strength Cond Res* 22: 182–191. 2008.
19. Samaras, TT. Advantages of shorter human height. In: *Human Body Size and the Laws of Scaling: Physiological, Performance, Growth, Longevity and Ecological Ramifications.* T. Samaras, ed. New York, NY: Nova Science Publishers, Inc., 2007. pp. 47–61.
20. Smirniotou, A, Katsikas, C, Paradisis, G, Argeitaki, P, Zacharogiannis, E, and Tziortzis, S. Strength-power parameters as predictors of sprinting performance. *J Sports Med Phys Fitness* 48: 447–454, 2008.
21. Spierer, DK, Petersen, RA, and Duffy, K. Response time to stimuli in division I soccer players. *J Strength Cond Res* 25: 1134–1141, 2011.
22. Spierer, DK, Petersen, RA, Duffy, K, Corcoran, BM, and Rawls-Martin, T. Gender influence on response time to sensory stimuli. *J Strength Cond Res* 24: 957–963, 2010.
23. Tønnessen, E. *Hvorfor ble de beste best? In Department of Physical Performance and Coaching.* Oslo, Norway: The Norwegian School of Sport Sciences, 2009.
24. Vickers, JN and Williams, AM. Performing under pressure: The effects of physiological arousal, cognitive anxiety, and gaze control in biathlon. *J Mot Behav* 39: 381–394, 2007.

Osteoporos Int (2015) 26:35–47
DOI 10.1007/s00198-014-2805-3

ORIGINAL ARTICLE

Preservation of volumetric bone density and geometry in trans women during cross-sex hormonal therapy: a prospective observational study

E. Van Caenegem · K. Wierckx · Y. Taes · T. Schreiner · S. Vandewalle · K. Toye · J.-M. Kaufman · G. T'Sjoen

Received: 27 December 2013 / Accepted: 1 July 2014 / Published online: 7 November 2014

Abstract
Summary Although trans women before the start of hormonal therapy have a less bone and muscle mass compared with control men, their bone mass and geometry are preserved during the first 2 years of hormonal therapy, despite of substantial muscle loss, illustrating the major role of estrogen in the male skeleton.
Purpose The aim of this study is to examine the evolution of areal and volumetric bone density, geometry, and turnover in trans women undergoing sex steroid changes, during the first 2 years of hormonal therapy.
Methods In a prospective observational study, we examined 49 trans women (male-to-female) before and after 1 and 2 years of cross-sex hormonal therapy (CSH) in comparison with 49 age-matched control men measuring grip strength (hand dynamometer), areal bone mineral density (aBMD), and total body fat and lean mass using dual X-ray absorptiometry (DXA), bone geometry and volumetric bone mineral density, regional fat, and muscle area at the forearm and calf using peripheral quantitative computed tomography. Standardized treatment regimens were used with oral estradiol valerate, 4 mg daily (or transdermal 17-β estradiol 100 μg/24 h for patients >45 years old), both combined with oral cyproterone acetate 50 mg daily.
Results Prior to CSH, trans women had lower aBMD at all measured sites (all $p<0.001$), smaller cortical bone size (all $p<0.05$), and lower muscle mass and strength and lean body mass (all $p<0.05$) compared with control men. During CSH, muscle mass and strength decreased and all measures of fat mass increased (all $p<0.001$). The aBMD increased at the femoral neck, radius, lumbar spine, and total body; cortical and trabecular bone remained stable and bone turnover markers decreased (all $p<0.05$).
Conclusions Although trans women, before CSH, have a lower aBMD and cortical bone size compared with control men, their skeletal status is well preserved during CSH treatment, despite of substantial muscle loss.

Keywords Bone transsexual gender dyspohria sex steroids prospective

Introduction

Sex steroids determine bone geometry during puberty. Men develop larger bones than women with a greater periosteal (outer) and endosteal (inner) circumference of the cortex, resulting in the sexual dimorphism of bone. This difference is mainly due to periosteal apposition during puberty on which testosterone has a positive influence [1]. Mechanical loading, through muscle mass and physical activity, has been shown to be an important factor in the acquisition of bone geometry in adulthood [2, 3]. Estrogens could increase the sensitivity of bone for mechanical stimuli, and mechanical stimuli are in their turn modulated by androgens, depending on GH-IGF1 action. However, there is conflicting data on this interaction of estrogens with the periosteal surface with evidence for both stimulatory and inhibitory effects [4, 5].

E. Van Caenegem (✉) · K. Wierckx · Y. Taes · S. Vandewalle · K. Toye · J.-M. Kaufman · G. T'Sjoen
Department of Endocrinology, Ghent University Hospital, De Pintelaan 185, 9000 Ghent, Belgium
e-mail: eva.vancaenegem@ugent.be

E. Van Caenegem · K. Wierckx · T. Schreiner · G. T'Sjoen
European Network for the Investigation of Gender Incongruence (ENIGI), Ghent, Belgium

T. Schreiner
Department of Endocrinology, Rikshospitalet, Oslo University Hospital, Oslo, Norway

G. T'Sjoen
Center for Sexology and Gender problems, Ghent University Hospital, Ghent, Belgium

Trans women (male-to-female trans persons) undergo hormonal treatment reducing the endogenous testosterone and inducing higher estrogen levels, which is referred to as cross-sex hormonal therapy (CSH). This way, masculine characteristics, e.g. hair growth, are suppressed and feminization is induced. In addition, sex reassignment surgery can be performed, which includes vaginoplasty and orchiectomy. The hormonal treatment reduces muscle mass and increases fat mass [6]. These changes during CSH in trans women are a model that may help to understand the differential effects of mechanical loading, fat and muscle mass, and sex steroid milieu in the bone. In an earlier cross-sectional study of our group, bone geometry was assessed in trans women after median 8 years of CSH and sex reassignment surgery (SRS). We found a lower areal and volumetric bone mineral density (aBMD and vBMD, respectively) and smaller bone size related to lower muscle mass and strength compared with male controls [7, 8]. Moreover, in a second study in trans women, at the age of peak bone mass and prior to any kind of hormonal treatment, we observed a lower bone mineral density and a smaller cortical bone area and thickness in relation to lower muscle mass compared with age-matched control males [9]. Prospective studies on the effects of CSH and CSH-induced muscle loss in bone geometry are still lacking. Other research using classical dual X-ray absorptiometry (DXA) found a maintained aBMD after 2 or more years of CSH in trans women [10–18]. In the latter studies, a variety of treatment regimens was used, some of which are no longer applied for safety reasons [19], e.g., ethinyl estradiol [11, 14, 18]. Also, treatment regimens with or without anti-androgens were described which hampers clear comparisons (cyproterone acetate in [14, 18] or GnRH-analogues in [10, 12, 13]).

In this prospective observational study, we examine bone mass using DXA and bone geometry using peripheral quantitative CT-scan (pQCT) in trans women before the start of CSH and after 1 and 2 years of CSH in relation to body fat, lean mass, and physical activity. This study is an extension on former research [9]: we doubled the group of trans women and followed all persons prospectively. Cross-sex hormonal therapy was initiated with a standardized treatment protocol with anti-androgens and estrogens.

Materials and methods

Study design and population

All trans women were diagnosed with gender dysphoria (DSM-5, 302.85; ICD-10, F64.0) and were recruited from the center for sexology and gender problems at the Ghent University Hospital, Belgium. All were treated following the World Professional Association for Transgender Health standards of care [20]. This research is part of the 'European Network for the Investigation of Gender Incongruence' (ENIGI), a collaboration of four major West European gender identity clinics (Amsterdam, Ghent, Hamburg, and Oslo), a study group created to obtain more transparency in diagnostics and treatment of gender dysphoria [21].

Between February 2010 and August 2012, all patients diagnosed with gender dysphoria and referred to our departments were invited to participate in this prospective study (trans women; n=87). After screening by thorough medical history and determination of serum sex steroids, 37 persons were excluded resulting in a total population of 50 trans women. Reasons for exclusion were previous hormonal therapy (n=22), unwilling (n=10), medical (gastric bypass; n=1), and other (n=4). One patient had subclinical hypergonadotrophic hypogonadism due to a primary testicular problem and was excluded from further analyses. A final number of 49 trans women who had never used any kind of cross-sex hormonal treatment nor anti-androgen therapy, and thus before SRS, was included. All participants were Caucasian. A male control population was used, matched for age (±2 years, median=1 year). These were healthy men recruited from communities around Ghent or who responded on posters spread at the Ghent University Hospital and on its website and in schools.

All participants were in good physical health and completed questionnaires about previous illness and medication use, current and past smoking habits, and physical activity by recording the weekly frequency of sports, recreational, and/or working activities (using Baecke's questionnaire [22]). At the start of the study, three trans women used a 5-alpha reductase inhibitor to reduce hair loss (finasteride 5 mg, n=2) and for benign prostate hypertrophy (dutasteride 0.5 mg, n=1). All three had serum testosterone within the normal ranges and only the latter had slightly elevated FSH and LH (11 and 16 U/l, respectively).

The trans women were evaluated after 1 and 2 years. In the meantime, structured clinical visits were conducted at 3, 6, 9, and 18 months. After 1-year follow-up, three trans women dropped out: two were lost to follow-up and one decided to stop hormones. The use of 5-alpha reductase inhibitors was ceased at baseline visit. One person was not compliant, with testosterone levels remaining within the normal male range, and was excluded for the 1- and 2-year evaluation. One trans woman skipped the bone examinations at 1 year (not the clinical evaluation), but did manage to do the visit at 2 years. This data is part of a large prospective study (ENIGI); at time of this analysis, all 49 subjects had completed 1-year follow-up and 29 had completed 2-year follow-up.

Standardized treatment regimens were started after the baseline visit with oral estradiol valerate, 4 mg daily (n=34) (Prognova ®, Bayer, Germany) or transdermal 17-β estradiol 100 μg/24 h for patients older than 45 years old (n=15) (Dermestril ®, Besins, Belgium), both combined with oral cyproterone acetate 50 mg daily (Androcur ®, Bayer, Germany). Transdermal estrogens were used in older trans women as this would have a lower thromboembolic risk [23]. When

psychologically indicated, trans women started with cyproterone acetate 50 mg alone, without estrogens, and estrogens were then associated after a median of 25 weeks (IQR 16–31). The latter group will be referred to as the "IAAM"-group (initial anti-androgens monotherapy) (n=18), whereas trans women who received combined anti-androgens with estrogens from the start are referred to as the "AA + E"-group (anti-androgens and estrogens combined) (n=31). Vitamin D supplements were used by a single trans woman at the baseline, but not by the control persons. None of the subjects used calcium supplements or bone-active drugs like bisphosphonates or SERMs at the baseline. The study protocol was approved by the ethics review board of the Ghent University Hospital, registered with clinicaltrials.gov (identifier: NCT01072825) and all participants gave written informed consent.

Body composition, muscle strength, and areal bone mineral density

Body weight and anthropometrics were measured in light indoor clothing without shoes. Standing height was measured using a wall-mounted Harpenden stadiometer (Holtain, Ltd., Crymuch, UK).

Grip strength at the dominant hand was measured using an adjustable hand-held standard grip device (JAMAR hand dynamometer, Sammons and Preston, Bolingbrook, IL, USA). The maximum strength of three attempts was assumed to best reflect the current status and history of their musculoskeletal adaptation and was expressed in kilograms (kg).

Body fat and lean mass, bone mineral content (BMC), bone area, and areal bone mineral density (aBMD) at the whole body, lumbar spine, non-dominant forearm, and left proximal femur (total hip and femoral neck region) were measured using dual X-ray absorptiometry (DXA) with a Hologic Discovery device (Software Version 11.2.1, Hologic, Inc., Bedford, MA, USA). The coefficient of variation for both spine and whole-body calibration phantoms was less than 1 %, as calculated from daily and weekly measurements, respectively.

Volumetric bone parameters and cross-sectional muscle and fat area

A pQCT device (XCT-2000, Stratec Medizintechnik, Pforzheim, Germany) was used to evaluate the cortical volumetric bone parameters at the dominant midradius and tibia (at 66 % of bone length) and trabecular bone parameters at the metaphysis (at 4 % of bone length) of the dominant radius. Over 90 % of the scans were performed by a single operator. Procedure details were as described previously [5]. Scans with large movement artifacts (n=2) and suspected position error (>10 % variation of total bone area at radius 4 %) (n=1) were excluded.

Biochemical determinations

Venous blood samples were obtained between 08.00 and 10.00 h after overnight fasting. All samples were stored at −80 °C until analysis.

Testosterone (T), estradiol (E2), estrone (E1), and androstenedione were determined using tandem mass spectrometry on an AB Sciex 5500 triple-quadrupole mass spectrometer (AB Sciex, Toronto Canada). Serum limit of quantification was 0.3 pg/ml for E2 and 0.5 pg/ml for E1, and the inter-assay CV was 4 % at 21 pg/ml for E2 and 7.6 % at 25 pg/ml for E1. Serum limit of quantification was 1 ng/dl (35 pmol/l) for T, and the inter-assay CV was 6.5 % at 3 ng/dl [24]. Commercial immunoassays kits were used to determine serum concentrations of sex hormone-binding globulin (SHBG), luteinizing hormone (LH), follicule stimulating hormone (FSH), 25-hydroxyvitamin D (25(OH)D) (Modular, Roche Diagnostics, Mannheim, Germany), and dehydroepiandrosterone sulfate (DHEAS), markers of bone turnover viz. C-terminal telopeptides of type I collagen (CTX) (bone resorption), and procollagen 1 aminoterminal propeptide (P1NP) and osteocalcin (OC) (bone formation) (Cobas 411, Roche Diagnostics, Mannheim, Germany). Radio-immunassays were used for leptin (Bio-connect diagnostics, Huissen, the Netherlands) and insulin-like growth factor 1 (IGF1) (Cisbio bioassays, Codolet, France). The intra- and inter-assay coefficients of variation for all assays were less than 10 %.

Statistical analysis

Descriptives are expressed as mean and standard deviation or median (first to third quartile), when criteria for normal distribution were not fulfilled. P values <0.05 were considered to indicate statistical significance; all tests were two-tailed. Comparison of general, anthropometric, biochemical, and hormonal determinations, bone and body composition before cross-sex hormonal therapy, between trans women and age-matched controls and between trans women of the AA + E and the IAAM group was performed using an independent sample t tests (or Mann–Whitney U test, when criteria for normal distribution were not fulfilled). Repeated measurements in trans women before and after 1 and 2 years of treatment were analyzed using mixed models to allow inclusion of all data from each subject despite missing values. Dependents in these models were the bone and body composition parameters. Study visit (defined as a categorical variable) was the main independent variable. Other independent variables, fixed factors (hormonal treatment protocol AA + E or IAAM) or covariates (e.g., age, weekly sports activity, muscle CSA, T, E2, PTH, 25(OH)D), were added to the model in specific questions. The p value for the fixed effect of the independent variable "study visit" is given in Tables 3 and 4. The p value for the fixed effect of the other independent variables,

covariates, is given in the results section "Influence of covariates on bone evolution during CSH".

Results

Before hormonal therapy: Comparison with age-matched controls

General characteristics, hormonal and biochemical determinations

Trans women at inclusion were median 30 years old with a wide range (minimum 17 and maximum 67) (Table 1). Mean body weight, height, and BMI were similar in trans women and control men. There were 22 % active smokers in this group compared with 16 % in controls (n.s.) and the amount of pack years was similar.

The T, E2, E1, LH, and SHBG were comparable in both groups. We observed a slightly higher FSH in trans women versus control men, which remained borderline significant after exclusion of 5-alpha reductase users ($n=3$). All participants had T levels within normal male range and all were clinically euthyroid.

Bone turnover markers, 25(OH) vitamin D status and leptin

Trans women had a significantly lower serum 25(OH)D and higher PTH than control men, which remained significant after adjustment for fat mass and month of visit (Table 1). Vitamin D insufficiency (defined as 25(OH)D<20 ng/ml) was prevalent in 67 % of trans women versus 35 % control males (chi-square test $p=0.001$). Hyperparathyroidism (PTH> 65 ng/l) was found in four vitamin D-deficient trans women versus none in the control group (chi-square test $p=0.033$).

Areal bone mineral density using DXA

Trans women had a significantly lower aBMD at the lumbar spine, hip, femoral neck, and radial forearm compared with the control men (Table 2). The prevalence of osteoporosis based on male reference ranges (following 'classical' WHO-criteria as originally proposed for postmenopausal women: defined as a T-score ≤2.5 SD) was 18 % at the lumbar spine and 11 % based on female reference ranges (as suggested to define osteoporosis in men [25]) in trans women versus 4 and 2 %, respectively, in the male control group (chi-square test, $p=0.025$).

Volumetric bone parameters at the upper and lower limb using pQCT

At both the radius and the tibia, the periosteal, outer cortical, circumference was smaller and the cortical bone area and thickness ($p=0.002$) were smaller compared with control men (Table 3, Fig. 1). Trans women had a lower trabecular vBMD at the radius ($p=0.013$), and a markedly lower polar strength strain index (SSI, as a measure of bone strength) ($p=0.017$) was observed at both cortical sites and controls.

Body composition and physical activity

Trans women presented with a significantly lower body lean mass (−4 %) and a lower grip strength and muscle mass, reflected by the muscle cross-sectional area (CSA) at the forearm compared with control men (Table 4). We also observed a tendency towards higher measures of fat mass despite the lower mean total body weight (n.s.) in trans women. The weekly sports activity was significantly lower in trans women compared with control men.

During hormonal therapy: Follow-up after 1 and 2 years of CSH

Follow-up and changes in sex steroids and hormonal levels

Serum T decreased and E2 increased significantly in trans women and were both within the normal female ranges at the 1- and 2-year time points of CSH (Table 1). A higher serum E2 (114 vs. 57 pg/ml) and lower E1 (107 vs. 344 pg/ml) were measured in transdermal estrogen users versus oral estrogen users at year one and two of CSH. Furthermore, there were no significant differences between the transdermal and oral estrogen users (after adjustment for age). SHBG was significantly increased at the 1-year time point and further increased at the 2-year treatment time point. Precursor androgens, androstenedione, and DHEAS decreased as well during treatment. Gonadotropines were significantly decreased at 1 year and remained suppressed after 2 years of treatment in those who did not undergo SRS during year two. In trans women who underwent SRS ($n=11$) and quit cyproterone acetate treatment afterwards, gonadotropines were higher and T lower in year two compared with trans women who did not undergo SRS yet ($n=17$) (Mann–Whitney U test $p_{FSH,\ LH}<0.001$ and $p_T=0.029$). Leptin was increased at 1 year and remained like this after 2 years. Body fat mass was a significant positive predictor of serum leptin levels at both time points (mixed model, dependent serum leptin and independent fixed factors, visit and body fat mass; body fat mass $p<0.001$). After 1 year, two active smokers quit their tobacco use and four ex-smokers took up smoking again during the first ($n=3$) and second ($n=1$) year.

Table 1 General characteristics, hormones, and bone turnover markers in control men and trans women before and after 1 and 2 years of CSH

	Control men (n=49)	Trans women		
		before CSH (n=49)	1 year CSH (n=44)	2 years CSH (n=29)
Age (years)	33±12	33±12	–	–
Weight (kg)	78.3±10.8	74.7±14.3*	76.5±14.2	80.8±15.9
Height (cm)	178.9±5.6	178.4±5.9	–	–
Pack year	0 (0–8)	19.11	–	–
Alcohol (drinks/week)	10 (3–16)***	2 (0–7)	1 (0–4)	3 (0–5)
Fracture prevalence(%)[a]	16	22	–	–
Sport index[b]	3.1±1.1*	2.6±1.1	2.5±1.2	2.5±1.1
Leisure time index[b]	2.9±0.7	3.1±0.6	3.0±0.7	3.1±0.7
Work index[b]	2.6±0.7	2.7±0.8	2.5±0.8	2.6±0.9
Total physical activity[b]	8.7±1.5	8.3±1.6	8.0±1.6	8.1±1.9
Testosterone (ng/dl)	515 (445–616)	546 (451–643)***	13 (10–15)	13 (10–24)
Estradiol (pg/ml)	21.0 (15.8–25.1)	20.8 (17.2–28.3)***	61.1 (46.5–85.8)	52.4 (36.2–75.1)
Estrone (pg/ml)	33.7 (27.9–44.5)	35.0 (26.5–44.3)***	262.1 (72–379.6)	193.2 (67.9–343.2)
LH (U/l)[c]	4 (3–6)	5 (4–7)***	<0.1 (<0.1–<0.1)	1 (<0.1–12)
FSH (U/l)[d]	3 (2-5)*	4 (3-6)***	<0.1 (<0.1–<0.1)	2 (<0.1–11)
SHBG (nmol/l)	35.6±(30.5–42.3)	38.5±(26.1–48.6)***	46.8±22.2	55.1±23.4
Androstenedione (ng/dl)	89 (76–115)	108 (93–125)***	61 (53–81)	61 (47–88)
DHEAS (μg/dl)	353±144	333±138**	274±125	309±180
TSH (mU/l)	2.3±1.1	2.4±1.3	2.4±1.3	2.3±1.1
Leptin (ng/ml)	4.2±3.2	5.9±4.2***	13.1±10.6	13.0±7.9
25(OH)D (ng/ml)	23±7***	16±8**	19±10	21±8
PTH (pg/ml)	36±12***	44±15*	39±14	40±1
P1NP (μg/l)	64.1±31.1	61.3±34.1	56.6±31.2	47.8±26.7
Osteocalcin (ng/ml)	23.1±7.0	25.7±10.3**	23.9±10.8	20.7±9.1
CTX (ng/ml)	0.52±0.26	0.49±0.25***	0.45±0.22	0.32±0.18
IGF1 (ng/ml)	251±90	240±78	275±74	238±66***

Descriptives are expressed as mean±SD or as median (1st–3rd quartile) when not normally distributed

Variables were compared between control men and trans women before CSH using independent t tests or Mann–Whitney U tests when not normally distributed. Repeated measures in trans women were tested using mixed models, the p value for the fixed effect of study visit (considering both years) is given

Conversions of conventional units to SI units: testosterone nmol/l; multiply ng/dl by 0.0347, estradiol nmol/l; multiply pg/ml by 3.671, estrone nmol/l; multiply pg/ml by 3.699, androstenedione nmol/l; multiply ng/dl by 0.0349, DHEAS nmol/l; multiply μg/dl by 27.14, leptin nmol/l; multiply ng/ml by 0.0625. 25(OH) vitamin D nmol/l; multiply ng/ml by 2.496, PTH ng/l; multiply pg/ml by 1, P1NP nmol/l; multiply μg/l by 0.0286, CTX ng/l; multiply ng/ml by 1000, IGF1 μg/l; multiply ng/ml by 1

[a] Using chi-square test

[b] Measured by the Baecke questionnaire on physical activity [22]

[c] Detection limit LH-assay=0.1 U/L

[d] Detection limit FSH-assay=0.1 U/L

Repeated measures in trans women were tested using mixed models, the p value for the fixed effect of study visit is given (*)

***$p \leq 0.001$

**$0.001 < p \leq 0.01$

*$0.01 < p \leq 0.05$

Bone turnover markers, 25-hydroxyvitamin D status and IGF1

CTX, P1NP, and OC decreased after the first and second year (mean decrease at year two was P1NP −18 %, OC −12 %, and CTX −25 %) (Table 1). Parathyroid hormone decreased after 1 year and 25(OH)D status ameliorated significantly between the 1- and 2-year visit. Vitamin D supplements were used by 11 and 5 participants at year one and two, respectively, which

Table 2 Areal bone parameters in control men and trans women before and after 1 and 2 years of CSH

	Control men	Trans women		
	(n=49)	before CSH (n=49)	1 year CSH (n=44)	2 years CSH (n=29)
Lumbar Spine				
Bone Area (cm^2)	70±7	67±6	68±7	68±7
BMC (g)	73±11***	64±13***	66.7±14.2	67±14
aBMD (g/cm^2)	1.040±0.113**	0.952±0.150***	0.983±0.156	0.982±0.136
Total Body				
Bone Area (cm^2)	2,340±150*	2,258±181***	2,282±183	2,315±200
BMC (g)	2,823±375***	2,466±362***	2,500±384	2,562±399
aBMD (g/cm^2)	1.205±0.102***	1.088±0.086**	1.090±0.090	1.100±0.090
Femoral Neck				
Bone Area (cm^2)	6±0.3	6±0.4	6±0.4	6±0.4
BMC (g)	5±1***	5±1*	5±1	5±0.9
aBMD (g/cm^2)	0.928±0.154***	0.795±0.119***	0.807±0.129	0.807±0.123
Total Hip				
Bone Area (cm^2)	44±4	43±4	43±4	43±4
BMC (g)	48±8***	41±8	41±8	42±9
aBMD (g/cm^2)	1.089±0.153***	0.947±0.134	0.952±0.141	0.955±0.145
Radius				
Bone Area (cm^2)	29±3***	17±3.9	17±1.7	17±1.8
BMC (g)	19±2***	11±2.3***	10±1.7	10±1.8
aBMD (g/cm^2)	0.639±0.052*	0.614±0.058**	0.621±0.057	0.622±0.062

Descriptives are expressed as mean±SD or as median (first–third quartile) when not normally distributed

Variables were compared between control men and trans women before CSH using independent t tests or Mann–Whitney U tests when not normally distributed. Repeated measures in trans women were tested using mixed models, the p value for the fixed effect of study visit (considering both years) is given

***$p \leq 0.001$

**$0.001 < p \leq 0.01$

*$0.01 < p \leq 0.05$

Table 3 Volumetric bone parameters at the tibia in control men and trans women before and after 1 and 2 years of CSH

	Control men (n=49)	Trans women		
		before CSH (n=49)	1 year CSH (n=44)	2 years CSH (n=29)
Tibia				
Cortical vBMD (mg/cm^3)	1,112±32	1,111±24	1,113±24	1,116±28
Cortical bone area (mm^2)	364±49***	330±46	329±48	321±48
Cortical thickness (mm)	4.4±0.7**	4.1±0.5	4.1±0.5	4.0±0.5
Periostal circumference (mm)	97±8*	94±7	93±6	94±8
Endosteal circumference (mm)	69±11	68±7	68±7	69±8
Polar SSI (mm^3)	3,033±495***	2,671±536	2,624±513	2,595±546

Descriptives are expressed as mean±SD or as median (first–third quartile) when not normally distributed

Variables were compared between control men and trans women before CSH using independent t tests or Mann–Whitney U tests when not normally distributed. Repeated measures in trans women were tested using mixed models, the p value for the fixed effect of study visit (considering both years) is given

***$p \leq 0.001$

**$0.001 < p \leq 0.01$

*$0.01 < p \leq 0.05$

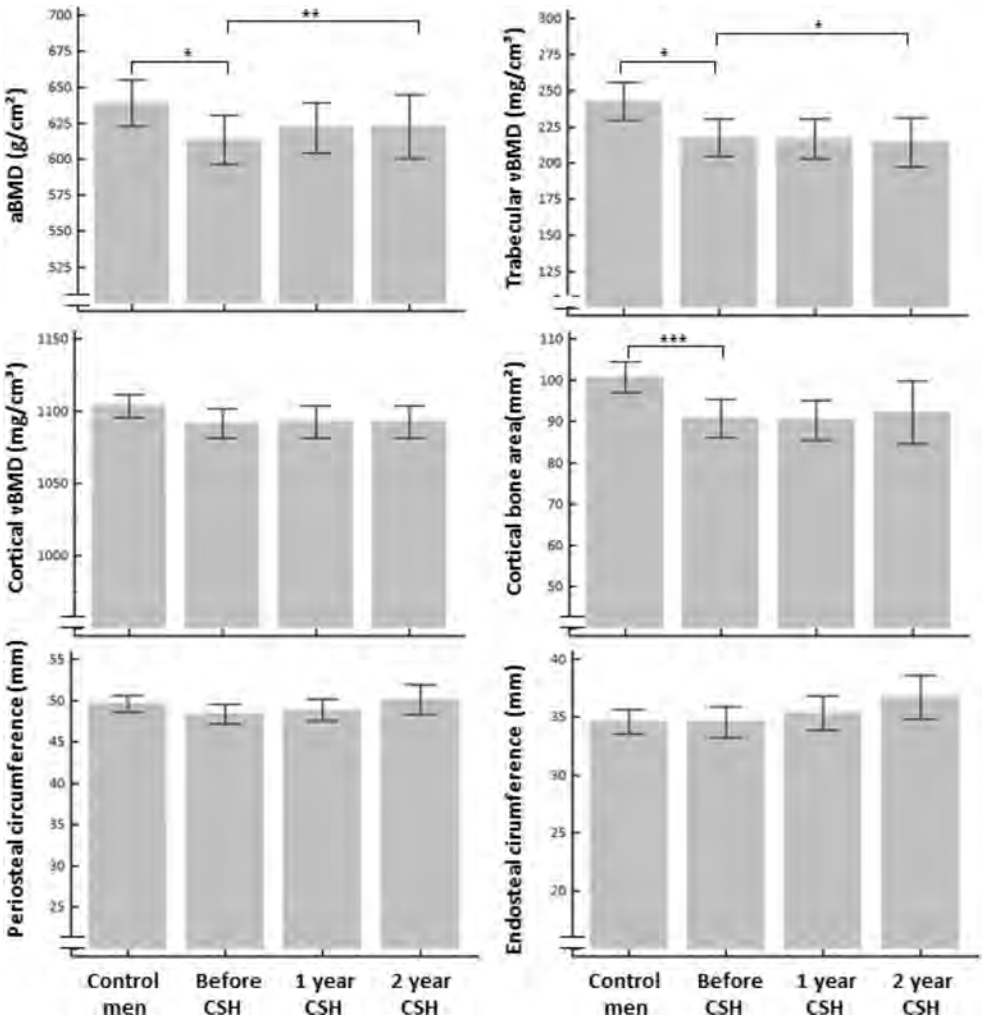


Fig. 1 Areal and volumetric bone parameters at the radius in trans women and control men

were combined with calcium in respectively four and three trans women. Bisphosphonates or other bone-active drugs were not used during the 2 years of CSH. IGF1 increased significantly after 1 year, but decreased again after the first year. The type of estrogen (oral vs. transdermal) did not have an impact on the evolution of IGF1 over time with and without adjustment for age (data not shown, p=n.s.).

Areal bone mineral density using DXA

We observed an increase in aBMD at the femoral neck, lumbar spine, radius, and whole body after 1 and 2 years of hormonal therapy (+1.8, +3.2, +1.1, and +0.8 % at year one, respectively) (Table 2). At the total hip, no changes were observed. In a sensitivity analysis, we corrected for fat mass by normalizing the measured aBMD at each site for the corresponding regional fat (aBMD divided by the measured fat mass by DXA in that region). This showed similar results (data not shown).

Volumetric bone parameters at the upper and lower limb using pQCT

No changes in cortical bone parameters over time were noted (Table 3, Fig. 1). Also, for the ratio of endosteal over periosteal circumference, no differences were observed at the 1- or 2-year time points (data not shown). Trabecular vBMD appeared to decrease over time (Fig. 1), although trabecular area remained stable.

Body composition and physical activity

All measures of fat mass, including subcutaneous fat mass, increased during the first and second year of cross-sex hormonal treatment (+25 % fat body mass in year one) (Table 4). Trans women lost lean body mass (−4 % in year one), grip strength and muscle CSA of the forearm and calf mainly during the first year. Hip circumference augmented significantly during the first year and consequently WHR decreased. Waist circumference remained stable. Weekly total physical activity including sports,

Table 4 Body composition in control men, trans women before and after 1 and 2 years of CSH

	Control men (n=49)	Trans women		
		before CSH (n=49)	1 year CSH (n=44)	2 years CSH (n=29)
Measures of fat mass				
Body Mass Index (kg/m^2)	24.5±3.4	23.5±4.2***	24±4.3	25.0±4.3
Waist circumference (cm)	86±9	85±12	83±11	85±11
Hip circumference (cm)	98±6	97±8***	100±8	101±8
Waist-hip ratio	0.9±0.1	0.9±0.1***	0.8±0.1	0.8±0.1
Fat body mass (kg)[a]	14.1±5.7	14.9±6.6***	18.9±6.9	21.1±7.4
Fat percentage (%)[a]	17.6±5.4	19.1±5.6***	24.2±5.0	25.6±4.9
Forearm fat CSA (mm^2)[b]	769±426	795±462***	1,245±496	1,382±490
Calf fat CSA (mm^2) [b]	1,436±503	1,687±883***	2,228±766	2,391±762
Measures of muscle mass				
Lean body mass (kg) [a]	61.3±6.8*	57.4±8.7***	55.1±8.7	57.1±9.8
Lean percentage (%)[a]	78.5±5.1	77.5±5.3***	72.5±4.6	71.7±4.1
Forearm muscle CSA (mm^2) [b]	4,512±579***	3,999±746***	3,664±783	3,825±867
Calf muscle CSA (mm^2) [b]	8,233±1,498	7,742±1,361**	7,623±1,479	7,448±1,390
Grip strength (N/kg)	49±6***	42±9***	39±9	38±10

CSA cross-sectional area. [a] Measured with DXA. [b] measured with pQCT

Descriptives are expressed as mean±SD or as median (first–third quartile) when not normally distributed

Variables were compared between control men and trans women before CSH using independent t tests or Mann–Whitney U tests when not normally distributed. Repeated measures in trans women were tested using mixed models, the p value for the fixed effect of study visit (considering both years) is given

***p≤0.001

**0.001<p≤0.01

*0.01<p≤0.05

physical activity during leisure time and at work, did not change significantly over time.

Influence of covariates on bone evolution during CSH

Adjusting the mixed models for age, BMI, body fat mass, or leptin did not alter any of the observed changes in bone or body composition over 2 years of CSH. E2 and T (at year one) were not associated with the evolutions in bone or body composition.

Including the weekly sports activity or muscle CSA at the forearm or calf after 1 year of therapy in the model did not alter any of the observed changes over time in bone or body composition. Weekly sports activity at year one was however positively associated with trabecular vBMD, cortical bone area, bone size (periosteal and endosteal circumference), and polar SSI at the radius and tibia and nearly all areal bone parameters at all measured sites and the body lean mass, grip strength, and muscle area at the forearm and calf (all p<0.05). The latter muscle mass parameters decreased less in time with a higher weekly sports activity. No interactions of sports activity with year of visit were noted. Positive relationships were also found between muscle CSA at the forearm and calf at 1 year of therapy and areal and volumetric bone parameters (all p≤0.004).

Adjusting for the 25(OH)D status and PTH (at year one) did not alter the observed effects in bone over time. All bone turnover markers at year one were negatively associated with aBMD at the spine, whole body, and total hip (all p≤0.026) as well as cortical bone area, cortical thickness, and SSI (radius and tibia, all p≤0.038).

Fat body mass after the first year was significantly associated with BMC and aBMD at the femoral neck, total hip and radius, whole body BMC (all p<0.05), but adjustment for fat mass did not change the observed evolution in any of the bone parameters.

The effect of initial anti-androgens in monotherapy for a short period (IAAM) versus anti-androgens plus estrogens from start (AA + E)

The group who had anti-androgens in monotherapy for a median of 25 weeks before the association of estrogens (IAAM, n=18) were a median of 10 years older than those without (AA + E, n=31) (median age 27 vs. 37 years, p=n.s.). Body weight, height, BMI, areal and volumetric bone parameters, and body composition were similar in both groups at baseline (data not shown). Total hip aBMD, P1NP, CTX, IGF1, and 25(OH)D were higher in the younger subjects of

AA + E-group (all $p<0.05$), which remained significant after adjustment for age, pack years, serum LH, and T.

In the first year of CSH, bone turnover markers increased in the IAAM-group, whereas a decrease was observed in the AA + E-group (significant interaction of type of treatment protocol and P1NP, $p=0.012$) (Fig. 2). Body fat mass ($p=0.040$), calf fat CSA ($p=0.006$), SHBG ($p=0.017$), and leptin ($p=0.035$) had a greater increase in the AA + E-group (Fig. 2). Apart from the latter differences, adjustment for treatment protocol did not influence the described changes in body composition and bone parameters under cross-sex hormonal therapy (data not shown).

Discussion

To our knowledge, this is the first prospective study on bone geometry in trans women, using pQCT, before and after 1 and 2 years of CSH. Studies in trans women provide a unique opportunity to examine the effects of sex steroids independent of sex chromosome determinants. We found that trans women, before any kind of hormonal therapy and sex reassignment surgery, already have a lower aBMD, a smaller bone size, and a lower muscle mass compared with age-matched control men. During the first 2 years of CSH, the bone turnover markers decreased and aBMD increased significantly. Trabecular and cortical bone parameters and bone size remained mainly stable during CSH in trans women.

The lower aBMD and lean body mass and higher fat mass in trans women before the start of any kind of therapy is in line with in a Norwegian study [11] and confirms our earlier report on trans women before CSH, which consisted of a smaller subgroup (only trans women at the age of peak bone mass) of the currently included trans women [9]. The lower weekly sports activity could have contributed to the lower muscle mass and strength and to a lower peak bone mass of trans women and both are important factors in building stronger bones during childhood [26] and adolescence [3, 27]. In addition, the lower 25(OH)D status and higher PTH could have contributed to the lower aBMD in trans women versus control males prior to CSH [28, 29]. The negative influence of smoking in bone [30] is not likely to explain the differences in bone between trans women and control men as the prevalence of active smokers and pack years is similar in both groups. Earlier research from our group showed lower aBMD, trabecular vBMD, and smaller bone size in relation to lower muscle mass and strength after a median of 8 years of CSH and SRS compared with male controls [7, 8]. One could hypothesize that the observed smaller bone geometry and lower aBMD and vBMD versus control men might have been present before the start of CSH and SRS due to differences in lifestyle.

Our prospective results are in line with the previous research using classical bone densitometry, which also found an maintained aBMD after 2 or more years of CSH [11, 13, 14, 18] or increase in aBMD at the lumbar spine after a minimum of 2 years of CSH [10, 12, 15–17]. In the latter studies, higher dosages of estrogens were used than the currently used protocol: 17-beta estradiol valerate 6 mg daily orally [12], ethinyl estradiol 35–100 μg daily orally [16], 10 mg estradiol valerate IM/10 days [13], or combinations [17] with [12, 13, 16] or without anti-androgens or gonadectomy [15].

A first explanation for the stable bone geometry and increased aBMD at the lumbar spine, femoral neck, radius, and whole body can be found in estrogen therapy, as estrogens are known to slow down bone resorption through direct and indirect effects on osteoclast formation, activity, and lifespan and inhibition of osteocyte apoptosis and maintain bone formation through direct effects on osteoblasts [31]. Lower bone resorption was indeed observed after 1 and 2 years of CSH in our group of trans women and in other cohorts [11, 14, 17, 18]. In prostate cancer patients with or without androgen deprivation therapy (ADT), estrogen (but not testosterone) was inversely associated with bone turnover markers [32] and estrogen therapy seemed to protect bones and reduce bone turnover [33, 34]. Furthermore, treatment with selective estrogen-modulator toremifene in prostate cancer patients on ADT reduced fracture incidence and bone turnover rates [35]. We also observed a decreased P1NP after 2 years of CSH, which was also previously observed in trans women using CSH [7, 11, 18]. The decreased bone turnover markers in trans women after 2 years is in agreement with a short-term experiment with sex steroid suppression in adult men followed by selective replacement of either estrogen, testosterone, or both [36]. Increased bone resorption in the absence of the two hormones and no changes in men receiving both hormones was observed [36], and this is also supported by the increased CTX and P1NP after 1 year in the IAAM-group of trans women who underwent a short period of hypogonadism without estrogen replacement. When replacing separately both hormones in the trial, the reduction in bone resorption was much larger with estrogen than with testosterone. Therefore, it was concluded that it is primarily estrogen rather than testosterone that regulates the process of bone resorption [36], which is in line with our results. We hypothesize that estrogen therapy induced a lower bone turnover leading to an increase in aBMD in trans women (by estrogen-mediated filling of the remodeling space), as supported by the observed inverse association of the bone turnover markers at year one with the changes in aBMD during CSH.

Secondly, we observe a preserved bone mass and bone geometry from baseline, despite the hypoandrogenic status and drastic loss of muscle mass during treatment. Weekly sports activity in spite of the net muscle loss still had a clear positive influence on areal and volumetric bone parameters and bone size and even on the increase in aBMD after 1 year at the total body.

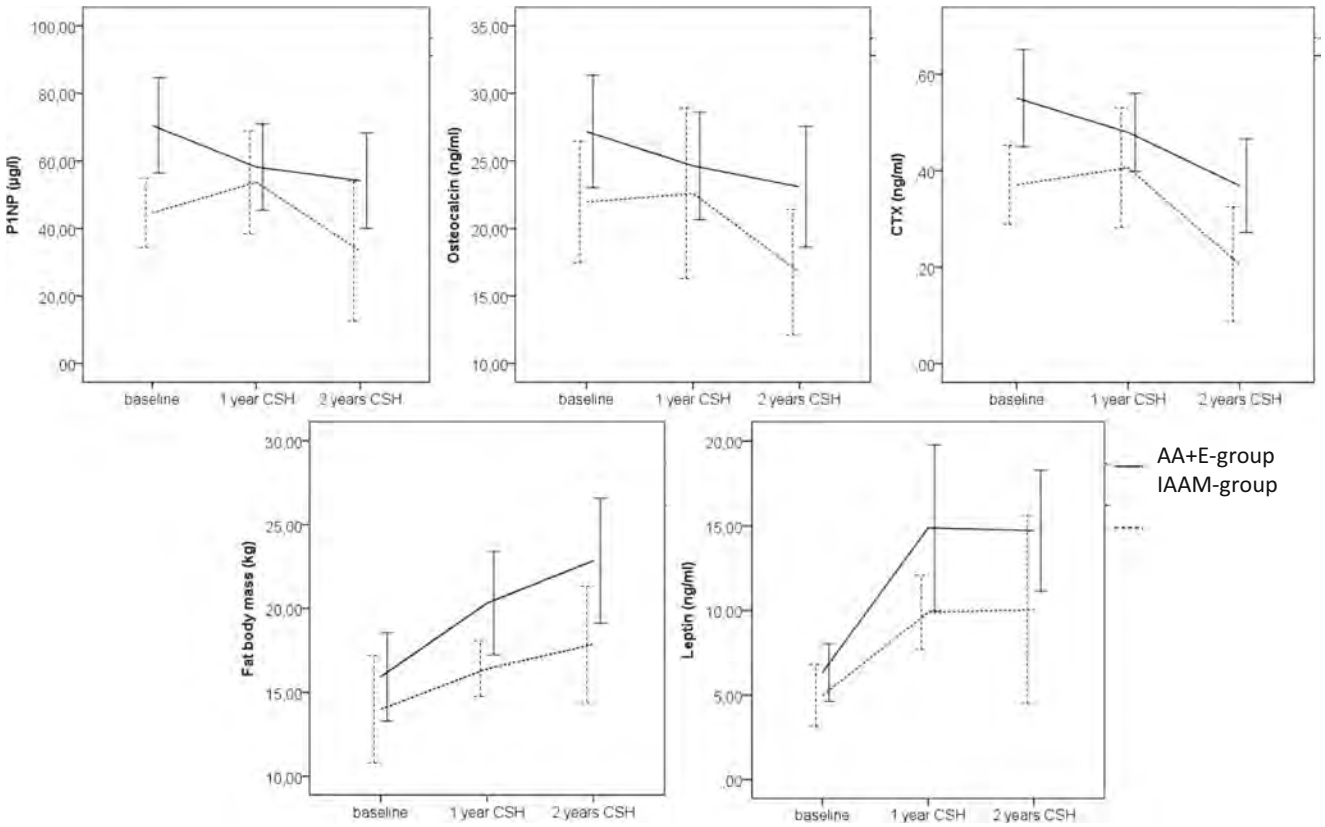


Fig. 2 Evolution of bone turnover markers, leptin, and total body fat mass in the AA + E-group and the IAAM-group after 1 and 2 years of CSH

Furthermore, the increased 25(OH)D and decreased in PTH is mostly likely due to counseling during treatment and prescription of vitamin D supplements. Recently, vitamin D supplementation in adults has been found to have a weighted mean difference of −0.3 to 0.8 % change in aBMD [37], which is lower than the observed increases in aBMD in our study. No associations with 25(OH)D, PTH, and bone parameters were found, so it seems unlikely that merely vitamin D supplementation would responsible for the observed increased aBMD in trans women.

Based on observations of positive associations of endogenous estrogen levels with vBMD and negative associations with the endosteal circumference in radius and/or tibia in large male cohorts, one might have expected an estrogen-mediated increase in cortical vBMD or a narrowing of the endosteal cortex [5, 38]. We found, however, stable volumetric bone measurements over time in trans women. Several explanations can be put forward. Firstly, the follow-up might have been too short to detect bone geometry changes. Although pQCT is considered to have high reproducibility [39], a recent paper assessing monitoring time intervals for longitudinal studies in postmenopausal women suggested a minimum interval of 2–6 years for changes at the radius and tibia [40]. Secondly, the group might have been too small to detect bone geometry changes. Positioning is extremely important in pQCT measurements, and special care has been taken to account for this (single operator, thorough positioning, scout view), although this could introduce extra variability, which is less in aBMD [41]. Another explanation might be that classical bone densitometry using DXA is two-dimensional and can be influenced by soft tissue superposition, whereas pQCT is not. Trans women indeed gain fat mass during follow-up. Large changes in body weight, BMI, and fat percentage can induce artefacts [42]. Adjustment of our models for total body fat mass did however not change the observed increase in aBMD over time and aBMD normalized for the respective regional fat mass showed a similar increase after 1 and 2 years of CSH. Bone turnover markers were also correlated with aBMD changes, suggesting a true effect on bone mass.

The stable areal and volumetric bone parameters support that the current treatment protocol with a lower dosage of estrogens and physiological female serum estradiol levels is a safe protocol for bone protection even with the combined use of cyproterone acetate 50 mg daily. Furthermore, treatment with anti-androgens alone for a short period before the combination with estrogen therapy does not seem to have an impact on bone compared with combined anti-androgens and estrogen treatment. Nonetheless, bone turnover markers did increase during the first year in the IAAM-group, whereas a decrease was observed in the AA + E-group. Given the observed high prevalence of osteoporosis before hormonal therapy and detrimental effects of long-term hypogonadism [32], prolonged used of anti-androgens in monotherapy is not

advisable. The higher bone turnover markers and IGF1 in treatment protocol AA + E versus IAAM at baseline were probably due to the age difference [29, 43].

IGF1 was higher in the first year of CSH in trans women in both transdermal as oral estrogen users. This increase seemed estrogen-modulated as increased IGF1 has also been described in orchidectomized mice using estradiol, whereas no changes were seen in orchidectomized mice with or without DHT treatment [44]. Serum IGF1 levels were independent of the route of administration of estrogens, in contrast with the previously described effects in postmenopausal women on estrogen therapy which could also be contributed to the type of estrogens used [45].

The observed body composition changes with increased fat mass and decreased muscle mass and strength are in line with the previous results in trans women using MRI [6] and DXA [7, 13] and pQCT [7] as is the increased serum leptin [46]. In particular, we observed more subcutaneous fat mass and an increased hip circumference, already after 1 year of CSH. We also observed a slower increase in fat mass and serum leptin in trans women who initially received cyproterone acetate for a short period alone, indicating the role of estrogens in the accrual of fat mass. A recently published short-term trial in adult men using sex steroid suppression followed by the replacement of testosterone with or without inhibition of estrogen synthesis by aromatase inhibitors however showed that estrogen deficiency contributed to the increased fat mass, independent of the dose of testosterone substitution [47]. A potential explanation of our findings may be found in the higher serum estradiol levels compared with males as divergent effects of estrogens depending on estrogen dosage cannot be ruled out.

Our study has several limitations. Firstly, gender dysphoria is a rare condition, and large-scaled samples are difficult to obtain with implications for power. Moreover, we describe trans women of a broad age range: some might not have fully reached peak bone mass, while others are already middle-aged. Adjustment for age did, however, not alter any of the results. Secondly, the effect of aging cannot be ruled out as we did not follow the control group prospectively. The changes in markers of bone turnover are however greater than expected during aging [43, 48]. Thirdly, trans women use pharmacological doses of estrogen and the observed effects can differ of those of endogenous estrogen in males. The strengths of our study are the standardized treatment protocol, the state of the art methods to measure sex steroids (tandem mass spectrometry), and the relatively large sample despite the rarity of the condition.

Conclusion

We conclude that whereas trans women have a lower aBMD and cortical bone size compared with control men before any kind of hormonal treatment, probably related to a different, more sedentary lifestyle, their skeletal status is well preserved during CSH treatment. During 2 years of CSH treatment, bone turnover decreased, aBMD increased, and bone geometry was stable in trans women, despite a substantially decreased muscle mass and strength, which is a further illustration of the major role of estrogens for preservation of the integrity of the male skeleton. Nevertheless, a longer follow-up might be needed to further detail the effects of CSH treatment on bone geometry.

Acknowledgments The authors are indebted to Griet De Cuypere, MD, PhD; Gunter Heylens, MD; Els Elaut, MSc, Birgit Van Hoorde, MSc; Steven Weyers, MD, PhD; Piet Hoebeke, MD, PhD; Stan Monstrey, MD, PhD; for referral of participants and to Jens Jacobeit, MD (Endokrinologikum, Hamburg, Germany) and Mick van Trotsenburg, MD (Vrije Universiteit, Amsterdam, the Netherlands) for their contribution to the development of the ENIGI endocrinological protocol. We thank all volunteers who participated as study subjects. We also thank Veronique Van den Bossche and Kathelyne Mertens for their excellent technical assistance. This work was supported in part by Grant G.0867.11 from the Research Foundation Flanders; Eva Van Caenegem, Sara Vandewalle, and Katrien Wierckx are holders of a PhD fellowship respectively from the Research Foundation Flanders (Eva Van Caenegem and Sara Vandewalle) and Ghent University (Katrien Wierckx).

Conflicts of interest Eva Van Caenegem, Katrien Wierckx, Youri Taes, Thomas Schreiner, Sara Vandewalle, Kaatje Toye, Jean-Marc Kaufman, and Guy T'Sjoen declare that they have no conflict of interest.

References

1. Seeman E (2001) Clinical review 137: Sexual dimorphism in skeletal size, density, and strength. J Clin Endocrinol Metab 86:4576–84
2. Frost HM (1987) Bone "mass" and the "mechanostat": a proposal. Anat Rec 219:1–9
3. Nilsson M, Ohlsson C, Oden A, Mellstrom D, Lorentzon M (2012) Increased physical activity is associated with enhanced development of peak bone mass in men: a five-year longitudinal study. J Bone Miner Res 27:1206–14
4. Callewaert F, Sinnesael M, Gielen E, Boonen S, Vanderschueren D (2010) Skeletal sexual dimorphism: Relative contribution of sex steroids, GH-IGF1, and mechanical loading. J Endocrinol 207:127–34
5. Lapauw B, Taes Y, Bogaert V, Vanbillemont G, Goemaere S, Zmierczak HG, De Bacquer D, Kaufman JM (2009) Serum estradiol and not testosterone influences volumetric bone mineral density and modulates the impact of physical activity on bone size at the age of peak bone mass—a study in healthy male siblings. J Bone Miner Res 24:1075–85
6. Elbers JM, Asscheman H, Seidell JC, Gooren LJ (1999) Effects of sex steroid hormones on regional fat depots as assessed by magnetic resonance imaging in transsexuals. Am J Physiol 276:E317–E325
7. Lapauw B, Taes Y, Simoens S, Van Caenegem E, Weyers S, Goemaere S, Toye K, Kaufman JM, T'Sjoen GG (2008) Body composition, volumetric and areal bone parameters in male-to-female transsexual persons. Bone 43:1016–21
8. T'Sjoen G, Weyers S, Taes Y, Lapauw B, Toye K, Goemaere S, Kaufman JM (2009) Prevalence of low bone mass in relation to

estrogen treatment and body composition in male-to-female transsexual persons. J Clin Densitom 12:306–13
9. Van Caenegem E, Taes Y, Wierckx K, Vandewalle S, Toye K, Kaufman JM, Schreiner T, Haraldsen I, T'Sjoen G (2013) Low bone mass is prevalent in male-to-female transsexual persons before the start of cross-sex hormonal therapy and gonadectomy. Bone 54:92–97
10. Dittrich R, Binder H, Cupisti S, Hoffmann I, Beckmann MW, Mueller A (2005) Endocrine treatment of male-to-female transsexuals using gonadotropin-releasing hormone agonist. Exp Clin Endocrinol Diabetes 113:586–92
11. Haraldsen IR, Haug E, Falch J, Egeland T, Opjordsmoen S (2007) Cross-sex pattern of bone mineral density in early onset gender identity disorder. Horm Behav 52:334–43
12. Mueller A, Dittrich R, Binder H, Kuehnel W, Maltaris T, Hoffmann I, Beckmann MW (2005) High dose estrogen treatment increases bone mineral density in male-to-female transsexuals receiving gonadotropin-releasing hormone agonist in the absence of testosterone. Eur J Endocrinol 153:107–13
13. Mueller A, Zollver H, Kronawitter D, Oppelt PG, Claassen T, Hoffmann I, Beckmann MW, Dittrich R (2011) Body composition and bone mineral density in male-to-female transsexuals during cross-sex hormone therapy using gonadotrophin-releasing hormone agonist. Exp Clin Endocrinol Diabetes 119:95–100
14. Lips P, Asscheman H, Uitewaal P, Netelenbos JC, Gooren L (1989) The effect of cross-gender hormonal treatment on bone metabolism in male-to-female transsexuals. J Bone Miner Res 4:657–62
15. Reutrakul S, Ongphiphadhanakul B, Piaseu N, Krittiyawong S, Chanprasertyothin S, Bunnag P, Rajatanavin R (1998) The effects of oestrogen exposure on bone mass in male to female transsexuals. Clin Endocrinol (Oxf) 49:811–14
16. Ruetsche AG, Kneubuehl R, Birkhaeuser MH, Lippuner K (2005) Cortical and trabecular bone mineral density in transsexuals after long-term cross-sex hormonal treatment: a cross-sectional study. Osteoporos Int 16:791–98
17. Sosa M, Jodar E, Arbelo E, Dominguez C, Saavedra P, Torres A, Salido E, de Tejada MJ, Hernandez D (2003) Bone mass, bone turnover, vitamin D, and estrogen receptor gene polymorphisms in male to female transsexuals: Effects of estrogenic treatment on bone metabolism of the male. J Clin Densitom 6:297–304
18. Van Kesteren P, Lips P, Gooren LJ, Asscheman H, Megens J (1998) Long-term follow-up of bone mineral density and bone metabolism in transsexuals treated with cross-sex hormones. Clin Endocrinol (Oxf) 48:347–54
19. Asscheman H, Giltay EJ, Megens JA, de Ronde WP, van Trotsenburg MA, Gooren LJ (2011) A long-term follow-up study of mortality in transsexuals receiving treatment with cross-sex hormones. Eur J Endocrinol 164:635–42
20. Coleman E, Bockting W, Botzer M, Cohen-Kettenis PT, De Cuypere G, Feldman J, Fraser L, Green J, Knudson G, Meyer W, Adler R, Brown G, Ehrbar R, Ettner R, Eyler E, Garofalo R, Karasic D, Lev AI, Mayer G, Meyer-Bahlburg H, Hall BP, Pfaefflin F, Rachlin K, Robinson B, Schechter L, Tangpricha V, van Trotsenburg M, Vitale A, Winter S, Whittle S, Wylie K, Zucker K (2011) Standards of care for the health of transsexual, transgender and gender nonconforming people. 7th edition. Int J Transgenderism 13:165–232
21. Kreukels BP, Haraldsen IR, De Cuypere G, Richter-Appelt H, Gijs L, Cohen-Kettenis PT (2012) A European network for the investigation of gender incongruence: the ENIGI initiative. Eur Psychiatry 27:445–50
22. Baecke JA, Burema J, Frijters JE (1982) A short questionnaire for the measurement of habitual physical activity in epidemiological studies. Am J Clin Nutr 36:936–42
23. Van Kesteren PJ, Asscheman H, Megens JA, Gooren LJ (1997) Mortality and morbidity in transsexual subjects treated with cross-sex hormones. Clin Endocrinol (Oxf) 47:337–342
24. Fiers T, Casetta B, Bernaert B et al (2012) Development of a highly sensitive method for the quantification of estrone and estradiol in serum by liquid chromatography tandem mass spectrometry without derivatization. J Chromatogr B Analyt Technol Biomed Life Sci 893–894:57–62
25. Kanis JA, Bianchi G, Bilezikian JP, Kaufman JM, Khosla S, Orwoll E, Seeman E (2011) Towards a diagnostic and therapeutic consensus in male osteoporosis. Osteoporos Int 22:2789–98
26. Gunter KB, Almstedt HC, Janz KF (2012) Physical activity in childhood may be the key to optimizing lifespan skeletal health. Exerc Sport Sci Rev 40:13–21
27. Delvaux K, Lefevre J, Philippaerts R, Dequeker J, Thomis M, Vanreusel B, Claessens A, Eynde BV, Beunen G, Lysens R (2001) Bone mass and lifetime physical activity in Flemish males: a 27-year follow-up study. Med Sci Sports Exerc 33:1868–75
28. Fujiyoshi A, Polgreen LE, Hurley DL, Gross MD, Sidney S, Jacobs DR Jr (2013) A cross-sectional association between bone mineral density and parathyroid hormone and other biomarkers in community-dwelling young adults: the CARDIA study. J Clin Endocrinol Metab 98:4038–46
29. Chaitou A, Boutroy S, Vilayphiou N, Munoz F, Delmas PD, Chapurlat R, Szulc P (2010) Association between bone turnover rate and bone microarchitecture in men: the STRAMBO study. J Bone Miner Res 25:2313–23
30. Kanis JA, Johnell O, Oden A, Johansson H, Eisman LC, Fujiwara S, KrogerH MCEV, Mellstrom D, Melton LJ, Pols H, Reeve J, Silman A, Tenenhouse A (2005) Smoking and fracture risk: a meta-analysis. OsteoporosInt 16:155–162
31. Khosla S, Oursler MJ, Monroe DG (2012) Estrogen and the skeleton. Trends Endocrinol Metab 23:576–81
32. Varsavsky M, Reyes-Garcia R, Garcia-Martin A, Rozas-Moreno P, Rocio GR, Munoz-Torres M (2014) Bone turnover markers in patients with prostate carcinoma: Influence of sex steroids levels. J Bone Miner Metab 32:65–70
33. Taxel P, Fall PM, Albertsen PC, Downset RD, Trahiotis M, Zimmerman J, Ohannessian C, Raisz LG (2002) The effect of micronized estradiol on bone turnover and calciotropic hormones in older men receiving hormonal suppression therapy for prostate cancer. J Clin Endocrinol Metab 87:4907–13
34. Eriksson S, Eriksson A, Stege R, Carlstrom K (1995) Bone mineral density in patients with prostatic cancer treated with orchidectomy and with estrogens. Calcif Tissue Int 57:97–99
35. Smith MR, Morton RA, Barnette KG, Sieber PR, Malkowicz SB, Rodrigez D, Hancock ML, Steiner MS (2013) Toremifene to reduce fracture risk in men receiving androgen deprivation therapy for prostate cancer. J Urol 189:S45–50
36. Falahati-Nini A, Riggs BL, Atkinson EJ, O'Fallon WM, Eastell R, Khosla S (2000) Relative contributions of testosterone and estrogen in regulating bone resorption and formation in normal elderly men. J Clin Invest 106:1553–60
37. Reid IR, Bolland MJ, Grey A (2014) Effects of vitamin D supplements on bone mineral density: a systematic review and meta-analysis. Lancet 11(383(9912)):146–55
38. Lorentzon M, Swanson C, Andersson N, Mellstrom D, Ohlsson C (2005) Free testosterone is a positive, whereas free estradiol is a negative, predictor of cortical bone size in young Swedish men: the GOOD study. J Bone Miner Res 20:1334–41
39. Rinaldi G, Wisniewski CA, Setty NG, Leboff MS (2011) Peripheral quantitative computed tomography: Optimization of reproducibility measures of bone density, geometry, and strength at the radius and tibia. J Clin Densitom 14:367–73
40. Duckham RL, Frank AW, Johnston JD, Olszynski WP, Kontulainen SA (2013) Monitoring time interval for pQCT-derived bone outcomes in postmenopausal women. Osteoporos Int 24:1917–22

41. Marjanovic EJ, Ward KA, Adams JE (2009) The impact of accurate positioning on measurements made by peripheral QCT in the distal radius. OsteoporosInt 20:1207–1214
42. Yu EW, Bouxsein M, Roy AE, Baldwin C, Cange A, Neer RM, Kaplan LM, Finkelstein JS (2013) Bone loss after bariatric surgery: Discordant results between DXA and QCT bone density. J Bone Miner Res
43. Goemaere S, Van Pottelbergh I, Zmierczak H, Toye K, Daems M, Demuynck R, Myny H, De Bacquer D, Kaufman JM (2001) Inverse association between bone turnover rate and bone mineral density in community-dwelling men >70 years of age: No major role of sex steroid status. Bone 29:286–91
44. Svensson J, Moverare-Skrtic S, Windahl S, Swanson C, Sjogren K (2010) Stimulation of both estrogen and androgen receptors maintains skeletal muscle mass in gonadectomized male mice but mainly via different pathways. J Mol Endocrinol 45:45–57
45. Leung KC, Johannsson G, Leong GM, Ho KK (2004) Estrogen regulation of growth hormone action. Endocr Rev 25:693–721
46. Elbers JM, Asscheman H, Seidell JC, Frolich M, Meinders AE, Gooren LJ (1997) Reversal of the sex difference in serum leptin levels upon cross-sex hormone administration in transsexuals. J Clin Endocrinol Metab 82:3267–70
47. Finkelstein JS, Lee H, Burnett-Bowie S-AAM, Pallais JC, Yu EW, Borges LF, Jones BF, Barry CV, Wulczyn KE, Thomas BJ, Leder BZ (2013) Gonadal steroids and body composition, strength, and sexual function in men. NEJM 369:1011–1022
48. Hochberg MC, Greenspan S, Wasnich RD, Miller P, Thompson DE, Ross PD (2002) Changes in bone density and turnover explain the reductions in incidence of nonvertebral fractures that occur during treatment with antiresorptive agents. J Clin Endocrinol Metab 87: 1586–92

BRIEFING BOOK

A REQUEST TO CONGRESS AND THE ADMINISTRATION

TO PRESERVE GIRLS' AND WOMEN'S SPORT

&

ACCOMMODATE TRANSGENDER ATHLETES

Prepared by The Women's Sports Policy Working Group (Revised as of February 3, 2021)
https://womenssportspolicy.org/
Contact: Donna Lopiano for additional information as needed (Donna.Lopiano@gmail.com or call 516-380-1213)

PRESERVING GIRLS' AND WOMEN'S SPORTS AND ACCOMMODATING TRANSGENDER ATHLETES

TABLE OF CONTENTS

WOMEN'S SPORTS POLICY WORKING GROUP

Donna de Varona, OLY. Two-time Olympic gold medalist in swimming. In 1965, UPI and AP voted her outstanding woman athlete in the world after she set 18 world records and fastest times. de Varona was a sports broadcasting pioneer as the youngest and one of the first women to work in the industry. As an Emmy recipient, she used her visibility to advise and advocate for the protection and promotion of Title IX as well as for the Ted Stevens Olympic and Amateur Sports Act. As the first President and Chair of the Board of the Women's Sports Foundation, she helped build the organization into a sustainable, influential entity. She has served on presidential commissions for five U.S. Presidents. Currently, de Varona is a member of the International Olympic Committee Communications Commission, and was recently voted onto the U.S. Olympic and Paralympic Committee Board of Directors.

Martina Navratilova, OLY. Former professional tennis player and coach. In 2005, *Tennis* magazine selected Navratilova the greatest female tennis player for the years 1975 - 2005. She is considered one of the best female tennis players of all time. Over her career, she won 18 Grand Slam singles titles, 31 Grand Slam women's doubles titles (an all-time record), and 10 Grand Slam mixed doubles titles, for a combined total of 59 major titles, marking the Open Era record for the most Grand Slam titles won by one player, male or female. Coached later in her career by the first trans-woman tennis player, Renée Richards, and long active in LGBTQ rights work and with the women's tennis tour, Navratilova is particularly well-positioned to contribute to thoughtful policy on the inclusion of trans women/girls in women's sport.

Donna A. Lopiano, Ph.D. President and founder of Sports Management Resources, LLC, Adjunct Professor of Sports Management, Southern Connecticut State University, former Chief Executive Officer of the Women's Sports Foundation (1992-2007), Director of Women's Athletics, University of Texas at Austin (1975-1992). President of The Drake Group—an organization focused on academic integrity in college sport. A Title IX sports pioneer, Lopiano specializes in gender equity in the educational and Olympic and elite sports spaces. As an athlete, she participated in 26 national championships in four sports and was a nine-time All-American at four different positions in softball, a sport in which she played on six national championship teams.

Nancy Hogshead-Makar, J.D., OLY, CEO Champion Women, civil rights lawyer, two-time Olympian, three-time gold medalist and one silver in swimming, U.S. National Team for eight years, 12 Halls of Fame, including the International Women's Sports Hall of Fame and the International Swimming Hall, 20 years of teaching Sports Law and Administration, current Professor at Rutgers University's Global Sports Business MS Program. Women's Sports Foundation - President 1991-1993, Legal Advisor, 2003-2010, Senior Director of Advocacy, 2010-2014.

Tracy Sundlun, CEO, Everything Running, Inc., Founding Board Member, National Scholastic Athletics Foundation. Co-Founder and Director of the National Scholastic (High School) Indoor & Outdoor Track & Field Championships (1984 – Present). Co-Founder, Rock 'n' Roll Marathon Series, at the time the largest running series in the world with over 500,000 participants annually in 29 events in 7 countries (1998 – 2016). Former club and collegiate track coach (including Georgetown, Colorado, USC), including for over 100 men and women in every event from 15 countries who have represented their country in the Olympic Games or other international competitions. Six-time Olympic Coach and Manager (1972 – 2016). Inducted into Running USA Hall of Champions.

Doriane Coleman, J.D. Professor of Law and Co-Director of the Center for Sports Law & Policy at Duke Law School; Senior Fellow at the Kenan Institute for Ethics and Associate of the Trent Center for Bioethics, Humanities & History of Medicine at Duke University & School Medicine; former collegiate and Swiss national champion in the 800 meters on the track. She has worked for years in domestic and international arenas on anti-doping policy and rules defining eligibility for the women's category. Her writing on sex in sport is widely referenced by policymakers considering the hard questions posed by trans and intersex inclusion in girls' and women's sport.

SUPPORTERS

Willie Banks, **OLY**, three-time Olympian and former world record holder in the triple jump

Joanna Harper, former elite marathoner, transgender athlete and researcher

Wendy Hilliard, nine-time member and two-time captain of Team USA in rhythmic gymnastics

Micki King, **OLY**, Olympic gold medalist and ten-time national champion in springboard and platform diving

Greg Louganis, OLY, four-time Olympic gold medalist in springboard and tower diving, second diver in history to sweep both diving events in consecutive Olympic Games, winner of 47 national and 13 world championships

Edwin Moses, **OLY**, two-time Olympic gold medalist , two-time World Champion, former world record holder, undefeated in the 400 meter hurdles for 10 years and 107 consecutive races

Benita Fitzgerald Mosley, **OLY**, Olympic gold medalist, two-time Olympian, and eight-time national champion in the 100 meters hurdles

Diana Nyad, one of the greatest ever long-distance swimmers credited with a record setting swim around Manhattan island and being first person to swim from Cuba to Florida without a shark cage

Renee Richards, former tennis player, one of the first professional athletes to identify as transgender

Sanya Richards-Ross, OLY, four-time Olympic gold medalist, six-time World Champion, ranked #1 in the world in the 400 meters from 2005 to 2009 and in 2012

Sally Roberts, three-time national wrestling champion, 2003 World Cup Champion, 2003 & 2005 World bronze medalist and a 2008 Olympic alternate

Lyn St. James, former Indycar and LeMans racecar driver, first woman to win Indianapolis 500 Rookie of the Year award, and one of *Sports Illustrated*'s "Top 100 Women Athletes of the Century"

Pam Shriver, OLY, Olympic gold medalist, winner of over 100 professional singles and doubles championships over 19 years, International Tennis Hall of Fame

Inge Thompson, **OLY**, ten-time national champion cyclist, three-time Olympian and two-time podium finisher at the Women's Tour de France

Champion Women, non-profit legal advocacy organization for girls and women in sports; harnessing the power of sport for social justice

The Drake Group, non-profit advocacy organization committed to defending academic integrity and protecting the health and well-being of athletes participating in collegiate sport

National Scholastic Athletics Foundation, non-profit organization created to fund competitive opportunities for high school track and field athletes and host the indoor and outdoor high school nationals

Wrestle Like A Girl, non-profit organization empowering girls and women using the sport of wrestling to become leaders in life

PRESERVING GIRLS' AND WOMEN'S SPORTS AND ACCOMMODATING TRANSGENDER ATHLETES

REQUEST

We ask Congress and the Administration to affirm Title IX's long-standing commitment to providing biological females[1] with equal experiences and opportunities in competitive sport, and to protecting their safety in contact sports, by permitting recipients of federal funds to continue to operate or sponsor separate athletic teams and events for males and females. In addition, we ask Congress and the Administration newly to provide for the participation of transgender girls and women within girls' and women's sports programs with appropriate accommodations if they have experienced all or part of male puberty (which is the scientific justification for separate sex sport). These accommodations should apply throughout interscholastic and intercollegiate sport, and the U.S. Olympic and Paralympic Movement. This request is limited to *competitive* interscholastic, intercollegiate, and developmental elite athletic programs. It does not affect physical education, intramurals, or recreational sports sponsored by municipalities, schools, and colleges.

THE PROBLEM

Girls'/Women's Competitive Sport Needs Protection and Trans Girls/Women Need to be Included with Appropriate Accommodations.

Sports have been continuously sex-segregated for over 100 years, across disciplines where male sex-linked advantages affect competitive opportunities for females. Congress passed Title IX in 1972 and approved its implementing regulations governing competitive sport in 1975, explicitly permitting girls' and women's sport to exist separate from boys' and men's sport. Law and sports policy makers understood that from the onset of male puberty, male bodies develop such that they are, as a group, faster, stronger, and more powerful than female bodies as a group. The performance gap between male and female athletes that emerges from that point typically ranges from 8-20% depending on the sport and event, and "up to 50% where explosive power and complex movement skills are pivotal."

Science not Ideology Dictates the Need for Sex Segregation in Sports.

If sports were not sex-segregated, female athletes would rarely be seen in finals or on victory podiums. Congress has long understood the benefits of sport and the benefits of having the opportunity to train, compete and win. Repeatedly, Congress has affirmed that it wanted both our sons and our daughters to realize those benefits, which are well-documented in the academic literature. Girls and women learn the benefits of teamwork in pursuit of conference, state and national championships; the self-esteem that flows from competent performance of physical skills; the life-changing power of competing against the best and standing on the podium; confidence borne of testing the limits of strength, speed, skill and reaction time; and the power of personal

[1] We use "female" throughout to denote a person's biological sex regardless of their gender identity. We use "trans(gender) girl/woman" throughout to denote a person born male who identifies as a girl/woman.

achievement and public recognition when setting school, meet and other records. And as sports double as an academic and social tool, these lessons and benefits reverberate well beyond the playing field throughout the lives of all female athletes.

The legislative history of Title IX is clear that Congress also understood that even when height, size, and weight are equal, males are incrementally stronger and generate more explosive force, so that if males and females are forced to compete against each other, the physical safety of females is differently at risk.

At the time Title IX's athletics regulations were passed, no one raised the issues of gender identity apart from biological sex, or whether trans girls/women with the post-pubescent advantages of biological males should be allowed to participate in the space created by Congress to protect the sport experiences of biological females. Today, however, trans girls/women are asking for the right to compete in girls' and women's sport, directly against female athletes, even when they retain some or all of their male sex-linked strength, power, and related performance advantages. For many people, the issue is not whether trans girls/women should be included in women's sport. Rather, it is whether female athletes can continue to be protected *and* trans girls/women accommodated within women's sports consistent with their gender identity.

States are Passing Conflicting Laws.

States have taken one of three general approaches to the issue of trans-inclusion in girls' and women's sport. Ten states expressly require males and females to participate in high school sports according to their birth sex, thereby prohibiting participation in girls' sports by trans girls, whether or not they have begun male puberty or have had hormone therapy. In contrast, seventeen states and the District of Columbia expressly require the inclusion of trans girls in girls' sports without regard to the extent to which they may retain the male-linked physical traits that otherwise justify excluding males from female sport on competitive fairness and safety grounds. Another seventeen states have adopted a policy similar to the NCAA rule, which allows trans girls to compete after taking gender-affirming hormones for a year. Finally, six states have no policy one way or the other regarding gender identity and sport.

Pending Legal Challenges.

None of the policies mentioned has been immune from a legal challenge. In Connecticut, one of the states that allows trans girls to compete in girls' sports without regard to whether they have experienced male puberty or are on gender affirming hormones, four female athletes – cis girls – and their mothers have sued their state high school athletic association. They contend that Connecticut's rule, which ignores biological sex and focuses on gender identity, has deprived them of school and state records they would otherwise have held, and from advancing in competitions, including qualifying for state and regional championships and becoming state champions, spots they would otherwise have won. Instead, the rule has allowed two trans girls to dominate their events. The Department of Education has concluded that the state's policy regarding transgender athletes violates Title IX's mandate of equal opportunity for both sexes since biological males are able to win in both the male and the female divisions. At the other end of the country and the political spectrum, Idaho has seen a similar legal battle erupt after it adopted a law mandating that athletic eligibility be based only on birth sex. To date there has been no approach that would include trans girls/women while preserving competitive opportunities for females.

The Cultural Battle Outside of the Courts has Not Allowed for Respectful Dialogue on Science, Policy, and Best Practices.

Transgender advocates accuse female athletes, their parents, and supporters of transphobia simply because they recognize the significance of sex in sport. Others seek unnecessarily to exclude all trans girls from all girls' sports regardless of whether they have experienced male puberty or are undergoing gender-affirming therapies. The conflicting positions have sparked a rhetorical battle about who will suffer more harm: trans girls who are prevented from competing as girls, or females who are forced to compete against athletes who have the male sex-linked advantages girls' and women's sport was designed to exclude. Throughout, surveys consistently show Americans want sports opportunities for girls and women. Only a minority of Americans – just 29% according to one recent survey – favor allowing transgender students to participate on the sports team consistent with their gender identity.

Conflicting Federal Legislative Proposals Take an "Either/Or" Approach.

Various bills in the 116th Congress would either require identical treatment of – no distinctions allowed between – females and trans girls (H.R. 5 – The Equality Act), or they would preclude all trans girls/women from participating in girls'/women's sports, which would be restricted to biological females (H.R. 5603, H.R. 8932 and S. 4649).

The Supreme Court in Bostock did not Resolve the Question of Separate Sex Sport.

Trans girls/women and their advocates argue that the Supreme Court's decision in Bostock v. Clayton County, 590 U.S.___, 140 S. Ct 1731 (2020), mandates the unconditional inclusion of trans girls/women in women's sports. This is misleading at best, as *Bostock* was about workplace discrimination under Title VII, not about sex segregation in competitive sport; in its decision, the Court expressly stated that it was defining "sex" to mean "biological sex" not "gender identity"; and it expressly reserved for another day – did not decide – the issue whether distinctions on the basis of sex are permissible for bathrooms, locker rooms, and sport.

THE SOLUTION

It is essential to protect girls' and women's sport. It is also good policy to be inclusive whenever possible without causing harm to the female sports competition and the individuals it is designed to protect. Congress and the Administration should make it clear that institutions governed by Title IX of the Education Amendments of 1972 (Title IX), the Ted Stevens Olympic and Amateur Sports Act (the Sports Act), and Title VII of the Civil Rights Act of 1964 (Title VII) will:

> (1) continue to be obligated to provide males and females with equal sporting opportunities on the basis of biological sex, and
>
> (2) be newly obligated to provide ways to include trans girls/women in girls'/women's sports that ensure competitive fairness and playing-safety without diminishing the protection of biological females.

This two-step approach protects the integrity of the existing competitive sport process in which millions of girls and women participate annually. It also incrementally and thoughtfully expands the development of additional sports opportunities for emerging trans girls/women.

Separate sex sport has always been an exception to our general non-discrimination laws. This exception is justified by real physical sex-linked differences that emerge from the onset of male puberty and that have significant implications for athletic performance and playing-safety. The lawfulness of this long-standing exception should be re-affirmed.

At the same time, the government should make it legally possible for trans girls/women to participate in girls'/women's sport in ways that do not affect competitive fairness and playing-safety. Because the onset of male puberty is the physical justification for separate sex sport, trans girls and women who have never experienced the onset of male puberty should be included without condition. Trans girls and women who have experienced the onset of male puberty should be accommodated in ways that recognize their male sex-linked advantages in strength, power, and endurance. Some – but not all – trans girls and women are on gender affirming hormones. Depending on the sport and the event, hormones may mitigate those advantages to some extent. However, the evidence is increasingly clear that hormones do not eliminate the legacy advantages associated with male physical development. Accommodations can and should take these legacy advantages into account.

Finally, it is important that there be national standards to ensure uniformity across the states. Competitive sport, i.e., sport that leads to records, titles, championships, and ultimately to Team USA, is an interconnected system comprised of high schools, colleges, and universities, and non-school club teams and programs. The former are governed by Title IX. The latter are generally under the jurisdiction of, or sponsored by, the U.S. Olympic and Paralympic Committee (USOPC) and/or regional and national sport governing bodies (NGBs). This integrated system is only local in the first instance, as teams and athletes move seamlessly from intra-state to interstate and international arenas as competition progresses. Inconsistent local, state, national, and international eligibility standards can create practical impediments to success for individual athletes, teams, and ultimately for the system as a whole.

Adopting international standards is a compromise, but given our safety, fairness, and inclusion goals, it would be a substantively sound and administratively efficient approach to national policy. The Olympic Movement has already committed itself both to protecting the female category and to scientific, evidence-based criteria for the conditional inclusion of trans girls/women in girls'/women's sport, including acceptable methods of accommodation. The standards are reviewed on an ongoing basis to ensure they are consistent with sport's policy goals and the best available scientific evidence on competitive fairness and safety. To date, no American sports organization or governing body has established a commitment to, or the capacity for, doing this work. Equally important is that pegging USA standards to those of the International Olympic Committee (IOC) and the International Federations (IFs) ensures that our country's athletes and teams – including our juniors who are Team USA's future - can move seamlessly from domestic to international competition, and that none of our elite athletes are ineligible at the outset.

TRANS INCLUSION CHART	PHYSICAL STATUS OF TRANS GIRL/WOMAN ATHLETE		
	Athlete has not experienced (any stage of) male puberty	**Athlete has experienced (any stage of) male puberty, is not on hormones, and/or has insufficiently mitigated legacy advantage**	**Athlete has experienced (any stage of) male puberty and is on hormones and has sufficiently mitigated legacy advantage**
PARTICIPATION SPORTS: INTRAMURALS, PHYS-ED, RECREATIONAL SPORT			
Non-contact sport (co-ed allowed with sport- and sex-specific safety rules)	**Included without conditions (no different requirements)**	**Included consistent with rules for co-ed play**	**Included consistent with rules for co-ed play**
Contact sport (co-ed allowed with sport- and sex-specific safety rules)	**Included without condition (no different requirements)**	**Included consistent with rules for co-ed play**	**Included consistent with rules for co-ed play**
COMPETITIVE SPORT: INTERSCHOLASTIC, INTERCOLLEGIATE, ELITE SPORT			
Non-contact sport	**Included without condition (no different requirements)**	**Included but no head-to-head competition e.g., separate heats, scoring, events, podiums**	**Included consistent with international rules e.g., current NCAA rule**
Contact sport	**Included without condition (no different requirements)**	**Included but no head-to-head competition and separate scoring, events, podium**	**Included consistent with international rules e.g., current NCAA rule**

Under this proposal, females who identify as boys and men or as gender fluid are always eligible to compete in the girls'/women's category so long as they are not on male gender affirming hormones. Sex-linked traits and classification of the activity as "participation sports" or "competitive sports" drive the conditions, not gender identity.

Existing rules for co-ed play can be adopted for this context as they represent evidence- and experience-based approaches that ensure both safety and competitive fairness for all participants. For example, the rules for co-ed intramural basketball mandate same sex player-to-player defense and boys/men are prohibited from rebounding while in the key. Similarly, co-ed volleyball,

requires the same number of players of each sex on the floor, an 8' rather than a 7'4" net, no blocking over the net or of serve, and at least one player from each sex must touch the ball prior to returning the ball over the net. These approaches can be borrowed when necessary to ensure fair, safe inclusion of trans girls/women in participation sports.

Trans women/girls who are not on gender affirming hormones, i.e., who have full male advantage, can choose to continue to compete in the boys'/men's category just as trans boys/men can and do sometimes choose to continue to compete in the girls'/women's category, and as females can and do sometimes choose to try out for the boys'/men's team. Or, they can choose to be accommodated within girls'/women's sport so long as it is not in head-to-head competition. This last condition means that a trans girl/woman not on hormones who chooses to compete in girls'/women's sport sometimes won't have a direct competitor, just as athletes in competitions sorted by weight classes sometimes lack an opponent and have to choose to win by forfeit or move up a weight class and compete with a disadvantage. Because of this, the choice to compete in the boys'/men's category or the girls'/women's category should always be up to the individual based on their personal circumstances and preferences. Either way, they have a place in sport.

The choice to peg eligibility to international rules is explained in the text immediately above the table. The current NCAA rule is superficially consistent with international rules; however, because the NCAA is not specific about the degree of mitigation required and does not monitor trans women for compliance, it is significantly flawed as administered. This problem is described below, in Q&A #26. The NCAA rule is noted here as an example of a domestic rule that, if properly administered, is a reasonable accommodation consistent with international rules.

MODEL STATUTORY AND REGULATORY LANGUAGE TO PRESERVE WOMEN'S SPORTS & ACCOMMODATE TRANS ATHLETES

The best way to solve the problem that is protecting girls' and women's sport while accommodating trans athletes is to enact standalone federal legislation. This approach would ensure clarity and consistency in the law's treatment of the issue by the federal government, the states, and sports governing bodies. The model language immediately below is thus for a standalone federal statute. Following that is language to amend the Title IX regulations governing separate sex sport, 34 C.F.R. § 106.41, and the Equality Act, in the event lawmakers prefer one or both alternative approaches. All three options are based in and compatible with the Title IX regulations. Approaching law reform related to girls'/women's competitive sport in this way ensures that the extensive web of related statutory, administrative, and case law that exists in this area is not unnecessarily voided by our proposed trans-inclusive law reform measures.

PROPOSED STANDALONE STATUTE

Schools receiving federal funds and sports governing bodies engaged in interstate commerce [covered entities][2] may operate or sponsor separate competitive sports teams and events based on biological sex[3] where group-based sex-linked traits affect playing-safety and competitive capacity.

(A) Covered entities shall provide equal athletic opportunities, treatment, services and benefits in kind, quality and availability to male and female athletes.[4]

(B) Covered entities may restrict eligibility for the female sport category if any sex-based differences would have a negative impact on the right of females to achieve equality of athletic opportunity.[5]

(1) With respect to competitive opportunities, if a covered entity provides male athletes and teams the opportunity to advance to invitational, conference, state, regional, national, and international competition in the boys' and men's division, it must provide a parallel opportunity to female athletes similarly to advance in the girls' and women's division.[6]

(C) Treatment of Transgender Athletes

Where a covered entity operates or sponsors separate sex teams and events, transgender athletes shall be accommodated as follows:

(1) Treatment of Transgender Boys and Men

(a) Trans boys/men who have not taken gender-affirming hormones may be included in girls' and women's sport without conditions or limitations.

---

2 Schools receiving federal funds are subject to Title IX of the Education Amendments of 1972. Sports governing bodies include public and private non-profit high school and age-group athletic associations, intercollegiate athletic associations, the U.S. Olympic and Paralympic Committee, and their member National Governing Bodies. These organizations may be subject to Title IX and/or the Ted Stevens Olympic and Amateur Sports Act.

3 See National Institutes of Health (NIH), Office of Research on Women's Health, Sex & Gender, https://orwh.od.nih.gov/sex-gender, last accessed on January 1, 2021 (explaining that "'[s]ex' refers to biological differences between females and males, including chromosomes, sex organs, and endogenous hormonal profiles. 'Gender' refers to socially constructed and enacted roles and behaviors which occur in a historical and cultural context and vary across societies and over time.")

4 U.S. Department of Education, Office for Civil Rights, Title IX 1979 Policy Interpretation, *available at* http://www.ed.gov/offices/OCR/docs/t9interp.html.

5 *See* McCormick v. School District of Mamaroneck, 370 F.3d 275 (2nd Cir. 2004) (holding that recipients must provide girls with equal opportunities to compete in championship games.); U.S. Department of Education, Office for Civil Rights, Title IX 1979 Policy Interpretation, *available at* http://www.ed.gov/offices/OCR/docs/t9interp.html.

6 *Id.*

(b) Trans boys/men who have begun taking gender-affirming hormones

(i) may compete in boys' and men's sport and in co-ed (open) sport, but

(ii) may not compete head-to-head against female athletes in girls' and women's sport.

(2) Treatment of Transgender Girls and Women

(a) Trans girls/women who have not begun male puberty do not have significant male sex-linked advantages and are to be included in girls' and women's sport without conditions or limitations.

(b) Trans girls/women who have experienced all or part of male puberty and who have sufficiently mitigated their male sex-linked advantages through surgery and/or gender affirming hormones consistent with the rules of their sport's international federations, may participate in girls'/women's sport without further conditions or limitations.

(c) Trans girls/women who have experienced all or part of male puberty and who have not at all, or only partially, mitigated their male sex-linked advantages according to the international federation standards in their sport

(i) may be included in girls'/women's sport but not in head-to-head competition against female athletes, and[7]

(ii) may be included in boys' and men's sport, and in co-ed (open) sport.

(D) The private medical information (PMI) necessary to determine an athlete's eligibility must be available to the relevant sports authority. The necessary information is limited to confirmation of the athlete's biological sex and of their hormone status over the relevant period of time.

(E) Policy, training and competition must encourage a safe, respectful, and affirming environment for all athletes.

[7] Head-to-head competition is when two or more athletes compete directly against one another other, for example in the same heat in the pool or on the track, or on the same court in basketball and volleyball. Trans girls/women who have not at all, or only partially, mitigated their male sex-linked advantages who want to be included under the girls'/women's sport umbrella must be accommodated by means that do not involve head-to-head competition. Acceptable accommodations should be developed by sports administrators who are experts in the affected sports and events, but they need not reinvent the wheel. Existing models for co-ed sport and weight and age divisions can be borrowed for this purpose. Examples of acceptable accommodations might include separate events, heats, divisions, or handicapping that permits separate scoring, separate teams, or separate recognition.

(F) This statute only applies to competitive sport, when the principal objective is to win individual or team championships, titles, medals, or prize money. It does not apply to recreational sport such as physical education classes or intramural events, the principal objective of which is to participate for health and enjoyment.

PROPOSED AMENDMENT TO THE EQUALITY ACT (H.R. 5 – 2019)

Amend SEC. 9. MISCELLANEOUS. as follows:

Within Section 1101. DEFINITIONS AND RULES., by inserting

"(4) SEX.—The term 'sex' includes—

"(A) biological sex, including the sex characteristics that account for the physical and physiological differences between males and females;[8]

"(B) sex stereotype;

"(C) pregnancy, childbirth, or a related medical condition;

"(D) sexual orientation; and

"(E) gender identity.~~; and~~

~~**"(D) sex characteristics, including intersex traits.**~~[9]

"Section 1106. RULES OF CONSTRUCTION.

"(A) Sex – Nothing in section 1101 or the provisions of a covered title incorporating a term defined or a rule specified in the section shall be construed –

"(1) To limit the protection against an unlawful practice on the basis of pregnancy, childbirth, or a related medical condition provided by section 701(k), or

"(2) To limit the obligation of programs and institutions covered by Title IX of the Education Amendments of 1972, the Ted Stevens Olympic and Amateur Sport Act, and Title VII of the Civil Rights Act of 1964 to provide separate opportunities on the basis of biological sex when this is necessary to protect the right of biological females to equality in competitive athletics, or

"(~~2~~)(3) To limit the protection against an unlawful practice on the basis of sex available under any provision of Federal law other than that covered title, prohibiting a practice on the basis of sex."

---

[8] See National Institutes of Health (NIH), Office of Research on Women's Health, Sex & Gender, https://orwh.od.nih.gov/sex-gender, last accessed on January 1, 2021 (explaining that "'[s]ex' refers to biological differences between females and males, including chromosomes, sex organs, and endogenous hormonal profiles. 'Gender' refers to socially constructed and enacted roles and behaviors which occur in a historical and cultural context and vary across societies and over time.")

[9] "Sex characteristics" as a term and category includes intersex traits which appears in (4) (A) above.

PROPOSED AMENDMENT TO THE TITLE IX REGULATIONS (34 C.F.R. § 106.41)

(a) *General.* No person shall, on the basis of sex, be excluded from participation in, be denied the benefits of, be treated differently from another person or otherwise be discriminated against in any interscholastic, intercollegiate, club or intramural athletics offered by a recipient, and no recipient shall provide any such athletics separately on such basis.

(b) *Separate teams.* Notwithstanding the requirements of paragraph (a) of this section, a recipient may operate or sponsor separate teams **based on biological sex** where selection for such teams is based upon competitive skill or the activity involved is a contact sport. However, where a recipient operates or sponsors **separate-sex teams and offers** a team in a particular sport for members of one sex but operates or sponsors no such team for members of the other sex, and athletic opportunities for members of that sex have previously been limited, members of the excluded sex must be allowed to try-out for the team offered unless the sport involved is a contact sport **implicating physical safety because of group-based differences in size, weight, strength, and explosive force.**[10] For the purposes of this part, contact sports include **but are not limited to** boxing, ~~**wrestling,**~~[11] rugby, ice hockey, football, basketball and other sports the purpose or major activity of which involves bodily contact.

(c) *Treatment of Transgender Athletes.*

(1) Because trans girls/women who have not begun male puberty do not have significant male sex-linked advantages, they shall be included in girls' and women's sport without conditions or limitations.

(2) Trans boys/men who have not taken gender-affirming hormones may be included in girls' and women's sport without conditions or limitations.

(3) Trans girls/women who have experienced all or part of male puberty and who have sufficiently mitigated their male sex-linked advantages – through surgery and/or gender affirming hormones consistent with the rules of their international federations – may participate in girls'/women's sport, including contact sport, without additional conditions or limitations.

(4) Trans girls/women who have experienced all or part of male puberty and who have not, or only insufficiently, mitigated their male sex-linked advantages according to the international federation standards in their sport may be accommodated within girls'/women's sports but not in head-to-head competition with female athletes.

---

[10] Because the contact sport exception is permissive not mandatory, schools may allow girls/women to try out for positions on boys'/men's contact sports teams. This is least controversial when the position at issue – as opposed to the sport in general – does not involve a high risk of significant physical impact. See, e.g., Vanderbilt kicker Sarah Fuller first woman to score in Power 5 football game, ESPN News Service, December 12, 2020.

[11] Wrestling was included on this list in the original Title IX regulations because it necessarily involves close physical contact. However, it is unlike the other contact sports on the list because it is scored for skill not for strength, size, and power, and because athletes compete in narrowly defined weight classes and are strictly monitored and penalized for physically dangerous moves. This is why wrestling and other grappling sports (e.g., judo) are safely conducted in co-ed contexts. We propose removing it from the list for these reasons and because its inclusion typically reflected cultural norms and not physical safety concerns.

Additional events, heats, divisions, or handicapping that permit separate scoring, separate teams, or separate recognition – e.g., according to existing models for different age and weight classes – should be considered as accommodations are developed for and in this category.

(5) The private medical information (PMI) relevant to determining an athlete's eligibility must be available to the governing sports authority but must be treated confidentially.

(6) Policy and training should encourage a safe, respectful, and affirming environment for all women and girls.

DEFINITIONS

ACCOMMODATION – The process of adapting or adjusting to someone or something without changing the underlying goal or design, e.g., in a workplace or educational program. In the context of sport, accommodation means adjusting an aspect of girls'/women's event to include trans girls with male sex-linked advantages in a way that does not diminish participation and competitive opportunities for females. Examples of accommodations already in use in sport include handicapping, separate heats, separate scoring and/or separate and equal teams. This list is not exhaustive.

ANTI-DOPING – The effort against doping in sport. Doping is the use of prohibited substances and methods. Prohibited Substances Lists in the United States are maintained by the United States Anti-Doping Agency (USADA) and the National Collegiate Athletic Association (NCAA). Testosterone is a steroid on both Prohibited Substances Lists. Its exogenous use by athletes is banned. The testosterone levels of international-caliber athletes are monitored by regular urine and blood tests to ensure they do not fluctuate beyond both their own naturally-occurring levels, and the normal group ranges for their sex.

CISGENDER (CIS) – An adjective that describes a person who is neither transgender nor gender fluid. It is also used to describe a person whose gender identity is consistent with their natal sex.

CIS MALE – A person whose biological sex is male who is neither transgender nor gender fluid.

CIS FEMALE – A person whose biological sex is female who is neither transgender nor gender fluid.

COMPETITIVE FAIRNESS – The state of play when the rules reflect — and events are conducted — consistent with the design of the sport. For example:

- Weight categories are fair when groups of comparably sized athletes are matched against one another. For example, a wrestling match is considered fair when the competitors compete in their narrowly defined weight classes and referees ensure that competitors' actions are authorized from within a range of permissible maneuvers.
- Age categories are fair when they recognize and mitigate competitive differences conferred on the body due to the age of the competitor.
- Similarly, sex segregated sport classifies athletes by their biological sex because of the significant performance gap between male athletes and female athletes, and to ensure that female athletes have the same competitive opportunities as their male counterparts. In this context, competitive fairness requires rules that protect the female category and the female athletes who reasonably rely on its integrity.

CONFIDENTIAL MEDICAL INFORMATION – Information, including protected health information (PHI), that is normally treated confidentially but is relevant for the determination of eligibility for sports participation and therefore shared in a limited way for this limited purpose.

FEMALE – An individual whose biological sex is female. Biological sex is sometimes referred to as natal sex. In contrast with males, females have ovaries, not testes; they make eggs, not sperm; and their endocrine system is estrogenic, not androgenic.

GENDER — Sometimes used as a synonym for sex; or to connote the complex relationship between physical sex-linked traits and one's internal sense of self as male, female, both, or neither; or one's sex-related expression.

GENDER AFFIRMING HORMONES – Medication taken by some trans people to counter their biological sex and affirm their gender identity. For example, trans girls/women may take estrogen to counter their male secondary sex traits and to feminize their bodies. Similarly, trans boys/men may take testosterone to counter their female secondary sex traits and to masculinize their bodies.

GENDER AFFIRMING SURGERY – Procedures undertaken by some trans people to construct or remove secondary sex traits to better reflect their gender identity, e.g., surgery to remove or construct breasts, and/or surgery to remove testes or ovaries and/or construct gender-conforming genitals.

GENDER IDENTITY – A person's deeply held inner sense of themselves as male, female, fluid, or neither. A person's gender identity may be different from their biological sex.

LEGACY ADVANTAGE - The permanent or long-lived physical effects of experiencing puberty in the male body. The term refers to the considerable size and strength advantages that remain even after hormone treatments or surgical procedures.

MALE – An individual whose biological sex is male. Biological sex is sometimes referred to as natal sex. In contrast with females, males have testes, not ovaries; they make sperm, not eggs; and their endocrine system is androgenic not estrogenic.

PERFORMANCE GAP – The percentage difference between male athletic performances and female athletic performances that result from biological sex-linked differences. Some individual females surpass some individual males, but depending on the sport and event, the gap between elite male performances and elite female performances overall generally ranges from 8-20%, and up to 50% in sports and events featuring explosive power. The very best elite female performances are regularly surpassed by non-elite male performances. Together with the commitment to sex equality, the substantial performance gap justifies separate sex teams and events.

PLAYING-SAFETY – The physical safety of athletes on the field of play.

PUBERTY – The period of sexual maturation and the development of fertility. Sexual maturation includes the development of secondary sex characteristics—the physical features associated with a male phenotype on the one hand, and a female phenotype on the other. In girls, the onset of puberty is generally between ages 8 and 13. In boys, it is generally between ages 9 and 14.

SEX ASSIGNED / RECORDED AT BIRTH - The designation of a newborn child's sex on their official birth record based on inspection of their external genitalia. This designation may be incorrect in the case of an infant with a difference of sex development (DSD) that affected the

development of their genitals. Sex recorded on birth certificates, passports, or drivers' licenses may or may not reflect biological sex and should not be determinative of eligibility for competition.

SEX / BIOLOGICAL SEX – Male or female, one of two classifications by which most organisms are grouped on the basis of their reproductive organs and functions. A person's sex also refers to the cluster of sex-linked characteristics or traits—i.e., chromosomal, gonadal, endocrinological (hormonal), and phenotypic characteristics, commonly used to distinguish males from females.

SEX-LINKED DIFFERENCES – Physical and physiological differences that are tied to being biologically male or biologically female. For purposes of sport, the main sex-linked differences are tied to the endogenous (natural) production in biological males of much higher levels of testosterone beginning from the onset of male puberty and continuously throughout the competitive athletic career.

SEX SEGREGATION – Refers to separating people by sex or by particular sex-linked traits such as testosterone. Formal sex segregation in competitive sports is constitutional because it is empowering not subordinating, and because it is the only way to ensure that females as a group have the same sports opportunities, experiences and successes as males as a group.

TESTOSTERONE / TESTOSTERONE RANGES – A hormone classified as an anabolic, androgenic steroid that builds tissue. In childhood, males and females produce about the same, small amount of testosterone. At the onset of puberty, the male testes begin to produce much more than the female ovaries. From that point forward, the normal female range[12] remains low and narrow, from 0.06 to 1.68 nmol/L, and the normal male range is relatively high and wide, from 7.7 to 29.4 nmol/L.

TRANSGENDER (TRANS) – An adjective describing a person whose gender identity is not the same as their biological sex. The person may or may not choose to transition medically through the use of gender-affirming hormones or surgery.

TRANS BOY/MAN – A person whose biological sex is female, while their gender identity is male; one who transitions from female to male.

TRANS GIRL/WOMAN—A person whose biological sex is male, while their gender identity is female; one who transitions from male to female.

UNCONDITIONAL INCLUSION – Inclusion in a category, classification, or group without preconditions, such as including a trans girl/woman in girls'/women's sport without first requiring her to reduce her male sex-linked advantages.

[12] We use the word "normal" throughout this document consistent with its standard scientific meaning, i.e., the normal range is the range within which almost all readings or levels occur. In medicine, the normal range is sometimes also referred to as the reference range.

FREQUENTLY ASKED QUESTIONS

I. ABOUT SCIENCE AND SEX

Q1. What is "biological sex"?

A1. Biological sex is the designation of an individual as male or female based on reproductive organs and associated primary and secondary sex characteristics. Biologically, they are either female with ovaries/eggs and an estrogenic endocrine system, or they are male with testes/sperm and an androgenic endocrine system.

Q2. What are sex differences?

A2. Sex differences are anatomical and physiological differences that are determined by or related to biological sex. Males on the one hand and females on the other have distinct genetic and chromosomal, gonadal, endocrinological, and phenotypic (external secondary) characteristics. The field of sex differences in biomedical research specifically studies these distinctions, which have implications not only for reproduction and sport, but also for immunology and cardiovascular health, among other things. As the Institute of Medicine has explained, "basic biochemical differences" exist between males and females even "at the cellular and molecular levels."

Q3. Why do we have separate sex sport?

A3. We have separate sex sport and eligibility criteria based on biological sex because this is the only way we can assure that female athletes have the same opportunities as male athletes not only to participate but also to win in competitive sport. We also separate males and females in contact sports for reasons related to on-the-field playing-safety. From the onset of male puberty, male bodies develop such that they are as a group faster, stronger, and more powerful than female bodies as a group. The performance gap between male and female athletes that emerges from that point typically ranges from 8-20%, but up to 50% depending on the sport and event. If we did not separate athletes on the basis of biological sex - if we used any other physical criteria - we would never see females in finals and on podiums.

Q4. Couldn't we have eligibility criteria for the two divisions (male and female) based on some different (other than sex) physical criteria?

A4. No. There are no other physical criteria that could be used to determine eligibility that would similarly assure sex equality in competitive sport. Based on those different criteria, e.g., matching leg length, wing span, height, weight, etc., males as a group would always outperform females as a group because their biological sex differences, primarily testosterone levels in the male range from the onset of puberty and throughout the athletic career. Team USA stars Missy Franklin and Ryan Lochte illustrate this point well. They are both multiple Olympic and world champions in swimming. Both had first class training, coaching, and support. Both are 6'2" with reported 6'4" wingspans. Both held the world record in the 200 meters backstroke. But had they raced each other on their best days, Lochte would have finished about a half lap ahead of Franklin. In 2012, the year Franklin set her world record, her time of 2:04.06 would have placed her in a tie for 50th in the U.S. men's Olympic Trials.

Q5. If a boy and a girl are the same height, weight, and body build, aren't they likely to be essentially the same athletically?

A5. No. Testosterone-driven sex differentiation at puberty results in males developing larger hearts and higher capacity for oxygen transport and carbohydrate processing, as well as different skeletal and muscular composition. All of these characteristics provide males with superior strength, speed, power, and endurance.

Q6. What do scientific experts estimate is the sports performance advantage of post pubescent males?

A6. Experts estimate the male advantage is normally between 8 and 20% depending on the sport and event, and up to 50% in sports and events featuring explosive power. For example: Team USA's best female high jumper is Vashti Cunningham, NFL star Randall Cunningham's daughter. She is regularly ranked among the top ten best female high jumpers in the world. Her best jump as a professional (6' 6 ½") is regularly surpassed by dozens of U.S. high school boys.

As the chart immediately below – comparing California high school performances – shows, this isn't a phenomenon exclusive to professionals. Because the performance gap emerges at the onset of male puberty, as a group, high school girls have no chance against high school boys as a group.

2019 CALIFORNIA REGIONAL HIGH JUMP RESULTS[13]

REGION	BEST HIGH SCHOOL BOY	BEST HIGH SCHOOL GIRL	% DIFFERENCE
Central	6'10"	5'10"	14.63%
Central Coast	6'6"	5'6"	15.38%
Los Angeles	6'2"	5'2"	16.22%
North Coast	6'10"	5'5"	20.73%
Northern	6'5"	5'6"	14.29%
Oakland	5'11"	4'10"	18.31%
Sac-Joaquin	6'8"	5'8 1/4"	14.69%
San Diego	6'8"	5'10 1/2"	11.88%
San Francisco	6'0"	4'10"	19.44%
Southern	7'0"	5'8 1/2"	18.45%

[13] This chart is based on data from Athletics.net, California High Jump Results, accessed on September 25, 2019.

Q7. Are advocacy groups correct when they say that it's a myth and an outdated stereotype that females can't compete with males?

A7. No. It is a fact - not myth or outdated stereotype - that starting from the onset of male puberty, i.e., starting in middle school, there is an average 8-20% performance gap between males and females, which reaches to 50% in some sports and events. The proposition that better resources and support for female athletes can change biological imperatives and competitive results is false. Some individual females can and will outperform some individual males. But even the very best female athletes are routinely surpassed not only by the very best male athletes but also by second tier male athletes. For example, the world records in the indoor men's and outdoor women's shot put are quite similar – 74' 10 1/2" for the men and 74'3" for the women. But the women's shot put is 8.8 lbs. while the men's is almost twice as heavy at 16 lbs. The same pattern holds for the women's world records in all of the races on the track from 100 meters to 10,000 meters. Indeed, not only are those records surpassed by many men each year, they are also surpassed by many high school boys. The pattern also holds for high school athletes who aren't yet superstars. With rare exceptions, from the onset of male puberty, even the best high school girls have no chance to succeed against high school boys.

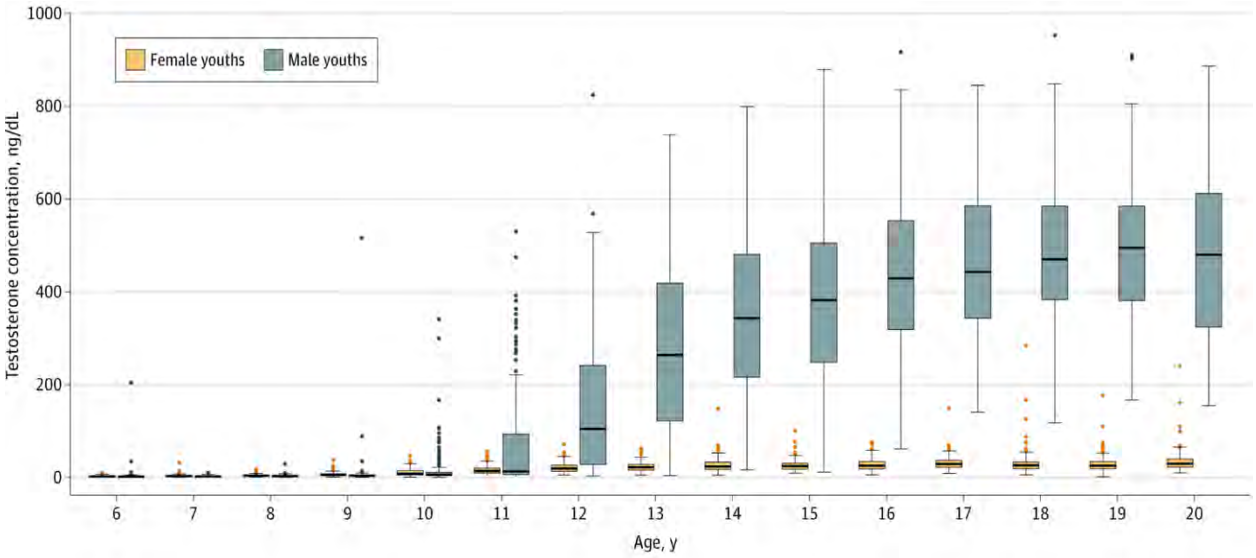
Q8. What does testosterone have to do with separate sex sport—why are we always hearing about testosterone in this context?

A8. Testosterone is an anabolic-androgenic steroid. *Anabolic* steroids build body tissue, including but not limited to bone and muscle tissue and red blood cells. *Androgenic* steroids are responsible for male sex differentiation, i.e., for the development of male primary sex characteristics (in utero), and male secondary sex characteristics (in puberty). Because of its body building and sex differentiation effects, testosterone produced *endogenously* (naturally within the human body) is the primary driver of the sex differences in athletic performance, i.e., of the performance gap between male and female athletes. Beginning at puberty, at approximately age 11, the male testes begin producing significantly more testosterone than they did earlier in childhood, and also significantly more than is ever produced by female ovaries. This increased production triggers the onset of male puberty, and thereafter builds and sustains the male body in the respects that matter for sports performance: speed, strength, power, and endurance. The *exogenous* use of testosterone (doping) is banned by all national and international sports organizations because of these anabolic effects.

Q9. What do people mean when they say that there is a "male range" and a "female range" for testosterone?

A9. Both males and females produce testosterone naturally in their bodies, males primarily in the testes and females primarily in the ovaries. Starting from the onset of male puberty, generally about age 11, testes begin to produce much more testosterone than ovaries. From that point forward, the normal female range is between 0.06 and 1.68 nanomoles per liter (nmol/L), and the normal male range is between 7.7 and 29.4 nmol/L. The gap between top of the female range and the bottom of the male range is 6.02 nmol/L. Converted to ng/dL – the metric typically used in medicine in the U.S. – the normal female range is from 1.73 to 48.45 ng/dL, the normal male range is from 222 to 848 ng/dL, and the gap between the top of the female range and the bottom of the male range is 173 ng/dL.

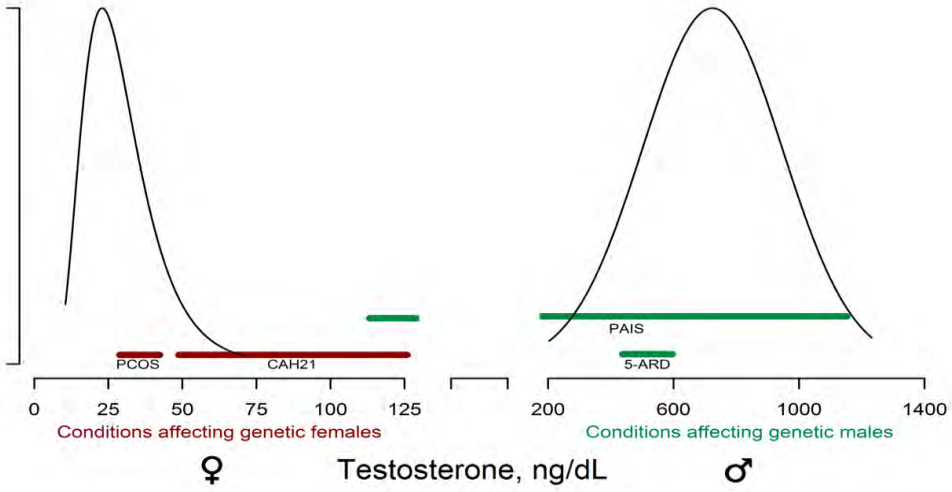
The figure below was published this year in the Journal of the American Medical Association (JAMA). It represents data from the U.S. National Health and Nutrition Examination Survey (NHANES). It shows the increase in testosterone concentration in male youth starting from age 11 onward, as well as the gap that emerges as a result between male and female testosterone levels.



(from J. Senefeld et al., JAMA Research Letter (2020))

Q10. Don't some healthy females produce testosterone in the "male" range?

A10. No. Although females do produce testosterone, mainly in their ovaries, healthy post-pubertal females never produce testosterone levels as high as post-pubertal males. Throughout childhood, up until the onset of male puberty, male and female testosterone levels are about the same; but from the onset of male puberty, male testes produce significantly more testosterone than female ovaries. From that point forward, normal female testosterone levels fall in a narrow range between 0.06 and 1.68 nanomoles per liter (nmol/L), and male levels fall in a broader range between 7.7 and 29.4 nmol/L. The gap between the normal male range and the normal female range is wide. As the following figure indicates, there is no overlap. Some biological females have higher than normal female testosterone levels, for example if they have polycystic ovaries, but again, no healthy female has a testosterone level even close to the normal male range.



This figure shows the normal female testosterone range on the left and the normal male range on the right. It also shows the abnormal testosterone ranges that can be produced by people with certain differences of sex development (DSDs). Some people with DSD prefer to describe themselves as intersex. The conditions marked in red are among those that affect genetic (biological) females. Those marked in green are two that affect (genetic) biological males. Those conditions are described further in the answer to Question 27 below.

Some advocates for trans and intersex athletes claim that there is an overlap in the normal ranges. This claim is not supported by the data or the current peer-reviewed literature. Their argument depends on the existence of a small number of outlier (abnormal) readings, i.e., on a small number of higher-than-normal female T readings and a small number of lower-than-normal male T readings. These abnormal readings are used by advocates to construct a "spectrum" that *appears* to negate the normal bimodal distribution by "filling in" the gap between the two ranges.[14] The figure above shows one way this optical effect can be achieved. It requires ignoring that more than 99% of the population has readings in the normal ranges, and then "filling in" the gap between those ranges with readings from the less than 1% of the population that has an intersex condition.

As the leading experts in the field have established, however, the overlap argument is not supported by the data points themselves, which do not distinguish between (1) doped and non-doped females; (2) females and males with differences of sex development; and (3) male readings taken at rest and following strenuous exercise—the latter has been established to lower normal levels temporarily. Additionally, they measure testosterone by immunoassay – which is inaccurate at lower testosterone concentrations in women – rather than by state-of-the-art methodology, i.e., by mass spectrometry.[15] Once those errors are corrected, the overlap disappears.

---

[14] Advocates may refer to this older paper to support their claim: Healy ML, et al., Endocrine profiles in 693 elite athletes in the postcompetition setting. Clin Endocr. 2014; 81(2): 294-305. PMID: 24593684.

[15] Handelsman DJ, Hirschberg AL, Bermon S. Circulating Testosterone as the Hormonal Basis of Sex Differences in Athletic Performance. Endocr Rev. 2018;39(5):803-29. Epub 2018/07/17. Clark RV, Wald JA, Swerdloff RS, Wang C, Wu FCW, Bowers LD, Matsumoto AM 2019 Large divergence in testosterone concentrations between

Q11. Don't elite female athletes have high testosterone levels—isn't this what makes them good athletes?

A11. No. Elite female athletes generally have testosterone levels within the normal female range, i.e., below 1.68 nmol/L. If they have the condition known as polycystic ovary syndrome (PCOS), they may have testosterone levels up to 3 nmol/L, or, in rare instances, up to 4.8 nmol/L. This is why some sports organizations, wishing to be inclusive of all possible healthy biological females, set their maximum testosterone level at 5 nmol/L.

Q12. Why have many sports organizations adopted a testosterone test for their eligibility standard for inclusion in women's sport?

A12. Testosterone is an excellent proxy for biological sex and a valid basis for an eligibility rule for the women's category for the following reasons:

- Testosterone is the primary driver of the sex differences in athletic performance;
- Sport already tests for and monitors testosterone levels as part of the normal anti-doping process; and
- Different sex testing protocols are more intrusive and, in some cases, less accurate.

No other single criterion so comprehensively addresses sport's and society's concerns about the testing protocol.

Q13. Why have some sports organizations adopted the testosterone level of 5 nmol/L as the upper limit for inclusion in the female category?

A13. Some sports organizations have adopted the level of 5 nmol/L as the upper limit for inclusion in the female category because it represents the outermost bounds that a healthy biological female – regardless of her legal or gender identity – can reach naturally. Almost all females, including elite athletes, have testosterone levels well below 5 nmol/L. The normal female range is between 0.06 and 1.68 nmol/L. Even females with the condition known as polycystic ovary syndrome (PCOS) – which can dramatically raise testosterone levels – only very occasionally reach 3 nmol/L, with rare readings up to 4.8 nmol/L. Setting the level at 5 nmol/L assures that no otherwise healthy biological female could be excluded by the standard. Given that 5 nmol/L is already high, however, some international federations are considering the lower limit of 3 nmol/L.

Q14. Why is only the female category policed for testosterone levels—why doesn't sport also set an upper limit for the male category?

A14. The female category was carved out from open (mixed or co-ed) sport as a protected space where females could compete only against each other and not also against males. It was designed specifically to exclude males, i.e., people with male sex-linked performance advantages. Testosterone is the primary driver of these sex-linked advantages. The male category is not policed because it does not need protection from itself; it was not designed to exclude or regulate males with natural male testosterone levels. Elite sport does, however, monitor testosterone levels in all athletes, male and female, for exogenous use of (i.e., doping with) androgens, including testosterone.

---

men and women: Frame of reference for elite athletes in sex-specific competition in sports, a narrative review. Clin Endocrinol (Oxf) 90:15-22.

Q15. Are advocacy groups correct when they say that there is no evidence that trans girls/women have an advantage over females in sport?

A15. No. They are wrong. Trans girls/women are biologically male. Consequently, unless they go on puberty blockers and then on gender affirming hormones before the onset of male puberty, they benefit from normal male sex development and differentiation. There is overwhelming evidence that individuals who are biologically male – however they identify – have an athletic advantage over individuals who are biologically female—however they identify. Gender identity is unrelated to athletic ability. Additionally, there is convincing evidence[16] that, depending on the task, skill, sport, or event, trans women maintain male sex-linked (legacy) advantages even after a year on standard gender-affirming hormone treatment.

Q16. Are advocacy groups correct when they say that any remaining advantages males have over females in sport are the result of cultural stereotypes and lesser opportunities for development, training, and competition?

A16. No. They are wrong. Although stereotypes and opportunities can affect the degree of the performance gap between the best females and the best males, the data and science are clear that for almost all sports and events the gap itself is biologically-based and immutable.

Q17. What does it mean physically or biologically to say that someone is "transgender"?

A17. A transgender person is currently defined as someone who identifies as other than their biological sex. For example, a trans girl/woman is someone who identifies as a girl/woman even though they are biologically male. A person does not need to take gender affirming hormones or have surgery to be considered transgender. Some transgender people are not on hormones and have not had surgery. Some transgender people take hormones but do not have surgery. And some transgender people do both. Whether a transgender person takes hormones, the level at which they choose to set their hormones, and whether they have surgery, are all matters of personal choice, medical advice, and/or opportunity.

Q18. Do all trans girls/women have a testosterone advantage?

A18. No. Those trans girls/women who *never* experience the onset of male puberty do not develop the secondary sex characteristics that are responsible for the performance gap between male and female athletes. Preventing male puberty involves taking puberty blockers before its onset, and thereafter transitioning to gender affirming hormones that keep testosterone levels consistently within the female range. In contrast, trans girls/women who go on blockers and/or gender affirming hormones and/or have a gonadectomy only *after* they experience some or all of

[16] (1) Roberts TA, Smalley J, Ahrendt D. Effect of gender affirming hormones on athletic performance in transwomen and transmen: implications for sporting organisations and legislators. Br J Sports Med. 2020. Epub 2020/12/09. (2) Hilton EN, Lundberg TR. Transgender Women in The Female Category of Sport: Perspectives on testosterone suppression and performance advantage. Sports Medicine. 2021;51:(in press) (PMID 33289906 and doi: 10.1007/s40279-020-01389-3). (3) Handelsman DJ, Hirschberg AL, Bermon S. Circulating Testosterone as the Hormonal Basis of Sex Differences in Athletic Performance. Endocr Rev. 2018;39(5):803-29. Epub 2018/07/17. See also (4) Gooren LJ, Bunck MC. Transsexuals and competitive sports. Eur J Endocrinol. 2004;151(4):425-9. (5) Wiik A, Lundberg TR, Rullman E, et al. Muscle Strength, Size, and Composition Following 12 Months of Gender-affirming Treatment in Transgender Individuals. J Clin Endocrinol Metab. 2020;105(3). Epub 2019/12/04.

male puberty retain a "legacy advantage" as a result of this experience. The degree of their legacy advantage depends on a combination of factors including: the extent to which they have experienced puberty; whether they had a gonadectomy (surgical removal of their testes); the levels at which they maintain their circulating testosterone; and the particular sport and event in which they compete.

Q19. What is meant by "legacy advantages" in the discussion of trans girls/women in girls/women's sport?

A19. Legacy advantages are the male sex-linked advantages that remain even after a trans girl/woman has gone on gender affirming hormones and/or gender affirming surgery. They are the benefits for sport of having gone through all or part of puberty as a male.

Q20. Does transgender inclusion have anything to do with doping or performance enhancing drugs (PEDs), and if so, what's the connection?

A20. Doping is the exogenous use – the taking – of prohibited performance enhancing drugs (PEDs), including testosterone and other body building androgens. These are among the substances that propelled the East German women to victory in the Olympic Games and World Championships in the 1970s and 1980s, costing clean American women and Team USA to lose out on medals they would otherwise have won. Some American athletes have also doped with androgens, but not in the systematic and state-sponsored way as the East Germans, and more recently the Russians. Trans girls/women who want to be included in girls'/women's events are not doping; that is, they aren't taking PEDs to compete. But their natural testosterone levels build strength, speed, and power in the same way that doping does; and because their natural levels are much higher than even those of doped female athletes, the effect on competition is the same or more overwhelming for the clean females in the field.

Q21. Do we have any data on the impact of trans girls with no medical intervention in girls' high school sports?

A21. Yes. The data that exist about trans girls with no medical intervention are consistent with the fact that they are biologically male. For example, based on its interpretation of the State of Connecticut's Equality Act, the Connecticut Interscholastic Athletics Conference (CIAC) permits trans girls to compete in girls' events even if they have not yet gone on puberty blockers or gender affirming hormones. (The CIAC places no physical or physiological conditions on their inclusion in girls' events). Two trans girls who used to compete on their schools' boys' teams moved to the girls' teams when they came out as trans. They immediately dominated their events at their conference, state, and regional competitions, even though their performances would have been insufficient to qualify them for post-season play had they competed in the boys' divisions. And although they started competing in girls' events before they began taking gender-affirming hormones, they continued to be among the best girls in their events even after they publicly stated they had started on puberty blockers and hormones. All told, just these two trans girls took “15 women’s state championship titles (titles held in 2016 by nine different Connecticut female

athletes) and . . . more than 85 opportunities to participate in higher level competitions from female track athletes in the 2017, 2018, and 2019 seasons."[17]

T MILLER – SPRINTS
55 meters indoors and 100 meters outdoors

GRADE	Hormone Status*	Event	Connecticut Boys' State Rankings	Connecticut Girls' State Rankings
9th	not on gender affirming hormones	Indoor-55m	662nd	32nd
		Outdoor-100m	326th	2nd
10th	not on gender affirming hormones	Indoor-55m	377th	5th
		Outdoor-100m	181st	1st
11th	not on gender affirming hormones	Indoor-55m	118th	1st
		Outdoor-100m	165th	1st
12th	on gender affirming hormones	Indoor-55m	335th	3rd
		Outdoor-100m	- / -	- / -

Miller competed on the boys' track team her freshman year and through the winter of her sophomore year. She came out publicly as transgender in the middle of 10th grade, and then switched to the girls' team for her remaining two-and-a-half years of eligibility.

Her hormone status for each season is derived from publicly-available information.[18] Because that information indicates she went on hormones for the first time only at the end of the 2019 outdoor season, i.e., sometime in May, and because her best time that year was run before then, she is listed here as "not on hormones" for the year.

[17] Verified Complaint for Declaratory and Injunctive Relief and Damages, Seoule et al. v. CIAC, Case No. 3:20-cv-00201, paragraph #77, filed in the United States District Court for the District of Connecticut (Feb. 12, 2020). These results are limited to conference, state and regional championships. They do not include all of the regular season or invitational events at which opportunities to move on through competitions to finals and/or wins and podium spots were affected.

[18] See, e.g., Beyond the Labels: Meet Terry Miller, Runner's Space.com, May 26, 2019, available at https://www.runnerspace.com/gprofile.php?mgroup_id=44531&do=news&news_id=576791 (implying that Miller attended NSAF Nationals as a spectator not a competitor in her junior year, 2019, because she was not eligible to compete there, and Miller herself suggesting that she began taking hormones only in the latter part of that same year).

The table shows rankings for the 55 meters indoors first, followed by the 100 meters outdoors. The rankings in blue font show the division she actually competed in, and the point at which she switched from the boys' to the girls' division. Simply by walking off of the track in the boys' events and walking onto the track in the girls' events, she went from barely being in the top 400 in the state to being #1 in the state.

The girls' rankings for her 9th grade year are those she would have achieved based on her times as run in boys' events. The boys' rankings for her sophomore, junior, and senior years are those she would have achieved based on her times as run in girls' events. There were no rankings for the 100 meters outdoors her 12th grade year (2020) because the season was cancelled due to COVID.

A YEARWOOD – SPRINTS
55 meters indoors and 100 meters outdoors

GRADE	Hormone Status*	Event	Connecticut Boys' State Rankings	Connecticut Girls' State Rankings
9th	not on gender affirming hormones	Indoor-55m	- / -	- / -
		Outdoor-100m	422nd	4th
10th	on gender affirming hormones	Indoor-55m	392nd	5th
		Outdoor-100m	470th	3rd
11th	on gender affirming hormones	Indoor-55m	194th	2nd
		Outdoor-100m	449th	5th
12th	on gender affirming hormones	Indoor-55m	170th	1st
		Outdoor-100m	- / -	- / -

Yearwood competed on the girls' team all four years in high school. She came out publicly as transgender in the 9th grade. Her hormone status for each season is derived from publicly available information.[19] The table shows rankings for the 55 meters indoors first, followed by the 100 meters

[19] See, e.g., Jeff Jacobs, As We Rightfully Applaud Yearwood, We Must Acknowledge Many Questions Remain, Hartford Current, June 17, 2017, available at https://www.courant.com/sports/hc-jacobs-column-yearwood-transgender-0531-20170530-column.html (reporting that Yearwood's father "said his daughter will begin consultations in June [2017] about hormonal treatment"). That was at the end of 9th grade. The fact that she competed at NSAF Nationals in 11th grade (March 2019) means that she was on hormones in 10th grade.

outdoors. The boys' rankings listed on the table are those she would have achieved based on her times run in girls' events. There were no rankings for the 100 meters outdoors her 12th grade year (2020), because the season was cancelled due to COVID.

We don't have statistics on the number of trans girls who have competed in girls' events in high school sports. However, it appears that, at least in the past, most were already on gender-affirming hormones by the time they sought to participate on girls' teams; trans advocacy groups seems generally to assume that this is the case when they speak to the issue. However, we are at a juncture in history where trans girls who are not on hormones are just beginning to ask to be included in girls' competitions. In part this is because the standard of care in trans-medicine now recommends that trans-kids "come out" socially before they transition medically; and many physicians now require that kids wait until they are 16 to go on gender-affirming hormones. For a trans girl, going out for a girls' school sports team is one way to come out socially. We are thus increasingly likely to face situations like that in Connecticut where trans athletes seek to compete in girls'/women's sport while not on hormones.

Q22. Do we have any data on the impact of trans boys with or without medical intervention in high school sports?

A22. Yes. The medical community now recommends that trans kids "come out" socially before they transition medically. While some trans girls have opted to go out for a girls' school sports team as one way to come out socially, this option is not so easily available to trans boys who, because they are biologically female, are unlikely to make a boys' team. As a result, some trans boys have chosen to come out socially while remaining on the girls' team. This has allowed them to continue to participate and to remain competitive in high school sport. Some trans girls have chosen this same path, coming out socially while remaining on the boys' team.

Q23. When post-pubescent trans girls take gender-affirming hormones, do their athletic performances decline? If so, does any performance or "legacy" advantage remain?

A23. Going on gender affirming hormones causes a decline in circulating levels of testosterone which, if consistently maintained over time, has some effect on athletic performance. This effect seems to be primarily on endurance, not on strength and power. The effect on speed seems to be dependent on the extent to which the event is endurance- as opposed to strength- and power-based. Thus, the nature and extent of the decline in male performance advantage, also known as the "legacy advantage", appears to depend on the sport and the event. It also depends on the extent to which the individual experienced male puberty before they began their physical transition, and on how high they choose to maintain their testosterone levels once they do go on gender affirming hormones. Regardless, as we explain in our answer to Question 15, the current state of the peer reviewed literature is that legacy advantages remain significant.

Q24. Why do some sports organizations and governing bodies – including the NCAA – require that trans girls/women reduce their testosterone levels for a year before they can compete in girls'/women's events?

A24. The NCAA, the IOC, and many international federations (IFs) and national governing bodies (NGBs), require trans girls/women to reduce their testosterone levels for at least a year before they can compete in girls'/women's events. This accommodation is a policy compromise,

based in the hypothesis that if a trans girl/woman reduces her testosterone levels into the female range and keeps her levels consistently within that range for at least a year, her male-linked advantages will decline to the point that it is fair to include her in girls'/women's competition. The hypothesis itself is based in the fact that trans girls/women are biologically male and that testosterone is the primary driver of the performance gap between male and female athletes. Just how much gender affirming hormones reduces her male sex-linked advantages and what "legacy advantages" remain is the subject of ongoing investigation.

Q25. Is there strong scientific evidence that trans girls/women have an unfair advantage over biological females even after a year of androgen-suppressing treatment?

A25. Yes. As our answer to Question 15 details, several peer-reviewed studies, including one based on data from the U.S. military, have confirmed that trans women retain their male sex-linked advantages even after a year on gender affirming hormones. This is especially the case for sports and events that are not endurance-based. Because of these retained advantages, USA Powerlifting and World Rugby have recently concluded that it isn't possible fairly and safely to include trans women in women's competition. Other international sports federations have rejected the International Olympic Committee's 2015 guidance suggesting that trans women be included in women's competition so long as they reduce their testosterone levels to the bottom of the male range (under 10 nmol/L). The latter federations, e.g., those that govern the sports of track and field, tennis, cycling, and rowing, have reduced the required testosterone level to within the female range.

Q26. Is the NCAA's testosterone rule for trans women athletes sufficient to ensure fairness to and the safety of the biological females in the field?

A26. No, not as currently administered. The NCAA rule is superficially similar to that of the IOC and other sports governing bodies in that it focuses on testosterone levels; however, as administered it currently lacks their rigor and detail. It provides only that trans women athletes need to be on gender affirming hormones for at least a year. It does not specify that they need to bring their testosterone levels into the female range; it does not require them to keep their levels consistently within that range; and it does not monitor their compliance. The hypothesis that reducing testosterone levels winds down the male performance advantage sufficient to ensure fairness to and safety for the female athletes in the field depends not only on getting those levels into the female range, but also maintaining them consistently within that range throughout the operative period. The NCAA rule has been properly criticized, including by trans women athletes and their coaches, for its lack of monitoring and guidance in these respects.

Q27. What if any is the relationship between intersex and trans athletes?

A27. Intersex conditions result from differences in biological sex development. They are also known as differences of sex development or DSDs. There are many different intersex conditions, but those that are relevant for sport all involve biological males – individuals with an XY karyotype, testes, and testosterone levels in the male range – whose sex development was atypical in some respect. For example, their external genitalia might not be fully formed or their androgen receptors may be less than typically sensitive. Athletes with such intersex conditions may be raised as male or female. People who are transgender do not generally consider themselves to be intersex.

The two are related in sport to the extent that they may both involve biological males with full or partial male advantage who seek eligibility to compete in girls'/women's sport.

The following table is illustrative. It is from an exhibit in the case brought by South African runner Caster Semenya against her international federation (the IAAF now World Athletics) at the Court of Arbitration for Sport (CAS) in Switzerland. Semenya is sometimes described as intersex. In 2019, CAS upheld the federation's eligibility rules for the women's category. Those rules require affected athletes to verifiably reduce their testosterone levels to within the normal female range for a 12-month period before they can compete in that category. Switzerland's Supreme Court affirmed the CAS decision in 2020.

COMPARING BIOLOGICAL SEX TRAITS
FOR PURPOSES OF GIRLS' AND WOMEN'S SPORT
(from IAAF Exhibit in Semenya and ASA v. IAAF)

	Typical Male	Person with 5-ARD (not on hormones)	Person who is Transgender MTF (not on hormones	Typical Female
Chromosomes	46 XY	46 XY	46 XY	46 XX
Gonads and Gametes	Testes & Sperm	Testes & Sperm	Testes & Sperm	Ovaries & Eggs
Endocrine system	Androgenic	Androgenic	Androgenic	Estrogenic
Sex hormones	Testosterone levels in male range	Testosterone levels in male range	Testosterone levels in male range	Testosterone levels in female range
Primary sex characteristics (develop in utero)	Testes, epididymis & vas deferens, prostate	Testes, epididymis & vas deferens, vestigial prostate	Testes, epididymis & vas deferens, prostate	Ovaries, fallopian tubes, uterus, vagina
Virilisation on puberty	Yes	Yes	Yes	No
Secondary sex characteristics (develop at puberty)	Male	Male	Male	Female
External genitalia	Penis, scrotum	Varies	Penis, scrotum	Clitoris, labia
Legal sex	Male	Varies	Varies	Female
Gender Identity	Male	Varies	Female	Female

II. ABOUT CURRENT LAW ON SEX AND SPORT

Q28: What law or laws currently provide for separate sex sport?
A28: Separate sex sport is regulated by a combination of statutes, regulations, and caselaw. This includes the Ted Stevens Olympic and Amateur Sports Act, Title IX and its regulations, the Equity in Athletics Disclosure Act, and court decisions interpreting their terms.

Q29: Are advocacy groups correct when they say that the law affords females the right to participate, not the right to win and set records, in sport?
A29: No. They are wrong. The point of the laws that create and regulate separate sex sport is to ensure that females have the same opportunities as males not only to participate but also to succeed. In addition to competing, this includes the fair ability to win and set records in regional, national, and international competitions. No male or female has an individual legal right to win or set records in their respective divisions, but as a class, females have the legal right to win and set records in girls' and women's sport, just as males that have that right in boys' and men's sport.

Q30: How would the redefinition of "sex" in federal law to include gender identity affect the legal status quo? For example, would it allow schools and sports organizations including the NCAA and USOPC to continue to maintain separate sex sport?
A30: The re-definition of "sex" to include "gender identity" in a law that prohibits discrimination "on the basis of sex" would mean that programs receiving federal funds and operating in interstate commerce could not lawfully distinguish a biological female from a trans girl/woman. This would make it prima facie unlawful to do what is currently permitted, i.e., to have teams and events that are separated on the basis of biological sex. It would also make it prima facie unlawful to use testosterone – a sex-linked trait – as an eligibility criterion for inclusion in girls' and women's elite sport, e.g., as is currently required by the NCAA, the USOPC, the IOC, and the international sports federations. Both separate sex sport itself and eligibility criteria based on biological sex and sex-linked traits like testosterone are currently lawful exceptions to general prohibitions on sex discrimination. For this to remain the case, the Equality Act would need to be amended to provide for an express exception for sport.

Q31: Why do proponents of the Equality Act (EA) assert that the redefinition of sex won't affect girls' and women's sport?
A31. Many of the EA's advocates argue that the proposed EA Act won't affect Title IX, without explaining why. Alternatively, others argue that, even if it does, Congress could restore separate sex sport after the EA's enactment, through specific legislation or regulations addressing sport. Restoring separate sex sport after the EA's enactment is highly unlikely as a matter of standard legal analysis, legislative history, and politics.

The EA is designed to amend the Civil Rights Act of 1964. The definitions in that statute have been and will continue to be the basis for interpreting or defining the same words as used in all other civil rights legislation. That is, Congress cannot re-define "sex" in the principal statute and not have that definition apply directly or indirectly to the use of that term in other legislation. In

fact, many of the EA proponents intend precisely this—make the change to the definition in the principal legislation, and this will automatically change the definition in related legislation. Moreover, as a matter of standard legal analysis, absent a legislative carve out for sport – i.e., an explicit acknowledgement of an exception – any newly enacted, categorical prohibition on discrimination between biological females and trans girls/women would be presumed to supersede any earlier legislation to the contrary, including Title IX.

The legislative history of the EA makes clear that its proponents intend for it to apply to sport with no conditions or exceptions and thus, to prohibit any distinctions between biological females and trans girls/women. At the House Judiciary Committee Hearings, both the witnesses and Democrats on the Committee insisted that trans girls/women be included in girls' and women's sport without any conditions because "trans girls are girls, trans women are women, period." And on the floor of the House, a bill was rejected by a vote of 181 to 228 that would have retained the longstanding exception in Title IX for separate sex sport based on biology. (Specifically, Congressman Steube proposed legislation providing that, "Nothing in this Act or any amendment made by this Act may be construed to diminish any protection under Title IX of the Education Amendments of 1972.") The Equality Act then passed the House by a vote of 236 to 173. This legislative history would be instructive in the future were the question to arise whether Congress intended to permit or preclude distinctions on the basis of biological sex.

The natural experiment with state versions of the EA also make clear that an explicit exception is necessary to maintain sex segregated sports and spaces. In those contexts, trans advocates argue that under the state EAs, it is impermissible to separate or in any way differently to treat trans girls within girls' sport. They make these arguments even though state legislatures did not consider sports as they were enacting their EA legislation.

Q32. How does the recent Supreme Court decision in *Bostock v. Clayton County* (2020) affect separate sex sport—does it prohibit all distinctions on the basis of sex, including in sport?

A32. In *Bostock*, the Supreme Court ruled that "sex" in Title VII means "biological sex." Contrary to what many proponents of the EA argue, *Bostock* did not define (or re-define) "sex" to include "gender identity." Rather, it held that Title VII's general prohibition of discrimination "on the basis of sex" precludes discrimination that takes into account a transgender employee's sex and gender identity. Firing a person because they are transgender – i.e., because their gender identity is nonconforming – requires taking their sex into account, and this is prohibited by Title VII. Because the case involved Title VII's general non-discrimination provision, not an existing exception that allows taking sex into account, the Court wrote that it was leaving the lawfulness of exceptions – including in bathrooms, locker rooms, and sport – for another day. *Bostock* explicitly did not rule on the lawfulness of the current scheme under Title IX and the other sport statutes.

Proponents of the EA nevertheless assert that *Bostock* applies to sport, completely ignoring the Court's express pronouncement to the contrary. Specifically, in cases pending in the lower courts, they argue that *Bostock* supports the redefinition of "sex" to include "gender identity", and that the decision requires the inclusion of trans girls/women in girls' and women's Title IX sport. Notably, however, they are inconsistent in their application of *Bostock* to the question whether administrators can lawfully distinguish biological females from transgender women and girls. For

example, in pending federal cases in Connecticut and Idaho, advocates for transgender athletes argue that their inclusion in girls' high school sports must be full and unconditional, without regard to whether they are on gender affirming hormones. However, in the Idaho matter, in which college sports are also at issue, they don't challenge the NCAA rule which distinguishes female athletes from transgender athletes by requiring trans women to undergo a year of gender affirming hormone treatments before they can compete in women's sport. They support the position that in college, conditions on transgender inclusion are permissible. This distinction between high school and college may make good policy sense; but it is an acknowledgement of the continued lawfulness not only of the NCAA rule, but also more generally of what that rule represents, i.e., the lawfulness of distinctions on the basis of sex in sport. It is also a tacit acknowledgment of the fact that – as the Supreme Court itself announced – *Bostock* is not dispositive in this area.

Q33. Does the law currently allow schools to distinguish females from trans girls/women? Can accommodations be developed that lawfully provide for their conditional inclusion in girls'/women's sport?

A33: The sex exception to general nondiscrimination law requires the exclusion of biological males from most girls' and women's sport. There is no case yet that finally resolves the question whether an exception to this general rule should be made for biological males who identify as women and girls. It is standard practice, however, for the courts to permit (and sometimes even to require) accommodations when there are good reasons for doing so, and when this is possible without imposing an undue burden. Thus, accommodations that would allow trans girls/women to compete in girls'/women's sport should be permissible so long as they meet these standard criteria.

III. ABOUT POLICY

Q34: What are the principles that the Women's Sport Policy Working Group used to develop its approach to trans-inclusion in girls'/women's sport?

A34. The principles that guided the Working Group in the development of its approach to trans inclusion in girls'/women's sport are the following:

1. Women's sport is designed to provide a space where biological females – whatever their gender identity – can compete only against each other and not also against biological males—whatever their gender identity. The design is based in compelling data and scientific evidence on the immutable performance gap between male athletes and female athletes. This separate sex space is worth preserving and protecting. Girls' and women's participation in competitive sport nurtures individual health and development, contributes to the welfare of the community, and powers society's perception of the strength and value of women and girls.

2. Trans girls and women are biologically male and so per the design would normally be excluded. However, because their inclusion could also produce real value both for the individuals concerned and for society, we should work to avoid unnecessary distinctions and exclusions.

3. Physical sex-linked differences between males and females are largely determined from the onset of male puberty; it is these differences that justify separate sex sport, and thus, they must be taken into account in developing responsible policy for girls' and women's sport.

4. Protocols for co-ed sports are instructive, as is the related tradition in law and policy of looking for ways to accommodate rather than to exclude when this is possible without doing harm to an otherwise valuable institutional design. Being transgender does not change the fact of one's biological sex. Where it is recognized in existing co-ed sports policy that sex is relevant to fairness and safety, it cannot be ignored simply because an individual identifies as transgender. Similarly, where existing co-ed sports policy recognizes that sex is not relevant to fairness and safety, the goal should be unconditional inclusion.

5. Specifically rejected as guiding principles are the unscientific, politically driven mantras that claim that:

   - "sex-linked differences including testosterone levels are indistinguishable from other differences like height, weight, wingspan, and foot shape";
   - "the performance gap between male and female athletes is based in myth, stereotype, and cultural inequities";
   - "the physical legacy advantages associated with developing as a biological male don't exist or matter to sports performance";
   - "there is no evidence that transwomen and girls have a competitive advantage over females"; and
   - "females only have the right to participate not to win".

   These patently false claims have no place in a serious discussion of the policy question whether and how to include transgender athletes in girls' and women's sport.

methodology for the development of soc8

Development of Chapters

Refine the topics and review questions

conduct the systematic reviews

Drafting of the recommendation statements

Delphi method to approve recommendations

grading

adding supportive text

editing

dissemination

2.3 Development of Chapters

The Guideline Steering Committee, in discussion with chapter members, determined the chapters for inclusion in the Standards of Care, based on the previous editions of the SOC. Four new chapters were added. The chapters in the Standards of Care 8th Version are:

1. Global Applicability of the Standards of Care
2. Terminology – Diagnostic Criteria
3. Epidemiologic Considerations
4. Overview of Therapeutic Approaches for Gender Health
5. Assessment, Support and Therapeutic Approaches for Children
6. Assessment, Support and Therapeutic Approaches for Adolescents with Gender Variance/Dysphoria **NEW**
7. Assessment for Adults with Gender Variance/Dysphoria
8. Assessment, Support and Therapeutic Approaches for Non-Binary Individuals **NEW**
9. Managing Mental and Behavioral Health Conditions in Adults
10. The Role of Primary Care in Gender Health
11. Hormone Therapy for Adolescents and Adults
12. Sexual Health Across the Lifespan **NEW**
13. Reproductive Health for Adolescents and Adults
14. Voice and Communication Therapy
15. Surgery for Adolescents and Adults and Postoperative Care and Follow-Up
16. Applicability of the Standards of Care to People Living in Institutional Environments
17. Applicability of the Standards of Care to People with Intersex Conditions
18. Applicability of the Standards of Care to Eunuchs **NEW**
19. Competency, Training, Education **NEW**
20. Ethics **NEW**

Back to top

2.4 Methodology of the SOC8

Several stages have been followed to develop the SOC8:

1. Based on the SOC7 the topics were reviewed and main questions were developed
2. Systematic literature reviews were conducted where appropriate
3. Draft recommendations statements were developed by the chapters and reviewed by the chairs
4. Delphi process was followed in order to approve recommendations
5. Approved recommendations were Graded
6. Supportive text was added for each recommendation
7. Independent checking of references of chapters.

The SOC8 chairs have approved every step above.

2.4.1 Refine the Topics and Review Questions

2.4.2 Conduct the Systematic Reviews

2.4.3 Drafting of the Recommendation Statements

2.4.4 Approval of the recommendation following the Delphi Method

2.4.5 Grading

2.4.6 adding Supportive text

2.5 Editing of the SOC8

An independent professional editor has been commissioned in order to edit the whole SOC8 in order for the SOC8 to read as written by one person.

2.6 Distribute Standards of Care for review

Once ready, the whole draft will be distributed via this website for comments from the WPATH members, the WPATH Global Trans Advisory Council, and open for comments for a period of one month. The comments will be reviewed by the SOC8 chairs and chapter leads and if appropriate and necessary changes will be made.

2.7 Disseminate the Standards of Care

The Standards of Care will be printed in a special edition of the International Journal of Transgender Health; this edition will be **open access**.

2.8 Plan to Update

The field of transgender health is rapidly evolving. Small adaptations/changes/addendums to the Standards of Care version 8 may take place. However, should new data become available that will significantly affect specific recommendations a revision of the SOC8 will be considered by a newly to be formed SOC8 Revision Committee. The SOC8 Revision Committee will be recruited and established after the completion and publication of the SOC8.

Back to top

Standards of Care Version 8 home
Learn about the History and Purpose of the SOC8
Learn about Establishing the SOC8 Revision Committee

Learn about the SOC8 Chapters
Meet the Chairs and the Lead Evidence Team

WPATH

World Professional Association for Transgender Health

- Contact
- Media Inquiry Form
- Terms
- Donate

Powered By ConferenceBrain

Login | Register

- About
  - Mission and Vision
  - EC and Board of Directors
  - Committees
  - Regional Organizations
  - Awards and History
- Conferences
  - WPATH 27th Scientific Symposium
- Global Education Institute
  - Upcoming Courses
  - GEI Global Impact
- Standards of Care 8
- Resources
  - Find a WPATH Member Provider
  - Career Center
  - Seniors, Youth, Schools, and Families
  - Public Documents
  - Health Plan Providers Training
  - CA SB855 Materials
- Publications
  - Standards of Care
  - International Journal of Transgender Health
- Membership
  - Membership Information
  - Student Membership
  - Apply for Membership
  - Member Directory
  - Lower Income Economy Countries
- Donate

Heydari *et al. Cell & Bioscience* (2022) 12:4
https://doi.org/10.1186/s13578-021-00741-y

Cell & Bioscience

REVIEW

Open Access

Y chromosome is moving out of sex determination shadow

Raheleh Heydari[1†], Zohreh Jangravi[2†], Samaneh Maleknia[1], Mehrshad Seresht-Ahmadi[3], Zahra Bahari[4], Ghasem Hosseini Salekdeh[5] and Anna Meyfour[1,6*]

Abstract

Although sex hormones play a key role in sex differences in susceptibility, severity, outcomes, and response to therapy of different diseases, sex chromosomes are also increasingly recognized as an important factor. Studies demonstrated that the Y chromosome is not a 'genetic wasteland' and can be a useful genetic marker for interpreting various male-specific physiological and pathophysiological characteristics. Y chromosome harbors male-specific genes, which either solely or in cooperation with their X-counterpart, and independent or in conjunction with sex hormones have a considerable impact on basic physiology and disease mechanisms in most or all tissues development. Furthermore, loss of Y chromosome and/or aberrant expression of Y chromosome genes cause sex differences in disease mechanisms. With the launch of the human proteome project (HPP), the association of Y chromosome proteins with pathological conditions has been increasingly explored. In this review, the involvement of Y chromosome genes in male-specific diseases such as prostate cancer and the cases that are more prevalent in men, such as cardiovascular disease, neurological disease, and cancers, has been highlighted. Understanding the molecular mechanisms underlying Y chromosome-related diseases can have a significant impact on the prevention, diagnosis, and treatment of diseases.

Keywords: Y chromosome, Sex differences, Cancer, Diseases, Male infertility, Inflammation, Neurodegenerative disorders, Germ cell tumors, Prostate cancer, Hepatocellular carcinoma

Background

The human Y chromosome is a haploid male-specific chromosome. It consists of about 60 million base pairs and approximately compromises 2% of the human genome [1]. From the evolution point of view, X and Y chromosomes started to evolve from a pair of ancestral autosomes about 25 million years ago [2]. About 95% of the Y chromosome is composed of the male-specific region of the Y chromosome (MSY), and the other 5% is two pseudoautosomal regions (PAR1 and PAR2) in two ends of this chromosome (Fig. 1). PAR1 and PAR2 with less than 3 Mb in length are the only regions that have maintained the ability to recombine with their X counterparts; therefore, MSY escapes meiotic recombination [3]. Based on evolutionary origin, euchromatic sequences of MSY are divided into three different classes: X-degenerate, X-transposed, and ampliconic sequences. X-degenerate sequences are single copy and broadly expressed genes which were evolved from ancestral autosomes to generate sex chromosomes. Their X homologs excessively escape X chromosome inactivation, thus researchers classified them as dose-sensitive and haplolethal genes. The X-transposed region is a result of a recent X-to-Y transposition that has preserved 99% similarity to their X chromosome sequences. Ampliconic sequences, as the largest part of the MSY, encode nine gene families which were acquired from diverse sources and then

*Correspondence: a.meyfour@sbmu.ac.ir
†Raheleh Heydari and Zohreh Jangravi contributed equally to this work
[1] Basic and Molecular Epidemiology of Gastrointestinal Disorders Research Center, Research Institute for Gastroenterology and Liver Diseases, Shahid Beheshti University of Medical Sciences, Tehran, Iran
Full list of author information is available at the end of the article

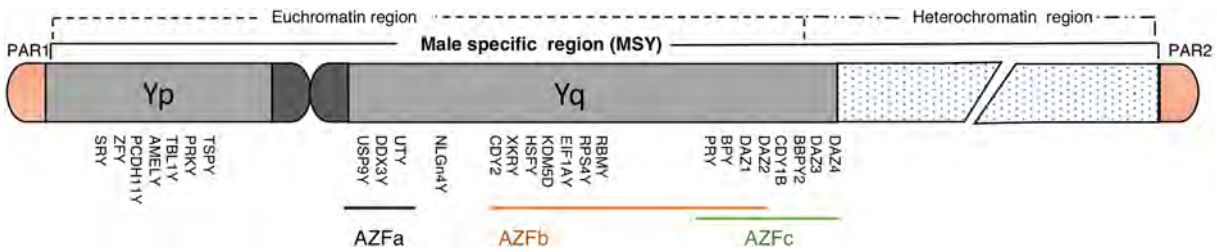


Fig. 1 Schematic representation of human Y chromosome. The location of the azoospermia factor (AZF) regions (a–c) and harboring genes have been shown on Yq. The location of some other genes on Yp has also been indicated

have undergone amplification [1]. Although the number of MSY genes and their X-homologs is small, they have remained conserved in the human genome over time due to their crucial functions [1]. The role of MSY genes in important cellular processes such as transcription regulation, translation, and protein stability in males is vital not only in sex determination but also in sex-dependent organ development [3]. It has been reported that testis, brain, heart, and kidney developments are associated with MSY genes expression [4, 5]. Despite extensive studies on the effect of these genes on the development pathways, some MSY genes have remained as missing proteins with no experimental protein evidence due to highly transient and spatio-temporal restricted expression patterns. For example, TBL1Y is a crucial protein in cardiac differentiation whose expression was first detected during embryonic stem cell differentiation into cardiomyocytes [6]. Furthermore, there are numerous reports on the direct linkage of MSY genes malfunction with several male-specific disorders, as well as gender differences in prevalence and severity of diseases [7–9]. The occurrence of these differences has been observed in both genetic and non-genetic disorders; for example, autism is four times more prevalent in males than females [7]. Although sex-related circulating hormones have been proposed as one of the causes of these differences, Y chromosome genes, with the cooperation of these hormones or independently, may be responsible for above mentioned sexual disparities [8, 10, 11]. In this article, the role of Y chromosome in male-specific diseases (male infertility and prostate cancer (PC), and the ones which primarily affect men such as cardiovascular diseases, inflammatory diseases, and various types of cancers has been reviewed (Fig. 2).

Y chromosome in male infertility

Infertility affects an estimated 15% of couples worldwide and male factors are responsible for about 40% of infertile cases [12]. It has been estimated that more than 2.5% of the male infertility cases occur due to chromosomal abnormalities, among which 1.14% are referred to as sex chromosomal abnormalities [13], such as the structural chromosomal abnormalities of the long arm of the Y chromosome (Yq) [14].

The role of the Y chromosome in male infertility has been extensively studied (for review, see refs [3, 15]). Current knowledge of the function of MSY genes in spermatogenesis is mainly based on the reported microdeletions in the Y chromosome of infertile men. The azoospermia factor (AZF) region which is located in Yq, harbors three subregions (AZFa, AZFb, and AZFc) (Fig. 1) involved in sperm development and function [1, 16]. Almost 25–55% of males with different testicular pathologies such as sperm maturation arrest, sertoli cell-only syndrome (SCOS), hypospermatogenesis, and 5–25% of males with severe oligozoospermia or azoospermia show microdeletions in AZF regions [3]. In addition to AZF complete deletion, the association of partial AZFc deletion such as b1/b3, b2/b3, and gr/gr with male infertility has also been reported [3]. AZFc and AZFb deletions transpired in about 60% and 6–10% of azoospermic patients, respectively, although these statistics can vary in distinct human populations. Recently, the association of AZFb deletions with variable testicular pathologies including meiotic arrest, cryptozoospermia, severe oligozoospermia, or oligoasthenoteratozoospermia has been comprehensively reviewed by Vogt et al. (see [17]).

Deletion in the *DDX3Y* and *USP9Y* genes located in the AZFa region is highly associated with the SCOS phenotype [18]. RNA-binding motif protein Y chromosome (RBMY) and PTPN13-like protein Y-linked (PRY) are the main players during spermatogenesis. RBMY is expressed in spermatogonia, spermatocytes, and round spermatids, indicating its important function as a testis-specific splicing factor in germ cell development. Furthermore, it has been reported that complete meiotic arrest is caused by deletions in *RBMY* and *PRY* genes [16]. *PRY* encodes a tyrosine phosphatase protein which is involved in the apoptosis process required to remove sperm cells carrying chromosomal abnormalities [16, 19].

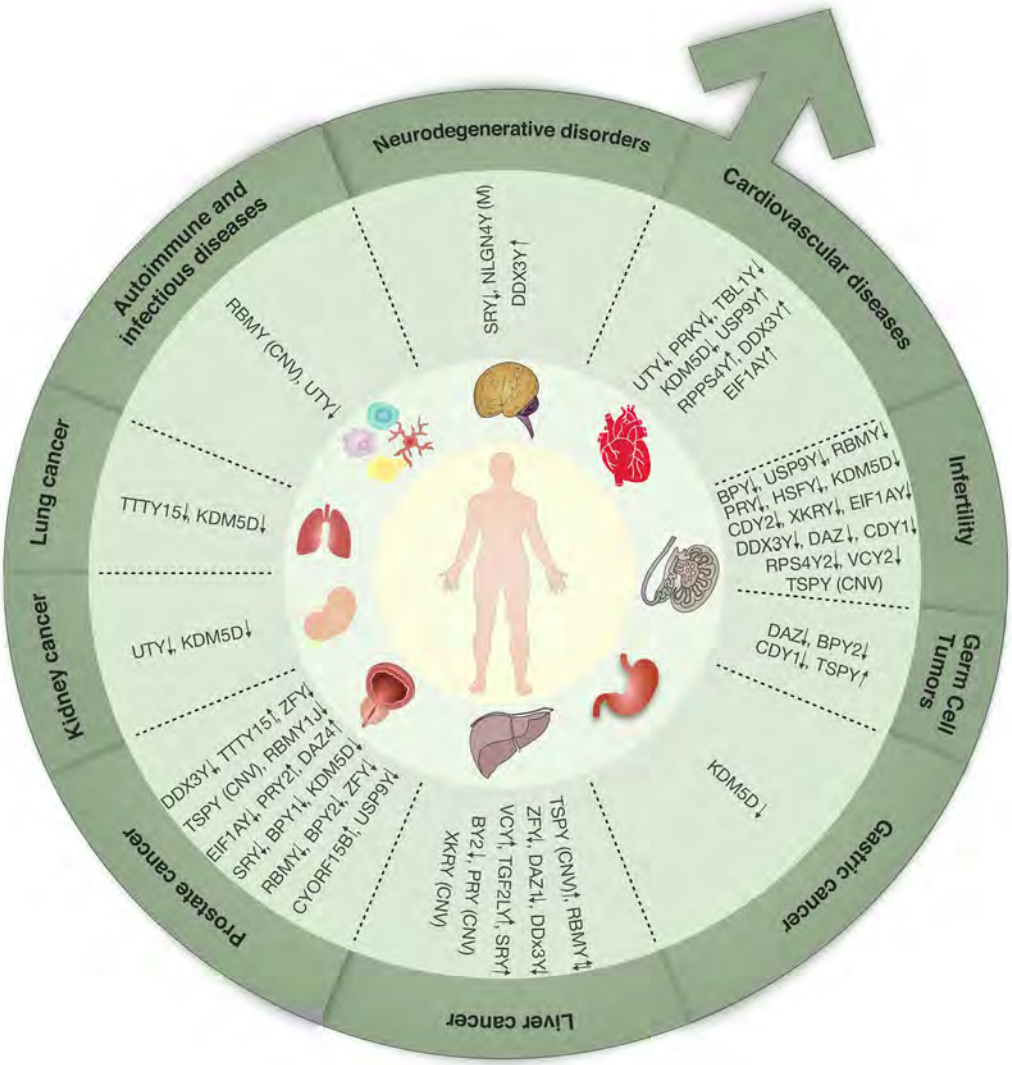


Fig. 2 Overview of Y chromosome genes in different diseases. The map shows Y chromosome genes whose expression and/or function have been confirmed in different diseases. CNV, Copy number variation; M, mutation; ↓, down-regulation; ↑, up-regulation

Alteration in the expression of *HSFY*, *KDM5D/SMCY*, *CDY2*, *XKRY*, *EIF1AY,* and *RPS4Y2*, which are located in the AZFb region, may lead to deteriorated spermatogenesis [16, 20, 21].

DAZ, *CDY1*, *BPY2,* and *PRY* are some genes located in AZFc which can be directly related to the incidence of oligozoospermia and azoospermia [22]. *DAZ* genes encode RNA-binding proteins which are crucial in all stages of spermatogenesis. *DAZ* expression induces the differentiation of pluripotent stem cells (PSC) toward primordial germ cell-like cells and promotes their maturation [23, 24].

CDY1 consists of a chromodomain and a histone acetyltransferase catalytic domain, whose expression is required for histone-protamine replacement in late spermatid nuclei [25]. VCY2 is a highly positive charged protein that is highly expressed in spermatogonia, spermatocytes, and round spermatids. It interacts with ubiquitin-protein ligase E3A, whose expression has been confirmed in ejaculated human spermatozoa, showing their possible positive effect on the preservation of sperm fertility [26, 27]. Differentially expression of some genes on human Y-chromosome such as *HSFY*, *BPY* has been reported in maturation arrest (MA) patients compared to the normal group [28]. Ahmadi Rastegar et al. introduced an isoform level signature of MSY genes to discriminate among MA, SCOS, and normal testicular tissues, which can be considered as a diagnostic marker

for the presence of mature sperm cells in candidate azoospermia men for surgery [28].

In addition to deletion and partial deletions, copy number variations (CNV) of Y chromosome genes can also cause spermatogenesis failure and male infertility. A high-throughput ligation-dependent probe amplification (HLPA) assay was designed to analyze CNVs in the 115 genomic loci covering the Y chromosome. The findings revealed that men with low sperm concentration (LSC) have lower copy numbers for heterochromatic sequences compared with the normal semen group [29]. Chen et al. for the first time, reported that ultra-low relative copy number (RCN) type and low RCN type in Yq12 are more related to male infertility [29]. In contrast, the relation between the additional copy numbers of *TSPY* and spermatogenic failure has also been observed [30].

Chromosomal microarray analysis (CMA) of Y-linked CNVs showed that both CNV size and the involvement of spermatogenesis-related genes determine the clinically relevant CNVs in infertile men [31]. Insufficient copy numbers of the *RBMY* gene can result in asthenozoospermia [32]. CNVs of *DAZ*, *CDY1*, and *BPY2* are correlated with decreased total motile sperm count and lead to azoospermia and moderate/severe oligozoospermia phenotypes [31, 33]. However, screening and detection of mosaic loss of chromosome Y (mLOY) by semi-quantitative multiplex polymerase chain reaction (PCR) and droplet digital PCR showed the infrequency of leukocyte mLOY in young men with spermatogenic failure [34].

Y chromosome in neurodevelopmental and neurodegenerative disorders

About 7 million people worldwide die each year from brain-related diseases. In addition to the high cost of treatment, neurological diseases strongly affect the presence of these patients in social activities and their quality of life. The human brain, as the most complex organ, is affected by sex differences in all anatomical, functional, and biochemical aspects [35]. Sexual dimorphism plays critical roles in various parameters such as brain area volume, cell number and cytoarchitecture, neural functions, synaptic connectivity, perception, cognition, and memory at all stages of development [36, 37]. In sexually dimorphic non-gonadal tissues such as the brain, it has been shown that the development of neurons in the brain is influenced by a regulated combination of the secretion of sex hormones such as testosterone in men and estrogen in women and the function of X and Y chromosomes, which exert sex-specific effects on the development and differentiation of XX and XY neurons [38]. Furthermore, it has been found that the brain cells of men and women, independent of the secretion of sex hormones, follow distinct transcriptional patterns, which can be the cause of differences in the brain developmental pathways, brain function, and behaviors of males and females [39]. The BrainSpan atlas (www.brainspan.org) showed the transcriptional expression of several Y chromosome genes during various stages of male brain development (e.g. *SRY*, *RPS4Y1*, *ZFY*, *PCDH11Y*, *TBL1Y*, *PRKY*, *USP9Y*, *DDX3Y*, *UTY/KDM6C*, *TMSB4Y*, *NLGN4Y*, *HSFY*, *TXLNGY*, *KDM5D* and *EIF1AY*). Furthermore, Vakilian et al. demonstrated that the expression of several MSY genes including *RBMY1*, *EIF1AY*, *DDX3Y*, *HSFY1*, *BPY2*, *PCDH11Y*, *UTY*, *RPS4Y1*, *USP9Y*, *SRY*, *PRY*, and *ZFY* was significantly overexpressed during neural cell differentiation of NTERA-2, a human embryonal carcinoma cell line [40]. They also showed that DDX3Y knockdown inhibited neural cell differentiation of NTERA-2 through cell growth arrest at the G1/S phase and overexpression of pro-apoptotic proteins [40]. There is ample evidence that the prevalence, susceptibility, and progression to deficits in the dopamine system such as Parkinson's disease (PD), attention-deficit hyperactivity disorder (ADHD), schizophrenia, and autism spectrum disorders (ASD), are higher in males than females [37].

Simunovic et al. performed a gene expression analysis on laser microdissected dopamine (DA) neurons from postmortem brains of sporadic PD male and female patients and showed that the major cellular pathways involved in PD pathogenesis such as oxidative phosphorylation, apoptosis, and synaptic transmission were more down-regulated in males compared to females. Results provided strong evidence on sex-specific dysregulation of gene expression in the pathogenesis of sporadic PD [41].

Dewing et al. showed the Y chromosome-linked, male-determining gene *SRY*, which is dominantly expressed in dopamine-abundant regions of the adult brain, directly regulates male-specific brain function. It modulates dopamine biosynthesis and subsequently affects voluntary movement in the male rodents, so it may increase the risk of disorders such as ASD and PD in males [42]. Lee et al. showed that nigral Sry expression persistently increased in animal and cell culture PD models and led to a male-specific mechanism of DA cell death. Reduction of nigral Sry expression by antisense oligonucleotides induced male-specific protective effects through the inhibition of DNA damage, mitochondrial degradation, and neuroinflammation in PD models [43]. These findings indicated that aside from the protective effects of female sex hormones, *Sry* up-regulation may also explain male bias in PD. In addition, it has been observed that mutations in Y chromosome genes such as NLGN4Y may be involved in the development of inherited diseases such as ASD [8]. NLGN4Y is a male-specific cell adhesion molecule belonging to the neuroligin (NLGN) family that plays a critical role in the formation of functional synapses

and regulation of synaptic activity [44]. Thus, mutation or any failure in its translation or function of NLGN4Y can lead to the development of ASD. The *NLGN4Y* mutation in XY men, as well as the increased NLGN4Y expression in males with XYY have been reported to be directly associated with autism [45]. Bioinformatic analyses on ChIP-seq/chip and gene expression datasets have shown that SRY/SOX3 target genes regulate sex-specific developmental processes such as neurodevelopment and potentially could contribute to sex-biased neurodevelopmental disorders. Furthermore, exclusive SRY or SOX3 target genes were found to be more associated with the late gestational and postnatal periods. Analysis of co-expressed networks of SOX3/SRY target genes provided new evidence for the regulatory role of SOX3 in both sexes while SRY exclusively contributes to ASD male predisposition [46]. Loss of chromosome Y (LOY), a mosaic aneuploidy which mainly detected in circulating white blood cells, has been considered as one of the underlying causes of aging-related diseases [47, 48]. Dumanski et al. applied SNP-array and whole-genome next-generation sequencing (WGS) to detect and validate the level of LOY mosaicism in three independent studies of different types including a case–control study and two prospective studies. Results indicated that men with LOY in blood cells are more susceptible to Alzheimer's disease (AD) [47]. Defective immunosurveillance as a result of extreme down-regulation of chromosome Y (EDY) could be a possible explanation for the association between LOY in blood cells and disease processes in other tissues [48].

The association of the Y chromosome and other neurological disorders such as ADHD and schizophrenia have also been investigated in a few studies [49, 50]. Although there are pieces of evidence that indicate the role of PCDH11Y in the susceptibility to psychiatric disorders, they were not supported by the study of Durand et al. in which the frequency of two *PCDH11Y* variants (F885V and K980) in males with different psychiatric disorders such as schizophrenia and ADHD was studied and no significant differences were observed in comparison with control populations [51].

Y chromosome in cardiovascular diseases

Cardiovascular diseases (CVDs) are a group of conditions affecting the heart and blood vessels and are the leading cause of death globally [52]. Studies have been shown gender-specific phenotypes in cardiac physiology and pathophysiology [52, 53]. Incidence, frequency, and severity of coronary artery disease (CAD), as the most common type of CVD, is higher in men than in women [52]. The reason for sexual dimorphism in the prevalence of CVD is not fully understood, however, there are strong shreds of evidence that sex-specific hormones might impact the cardiac homeostasis in females and males resulting in estradiol-related protection and testosterone-associated vulnerability, respectively [53, 54]. In addition to the effective mechanisms of hormones and their receptors, these sexual dimorphisms might also be induced by sex chromosomes [55]. Studies have shown that the differences between the sequence, expression, and regulatory roles of sex chromosomal genes, independent of the gonad and its hormonal influence, result in cell autonomous sexual dimorphism [56, 57]. Comparison of the heart function between two mouse strains, C57BL/6J and C57BL/6J.$Y^{A/J}$ (a chromosome-substituted C57BL/6J line in which the original MSY had been substituted for that from A/J mice) revealed that androgens alone are not sufficient to exert male-specific phenotypes in certain cardiac functions such as circadian rhythms and myocardial functional reserve. Genetic material from MSY was considered as a mandatory element to complete the functions of androgens [58].

A positive correlation has been reported between men diagnosed with Y polysomy and the risk of circulatory system death [59]. Population-based studies showed the risk of CAD increases in carriers of haplogroup I1 compared to other Y chromosome haplogroups, showing pleiotropic effects of the Y chromosome on susceptibility to CAD [5, 10]. The risk of atherosclerotic plaque and femoral artery bifurcations also increases in haplogroup k compared to other ones [60]. Genotyping of 11 MSY markers in three cohorts including 3233 British men showed the association between haplogroup I and increased risk of CAD [61]. Transcriptome-wide analysis of macrophages derived from 134 patients with premature myocardial infarction and 121 normal controls led to identify 30 differentially expressed pathways between haplogroup I1 and carriers of other haplogroups which were majorly involved in immunity, confirming the important role of inflammation in the pathogenesis of CAD [61]. Bloomer et al. showed that the expression of *UTY* and *PRKY* decreased in macrophages derived from men with haplogroup I lineage [62]. Down-regulation of *UTY* in macrophages led to changes in the expression of 59 CAD-related pathways [10]. These observations along with the animal model study indicate that inflammation can be considered as a missing link between the Y chromosome and CAD [5, 63].

The association of MSY genes with the risk factors of CVD including hypertension, circulating total cholesterol, LDL, and paternal history of cardiac diseases has been shown using gene single nucleotide polymorphisms [64–66]. Transcriptome analyses of heart tissues from healthy individuals and patients with non-ischemic cardiomyopathy and new-onset heart failure showed the differences in expression levels of sex chromosome genes.

Y-chromosome-related transcripts including USP9Y, DDX3Y, RPS4Y1, and EIF1AY were overexpressed in males while the expression of two X-linked genes *XIST* and *ZFX* increased in females with new-onset heart failure [67]. The onset of sex-biased protein expression and sex disparities in heart tissue was observed in the early stages of embryonic development, before the gonad formation [68]. By taking advantage of human PSC, Meyfour et al. showed that Y chromosome genes are differentially expressed during cardiac development [6]. Among them, TBL1Y was overexpressed at the cardiac mesoderm stage, an opposite expression pattern to what was observed for its X counterpart, TBL1X [6]. The association of functional null mutations of *TBL1Y* with non-syndromic coarctation of the aorta confirmed its important role in the pathophysiology of CVD [69]. The necessity of KDM5D expression has been reported during the differentiation of human embryonic stem cells into cardiomyocytes. Down-regulation of KDM5D interrupted the cardiac differentiation by inhibiting cell cycle progression [70].

UTY and UTX/KDM6A belong to a subfamily of JmjC domain-containing proteins that catalyze the demethylation of N^{ϵ}-methylated histone 3 lysine 27 (H3K27), an important mark for transcriptional repression [71]. Wang et al. reported that *UTX* knockout (KO) male embryonic stem cells (ESC) showed severe defects in mesoderm differentiation and induction of Brachyury [72]. Regarding the derivation of cardiomyocytes from mesoderm and the regulatory role of Brachyury in ESC differentiation into the mesoderm, the role of UTX in cardiac development can be concluded [72]. They also indicated that *UTY* can partially compensate *UTX* deficiency because male *UTX* KO mouse embryos expressed normal levels of UTY and survived until birth, while female *UTX* KO mice showed defects in embryonic cardiac development and Brachyury expression which led to embryonic lethality [72]. Lee et al. also confirmed the expression of *UTY* and *UTX* in developing mouse embryos [73]. Cardiac differentiation of *UTX*-null ESC revealed that UTX was necessary to activate the cardiac-specific gene program and expression of *UTY* was dependent on *UTX*. However, unlike male ESC, the expression of UTY was detected in *UTX*-null $UTX^{\Delta/y}$ male embryos, indicating the independent expression of *UTY* from *UTX* in male mouse embryos. This study also supported the compensatory role of UTY for its X homolog [73].

Several studies showed that mutations in the BCL6 corepressor (*BCOR*) gene located on the X chromosome could be responsible for different diseases such as Lenz microphthalmia syndrome and oculofaciocardiodental (OFCD) in which cardiac defect is one of the main predominate phenotypes [74, 75]. Zhu et al. reported a 7-month-old boy with Lenz microphthalmia/OFCD syndrome that had multiple defects such as glaucoma, cerebral white matter hypoplasia, and congenital heart defect. Genetic analysis showed a novel missense mutation (c.G1619A; p.R540Q) in *BCOR* [76]. Overexpression of *BCORP1* has been reported during cardiac differentiation of ESC into cardiomyocytes [6] thus like its X counterpart, *BCOR* may contribute to congenital anomalies.

Y chromosome in autoimmune and infectious diseases

Y chromosome has been introduced as a regulatory element of immune cell transcriptome that is involved in susceptibility to autoimmune and infectious diseases [55, 77]. LOY analysis using SNP-arrays in sorted- and single-cells leukocytes showed that LOY in $CD4^+$ T cells, granulocytes, and NK cells can be associated with different diseases such as AD and cancer. Furthermore, RNA-sequencing (RNA-Seq) of leukocytes demonstrated the LOY-associated transcriptional effect (LATE) on autosomal genes. LATE genes were majorly involved in immune functions, explaining how LOY in immune cells increases the risk for diseases [78].

Using chromosome Y-substituted mouse strains, it was shown that variation in copy number of *Sly* and *Rbmy* on the Y chromosome plays a potential role in susceptibility to and severity of autoimmune diseases including experimental allergic encephalomyelitis and myocarditis [77]. Gene expression analysis of circulating naïve $CD4^+$ T cells from 37 patients with the clinically isolated syndrome (CIS), an early form of multiple sclerosis [79] and overlapping with data set obtained from $CD4^+$ T cells of chromosome Y-substituted mouse strains led to identifying 440 genes common between mouse and human which were involved in central dogma, providing further evidence of a profound effect Y chromosome on susceptibility to autoimmune disease [77].

The role of *UTY* and its X homolog *UTX* has been determined in the production of proinflammatory cytokines. Kruidenier et al. introduced a small molecule catalytic site inhibitor that could selectively target the function of H3K27-specific JMJ subfamily and reduce the expression of inflammatory cytokines in LPS-induced macrophages [80]. Dysregulation of *UTY* is a characteristic of Y chromosome haplogroup l [62]. *UTY* encodes a histocompatibility antigen that plays a crucial role in the rejection of male stem cell transplantation [81].

Sex differences in susceptibility to infectious diseases have also been considered as a major challenge in dealing with these types of diseases ignored [82]. Although the influence of the sex hormones on the immune system is undeniable, the influence of sex chromosomes on susceptibility to infectious disease cannot be ignored

[83]. Krementsov et al. showed that genetic variation in chromosome Y affects the survival following murine influenza A virus (IAV) infection. They showed that specific Y chromosome variants increase susceptibility to IAV in males and reinforce activation of proinflammatory IL-17-producing γδ T cells in lung tissue [83]. The association between Y chromosome variants and survival following infection with Coxsackievirus B3 virus (CVB3) has also been reported [84].

The AIDS progression and related death as well as resistance to highly active antiretroviral therapy (HAART) are faster in HIV-infected men with Y haplogroup I than other Y haplogroups [85]. A plethora of epidemiological studies has indicated sex disparities in COVID-19 vulnerability. The prevalence of infection and death is higher in men compared to women [86]. Delanghe et al. reported a marked correlation between COVID-19 prevalence and mortality with R1b-S116 haplotype frequency in the European population [87]. The correlation of the ancestry marker R1b1b2 with both infection and mortality of SARS-CoV-2 needs to be more investigated [88]. These results could be related to the regulatory role of Y chromosome genes in viral infections, and immune and inflammatory responses [5].

Y chromosome and cancer

The impact of sex differences on the risks, incidence, and progression of various cancers has been reported in numerous studies [89, 90]. Cook and colleagues showed the male preference of cancer mortality in different cancer types [91]. Aberrant expression of Y chromosome genes may explain some mechanisms responsible for such sex differences in susceptibility and incidence of cancers [92]. Tricarico and colleagues emphasized a role for the differential activity of X- and Y-linked tumor-suppressor genes in males and females. Enzymatic and non-enzymatic activities of these epigenetic modulators profoundly change the expression of target genes [89].

Liver cancer

Liver cancer is the fourth leading cause of global cancer death and the cause of over 700,000 death annually. Primary liver cancer is classified into different types; hepatocellular carcinoma (HCC, almost 85% of the cases), intrahepatic cholangiocarcinoma (ICC, 10–15% of cases), and also some other rare cases [93]. Etiologically HCC is correlated with a variety of factors like aflatoxin, smoking, heavy alcohol consumption, and especially Hepatitis B virus (HBV) infection [94]. It has been shown that sexual dimorphism is a risk factor for this disease and male cirrhotic patients are more susceptible to develop HCC than female patients [95]. HCC incidence in men is about 3–6 times more than in women [96], therefore, sex is a key factor for prognosis, aggressiveness, and treatment of this type of liver cancer. Although sex hormones like androgens and estrogens have been studied as a potential factor involved in hepatocyte development and enhancers of HCC proliferation, the exact molecular mechanism of this cancer is not fully understood. There is no reasonable evidence on the hormone response of HCC cells, and androgen and estrogen therapy did not indicate a beneficial effect on patients' survival [97]. Furthermore, there is some evidence about the effect of the androgen receptor (AR) on the progression of hepatocarcinogenesis in patients carrying HBV and HCV [98]. The liver is an organ with sexual dimorphism in immune response, mitochondrial function, membrane lipid composition, and gene expression (99). Park et al. investigated the detailed genomic alterations in 5 Korean HCC cell lines using comparative genomic hybridization (CGH). Results showed significant loss of DNA copy number of cancer-related genes on the Y chromosome such as *TSPY*, *XKRY*, *PRY* in comparison with normal samples [100].

Aberrant expression of *TSPY*, *RBMY*, *SRY*, *VCY,* and the other Y chromosome genes has been reported to be involved in hepatocellular carcinogenesis [92, 101–104]. Kido and colleagues showed that *TGF2LY* and *VCY* were up-regulated in about 30% of HCC patients, while *DDX3Y*, *ZFY*, and *DAZ1* were down-regulated in about 70% of patients [92].

Dual functional roles of the Y-linked RBMY have been reported in hepatocarcinogenesis in different studies [101, 103, 105–107]. Tsuei et al. showed the expression of one to four different transcripts of RBMY including wild type and variants with N-terminal RRM deletion, C-terminal SRGY (serine–arginine–glycine–tyrosine) boxes deletion, or deletion of both domains in males with HCC and hepatoblastoma. Given that RBMY is a male germ cell-specific RNA-binding protein and it is not expressed in non-tumor liver counterparts, cirrhotic liver, and the other cancers, *RBMY* could be introduced as a new male-specific oncogene for liver cancer [103].

Western blot and immunohistochemistry (IHC) analyses of animal and human tissues showed that the Ser/Thr phosphorylated RBMY was only expressed in the cytoplasm of human and rodent fetal hepatocytes while its expression was inactivated in mature cells. However, cytoplasmic expression of RBMY was also observed in hepatic cancer stem cells and significantly associated with a poor prognosis and decreased survival rate in HCC patients. Mechanistically, cytoplasmic expression of RBMY leads to inactivation of glycogen synthase kinase 3β, translocation of β-catenin to the nucleus, and abnormal activation of the Wnt/β-catenin signaling pathway, thus facilitating the proliferation and cell cycle progression in HCC cells [106]. Down-regulation of *RBMY*

reduced the transformation and anti-apoptotic ability of HepG2 cells, while expression of *RBMY* induced hepatocarcinogenesis in transgenic mice. In fact, RBMY increased AR activity and induced carcinogenic effects in hepatoma cell lines and human HCC tissues through down-regulation of AR inhibitory variant AR45. Therefore, regulation of AR activity can be considered as another mechanism of RBMY involvement in hepatocarcinogenesis [105].

Data mining of IHC analyses of HCC specimens also confirmed the oncogenic properties of *RBMY* in HCC, however, overexpression of *RBMY* in an HCC cell line HuH-7 and a hepatoblastoma cell line HepG2 showed an inhibitory effect on cell proliferation. Overexpression of *RBMY* in HuH-7 cell line led to down-regulation of the RAS/RAF/MAP and PIP3/AKT signaling pathways and abolished HCC development in a mouse liver cancer model. Altogether, Kido et al. concluded that the expression levels and spatiotemporal patterns of RBMY define the tumor-suppressing or oncogenic roles of RBMY during oncogenic processes. It seems that RBMY functions as a male-specific tumor suppressor at early stages of HCC development and can suppress cell proliferation and pro-oncogenic pathways. However, after surviving and adapting tumor cells in proliferative mode, RBMY acts as a proto-oncogene and induces its chronic effects to promote HCC progression [107].

The association of *TSPY* as a proto-oncogene and inhibitor of anti-oncogenic genes has been reported with a poor prognosis of HCC in men [102]. *TSPY* gene is located within the gonadoblastoma locus on the Y chromosome (GBY) with over 30 tandemly repeats, which increases the risk of gonadoblastoma development in XY patients with disorders of sexual development [108–111]. Expression of this gene is frequently observed in some somatic cancers such as liver cancer [92, 112]. Its overexpression resulted in increased cell proliferation and tumor growth in HCC cases through the suppression of anti-oncogenic genes [111, 113]. It has been reported that cell-cycle regulators and cell division factors like BUB1, CDC25B, CDC45, CENPA, PRC1, PRIM1, RRM2, SPC24, and growth factor receptors like ADGRD1 and HMMR are up-regulated by the *TSPY* gene [102]. Shirakawa and colleagues showed the co-expression of *TSPY* and Glypican-3 (*GPC3*) as a sensitive and specific biomarker of HCC [114]. Kido and colleagues identified a TSPY co-expression network (TCN) which activated in 30% of males with HCC [101]. Ectopic activation of *TSPY* and/or inactivation of its X homolog (*TSPX*) as a tumor suppressor could explain sexual dimorphisms in HCC [112].

There is also some evidence on the oncogenicity of the SRY and the formation of cancer stem cells in male HCC [115]. In an in vivo study, down-regulation of *Sry* resulted in lower malignancy, invasiveness, and tumorigenesis of rHCC cells via the inhibition of Sgf29. Sgf29 is a subunit of the SAGA (Spt–Ada–Gcn5 acetyltransferase) complex that is required to bind tri-methylated lysine-4 of histone H3 (H3K4me3). Studies have been revealed that SRY is the upstream regulator of this gene and up-regulation of *Sgf29* induces tumorigenicity and metastasis through the c-Myc-mediated malignant transformation [116, 117]. Furthermore, data mining of RNA-Seq data of 27 male tumor/non-tumor paired samples from The Cancer Genome Atlas (TCGA) showed that *DAZ1* and *BPY2* are frequently down-regulated in HCC patients [92].

Prostate cancer

Prostate cancer (PC) is a genetically heterogeneous disease and genetic factors play crucial roles in the development of this cancer [118]. It is the second most common cancer and the fifth leading cause of malignancy in men worldwide [119]. It is also the cause of over 3% of death caused by cancer among men [120]. Changes in the Y chromosome genes are likely to be involved in the development and progression of PC. Genomic instability of Y chromosome specific repeated DNA family (DYZ1) has been observed in individuals with PC and it can be used as a marker [121]. Although Y losses occur at high rates in most cancer types, LOY is a rare event in PC [122]. Loss of *SRY*, *ZFY*, *BPY1*, *KDM5D*, *RBMY*, *BPY2,* and other MSY genes has been observed in high grades and advanced stages of this disease [123, 124]. Array-based CGH on prostate tumors and PC cell lines showed that loss of *TSPY* gene copies is associated with an increased risk of PC [125]. Furthermore, specific loci on the Y chromosome can be associated with PC. Some loci increase the incidence of PC and some others decrease it [126–129]. In a study conducted by Nargessi et al., four Y-linked short tandem repeats (STRs), including DYS388, DYS435, DYS437, and DYS439 were genotyped in Malaysian males with PC and healthy controls using a Genetic Analysis System. Results revealed that allele 12 of DYS388, allele 14 of DYS439, or haplotype CAAA are associated with susceptibility to develop PC, and Y-lineages with allele 10 of DYS388 or haplotype AABC are more resistant to the disease. Therefore, these DYS loci as well as haplotypes could be used as a screening method to predict PC susceptibility among Malaysian males [126]. Furthermore, six aberrant DNA methylation sites on the Y chromosome were found in PC tissues, of which cg05163709 site methylation was significantly correlated with PC and was presented as a potential diagnostic biomarker with high specificity and sensitivity [129].

Reverse transcription-PCR (RT-PCR) analysis of Y chromosome genes in a panel of samples diagnosed with

low/high grade prostate adenocarcinoma and benign prostatic hyperplasia (BPH), as well as PC cell lines showed the differential expression patterns of Y chromosome genes such as *SRY*, *PRY*, *TSPY*, *RBMYIH*, *SMCY*, *ZFY* and *EIF1AY* in PC [130, 131]. These results indicated that Y chromosome genes might be either involved in or influenced by oncogenic processes governing PC development and progression. Co-expression networks were reconstructed using an available microarray dataset on normal and different stages of PC tissues, which was deposited with the NCBI Gene Expression Omnibus (GEO), to investigate the role of Y chromosome genes in PC biology. Network analysis led to identify 18 PC-related pathways in which 22 Y chromosome genes were enriched. *CYORF15B*, *DAZ4,* and *PRY2* were up-regulated while *RBMY1J*, *USP9Y*, *DDX3Y,* and *KDM5D* showed an opposite expression pattern and decreased during PC progression [132].

TSPY, a Y chromosome-linked oncogene, is frequently activated in PC and its expression is correlated with the poor prognosis of PC [133, 134]. It can shorten the G2/M stage and accelerate cell proliferation [135]. Other studies showed that AR binds to the *TSPY* promoter and enhances its transcription through the regulation of DNA methylation in PC cells [133]. TSPY and its X-located homolog (TSPX) competitively bind to the AR and play opposing roles in the transactivation functions of AR and AR-Variants which can explain the pathogenesis of male-specific PC as well as sexual dimorphisms in the health and diseases of men [111].

KDM5D, a male-specific histone demethylase has been introduced as a tumor suppressor gene in different studies [136–139]. The knockdown of two different isoforms of KDM5D using the short interfering RNA (siRNA) approach confirmed its tumor suppressor role in a PC cell line [136].

Fluorescence in situ hybridization (FISH) analysis showed that the deletions on Y chromosome could explain the cause of decreased KDM5D expression in PC cells. Furthermore, *KDM5D* knockdown led to aggressive PC by altering the expression of target genes such as cell cycle regulators. ChIP-sequencing and motif analyses of KDM5D-binding sites confirmed that KDM5D as a chromatin modifier binds to promoter regions with co-enrichment of the motifs of critical transcription factors such as the E2F family and MYBL2 that regulate the cell cycle [139]. Furthermore, KDM5D levels were highly reduced in metastatic prostate tumors compared with normal tissues and primary prostate tumors. KDM5D suppresses invasion-associated genes including *MMP1*, *MMP2*, *MMP3*, *MMP7*, and *Slug* in PC cells in vivo and in vitro through H3K4 demethylation [140]. Komura et al. showed the crucial role of AR signaling in the sensitivity of PC cell lines, LNCaP and LAPC4, to docetaxel. Docetaxel is prescribed as an important treatment option for patients with metastatic castration-resistant PC. RNA-Seq followed by functional analyses revealed that KDM5D can be considered as a potential mediator of docetaxel sensitivity in the presence of dihydrotestosterone (DHT). Mechanistically KDM5D binds to AR and controls its transcriptional activity by demethylating H3K4me3 active transcriptional marks [138]. They showed that serine/threonine protein kinase ATR inhibitors could be a new therapeutic approach in aggressive prostate tumors deficient in KDM5D [139].

LNCaP and LAPC4 long noncoding RNAs (lncRNAs) have been shown to get involved in critical physiological and pathological processes, such as PC [141]. In a recent study, it has been shown that lncRNA TTTY15 located in the AZFa region of the Y chromosome, is significantly up-regulated in PC tissues compared to normal ones [142]. Moreover, knockout of Y-chromosomal lncRNA TTTY15 using CRISPR/Cas9 technologies resulted in the inhibition of prostate cell growth and migration in vitro and in vivo, introducing lncRNA TTTY15 as a therapeutic target for PC [142].

Germ cell tumors

Testicular germ cell tumor (TGCT) is the most common malignancy in men ages 15 to 35 [143].

Understanding the etiology and pathogenesis of TGCTs has received considerable attention due to their rising rate [144]. TGCTs have two distinct subtypes including seminomatous and non-seminomatous groups which are characterized by different molecular and histological patterns [145]. The involvement of genetic factors on the Y chromosome in the development of TGCTs has been investigated in various studies [146, 147]. Y microdeletion gr/gr has been introduced as a rare, low-penetrance allele that confers susceptibility to TGCT. Results from the largest European study showed that the *gr/gr* deletion increases the risk of TGCT independently from altered spermatogenesis [146]. *gr/gr* deletion is accompanied by loss of *DAZ*, *BPY2*, and *CDY1* genes, indicating the suppressor roles of these genes in TGCT development [148]. Quantitative PCR (qPCR) by 15 probes spanning the Y chromosome on blood- and buccal-derived DNA samples from two case–control studies including the Familial Testicular Cancer Study (FTC) and the Servicemen's Testicular Tumor Environmental and Endocrine Determinants Study (STEED) revealed the possible association between mLOY and familial TGCT. However, there was not a significant difference in target to reference (T/R) ratio between TGCT cases and controls for the STEED samples. *ZFY*, *AMELY*, *USP9Y*, *DDX3Y*, *TMSB4Y*, *NLGN4Y*, *CYorf15A*, *CYorf15B*, and *EIF1AY*

were markers that showed a significant T/R ratio in FTC samples [149].

Gonadoblastoma is the precursor of invasive TGCTs in dysgenetic gonads. GBY known as an oncogenic locus appears to be involved in TGCT development. IHC results showed that *TSPY* is ectopically and abundantly expressed not only in gonadoblastoma tissues, but also in TGCTs, including the premalignant precursor (carcinoma in situ), seminoma, and non-seminomas [150]. Several studies reported aberrant expression of *TSPY* as the putative gene of GBY, in gonadoblastomas and TGCTs [113, 150, 151]. Protein and gene expression analyses in TGCT tissues compared to normal samples indicated that TSPY co-expressed with proliferative markers such as Ki-67, cyclin B1, and germ cell tumor markers such as PLAP**,** OCT4, and c-KIT [150]. Overexpression of *TSPY* in Hela and NIH3T3 cells as well as in vivo analyses showed that TSPY could enhance cell proliferation and tumorigenesis. Mechanistically, the shortening of the G_2/M transition in overexpressed *TSPY* cells was associated with an early degradation of the mitotic cyclin B1. Degradation of cyclin B1 is required to exit mitosis [113]. Promoter assay and functional domain analyses showed that TSPY is co-localized with AR in the promoters of the endogenous androgen-responsive genes and exacerbates AR functions. Considering the role of AR in cancer progression, the oncogenic role of *TSPY* can be concluded [111].

Other cancers

Although mLOY has been reported as the most frequent somatic variant in different cancers in males, its phenotypic consequences are complex and ambiguous [152]. In a population-based study, Qin et al. proposed a "two-sides" model for the role of LOY in lung cancer in which genetically defined mLOY decreased the risk of lung cancer and predicted a better prognosis while aberrant LOY caused by environmental factors like smoking exerted an effect on lung carcinogenesis [153].

Transcriptome analysis of tumor and matched unaffected pulmonary tissues from patients with non-small-cell lung cancer (NSCLC) led to identify sex-specific co-expression networks. Results showed that partial losses of the Y chromosome, particularly *KDM5D* deficiency at the heart of the co-expression network increases the risk of death in NSCLC male patients, thus may contribute to sexual dimorphism in lung cancer [154]. Analysis of sequencing read coverage of 20 MSY genes and RNA-seq data obtained from normal and tumor tissues also showed that the expression of epigenetic modifiers KDM5D and/or KDM6C is reduced due to LOY in clear cell renal cell carcinoma (ccRCC) [155]. The mechanism of action of *KDM5D* as a tumor suppressor gene has been investigated in gastric cancer (GC) [156]. IHC staining of GC and normal tissues showed the decreased expression of KDM5D in GC. Down-regulation of *KDM5D* accelerated the migration and invasion of GC cells by activating epithelial–mesenchymal transition (EMT). Mechanistically, decreased expression of *KDM5D* induces the expression of cullin 4A (*CUL4A*), which in turn leads to the overexpression of *ZEB1* (EMT inducer) and down-regulation of *p21* and *p53* [156]. It seems that upstream oncogenic factors such as ETS variant 4 (ETV4) and mirR-4661-5p decrease the expression of *KDM5D*, which subsequently results in the activation of downstream oncogenes such as Methionyl-tRNA synthetase 2 (*Mars2*) in GC [157, 158].

DNA analyses of the peripheral blood sample from male patients with PC and colorectal cancer (CC), and healthy controls showed that LOY is a more significant predictor of cancer presence than age [159, 160]. In addition, Agahozo et al. performed ICH and FISH analyses using an X and Y probe to evaluate the prevalence of LOY in male breast cancer (BC). They introduced LOY as an early indicator of male breast carcinogenesis, particularly in estrogen-receptor (ER) and progesterone receptor (PR) negative tumors [161]. Westra et al. assessed the risk of Barrett's esophagus (BE) and esophageal adenocarcinoma (EAC) development among six Y-chromosomal haplogroups including DE, F (xJ, xK), K (xP), J, P (xR1a), and R1a in a white male population. F haplogroup was found to predispose BE patients to cancer development while R1a and K haplogroups were determined as protective factors against BE development [162]. The correlation of LOY with poor prognosis of EAC was also demonstrated by using Y chromosome specific fluorescence in-situ probes [163]. In addition to the above-mentioned studies, a significantly higher frequency of LOY has been reported in blood cells of patients with colorectal, head and neck, bladder, leukemia, and pancreatic cancers compared to healthy individuals [160, 164–167].

The role of Y chromosome-linked noncoding RNAs in cancer progression and suppression is also remarkable. Low *TTTY15* expression resulted in down-regulation of TBX4 and a worse prognosis of NSCLC patients [168]. Brownmiller and colleagues also showed the role of Y chromosome lncRNAs in the radiation response of male NSCLC cells [169]. A Y-linked lncRNA, LINC00278 that encodes a Yin Yang 1 (YY1)-binding micropeptide (YY1BM), is down-regulated in male esophageal squamous cell carcinoma (ESCC). YY1BM, which functions as a tumor suppressor, binds to multifunctional transcription factor YY1 and blocks its interaction to AR, thus decreasing the expression of Eukaryotic Elongation Factor 2 Kinase (eEF2K) and inducing apoptosis in ESCC. Cigarette smoking

negatively affects m6A modification of LINC00278 and YY1BM translation and leads to male ESCC progression [170].

Conclusion

Although most sex differences in occurrence and prevalence of diseases have been associated with the function of sex hormones, molecular studies have assigned a hormone-independent role to the differential expression of genes, especially those located on sex chromosomes. Y chromosome genes independently and/or in conjunction with sex hormones, beyond their X-linked collective tasks determine the male-specific characteristics. In this review, we highlighted major recent findings on the contribution of Y chromosome genes to disease susceptibility to various human diseases and showed that how LOY and translation/function failure of Y chromosome genes can affect the pathogenesis of male-specific diseases.

Despite the vast investigation, little knowledge exists on the molecular mechanisms involved in these sex disparities. This might have been originated from the biological limitations and/or experimental issues such as low expression of MSY genes in rare organs or cell types, high similarity with their X counterparts, hormone effects on intracellular processes, and the absence of mixed-sex experimental groups in cellular, animal, and human studies. In the human Y chromosome proteome project, as a part of the Chromosome-Centric Human Proteome Project (C-HPP), the function of MSY proteins was explored in organ development by taking advantage of PSCs, which are capable of differentiation into all cell types of the human body [171]. We believe that hormone-free systems like PSC and their derivatives as well as organoids, which are in vitro generated copies of human organs, can facilitate the mechanistic studies to explore the role of Y chromosome genes in health and disease and provide novel insights into gender disparity and sex-specific therapeutic strategies for diseases.

Acknowledgements

The authors would like to thank the Research Institute for Gastroenterology and Liver Diseases, Shahid Beheshti University of Medical Sciences for its support.

Authors' contributions

RH and ZH performed the literature search and wrote the manuscript. SM, MSA, ZB, and GHS revised the manuscript. AM wrote and supervised the manuscript. All authors read and approved the final manuscript.

Funding

This study was supported by grants from the Research Institute for Gastroenterology and Liver Diseases, Shahid Beheshti University of Medical Sciences.

Availability of data and materials

Not applicable.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no conflict of interest.

Author details

[1]Basic and Molecular Epidemiology of Gastrointestinal Disorders Research Center, Research Institute for Gastroenterology and Liver Diseases, Shahid Beheshti University of Medical Sciences, Tehran, Iran. [2]Department of Biochemistry, Faculty of Medicine, Baqiyatallah University of Medical Sciences, Tehran, Iran. [3]Department of Basic Science and Advanced Technologies in Biology, University of Science and Culture, Tehran, Iran. [4]Department of Physiology and Medical Physics, Faculty of Medicine, Baqiyatallah University of Medical Sciences, Tehran, Iran. [5]Department of Molecular Sciences, Macquarie University, Macquarie Park, NSW, Australia. [6]Department of Stem Cells and Developmental Biology, Cell Science Research Center, Royan Institute for Stem Cell Biology and Technology, ACECR, Tehran, Iran.

Received: 1 September 2021 Accepted: 21 December 2021
Published online: 04 January 2022

References

1. Skaletsky H, Kuroda-Kawaguchi T, Minx PJ, Cordum HS, Hillier L, Brown LG, et al. The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. Nature. 2003;423(6942):825–37.
2. Bellott DW, Hughes JF, Skaletsky H, Brown LG, Pyntikova T, Cho T-J, et al. Mammalian Y chromosomes retain widely expressed dosage-sensitive regulators. Nature. 2014;508(7497):494–9.
3. Colaco S, Modi D. Genetics of the human Y chromosome and its association with male infertility. Reprod Biol Endocrinol. 2018;16(1):14.
4. Meyfour A, Pooyan P, Pahlavan S, Rezaei-Tavirani M, Gourabi H, Baharvand H, et al. Chromosome-centric human proteome project allies with developmental biology: a case study of the role of Y chromosome genes in organ development. J Proteome Res. 2017;16(12):4259–72.
5. Maan AA, Eales J, Akbarov A, Rowland J, Xu X, Jobling MA, et al. The Y chromosome: a blueprint for men's health? Eur J Hum Genet. 2017;25(11):1181–8.
6. Meyfour A, Ansari H, Pahlavan S, Mirshahvaladi S, Rezaei-Tavirani M, Gourabi H, et al. Y chromosome missing protein, TBL1Y, may play an important role in cardiac differentiation. J Proteome Res. 2017;16(12):4391–402.
7. Skuse DH. Imprinting, the X-chromosome, and the male brain: explaining sex differences in the liability to autism. Pediatr Res. 2000;47(1):9.
8. Serajee FJ, Mahbubul HA. Association of Y chromosome haplotypes with autism. J Child Neurol. 2009;24(10):1258–61.
9. Goulmy E, Termijtelen A, Bradley B, Van Rood J. Y-antigen killing by T cells of women is restricted by HLA. Nature. 1977;266(5602):544–5.
10. Eales JM, Maan AA, Xu X, Michoel T, Hallast P, Batini C, et al. Human Y chromosome exerts pleiotropic effects on susceptibility to atherosclerosis. Arterioscler Thromb Vasc Biol. 2019;39(11):2386–401.
11. Arnold AP. The end of gonad-centric sex determination in mammals. Trends Genet. 2012;28(2):55–61.
12. Al-Achkar W, Wafa A, Moassass F. Cytogenetic abnormalities and Y-chromosome microdeletions in infertile Syrian males. Biomed Rep. 2013;1(2):275–9.
13. Mierla D, Malageanu M, Tulin R, Albu D. Prevalence of chromosomal abnormalities in infertile couples in Romania. Balkan J Med Genet. 2015;18(1):23–30.
14. Kalantari H, Asia S, Totonchi M, Vazirinasab H, Mansouri Z, Moradi SZ, et al. Delineating the association between isodicentric chromosome Y and infertility: a retrospective study. Fertil Steril. 2014;101(4):1091–6.

15. Bansal SK, Gupta G, Rajender S. Y chromosome b2/b3 deletions and male infertility: a comprehensive meta-analysis, trial sequential analysis and systematic review. Mutat Res Rev Mutat Res. 2016;768:78–90.
16. Vogt PH. Azoospermia factor (AZF) in Yq11: towards a molecular understanding of its function for human male fertility and spermatogenesis. Reprod Biomed Online. 2005;10(1):81–93.
17. Vogt PH, Bender U, Deibel B, Kiesewetter F, Zimmer J, Strowitzki T. Human AZFb deletions cause distinct testicular pathologies depending on their extensions in Yq11 and the Y haplogroup: new cases and review of literature. Cell Biosci. 2021;11(1):60.
18. Krausz C, Degl'Innocenti S, Nuti F, Morelli A, Felici F, Sansone M, et al. Natural transmission of USP9Y gene mutations: a new perspective on the role of AZFa genes in male fertility. Hum Mol Genet. 2006;15(18):2673–81.
19. Stouffs K, Lissens W, Verheyen G, Van Landuyt L, Goossens A, Tournaye H, et al. Expression pattern of the Y-linked PRY gene suggests a function in apoptosis but not in spermatogenesis. Mol Hum Reprod. 2004;10(1):15–21.
20. Sato Y, Yoshida K, Shinka T, Nozawa S, Nakahori Y, Iwamoto T. Altered expression pattern of heat shock transcription factor, Y chromosome (HSFY) may be related to altered differentiation of spermatogenic cells in testes with deteriorated spermatogenesis. Fertil Steril. 2006;86(3):612–8.
21. Vinci G, Raicu F, Popa L, Popa O, Cocos R, McElreavey K. A deletion of a novel heat shock gene on the Y chromosome associated with azoospermia. Mol Hum Reprod. 2005;11(4):295–8.
22. Reijo R, Lee T-Y, Salo P, Alagappan R, Brown LG, Rosenberg M, et al. Diverse spermatogenic defects in humans caused by Y chromosome deletions encompassing a novel RNA-binding protein gene. Nat Genet. 1995;10(4):383–93.
23. Reynolds N, Cooke HJ. Role of the DAZ genes in male fertility. Reprod Biomed Online. 2005;10(1):72–80.
24. Kee K, Angeles VT, Flores M, Nguyen HN, Reijo Pera RA. Human DAZL, DAZ and BOULE genes modulate primordial germ-cell and haploid gamete formation. Nature. 2009;462(7270):222–5.
25. Lahn BT, Tang ZL, Zhou J, Barndt RJ, Parvinen M, Allis CD, et al. Previously uncharacterized histone acetyltransferases implicated in mammalian spermatogenesis. Proc Natl Acad Sci. 2002;99(13):8707–12.
26. Tse J, Wong E, Cheung A, O W, Tam P, Yeung W. Specific expression of VCY2 in human male germ cells and its involvement in the pathogenesis of male infertility. Biol Reprod. 2003;69(3):746–51.
27. Wong EY, Jenny Y, Yao K-M, Tam P-C, Yeung WS. VCY2 protein interacts with the HECT domain of ubiquitin-protein ligase E3A. Biochem Biophys Res Commun. 2002;296(5):1104–11.
28. Ahmadi Rastegar D, Sharifi Tabar M, Alikhani M, Parsamatin P, Sahraneshin Samani F, Sabbaghian M, et al. Isoform-level gene expression profiles of human Y chromosome azoospermia factor genes and their X chromosome paralogs in the testicular tissue of non-obstructive azoospermia patients. J Proteome Res. 2015;14(9):3595–605.
29. Chen S, Zhang Q, Chu L, Chang C, Chen Y, Bao Z, et al. Comprehensive copy number analysis of Y chromosome-linked loci for detection of structural variations and diagnosis of male infertility. J Hum Genet. 2021. https://doi.org/10.1038/s10038-021-00973-3.
30. Vodicka R, Vrtel R, Dusek L, Singh AR, Krizova K, Svacinova V, et al. TSPY gene copy number as a potential new risk factor for male infertility. Reprod Biomed Online. 2007;14(5):579–87. https://doi.org/10.1016/s1472-6483(10)61049-8.
31. Zhu Y, Hu L, Cao D, Ou X, Jiang M. Chromosomal microarray analysis of infertile men with azoospermia factor microdeletions. Gene. 2020;735:144389.
32. Yan Y, Yang X, Liu Y, Shen Y, Tu W, Dong Q, et al. Copy number variation of functional RBMY1 is associated with sperm motility: an azoospermia factor-linked candidate for asthenozoospermia. Hum Reprod. 2017;32(7):1521–31.
33. Noordam MJ, Westerveld GH, Hovingh SE, van Daalen SK, Korver CM, van der Veen F, et al. Gene copy number reduction in the azoospermia factor c (AZFc) region and its effect on total motile sperm count. Hum Mol Genet. 2011;20(12):2457–63.
34. Suzuki E, Kobori Y, Katsumi M, Ushijima K, Uchiyama T, Okada H, et al. Copy-number analysis of Y-linked loci in young men with non-obstructive azoospermia: implications for the rarity of early onset mosaic loss of chromosome Y. Reprod Med Biol. 2020;19(2):178–81.
35. Grgurevic N, Majdic G. Sex differences in the brain—an interplay of sex steroid hormones and sex chromosomes. Clin Sci. 2016;130(17):1481–97.
36. Johansson MM, Lundin E, Qian X, Mirzazadeh M, Halvardson J, Darj E, et al. Spatial sexual dimorphism of X and Y homolog gene expression in the human central nervous system during early male development. Biol Sex Differ. 2016;7(1):1–17.
37. Loke H, Harley V, Lee J. Biological factors underlying sex differences in neurological disorders. Int J Biochem Cell Biol. 2015;65:139–50.
38. Carruth LL, Reisert I, Arnold AP. Sex chromosome genes directly affect brain sexual differentiation. Nat Neurosci. 2002;5(10):933–4.
39. Sekido R. The potential role of SRY in epigenetic gene regulation during brain sexual differentiation in mammals. Adv Genet. 2014;86:135–65.
40. Vakilian H, Mirzaei M, Sharifi Tabar M, Pooyan P, Habibi Rezaee L, Parker L, et al. DDX3Y, a male-specific region of Y chromosome gene, may modulate neuronal differentiation. J Proteome Res. 2015;14(9):3474–83.
41. Simunovic F, Yi M, Wang Y, Stephens R, Sonntag KC. Evidence for gender-specific transcriptional profiles of nigral dopamine neurons in Parkinson disease. PLoS ONE. 2010;5(1):e8856.
42. Dewing P, Chiang CW, Sinchak K, Sim H, Fernagut P-O, Kelly S, et al. Direct regulation of adult brain function by the male-specific factor SRY. Curr Biol. 2006;16(4):415–20.
43. Lee J, Pinares-Garcia P, Loke H, Ham S, Vilain E, Harley VR. Sex-specific neuroprotection by inhibition of the Y-chromosome gene, SRY, in experimental Parkinson's disease. Proc Natl Acad Sci USA. 2019;116(33):16577–82.
44. Bottos A, Rissone A, Bussolino F, Arese M. Neurexins and neuroligins: synapses look out of the nervous system. Cell Mol Life Sci. 2011;68(16):2655–66.
45. Zhang C, Milunsky JM, Newton S, Ko J, Zhao G, Maher TA, et al. A neuroligin-4 missense mutation associated with autism impairs neuroligin-4 folding and endoplasmic reticulum export. J Neurosci. 2009;29(35):10843–54.
46. Tahira AC, Barbosa AR, Feltrin ASA, Gastaldi VD, de Toledo VHC, de Carvalho Pereira JG, et al. Putative contributions of the sex chromosome proteins SOX3 and SRY to neurodevelopmental disorders. Am J Med Genet B Neuropsychiatr Genet. 2019;180(6):390–414.
47. Dumanski JP, Lambert J-C, Rasi C, Giedraitis V, Davies H, Grenier-Boley B, et al. Mosaic loss of chromosome Y in blood is associated with Alzheimer disease. Am J Hum Genet. 2016;98(6):1208–19.
48. Caceres A, Jene A, Esko T, Perez-Jurado LA, Gonzalez JR. Extreme downregulation of chromosome Y and Alzheimer's disease in men. Neurobiol Aging. 2020;90:150.e1-e4.
49. Bache WK, DeLisi LE. The sex chromosome hypothesis of schizophrenia: alive, dead, or forgotten? A commentary and review. Mol Neuropsychiatry. 2018;4(2):83–9.
50. Crow TJ. The XY gene hypothesis of psychosis: origins and current status. Am J Med Genet B Neuropsychiatr Genet. 2013;162B(8):800–24.
51. Durand CM, Kappeler C, Betancur C, Delorme R, Quach H, Goubran-Botros H, et al. Expression and genetic variability of PCDH11Y, a gene specific to Homo sapiens and candidate for susceptibility to psychiatric disorders. Am J Med Genet B Neuropsychiatr Genet. 2006;141B(1):67–70.
52. Molina E, Clarence EM, Ahmady F, Chew GS, Charchar FJ. Coronary artery disease: why we should consider the Y chromosome. Heart Lung Circ. 2016;25(8):791–801.
53. Regitz-Zagrosek V, Oertelt-Prigione S, Seeland U, Hetzer R. Sex and gender differences in myocardial hypertrophy and heart failure. Circ J. 2010;74(7):1265–73.
54. Blenck CL, Harvey PA, Reckelhoff JF, Leinwand LA. The importance of biological sex and estrogen in rodent models of cardiovascular health and disease. Circ Res. 2016;118(8):1294–312.
55. Regitz-Zagrosek V, Kararigas G. Mechanistic pathways of sex differences in cardiovascular disease. Physiol Rev. 2017;97(1):1–37.
56. Arnold AP. Y chromosome's roles in sex differences in disease. Proc Natl Acad Sci USA. 2017;114(15):3787–9.
57. Snell DM, Turner JMA. Sex chromosome effects on male-female differences in mammals. Curr Biol. 2018;28(22):R1313–24.

58. Praktiknjo SD, Picard S, Deschepper CF. Comparisons of chromosome Y-substituted mouse strains reveal that the male-specific chromosome modulates the effects of androgens on cardiac functions. Biol Sex Differ. 2016;7:61.
59. Higgins CD, Swerdlow AJ, Schoemaker MJ, Wright AF, Jacobs PA. Mortality and cancer incidence in males with Y polysomy in Britain: a cohort study. Hum Genet. 2007;121(6):691–6.
60. Voskarides K, Hadjipanagi D, Papazachariou L, Griffin M, Panayiotou AG. Evidence for contribution of the y chromosome in atherosclerotic plaque occurrence in men. Genet Test Mol Biomarkers. 2014;18(8):552–6.
61. Charchar FJ, Bloomer LD, Barnes TA, Cowley MJ, Nelson CP, Wang Y, et al. Inheritance of coronary artery disease in men: an analysis of the role of the Y chromosome. Lancet. 2012;379(9819):915–22.
62. Bloomer LD, Nelson CP, Eales J, Denniff M, Christofidou P, Debiec R, et al. Male-specific region of the Y chromosome and cardiovascular risk: phylogenetic analysis and gene expression studies. Arterioscler Thromb Vasc Biol. 2013;33(7):1722–7.
63. Khan SI, Andrews KL, Jennings GL, Sampson AK, Chin-Dusting JPF. Y chromosome, hypertension and cardiovascular disease: is inflammation the answer? Int J Mol Sci. 2019;20(12):2892.
64. Shankar RR, Charchar FJ, Eckert GJ, Saha C, Tu W, Dominiczak AF, et al. Studies of an association in boys of blood pressure and the Y chromosome. Am J Hypertens. 2007;20(1):27–31.
65. Charchar FJ, Tomaszewski M, Lacka B, Zakrzewski J, Zukowska-Szczechowska E, Grzeszczak W, et al. Association of the human Y chromosome with cholesterol levels in the general population. Arterioscler Thromb Vasc Biol. 2004;24(2):308–12.
66. Ellis JA, Stebbing M, Harrap SB. Association of the human Y chromosome with high blood pressure in the general population. Hypertension. 2000;36(5):731–3.
67. Heidecker B, Lamirault G, Kasper EK, Wittstein IS, Champion HC, Breton E, et al. The gene expression profile of patients with new-onset heart failure reveals important gender-specific differences. Eur Heart J. 2010;31(10):1188–96.
68. Shi W, Sheng X, Dorr KM, Hutton JE, Davies HA, Andrade TD, et al. Cardiac sex differences are established prior to gonad formation. bioRxiv. 2020.
69. Tagariello A, Breuer C, Birkner Y, Schmidt S, Koch A, Cesnjevar R, et al. Functional null mutations in the gonosomal homologue gene TBL1Y are associated with non-syndromic coarctation of the aorta. Curr Mol Med. 2012;12(2):199–205.
70. Meyfour A, Pahlavan S, Ansari H, Baharvand H, Salekdeh GH. Down-regulation of a male-specific H3K4 demethylase, KDM5D, impairs cardiomyocyte differentiation. J Proteome Res. 2019;18(12):4277–82.
71. Walport LJ, Hopkinson RJ, Vollmar M, Madden SK, Gileadi C, Oppermann U, et al. Human UTY(KDM6C) is a male-specific N-methyl lysyl demethylase. J Biol Chem. 2014;289(26):18302–13.
72. Wang C, Lee JE, Cho YW, Xiao Y, Jin Q, Liu C, et al. UTX regulates mesoderm differentiation of embryonic stem cells independent of H3K27 demethylase activity. Proc Natl Acad Sci USA. 2012;109(38):15324–9.
73. Lee S, Lee JW, Lee SK. UTX, a histone H3-lysine 27 demethylase, acts as a critical switch to activate the cardiac developmental program. Dev Cell. 2012;22(1):25–37.
74. Hilton E, Johnston J, Whalen S, Okamoto N, Hatsukawa Y, Nishio J, et al. BCOR analysis in patients with OFCD and Lenz microphthalmia syndromes, mental retardation with ocular anomalies, and cardiac laterality defects. Eur J Hum Genet. 2009;17(10):1325–35.
75. Ng D, Thakker N, Corcoran CM, Donnai D, Perveen R, Schneider A, et al. Oculofaciocardiodental and Lenz microphthalmia syndromes result from distinct classes of mutations in BCOR. Nat Genet. 2004;36(4):411–6.
76. Zhu X, Dai FR, Wang J, Zhang Y, Tan ZP, Zhang Y. Novel BCOR mutation in a boy with Lenz microphthalmia/oculo-facio-cardio-dental (OFCD) syndrome. Gene. 2015;571(1):142–4.
77. Case LK, Wall EH, Dragon JA, Saligrama N, Krementsov DN, Moussawi M, et al. The Y chromosome as a regulatory element shaping immune cell transcriptomes and susceptibility to autoimmune disease. Genome Res. 2013;23(9):1474–85.
78. Dumanski JP, Halvardson J, Davies H, Rychlicka-Buniowska E, Mattisson J, Moghadam BT, et al. Immune cells lacking Y chromosome show dysregulation of autosomal gene expression. Cell Mol Life Sci. 2021;78(8):4019–33.
79. Corvol JC, Pelletier D, Henry RG, Caillier SJ, Wang J, Pappas D, et al. Abrogation of T cell quiescence characterizes patients at high risk for multiple sclerosis after the initial neurological event. Proc Natl Acad Sci USA. 2008;105(33):11839–44.
80. Kruidenier L, Chung CW, Cheng Z, Liddle J, Che K, Joberty G, et al. A selective jumonji H3K27 demethylase inhibitor modulates the proinflammatory macrophage response. Nature. 2012;488(7411):404–8.
81. Vogt MH, Goulmy E, Kloosterboer FM, Blokland E, de Paus RA, Willemze R, et al. UTY gene codes for an HLA-B60-restricted human male-specific minor histocompatibility antigen involved in stem cell graft rejection: characterization of the critical polymorphic amino acid residues for T-cell recognition. Blood. 2000;96(9):3126–32.
82. Klein SL, Hodgson A, Robinson DP. Mechanisms of sex disparities in influenza pathogenesis. J Leukoc Biol. 2012;92(1):67–73.
83. Krementsov DN, Case LK, Dienz O, Raza A, Fang Q, Ather JL, et al. Genetic variation in chromosome Y regulates susceptibility to influenza A virus infection. Proc Natl Acad Sci USA. 2017;114(13):3491–6.
84. Case LK, Toussaint L, Moussawi M, Roberts B, Saligrama N, Brossay L, et al. Chromosome Y regulates survival following murine coxsackievirus b3 infection. G3 Genes Genomes Genet. 2012;2(1):115–21.
85. Sezgin E, Lind JM, Shrestha S, Hendrickson S, Goedert JJ, Donfield S, et al. Association of Y chromosome haplogroup I with HIV progression, and HAART outcome. Hum Genet. 2009;125(3):281–94.
86. Vahidy FS, Pan AP, Ahnstedt H, Munshi Y, Choi HA, Tiruneh Y, et al. Sex differences in susceptibility, severity, and outcomes of coronavirus disease 2019: cross-sectional analysis from a diverse US metropolitan area. PLoS ONE. 2021;16(1):e0245556.
87. Delanghe JR, De Buyzere ML, De Bruyne S, Van Criekinge W, Speeckaert MM. The potential influence of human Y-chromosome haplogroup on COVID-19 prevalence and mortality. Ann Oncol. 2020;31(11):1582–4.
88. Ibrahim M, Salih A. The Y chromosome ancestry marker R1b1b2: a surrogate of the SARS-CoV-2 population affinity. Hum Genome Var. 2021;8(1):11.
89. Tricarico R, Nicolas E, Hall MJ, Golemis EA. X- and Y-linked chromatin-modifying genes as regulators of sex-specific cancer incidence and prognosis. Clin Cancer Res. 2020;26(21):5567–78.
90. Fajkovic H, Halpern JA, Cha EK, Bahadori A, Chromecki TF, Karakiewicz PI, et al. Impact of gender on bladder cancer incidence, staging, and prognosis. World J Urol. 2011;29(4):457–63.
91. Cook MB, McGlynn KA, Devesa SS, Freedman ND, Anderson WF. Sex disparities in cancer mortality and survival. Cancer Epidemiol Biomarkers Prev. 2011;20(8):1629–37.
92. Kido T, Lau YF. Roles of the Y chromosome genes in human cancers. Asian J Androl. 2015;17(3):373–80.
93. El-Serag HB. Epidemiology of viral hepatitis and hepatocellular carcinoma. Gastroenterology. 2012;142(6):1264-73.e1.
94. Yang JD, Hainaut P, Gores GJ, Amadou A, Plymoth A, Roberts LR. A global view of hepatocellular carcinoma: trends, risk, prevention and management. Nat Rev Gastroenterol Hepatol. 2019;16(10):589–604.
95. Tarao K, Ohkawa S, Shimizu A, Harada M, Nakamura Y, Ito Y, et al. The male preponderance in incidence of hepatocellular carcinoma in cirrhotic patients may depend on the higher DNA synthetic activity of cirrhotic tissue in men. Cancer. 1993;72(2):369–74.
96. Lui WY, Lin HL, Chau GY, Liu TY, Chi CW. Male predominance in hepatocellular carcinoma: new insight and a possible therapeutic alternative. Med Hypotheses. 2000;55(4):348–50.
97. Nagasue N, Kohno H. Hepatocellular carcinoma and sex hormones. HPB Surg. 1992;6(1):1–6.
98. Ruggieri A, Barbati C, Malorni W. Cellular and molecular mechanisms involved in hepatocellular carcinoma gender disparity. Int J Cancer. 2010;127(3):499–504.
99. Dhir RN, Dworakowski W, Thangavel C, Shapiro BH. Sexually dimorphic regulation of hepatic isoforms of human cytochrome p450 by growth hormone. J Pharmacol Exp Ther. 2006;316(1):87–94.
100. Park SJ, Jeong SY, Kim HJ. Y chromosome loss and other genomic alterations in hepatocellular carcinoma cell lines analyzed by CGH and CGH array. Cancer Genet Cytogenet. 2006;166(1):56–64.

101. Kido T, Lau YC. Identification of a TSPY co-expression network associated with DNA hypomethylation and tumor gene expression in somatic cancers. J Genet Genomics. 2016;43(10):577–85.
102. Kido T, Lau YC. The Y-linked proto-oncogene TSPY contributes to poor prognosis of the male hepatocellular carcinoma patients by promoting the pro-oncogenic and suppressing the anti-oncogenic gene expression. Cell Biosci. 2019;9:22.
103. Tsuei DJ, Hsu HC, Lee PH, Jeng YM, Pu YS, Chen CN, et al. RBMY, a male germ cell-specific RNA-binding protein, activated in human liver cancers and transforms rodent fibroblasts. Oncogene. 2004;23(34):5815–22.
104. Li S, Mo C, Huang S, Yang S, Lu Y, Peng Q, et al. Over-expressed Testis-specific Protein Y-encoded 1 as a novel biomarker for male hepatocellular carcinoma. PLoS ONE. 2014;9(2):e89219.
105. Tsuei DJ, Lee PH, Peng HY, Lu HL, Su DS, Jeng YM, et al. Male germ cell-specific RNA binding protein RBMY: a new oncogene explaining male predominance in liver cancer. PLoS ONE. 2011;6(11):e26948.
106. Chua HH, Tsuei DJ, Lee PH, Jeng YM, Lu J, Wu JF, et al. RBMY, a novel inhibitor of glycogen synthase kinase 3beta, increases tumor stemness and predicts poor prognosis of hepatocellular carcinoma. Hepatology. 2015;62(5):1480–96.
107. Kido T, Tabatabai ZL, Chen X, Lau YC. Potential dual functional roles of the Y-linked RBMY in hepatocarcinogenesis. Cancer Sci. 2020;111(8):2987–99.
108. Yin YH, Li YY, Qiao H, Wang HC, Yang XA, Zhang HG, et al. TSPY is a cancer testis antigen expressed in human hepatocellular carcinoma. Br J Cancer. 2005;93(4):458–63.
109. Salo P, Kaariainen H, Petrovic V, Peltomaki P, Page DC, de la Chapelle A. Molecular mapping of the putative gonadoblastoma locus on the Y chromosome. Genes Chromosomes Cancer. 1995;14(3):210–4.
110. Kido T, Lau YF. The human Y-encoded testis-specific protein interacts functionally with eukaryotic translation elongation factor eEF1A, a putative oncoprotein. Int J Cancer. 2008;123(7):1573–85.
111. Li Y, Zhang DJ, Qiu Y, Kido T, Lau YC. The Y-located proto-oncogene TSPY exacerbates and its X-homologue TSPX inhibits transactivation functions of androgen receptor and its constitutively active variants. Hum Mol Genet. 2017;26(5):901–12.
112. Kido T, Lo RC, Li Y, Lee J, Tabatabai ZL, Ng IO, et al. The potential contributions of a Y-located protooncogene and its X homologue in sexual dimorphisms in hepatocellular carcinoma. Hum Pathol. 2014;45(9):1847–58.
113. Oram SW, Liu XX, Lee TL, Chan WY, Lau YF. TSPY potentiates cell proliferation and tumorigenesis by promoting cell cycle progression in HeLa and NIH3T3 cells. BMC Cancer. 2006;6:154.
114. Shirakawa H, Kuronuma T, Nishimura Y, Hasebe T, Nakano M, Gotohda N, et al. Glypican-3 is a useful diagnostic marker for a component of hepatocellular carcinoma in human liver cancer. Int J Oncol. 2009;34(3):649–56.
115. Liu C, Ren YF, Dong J, Ke MY, Ma F, Monga SPS, et al. Activation of SRY accounts for male-specific hepatocarcinogenesis: implication in gender disparity of hepatocellular carcinoma. Cancer Lett. 2017;410:20–31.
116. Kurabe N, Katagiri K, Komiya Y, Ito R, Sugiyama A, Kawasaki Y, et al. Deregulated expression of a novel component of TFTC/STAGA histone acetyltransferase complexes, rat SGF29, in hepatocellular carcinoma: possible implication for the oncogenic potential of c-Myc. Oncogene. 2007;26(38):5626–34.
117. Kurabe N, Murakami S, Tashiro F. SGF29 and Sry pathway in hepatocarcinogenesis. World J Biol Chem. 2015;6(3):139–47.
118. Brothman AR, Maxwell TM, Cui J, Deubler DA, Zhu XL. Chromosomal clues to the development of prostate tumors. Prostate. 1999;38(4):303–12.
119. Rawla P. Epidemiology of prostate cancer. World J Oncol. 2019;10(2):63–89.
120. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68(6):394–424.
121. Yadav SK, Kumari A, Javed S, Ali S. DYZ1 arrays show sequence variation between the monozygotic males. BMC Genet. 2014;15:19.
122. Stahl PR, Kilgue A, Tennstedt P, Minner S, Krohn A, Simon R, et al. Y chromosome losses are exceedingly rare in prostate cancer and unrelated to patient age. Prostate. 2012;72(8):898–903.
123. Perinchery G, Sasaki M, Angan A, Kumar V, Carroll P, Dahiya R. Deletion of Y-chromosome specific genes in human prostate cancer. J Urol. 2000;163(4):1339–42.
124. Jangravi Z, Alikhani M, Arefnezhad B, Sharifi Tabar M, Taleahmad S, Karamzadeh R, et al. A fresh look at the male-specific region of the human Y chromosome. J Proteome Res. 2013;12(1):6–22.
125. Vijayakumar S, Hall DC, Reveles XT, Troyer DA, Thompson IM, Garcia D, et al. Detection of recurrent copy number loss at Yp11.2 involving TSPY gene cluster in prostate cancer using array-based comparative genomic hybridization. Cancer Res. 2006;66(8):4055–64.
126. Nargesi MM, Ismail P, Razack AH, Pasalar P, Nazemi A, Oshkoor SA, et al. Linkage between prostate cancer occurrence and Y-chromosomal DYS loci in Malaysian subjects. Asian Pac J Cancer Prev. 2011;12(5):1265–8.
127. Eeles RA, Olama AA, Benlloch S, Saunders EJ, Leongamornlert DA, Tymrakiewicz M, et al. Identification of 23 new prostate cancer susceptibility loci using the iCOGS custom genotyping array. Nat Genet. 2013;45(4):385–91, 91e1–2.
128. Han Y, Rand KA, Hazelett DJ, Ingles SA, Kittles RA, Strom SS, et al. Prostate cancer susceptibility in men of African Ancestry at 8q24. J Natl Cancer Inst. 2016;108(7):djv431.
129. Yao L, Ren S, Zhang M, Du F, Zhu Y, Yu H, et al. Identification of specific DNA methylation sites on the Y-chromosome as biomarker in prostate cancer. Oncotarget. 2015;6(38):40611–21.
130. Lau YF, Zhang J. Expression analysis of thirty one Y chromosome genes in human prostate cancer. Mol Carcinog. 2000;27(4):308–21.
131. Dasari VK, Goharderakhshan RZ, Perinchery G, Li LC, Tanaka Y, Alonzo J, et al. Expression analysis of Y chromosome genes in human prostate cancer. J Urol. 2001;165(4):1335–41.
132. Khosravi P, Gazestani VH, Asgari Y, Law B, Sadeghi M, Goliaei B. Network-based approach reveals Y chromosome influences prostate cancer susceptibility. Comput Biol Med. 2014;54:24–31.
133. Leng X, Liu M, Tao D, Yang B, Zhang Y, He T, et al. Epigenetic modification-dependent androgen receptor occupancy facilitates the ectopic TSPY1 expression in prostate cancer cells. Cancer Sci. 2021;112(2):691–702.
134. Lau YC, Li Y, Kido T. Battle of the sexes: contrasting roles of testis-specific protein Y-encoded (TSPY) and TSPX in human oncogenesis. Asian J Androl. 2019;21(3):260–9.
135. Li Y, Lau YF. TSPY and its X-encoded homologue interact with cyclin B but exert contrasting functions on cyclin-dependent kinase 1 activities. Oncogene. 2008;27(47):6141–50.
136. Jangravi Z, Tabar MS, Mirzaei M, Parsamatin P, Vakilian H, Alikhani M, et al. Two splice variants of Y chromosome-located lysine-specific demethylase 5D have distinct function in prostate cancer cell line (DU-145). J Proteome Res. 2015;14(9):3492–502.
137. Jangravi Z, Najafi M, Shabani M. Investigation of histone lysine-specific demethylase 5D KDM5D) isoform expression in prostate cancer cell lines: a system approach. Iran Biomed J. 2016;20(2):117–21.
138. Komura K, Jeong SH, Hinohara K, Qu F, Wang X, Hiraki M, et al. Resistance to docetaxel in prostate cancer is associated with androgen receptor activation and loss of KDM5D expression. Proc Natl Acad Sci USA. 2016;113(22):6259–64.
139. Komura K, Yoshikawa Y, Shimamura T, Chakraborty G, Gerke TA, Hinohara K, et al. ATR inhibition controls aggressive prostate tumors deficient in Y-linked histone demethylase KDM5D. J Clin Invest. 2018;128(7):2979–95.
140. Li N, Dhar SS, Chen TY, Kan PY, Wei Y, Kim JH, et al. JARID1D is a suppressor and prognostic marker of prostate cancer invasion and metastasis. Cancer Res. 2016;76(4):831–43.
141. Evans JR, Feng FY, Chinnaiyan AM. The bright side of dark matter: lncRNAs in cancer. J Clin Invest. 2016;126(8):2775–82.
142. Xiao G, Yao J, Kong D, Ye C, Chen R, Li L, et al. The long noncoding RNA TTTY15, which is located on the Y chromosome, promotes prostate cancer progression by sponging let-7. Eur Urol. 2019;76(3):315–26.
143. Batool A, Karimi N, Wu XN, Chen SR, Liu YX. Testicular germ cell tumor: a comprehensive review. Cell Mol Life Sci. 2019;76(9):1713–27.
144. Segal R. Surveillance programs for stage I nonseminomatous germ cell tumors of the testis. Urol Oncol. 2006;24(1):68–74.
145. Singla N, Lafin JT, Ghandour RA, Kaffenberger S, Amatruda JF, Bagrodia A. Genetics of testicular germ cell tumors. Curr Opin Urol. 2019;29(4):344–9.

146. Moreno-Mendoza D, Casamonti E, Paoli D, Chianese C, Riera-Escamilla A, Giachini C, et al. gr/gr deletion predisposes to testicular germ cell tumour independently from altered spermatogenesis: results from the largest European study. Eur J Hum Genet. 2019;27(10):1578–88.
147. Linger R, Dudakia D, Huddart R, Easton D, Bishop DT, Stratton MR, et al. A physical analysis of the Y chromosome shows no additional deletions, other than Gr/Gr, associated with testicular germ cell tumour. Br J Cancer. 2007;96(2):357–61.
148. Anderson PD, Lam MY, Poirier C, Bishop CE, Nadeau JH. The role of the mouse y chromosome on susceptibility to testicular germ cell tumors. Cancer Res. 2009;69(8):3614–8.
149. Machiela MJ, Dagnall CL, Pathak A, Loud JT, Chanock SJ, Greene MH, et al. Mosaic chromosome Y loss and testicular germ cell tumor risk. J Hum Genet. 2017;62(6):637–40.
150. Li Y, Tabatabai ZL, Lee TL, Hatakeyama S, Ohyama C, Chan WY, et al. The Y-encoded TSPY protein: a significant marker potentially plays a role in the pathogenesis of testicular germ cell tumors. Hum Pathol. 2007;38(10):1470–81.
151. Kersemaekers AM, Honecker F, Stoop H, Cools M, Molier M, Wolffenbuttel K, et al. Identification of germ cells at risk for neoplastic transformation in gonadoblastoma: an immunohistochemical study for OCT3/4 and TSPY. Hum Pathol. 2005;36(5):512–21.
152. Forsberg LA, Rasi C, Malmqvist N, Davies H, Pasupulati S, Pakalapati G, et al. Mosaic loss of chromosome Y in peripheral blood is associated with shorter survival and higher risk of cancer. Nat Genet. 2014;46(6):624–8.
153. Qin N, Li N, Wang C, Pu Z, Ma Z, Jin G, et al. Association of mosaic loss of chromosome Y with lung cancer risk and prognosis in a Chinese population. J Thorac Oncol. 2019;14(1):37–44.
154. Willis-Owen SAG, Domingo-Sabugo C, Starren E, Liang L, Freidin MB, Arseneault M, et al. Y disruption, autosomal hypomethylation and poor male lung cancer survival. Sci Rep. 2021;11(1):12453.
155. Arseneault M, Monlong J, Vasudev NS, Laskar RS, Safisamghabadi M, Harnden P, et al. Loss of chromosome Y leads to down regulation of KDM5D and KDM6C epigenetic modifiers in clear cell renal cell carcinoma. Sci Rep. 2017;7:44876.
156. Shen X, Hu K, Cheng G, Xu L, Chen Z, Du P, et al. KDM5D inhibit epithelial–mesenchymal transition of gastric cancer through demethylation in the promoter of Cul4A in male. J Cell Biochem. 2019;120(8):12247–58.
157. Gu J, Chu K. Increased Mars2 expression upon microRNA-4661-5p-mediated KDM5D downregulation is correlated with malignant degree of gastric cancer cells. Cell Biol Int. 2021;45(10):2118–28.
158. Cai L, Chen Q, Fang S, Lian M, Lian M, Cai M. ETV4 promotes the progression of gastric cancer through regulating KDM5D. Eur Rev Med Pharmacol Sci. 2020;24(5):2442–51.
159. Noveski P, Madjunkova S, Sukarova Stefanovska E, Matevska Geshkovska N, Kuzmanovska M, Dimovski A, et al. Loss of Y chromosome in peripheral blood of colorectal and prostate cancer patients. PLoS ONE. 2016;11(1):e0146264.
160. Asim A, Agarwal S, Avasthi KK, Sureka S, Rastogi N, Dean DD, et al. Investigation of LOY in prostate, pancreatic, and colorectal cancers in males: a case-control study. Expert Rev Mol Diagn. 2020;20(12):1259–63.
161. Agahozo MC, Timmermans MA, Sleddens HF, Foekens R, Trapman-Jansen AM, Schröder CP, et al. Loss of Y-chromosome during male breast carcinogenesis. Cancers. 2020;12(3):631.
162. Westra W, Rygiel A, Mostafavi N, De Wit G, Roes A, Moons L, et al. The Y-chromosome F haplogroup contributes to the development of Barrett's esophagus-associated esophageal adenocarcinoma in a white male population. Dis Esophagus. 2020;33(9):doaa011.
163. Loeser H, Wölwer CB, Alakus H, Chon S-H, Zander T, Buettner R, et al. Y chromosome loss is a frequent event in Barrett's adenocarcinoma and associated with poor outcome. Cancers. 2020;12(7):1743.
164. Minner S, Kilgué A, Stahl P, Weikert S, Rink M, Dahlem R, et al. Y chromosome loss is a frequent early event in urothelial bladder cancer. Pathology. 2010;42(4):356–9.
165. Veiga LC, Bergamo NA, Reis PP, Kowalski LP, Rogatto SR. Loss of Y-chromosome does not correlate with age at onset of head and neck carcinoma: a case-control study. Braz J Med Biol Res. 2012;45(2):172–8.
166. Hollows R, Wei W, Cazier JB, Mehanna H, Parry G, Halford G, et al. Association between loss of Y chromosome and poor prognosis in male head and neck squamous cell carcinoma. Head Neck. 2019;41(4):993–1006.
167. Shahrabi S, Khodadi E, Saba F, Shahjahani M, Saki N. Sex chromosome changes in leukemia: cytogenetics and molecular aspects. Hematology. 2018;23(3):139–47.
168. Lai IL, Chang YS, Chan WL, Lee YT, Yen JC, Yang CA, et al. Male-specific long noncoding RNA TTTY15 inhibits non-small cell lung cancer proliferation and metastasis via TBX4. Int J Mol Sci. 2019;20(14):3473.
169. Brownmiller T, Juric JA, Ivey AD, Harvey BM, Westemeier ES, Winters MT, et al. Y Chromosome LncRNA are involved in radiation response of male non-small cell lung cancer cells. Cancer Res. 2020;80(19):4046–57.
170. Wu S, Zhang L, Deng J, Guo B, Li F, Wang Y, et al. A novel micropeptide encoded by Y-linked LINC00278 links cigarette smoking and ar signaling in male esophageal squamous cell carcinoma. Cancer Res. 2020;80(13):2790–803.
171. Alikhani M, Karamzadeh R, Rahimi P, Adib S, Baharvand H, Salekdeh GH. Human proteome project and human pluripotent stem cells: odd bedfellows or a perfect match? J Proteome Res. 2020;19(12):4747–53.

Publisher's Note

"Nature vs. Nurture: Have Performance Gaps Between Men and Women Reached an Asymptote?"
by Millard-Stafford M, Swanson AE, Wittbrodt MT
International Journal of Sports Physiology and Performance

Note. This article will be published in a forthcoming issue of the *International Journal of Sports Physiology and Performance*. The article appears here in its accepted, peer-reviewed form, as it was provided by the submitting author. It has not been copyedited, proofread, or formatted by the publisher.

Section: Invited Commentary

Article Title: Nature vs. Nurture: Have Performance Gaps Between Men and Women Reached an Asymptote?

Authors: Mindy Millard-Stafford, Ann E. Swanson, and Matthew T. Wittbrodt

Affiliations: School of Biological Sciences, Georgia Institute of Technology, Atlanta, GA.

Journal: *International Journal of Sports Physiology and Performance*

Acceptance Date: February 8, 2018

DOI: https://doi.org/10.1123/ijspp.2017-0866

"Nature vs. Nurture: Have Performance Gaps Between Men and Women Reached an Asymptote?"
by Millard-Stafford M, Swanson AE, Wittbrodt MT
International Journal of Sports Physiology and Performance

INVITED COMMENTARY

NATURE VS. NURTURE: HAVE PERFORMANCE GAPS BETWEEN MEN AND WOMEN REACHED AN ASYMPTOTE?

Mindy Millard-Stafford, Ann E. Swanson, Matthew T. Wittbrodt

School of Biological Sciences, Georgia Institute of Technology, Atlanta, GA 30332

Address correspondence to: Mindy Millard-Stafford, PhD
Exercise Physiology Laboratory-BioSci
Georgia Institute of Technology
555 14th St.
Atlanta, GA 30332-0356
Email: mindy.millardstafford@ap.gatech.edu
Phone: (404) 894-6274
FAX: (404 894-9982

Running Head: Have Performance Gaps Reached an Asymptote?
Co-authors email and phone: annelizabethswanson@gmail.com**; (813) 546-5424**
mwittbrodt3@gatech.edu**; (231) 590-8940**

"Nature vs. Nurture: Have Performance Gaps Between Men and Women Reached an Asymptote?"
by Millard-Stafford M, Swanson AE, Wittbrodt MT
International Journal of Sports Physiology and Performance

ABSTRACT

Men outperform women in sports requiring muscular strength and/or endurance, but the relative influence of "nurture" versus "nature" remains difficult to quantify. Performance gaps between elite men and women are well-documented using world records in second, centimeter or kilogram sports. However, this approach is biased by global disparity in reward structures and opportunities for women. Despite policies enhancing female participation (Title IX legislation), USA women only closed performance gaps by 2 and 5% in Olympic Trial swimming and running, respectively, from 1972 to 1980 (with no change thereafter through 2016). Performance gaps of 13% in elite mid-distance running and 8% in swimming (~4 min duration) remain, the 5% differential between sports indicative of load carriage disadvantages of higher female body fatness in running. Conversely, sprint swimming exhibits a greater sex difference than sprint running suggesting anthropometric/power advantages unique to swim block starts. The ~40 y plateau in the performance gap suggests a persistent dominance of biological influences (e.g., longer limb levers, greater muscle mass, aerobic capacity, lower fat mass) on performance. Current evidence suggests women will not swim or run as fast as men in Olympic events, which speaks *against* eliminating sex segregation in these individual sports. Whether hormone reassignment sufficiently levels the playing field in Olympic sports for transgender females (born and socialized male) remains an issue to be tackled by sport governing bodies.

Key Words: athletics; track; swimming; gender; sex difference

"Nature vs. Nurture: Have Performance Gaps Between Men and Women Reached an Asymptote?"
by Millard-Stafford M, Swanson AE, Wittbrodt MT
International Journal of Sports Physiology and Performance

Introduction

Sport is one of few institutions where men and women typically (but not always) are placed in distinct categories to compete, a practice established during the Modern Era of the Olympic Games. The battle of the sexes in sport continues to be a topic of interest in both scientific literature and the media (e.g., under isolated competitions when women outperform male competitors). The sex difference in performance is highly cited over several decades, predominantly in sports measured objectively by time (e.g., running, swimming), centimeters and kilograms. However, elite athletes by their nature are "outliers" so any valid comparison of men and women must be equally representative of the population distribution (e.g. top 0.5-1%) to assess differences.

Elite men outperform women in sports requiring muscular strength/endurance and/or aerobic capacity. [1,2,3] Much of this performance gap is explained by biological sex differences, but environmental influences arguably play an important role although difficult to quantify. [4] It was predicted [5] women would eventually equal or exceed men in running (200 m through the marathon) based upon greater rates of female performance improvement extrapolated over time. This sparked an ongoing debate; [6,7] specifically, will women athletes surpass men, particularly as distance increases in sports where greater inherent body fatness poses less of a disadvantage?

The performance gap (i.e., % sex difference) currently appears relatively stable for international distance running [3,8] and swimming, [9] but may be increasing [10] in sprint events, particularly when observed within shorter specific time frames. Current theories also suggest if the distance becomes sufficiently long (e.g., ultra-marathon running, swimming), [6,11-12] women will close the gap. However, contributing factors are not solely biological, but also socio-cultural (e.g., norms of acceptance, access to coaching, motivation, participation opportunities). [13] Both female

"Nature vs. Nurture: Have Performance Gaps Between Men and Women Reached an Asymptote?"
by Millard-Stafford M, Swanson AE, Wittbrodt MT
International Journal of Sports Physiology and Performance

percentage of events (47.5%) and competitors (45.5%) continued to increase in the 2016 Olympic Games.

Objective of this Commentary

Current approaches to examine the sex difference in performance rely upon world records or international race databases limited by selection bias due to unequal cultural acceptance and sport opportunities for women globally.[4] Thus, an alternate approach (i.e., examining sport performances following a major societal shift within a nation) attenuates this bias and facilitates partitioning the effects of nature vs. nurture. Our goal was to focus upon the effects on the performance gap following "equal opportunity" legislation in the USA.

Methodological Approach

Following 1972 federal legislation (Title IX), sport opportunities for USA women increased beyond high school (e.g., due to college scholarships), although the deadline for institutions to comply was much later. Moreover, recent financial incentives from USA National Team/Olympic medal compensation and professionalism are similar for men and women in swimming and running. [4] Therefore, examination of American elite performances since 1972 (Title IX) could reflect relative contributions of socio-cultural factors (nurture) versus assumed constants of biological differences (nature) on the performance gap. Swimming is distinct from other competitive sports in the USA, because boys and girls traditionally train together from early ages (7-8 y) up through college, having similar access to coaching and competition.[4] Moreover, unlike distance running, women's Olympic swimming has been held for over a century (similar to men), and long considered a socially acceptable, "gender neutral" sport for girls. [14] Comparison to another Olympic weight-bearing sport (i.e., running) with events of similar duration/ intensity

"Nature vs. Nurture: Have Performance Gaps Between Men and Women Reached an Asymptote?"
by Millard-Stafford M, Swanson AE, Wittbrodt MT
International Journal of Sports Physiology and Performance

could account for other biologically-relevant differences in performance. Furthermore, any historical change would estimate the magnitude of the effect that socio-cultural factors played in male-female differences, while any plateau thereafter theoretically representing the extent to which biological factors persist.

The Olympic Trials represents the most highly competitive event in the USA with all top athletes competing in a single competition. Thus, this yields a representative sample distribution of elite performances (as opposed to a random individual swim or track meet) under similar environmental conditions (unlike world records). Performance times for top eight finishers were extracted from official archived results on the websites of USA Swimming[15] and USA Track and Field News.[16] Only 100 to 1500 m events in athletics were extracted due to incomplete historical data (several female events not added until 1984), and distance only up to 400 m swimming where both men and women compete in the Olympics (with change to occur in 2020). The % sex difference was calculated for each pairwise comparison (1st place male vs. 1st place female through 8th place) similar to other studies [7,10,17] using the following equation (where n = nth placing for a given event): Sex Difference (%) = [($Female_n$ (s) – $Male_n$ (s)) / $Male_n$ (s)] * 100.

Historical progression of the performance gap

Data are illustrated each year by distance with a locally weighted polynomial regression for swim and run events (Figure 1 A and B), with an early plateau clearly observed in both sports. Following Title IX, overall mean swim performance gap was higher in 1972 (13.2 ± 2.0%) and 1976 (12.4 ± 1.7%) versus all subsequent years (1980-2016), remaining stable thereafter at 11.2 ± 1.7% (across all swim strokes/distances a net change of 2%). Surprisingly, the swimming performance gap was actually lower in 1968 compared to 1972, possibly due to the Olympic Trials held in different pools and dates (influencing performance due to timing of taper/rest cycles). For

"Nature vs. Nurture: Have Performance Gaps Between Men and Women Reached an Asymptote?"
by Millard-Stafford M, Swanson AE, Wittbrodt MT
International Journal of Sports Physiology and Performance

running (Panel B), the performance gap was also higher in 1972 (17.3 ± 3.0%) and 1976 (14.2 ± 2.4%) compared to all subsequent years (1980-2016), remaining stable at 12.6 ± 2.0% (net change of 4.7%). Greater sex differences occurred in swim sprints (50/100m) but, conversely, lower in run sprints (100/200m).

Compared to men in the 1972 Olympic Trials, women swimmers in 2016 would have placed among the top 8 men in 1972 in all events (winning 100 m Breaststroke, 400 m Freestyle*) except* for 100 m Freestyle (Table 1). In contrast, for every running event *except* the marathon, 2016 female winners would have placed *last* in the 1972 field of male competitors (Table 2). Narrowing of the run performance gap ranged from 3% (200 m) to 7% (800 m). However, the 5 km (added later to the Olympics) showed no closure in the gap and in 10 km, the gap actually widened significantly by 3%. This apparent paradox compared to swimming is likely due to technological enhancements (e.g. pool construction, lane lines and swim suit/equipment designs) and rule changes/stroke techniques that dramatically improved overall swim performances since 1972 as compared to a relative flattening in track performances.

Mandating similar athletic opportunities for girls (Title IX) clearly narrowed the performance gap. We estimate increased opportunities for women in the U.S. closed the performance gap up to 5%, but at a magnitude of less than half (2%) in a sport (swimming) with a longer history of elite competition and social acceptance for women. Moreover, this narrowing in the gap occurred relatively soon (within 8 y), remaining stable at ~13 and 11% for run/swim sprints to mid-distance events, similar to world best times using regression models finding a stable gap (since 1983) averaging 10-11% for running (all distances through the marathon) and 9% for swimming.[9] As suggested previously[10], the question is whether current environmental influences (opportunity, reward structures) are minimally contributing to this gap, suggesting differences are

"Nature vs. Nurture: Have Performance Gaps Between Men and Women Reached an Asymptote?"
by Millard-Stafford M, Swanson AE, Wittbrodt MT
International Journal of Sports Physiology and Performance

now attributed predominantly to nature? At the elite level, significant sex differences persist, which speaks *against* eliminating sex segregation in these Olympic sports. [18]

Influence of event duration

Performance gaps in events of similar duration are compared in Figure 2 (averaged across 1980-2016) for < 30 s (50 m freestyle vs. 200 m run), ~60 s (100 m freestyle vs. 400 m run), ~120 s (200 m freestyle vs. 800 m run) and ~240 s (400 m freestyle vs. 1500 m run). Performance gaps differed by sport across all event durations but were lower in running 200 m compared to swimming 50 m freestyle by 2% and higher for running in events ≥ 60 s, widening to 5% at ~4 min. As distance increased, performance gaps in swimming became progressively smaller; however, this inverse relationship was absent in running, a consistent finding across world record run performances from 1500 m through the marathon. [3,8,19]

The smallest performance gaps (100 m run, 400 m freestyle) occurred on opposite ends of the sprint-endurance spectrum. Based on the 5% higher gap in running versus swimming in a 4 min Olympic event, fat mass carriage likely contributes ~5% of the running performance gap, representing ~37% of the total performance gap in the 1500 m. Interestingly, this corresponds to an experimental model where adding excess mass to men (normalizing to their female counterparts' body fatness) reduced the sex difference in run performance by 30% [20], a stronger predictor than run economy or cardiorespiratory capacity.[21]

After "removing" the impact of fat/load carriage, the residual male advantage in swimming 400 m remains at ~8% (Figure 2). Numerous factors may account for this. Lower absolute cardiorespiratory capacity due to: reduced hemoglobin mass, blood volume, heart/lung size relative to stature, and muscle mass likely explain female disadvantages [1,8] as well as shorter levers applying force (distance per stroke). Leg power off start/turns appear to contribute ~1% of the

"Nature vs. Nurture: Have Performance Gaps Between Men and Women Reached an Asymptote?"
by Millard-Stafford M, Swanson AE, Wittbrodt MT
International Journal of Sports Physiology and Performance

male advantage [22] in 400 m pool swimming, based upon slightly lower differences (~7%) in elite 10 km open water swimming. [17] Reduced sex differences as swim distance increases is consistent with others, [7,23] but contrasts with ultra-distance running where the magnitude of performance gaps may widen (up to ~17% faster for men). [24,25] A lower disadvantage for women in open water distance swimming[26,27] is attributed to enhanced center of buoyancy, [28] swimming economy,[19] greater mechanical efficiency [29] and lower underwater torque (tendency for feet to sink). [30] However, data from select ultra-distance events (i.e., non-representative samples)[26,27] led to speculation that if competitive events are long enough, women may eventually close the performance gap in distance running[11] and swimming.[12] Unlike predictions of eventual closure, [12] current evidence using representative elite populations[17] suggests a sex difference will persist.

One might assume performance gaps would be greatest in events requiring explosive muscular power/sprinting ability, but we found this only in swimming, not running. The greater gap in sprint swimming (13%) vs. running (10-11%) suggests anthropometric advantages associated with the start and/or upper body power in swimming contribute an additional ~2% beyond advantages for men assumed during sprint running. The gap was predicted to eventually close in sprints, [31] an interesting position given our lowest performance gap in the 100 m run. A small rise in performance gap of sprint events during the 1980s was attributed to improved drug testing, presumably reducing anabolic steroid abuse in women. [10] Physical characteristics advantages under the influence of testosterone (muscular hypertrophy/strength) are well known. [32] However, if power/strength advantages are the primary determinants of sprint performance, then in both swimming and running the shortest Olympic distances (50 m and 100 m, respectively) we might expect a greater male advantage, which was observed only in swimming and not running. Our data was consistent with the proposal that less inertia to be overcome by women to accelerate

"Nature vs. Nurture: Have Performance Gaps Between Men and Women Reached an Asymptote?"
by Millard-Stafford M, Swanson AE, Wittbrodt MT
International Journal of Sports Physiology and Performance

a smaller body mass on land might explain lower sex differences associated with the shortest running events on the track. [33]

A significant portion of the residual sex difference observed in elite sprint running (~10%) is likely due to greater male muscle mass to generate peak horizontal power, although few studies on anaerobic power differences are available.[34] Testosterone levels do not predict performance in sprint/power events in elite athletics. [35] Furthermore, sex is not binary with examples of genetic intersex conditions, which may be more prevalent in elite female Olympians. [18] The validity of sex testing for sport classification (based on testosterone levels) remains a controversial issue beyond the scope of this commentary. However, looming as a future issue is whether testosterone hormone reassignment (after some minimal time following "completion of anatomic changes") sufficiently levels the playing field in Olympic sports for transgender females (born and socialized male).[36,37] Longitudinal studies following post-pubertal hormone reassignment might quantify the impact of testosterone per se on performance. Although few IOC cases are under consideration, post-pubertal anthropometric advantages (stature, lever length) would presumably persist along with any potential socio-cultural "advantage" during growth and development.[36]

Nurture factors remaining to be quantified

In terms of other "nurture" factors, increasing age of elite athletes reflects motivation to train and remain competitive in sport. Men and women may peak at similar ages (late 20s), but top placing marathoners were older for women. [38] Older swimmers are returning due to professional sponsorships available (e.g., 41 y old female 2008 Olympic silver medalist). Recent reports [39] suggest 20 y is the peak age for international swimming with "little sex difference" consistent with Olympic medalists in swimming. [40] Although some key factors (increased professional opportunity, later competitive age) appear "equivalent" between the sexes, other possible "nurture"

"Nature vs. Nurture: Have Performance Gaps Between Men and Women Reached an Asymptote?"
by Millard-Stafford M, Swanson AE, Wittbrodt MT
International Journal of Sports Physiology and Performance

influences may remain, although difficult to quantify. Some sports may favor greater male professional sponsorships and create more limitations for women to return following extended lay-offs or other time challenges for training during childbearing years. Given the historical context (2% narrowing in swimming over 44 y), a reasonable assumption might be that no more than 2% of the current performance gap could still potentially be attributed to socio-cultural influences (e.g., in running or other sports).

Conclusion

Performance gaps between USA men and women stabilized within less than a decade after federal legislation provided equal opportunities for female participation but only modestly closed the overall gap in Olympic swimming by 2% (5% in running). Although performance gaps narrow as swimming distance increases, the opposite effect (lowest gap at shortest distance) occurs in running. The magnitude of each biological "advantage" for men is, therefore, not necessarily constant along sprint/endurance domains, further compounding the difficulty in quantifying socio-cultural influences on the performance gap. The sex difference in 400 m swimming (8%) compared to 1500 m running (13%) suggests a 5% mode differential is due to sex-specific fat during weight bearing. Other advantages include lever lengths and stature, particularly more evident in sprint swimming than running (added on top of the ability to generate greater peak velocity). Stable historical trends in a society accepting of female Olympic athletes suggest women will not swim or run as fast, primarily due to underlying biological differences. Future trends may prove otherwise if sex classification lines continue to blur.[18]

"Nature vs. Nurture: Have Performance Gaps Between Men and Women Reached an Asymptote?"
by Millard-Stafford M, Swanson AE, Wittbrodt MT
International Journal of Sports Physiology and Performance

ACKNOWLEDGEMENTS

We appreciate the assistance of Dr. Teresa Snow and Gabriel Mesa, Bradley Homza, Donjae Jones, and Megan Young.

Disclosure statement

No potential conflict of interest was reported by the authors.

"Nature vs. Nurture: Have Performance Gaps Between Men and Women Reached an Asymptote?"
by Millard-Stafford M, Swanson AE, Wittbrodt MT
International Journal of Sports Physiology and Performance

References

1. Cheuvront SN, Carter R, Deruisseau KC, Moffatt RJ. Running performance differences between men and women:an update. *Sports Med.* 2005;35(12):1017-1024.

2. Joyner MJ. Physiological limiting factors and distance running: influence of gender and age on record performances. *Exerc Sport Sci Rev.* 1993;21:103-133.

3. Sparling PB, O'Donnell EM, Snow TK. The gender difference in distance running performance has plateaued: an analysis of world rankings from 1980 to 1996. *Med Sci Sports Exerc.* 1998;30(12):1725-1729.

4. Capranica L, Piacentini MF, Halson S, Myburgh KH, Ogasawara E, Millard-Stafford M. The gender gap in sport performance: equity influences equality. *Int J Sports Physiol Perform.* 2013;8(1):99-103.

5. Whipp BJ, Ward SA. Will women soon outrun men? *Nature.*1992;355(6355):25.

6. Knechtle B, Rosemann T, Lepers R, Rust CA. Women outperform men in ultradistance swimming: the Manhattan Island Marathon Swim from 1983 to 2013. *Int J Sports Physiology Perform.* 2014;9(6):913-924.

7. Wolfrum M, Rust CA, Rosemann T, Lepers R, Knechtle B. Changes in breaststroke swimming performances in national and international athletes competing between 1994 and 2011 -a comparison with freestyle swimming performances. *BMC Sports Sci Med Rehab.* 2014;6:18.

8. Joyner MJ. Physiological limits to endurance exercise performance: influence of sex. *J Physiol.* 2017;595(9):2949-2954.

9. Thibault V, Guillaume M, Berthelot G, et al. Women and Men in Sport Performance: The Gender Gap has not Evolved since 1983. *J Sports Sci Med.* 2010;9(2):214-223.

10. Seiler S, De Koning JJ, Foster C. The Fall and Rise of the Gender Difference in Elite Anaerobic Performance 1952-2006. *Med Sci SportsExerc.* 2007;39(3):534-540.

11. Beneke R, Leithauser RM, Doppelmayr M. Women will do it in the long run. *Br J Sports Med.* 2005;39(7):410.

12. Tanaka H. The battle of the sexes in sports. *Lancet.* 2002;360(9326):92.

13. Deaner RO. Physiology does not explain all sex differences in running performance. *Med Science Sports Exerc.* 2013;45(1):146-147.

14. Schmalz DK, et al. Girlie girls and manly men: Children's stigma consciousness of gender in sports and physical activities. *J Leisure Res.* 2006;38(4):536-557.15.http://bit.ly/1VVnSzO

"Nature vs. Nurture: Have Performance Gaps Between Men and Women Reached an Asymptote?"
by Millard-Stafford M, Swanson AE, Wittbrodt MT
International Journal of Sports Physiology and Performance

16. https://trackandfieldnews.com/index.php/archivemenu/26-news/1145-ot-history-2012; http://trackfield.brinkster.net/OlympicTrials.asp?TourCode=T&Year=1972&Gender=M&TF=T&P=F&By=Y&Count=

17. Vogt P, Rüst CA, Rosemann T, Lepers R, Knechtle B. Analysis of 10 km swimming performance of elite male and female open-water swimmers. *SpringerPlus.* 2013;2:603.

18. Foddy B, Savulescu J. Time to re-evaluate gender segregation in athletics? *Br J Sports Med.* 2011;45(15):1184-1188.

19. Tanaka H, Seals DR. Age and gender interactions in physiological functional capacity: insight from swimming performance. *J Appl Physiol.*1997;82(3):846-851.20.Cureton KJ, Sparling PB. Distance running performance and metabolic responses to running in men and women with excess weight experimentally equated. *Med Sci Sports Exerc.* 1980;12(4):288-294.

21. Sparling PB, Cureton KJ. Biological determinants of the sex difference in 12-min run performance. *Med Sci Sports Exerc.* 1983;15(3):218-223.

22. Veiga S, Roig A. Effect of the starting and turning performances on the subsequent swimming parameters of elite swimmers. *Sports Biomech.* 2016; 31:1-11.

23. Wild S, Rust CA, Rosemann T, Knechtle B. Changes in sex difference in swimming speed in finalists at FINA World Championships and the Olympic Games from 1992 to 2013. *BMC Sports Sci Med Rehab.* 2014;6:25.

24. Coast JR, Blevins JS, Wilson BA. Do gender differences in running performance disappear with distance? *Can J Appl Physiol.* 2004;29(2):139-145.

25. Rust CA, Knechtle B, Rosemann T, Lepers R. Analysis of performance and age of the fastest 100-mile ultra-marathoners worldwide. *Clinics (Sao Paul).* 2013;68(5):605-611.

26. Rust CA, Knechtle B, Rosemann T, Lepers R. Women reduced the sex difference in open-water ultra-distance swimming La Traversee Internationale du Lac St-Jean, 1955-2012. *Appl Physiol Nutr Metab.* 2014;39(2):270-273.

27. Eichenberger E, Knechtle B, Knechtle P, et al. Sex difference in open-water ultra-swim performance in the longest freshwater lake swim in Europe. *J Strength Cond Res.* 2013;27(5):1362-1369.

28. McLean SP, Hinrichs RN. Sex differences in the centre of buoyancy location of competitive swimmers. *J Sports Sci.* 1998;16(4):373-383.

29. Lavoie JM, Montpetit RR. Applied physiology of swimming. *Sports Med.*1986;3(3):165-189.

"Nature vs. Nurture: Have Performance Gaps Between Men and Women Reached an Asymptote?"
by Millard-Stafford M, Swanson AE, Wittbrodt MT
International Journal of Sports Physiology and Performance

30. Pendergast DR, Di Prampero PE, Craig AB, Jr., Wilson DR, Rennie DW. Quantitative analysis of the front crawl in men and women. *J Appl Physiol: Respir, Environ, Exerc Physiol.* 1977;43(3):475-479.

31. Tatem AJ, Guerra CA, Atkinson PM, Hay SI. Athletics: momentous sprint at the 2156 Olympics? *Nature.* 2004;431(7008):525.

32. Lamb DR. Androgens and exercise. *Med Sci Sports.*1975;7(1):1-5.

33. Sandbakk Ø, Solli GS, Holmberg HC. Sex Differences in World-Record Performance: The Influence of Sport Discipline and Competition Duration. Int J Sports Physiol Perform. 2018; Jan 2:1-7.

34. Haugen T, Paulsen G, Seiler S, Sandbakk Ø. New Records in Human Power. *Int J Sports Physiol Perform.*2017; Sept 5:1-27 (Epub ahead of print).

35. Bermon S, Garnier PY. Serum androgen levels and their relation to performance in trackand field: mass spectrometry results from 2127 observations in male and female elite athletes. *Br J Sports Med.* 2017;51(17):1309-1314.

36. Genel M. Transgender Athletes: How can they be accommodated? *Curr Sports Med Rep. 2017;*16(1), 12-13.

37. Ljungqvist A, Genel, M. (2005). Essay: Transsexual athletes--when is competition fair? *Lancet, 2005;366*, S42-43.

38. Hunter SK, Stevens AA, Magennis K, Skelton KW, Fauth M. Is there a sex difference in the age of elite marathon runners? *Med Sci Sports Exerc.* 2011;43(4):656-664.

39. Allen SV, Hopkins WG. Age of Peak Competitive Performance of Elite Athletes: A Systematic Review. *Sports Med.* 2015;45(10):1431-1441.

40. Mazzilli FM. Increase in the Age of Olympic Swimmers in Modern Times. *J Strength Cond Res.* 2017;31(8):2208-2215.

"Nature vs. Nurture: Have Performance Gaps Between Men and Women Reached an Asymptote?"
by Millard-Stafford M, Swanson AE, Wittbrodt MT
International Journal of Sports Physiology and Performance

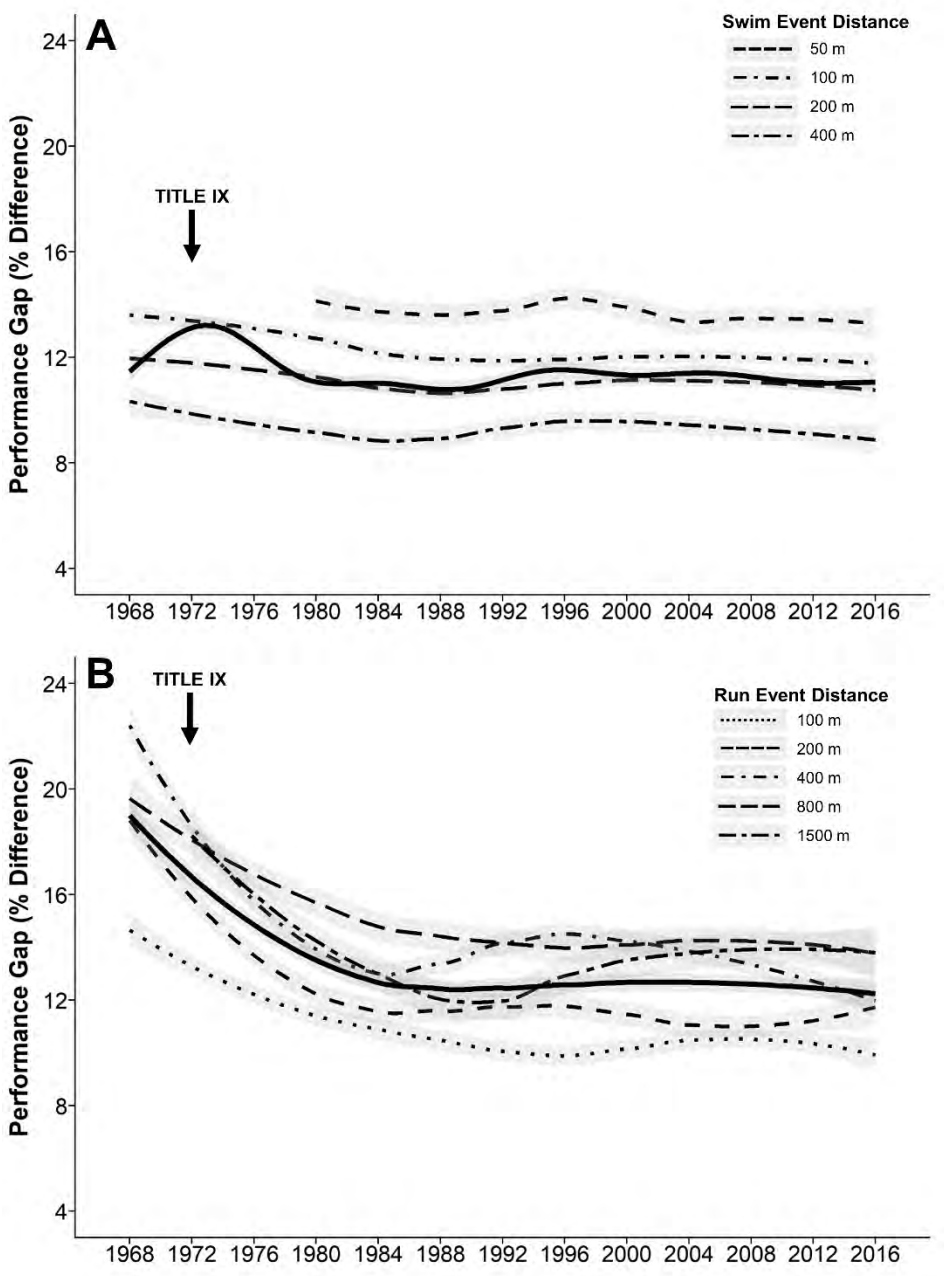


Figure 1. Performance gap in USA Olympic Trials (1968-2016) for top eight finalists over time with data smoothing function plotted by distance (Panel A: All swimming strokes combined and Panel B: Running events up to 1500 m). Solid lines denote the overall mean and shaded areas represented 95% Confidence Intervals. Stabilization occurred after 1980.

"Nature vs. Nurture: Have Performance Gaps Between Men and Women Reached an Asymptote?"
by Millard-Stafford M, Swanson AE, Wittbrodt MT
International Journal of Sports Physiology and Performance

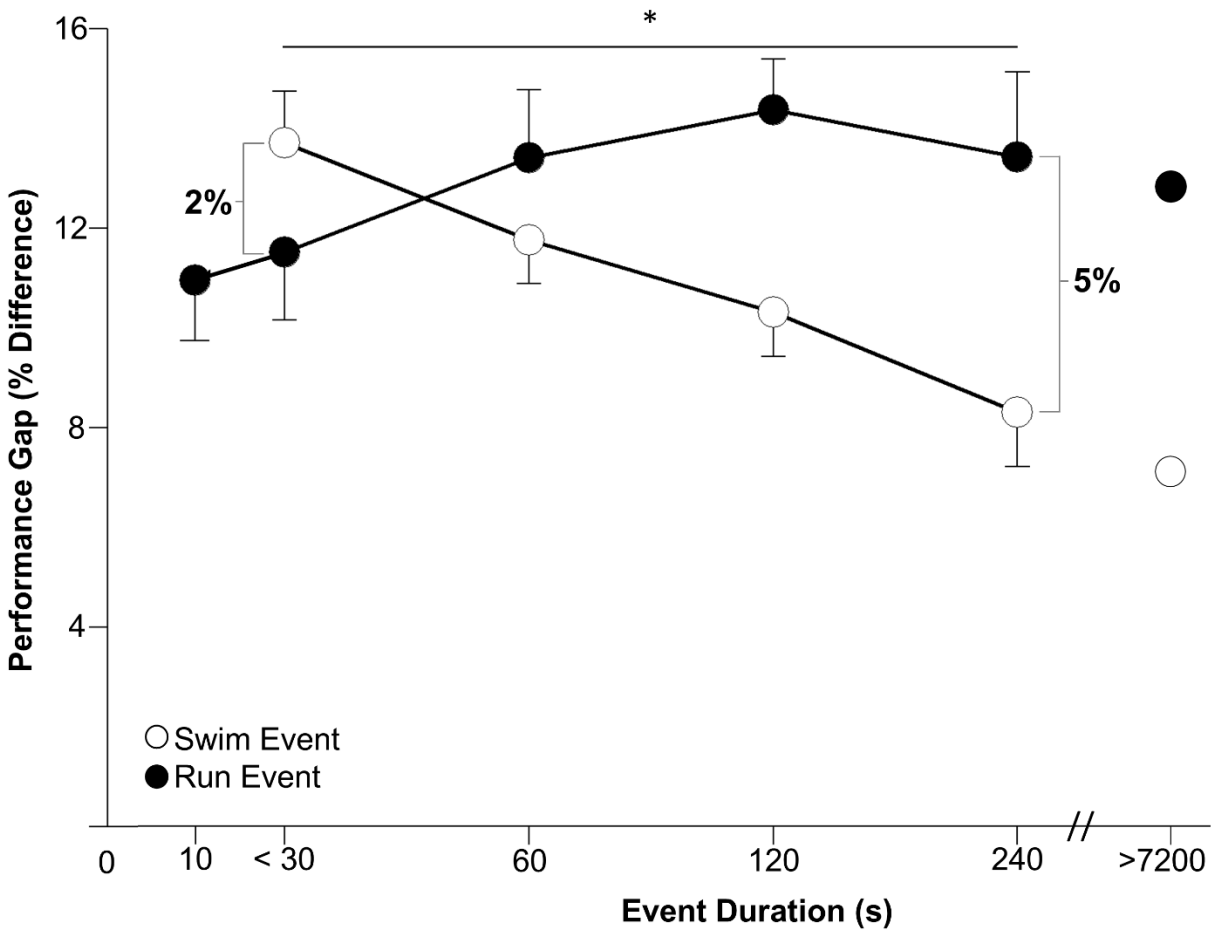


Figure 2. Mode-specific sex differences after historical stabilization (collapsed across USA Olympic Trial data years 1980-2016). Events compared by relative duration: < 30 s= 50 m Freestyle vs. 200 m Run; 60 s = 100 m Freestyle vs. 400 m Run; 120 s = 200 m Freestyle vs. 800 m Run; 240 s = 400 m Freestyle vs. 1500 m Run.

* indicates difference between freestyle swimming and running at each duration. Open and closed circles at 7200 s are published values.[3,17]

"Nature vs. Nurture: Have Performance Gaps Between Men and Women Reached an Asymptote?"
by Millard-Stafford M, Swanson AE, Wittbrodt MT
International Journal of Sports Physiology and Performance

Table 1. %Sex Difference by Swimming Event Over 44 Years (1972 vs. 2016)

SWIMMING	Mean (±SD) Sex Difference by Event (1972 vs. 2016)		1st Place Time (s)		2016 Top Woman's Place in Men's 1972 Race
Event	1972	2016	1972 Men	2016 Women	
50 m Free	14.3 ± 0.5%[A]	13.1 ± 0.5%*	23.07	24.28	70th
100 m Free	13.2 ± 0.4%	10.9 ± 0.6%*	51.91	53.28	9th
200 m Free	11.2 ± 0.6%	9.3 ± 0.6%*	113.58	114.88	5th
400 m Free	9.5 ± 0.8%	8.5 ± 1.5%	240.70	238.98	1st
100 m Back	14.6 ± 0.9%	11.8 ± 1.3%*	58.61	59.02	2nd
200 m Back	14.4 ± 1.3%	10.9 ± 0.5%*	126.57	126.90	3rd
100 m Breast	16.2 ± 1.0%	12.0 ± 0.9%*	65.99	65.20	1st
200 m Breast	13.5 ± 0.7%	12.3 ± 1.1%*	143.27	144.08	2nd
100 m Fly	14.9 ± 1.1%	12.7 ± 1.4%*	54.56	56.48	2nd
200 m Fly	12.7 ± 1.1%	11.4 ± 0.9%*	121.53	126.80	12th
200 m IM	13.0 ± 1.2%	10.7 ± 1.0%*	129.30	129.54	Tie 2nd
400 m IM	12.0 ± 1.1%	9.0 ± 0.8%*	270.81	273.73	3rd
Overall	**13.2 ± 2.0%**	**11.1 ± 1.5%**	-	-	-

[A] For 50 m Free, 1980 is the earliest year available for women so compared both groups from 1980

* Performance gap narrowed ($p < 0.05$) for all events except 400 m Free

"Nature vs. Nurture: Have Performance Gaps Between Men and Women Reached an Asymptote?"
by Millard-Stafford M, Swanson AE, Wittbrodt MT
International Journal of Sports Physiology and Performance

Table 2. %Sex Difference by Running Event Over 44 Years (1972 vs. 2016)

RUNNING	Mean (±SD) Sex Difference by Event (1972 vs. 2016)		1st Place Time (s)		2016 Top Woman's Place in Men's 1972 Race
Event	1972	2016	1972 Men	2016 Women	
100 m	13.4 ± 0.6%	9.1 ± 0.9%*	9.90	10.74	> 38th X
200 m	15.5 ± 0.8%	12.6 ± 1.0%*	20.40	22.25	>28th X
400 m	19.1 ± 1.3%	11.9 ± 0.7%*	44.01	49.68	>20th X
800 m	20.4 ± 0.6%	13.2 ± 0.6%*	104.30	119.10	>26th X
1500 m	18.3 ± 3.4%	13.9 ± 0.5%*	221.50	244.74	>24th X
5 km	11.8 ± 0.7%B	11.9 ± 0.7%	802.80	905.01	>18th X
10 km	9.8 ± 0.4%C	12.7 ± 0.5%†	1715.60	1901.62	>15th X
Marathon	15.5 ± 0.5%D	12.8 ± 0.7%*	8157.80	8900.00	20th
Overall	**17.3 ± 2.8%**	**12.1 ± 1.9%**	-	-	-

* Performance gap narrowed ($p < 0.05$).

† Performance gap widened ($p < 0.05$) from 1988 (earliest common available data for both men and women).

B For 5 km, 1996 is the earliest year available for women.

C For 10 km, 1988 is the earliest year available for women.

D For marathon, 1984 is the earliest year available for women.

X Place corresponds to the total number of male Olympic Trial competitors listed (or last place finish in 1972 for the 2016 top woman).

Transwomen in elite sport: scientific and ethical considerations

Taryn Knox,[1] Lynley C Anderson,[1] Alison Heather[2]

[1]Bioethics Centre, University of Otago, Dunedin, New Zealand
[2]Department of Physiology, University of Otago, Dunedin, New Zealand

Correspondence to
Associate Professor Lynley C Anderson, Bioethics Centre, University of Otago, Dunedin 9001, New Zealand; lynley.anderson@otago.ac.nz

Received 15 October 2018
Revised 21 February 2019
Accepted 9 April 2019
Published Online First 19 June 2019

ABSTRACT

The inclusion of elite transwomen athletes in sport is controversial. The recent International Olympic Committee (IOC) (2015) guidelines allow transwomen to compete in the women's division if (amongst other things) their testosterone is held below 10 nmol/L. This is significantly higher than that of cis-women. Science demonstrates that high testosterone and other male physiology provides a performance advantage in sport suggesting that transwomen retain some of that advantage. To determine whether the advantage is unfair necessitates an ethical analysis of the principles of inclusion and fairness. Particularly important is whether the advantage held by transwomen is a tolerable or intolerable unfairness. We conclude that the advantage to transwomen afforded by the IOC guidelines is an intolerable unfairness. This does not mean transwomen should be excluded from elite sport but that the existing male/female categories in sport should be abandoned in favour of a more nuanced approach satisfying both inclusion and fairness.

The International Olympic Committee (IOC) guidelines[1] that allow male-to-female transgender athletes to compete in the women's category at the elite level has raised significant debate.[2–7] These guidelines specify that transwomen athletes who have demonstrated total testosterone levels below 10 nmol/L for at least 12 months can compete in the women's division. The previous requirement (in the IOC's 2004 guidelines) for gender affirmation surgery has been removed.

A recent New Zealand (NZ) case has polarised opinion about the inclusion of transwomen in women's sport. Laurel Hubbard, a 39-year-old transwoman, competed in the 90 kg+ female category as a weightlifter in the 2018 Commonwealth Games. In 2017, Hubbard won silver medals at the weightlifting world championships—the first NZ weightlifter to win a medal at any world championships. Hubbard previously competed as an adult male and was a junior 105 kg+ NZ record holder. Responses to Hubbard's successes as a female competitor have ranged from support to dismay. Supporters claim she has every right to compete with the women after passing 'straightforward' hormone regulations, and that 'anyone who says otherwise is prejudiced or jealous'.[8] This fits with an inclusion-first policy that argues 'there is a fundamental human right for everyone to be recognised in the gender in which they identify'.[7] Conversely, one of Hubbard's competitors said, "we all deserve to be on an even playing field. If (the playing field) is not even, why are we doing the sport?".[8] Those who object to the IOC guidelines[1] argue they are unfair, specifically that transwomen who now meet those criteria have an advantage not available to cis-women, thus creating an uneven playing field[8 9]. These conflicting views highlight the tension between the principles of inclusion and fairness and raise the question as to whether or how transwomen should be included in elite women's sport.[10]

To determine whether the IOC[1] guidelines adequately address the fairness principle requires, among other things, a scientific understanding of whether transwomen have a performance advantage.

Here we provide a thorough examination of the science to show that elite male athletes have a performance advantage over their female counterparts due to physiological differences. By reasonable inference, the science of male physiology suggests the IOC guidelines may allow elite transwomen athletes to have a performance advantage in comparison with cis-women. On its own, this does not show that transwomen should not compete in the women's division. The arguments for inclusion must also be considered. One such argument is that transwomen should be able to compete in the women's division because even though they may have an advantage, the advantage is not unfair. After assessing the scientific and normative arguments, we conclude that the IOC guidelines are poorly drawn and do not adequately address the fairness principle in elite sport. Far from arguing that transwomen be excluded, we are in favour of a radical change to the outdated structure of the gender divisions currently used in elite sport.

TERMINOLOGY

Before the science is discussed, it is helpful to define terms.

Sex refers to a person's physical characteristics, including their reproductive system (ie, whether they have ovaries or testes), hormones, chromosomes (classically male XY and female XX) and external genitalia. Sex is most commonly determined on the basis of external genitalia. *Intersex* people (ie, those with differences of sexual development (DSD)) are born with reproductive anatomy, chromosomes and/or hormones that cannot be straightforwardly categorised as male or female.

The conflation of transgender and intersex people leads to confusion around effective testosterone versus testosterone levels.[3] This article solely discusses transwomen, and therefore our arguments are not complicated by androgen insensitivities in which receptors are not sensitive, or only partially sensitive to testosterone.[3 11–13] This review only requires discussion of testosterone levels and prior male physiology.

To cite: Knox T, Anderson LC, Heather A. *J Med Ethics* 2019;**45**:395–403.

In the scientific literature, the terms 'male' and 'female' are used to refer to biological sex. Outside of science, 'male' and 'man' are used interchangeably, as are 'female' and 'woman'. We follow both these conventions, except in cases in which doing so does not read well in everyday English.

Gender refers to one's sense of self as a man or a woman (or something else such as gender-neutral or gender-fluid). *Cis-women* are those women whose sex and gender align—they are born female and identify as such. Equally, a cis-man's sex and gender are both male. For a minority of people, sex and gender do not align. A person may be of the male sex but identify as a woman, or be of the female sex but identify as a man. These people are *transgender*.[i] A transgender person may or may not undergo transition, including social changes (coming out to friends and family, changing one's name, personal pronoun, style of dress) and medical intervention (hormone therapy, gender affirmation surgery).

A transwoman who has had full gender affirmation surgery (including testes removal) will have very low testosterone levels below 1 nmol/L.[14] These transwomen will have much less opportunity for a performance advantage in comparison to elite cis-women athletes. Henceforth, unless specifically stated, we use the term 'transwomen' to refer to those elite transwomen athletes who have testosterone levels from 6 to 10 nmol/L (primarily in cases in which a transwoman has retained her testes). Equally, we use the term 'cis-women' to refer to elite cis-women athletes.

Unless stated otherwise, this article solely concerns elite sport. We use Swann *et al's*[15] definition of *elite athlete* which includes athletes competing at national, professional or university (in USA) levels, either paid or unpaid[7 15 16].

Finally, *inclusion* refers to the idea that all athletes should be included.[7] *Fairness* in sport can be used in a broad sense in which inclusion is considered an aspect of fairness.[16] However, to clearly illustrate the conflict between the principles of fairness and inclusion in sport, for the purposes of this article we will not consider inclusion to be an aspect of *fairness*. Instead, we use *fairness* more narrowly to refer to the idea that all athletes must begin from roughly the same starting point.

DIFFICULTY TALKING ABOUT THIS TOPIC

Discussing transwomen's inclusion in elite women's sport is a difficult conversation. Even asking whether transwomen have an unfair performance advantage over cis-women might be offensive to some. Any discussion questioning the inclusion of transwomen may be dismissed as transphobic, prejudiced or coming from a lens of cis-normativity. We recognise the importance of including transpeople in society, including sport, and acknowledge the difficulties faced by many transpeople in establishing safe spaces in everyday life (Bagger[17 18]). However, it is critical that the guidelines attend to both fairness and inclusion to meet the needs of both transwomen and cis-women.

GUIDELINES

IOC guidelines

In 2015, the IOC updated its guidelines, declaring that transwomen athletes can compete in the women's division, if: (A) They have declared, for sporting purposes, their gender to be female for at least 4 years. (B) Their blood testosterone levels are below 10 nmol/L for at least 12 months prior to competition. (This is a general guideline and cases may be reviewed on an individual basis to determine whether 12 months is sufficient time to minimise advantage.)

The 2015 IOC guidelines are markedly different from the 2004 IOC guidelines which held that transgender people who had transitioned after puberty could participate in future Olympic competitions in line with their gender identity if they had: (A) Fully transitioned, that is, they had taken hormone treatment for a minimum of 2 years and had had genital affirmation surgery (including removal of testes for transwomen). (B) Lived in their experienced gender for a minimum of 2 years. (C) Legal recognition of their gender.

A significant difference from the 2004 to the 2015 policies is that genital affirmation surgery is no longer required, instead relying on hormone therapy to maintain testosterone levels under 10 nmol/L for 1 year.[5] This cap was selected because it is at the lower end of the testosterone level for young (< 40 years old) adult men.[19]

In the remainder of this paper, we ask whether 10 nmol/L is too high because the normal healthy female testosterone range is 0–1.7 nmol/L.[19 ii]

The 2004 IOC position is a heavily restricted inclusion model. The 2015 guidelines have moved along the spectrum to a more inclusive and less invasive approach. At the same time, the IOC maintains that the overriding sporting objective of the Olympics is fair competition.[1] Further restrictions may be implemented if it is 'necessary and proportionate to the achievement of that objective (fairness)' and it is possible that transathletes may be further restricted 'in light of any scientific or medical developments'.[1]

Canadian Centre for Ethics in Sport

In 2016, the Canadian Centre for Ethics in Sport (CCES) published guidelines pertaining to the inclusion of transpeople. The CCES noted seven principles of sport, two of which are important for our purposes, namely, 'include everyone' and 'fair play'.[7] The CCES considered transpeople's participation at all levels of sport and concluded that inclusion is the most important principle. However, they recognised that for high performance sport, the 'fair play' principle requires greater consideration (c.f. the 2015 IOC guidelines which claim that fair competition is the overriding sporting objective of the Olympics).[7] With this in mind, CCES provides the following policy guidelines.

1. Hormone therapy is not required unless it can be shown that it is a reasonable and bona fide requirement. The burden of proof is on the sporting organisation to demonstrate a need for hormone therapy.
2. Declaration of trans status is not required unless there is a justified reason. Transpeople are not required to disclose personal information beyond that required by cis-gender athletes.
3. Gender affirmation surgery is not required for any reason.[7]

CCES says that for most sports, there is insufficient evidence to require hormone therapy for transathletes. They say that while there is a 'persistent, ingrained assumption … that men are generally faster, stronger, and better at sport than women',[7] there is a lack of scientific evidence to directly and consistently connect endogenous testosterone levels with athletic performance.[7]

[i]An individual's sexuality (whether they are homosexual, heterosexual or something else) is irrelevant to whether a person is cisgender or transgender.

[ii]As measured by liquid chromatography-mass spectrometry (LCMS).

The CCES base their inclusive policy on the paucity of direct evidence concerning the competitive advantage of transwomen and use the absence of evidence as a green light for including transathletes. They maintain that the benefits of testosterone must be proven to require regulation.[7] In the following section, we show that there is evidence to suggest transwomen have a performance advantage over cis-women. We use the term *suggest* because there have been very few studies on the performance of transwomen athletes at the elite level. (Harper's[20] study does not consider elite athletes, and is limited insofar as it has a very small sample size.) Given the absence of evidence directly related to elite transwomen athletes, and as transwomen were previously biologically male, we extrapolate from evidence based on male physiology.

SCIENCE

Two fundamental assumptions emerge from the IOC guidelines.

1. High testosterone levels provide an all-purpose benefit in sport.
2. Transwomen with testosterone levels under 10 nmol/L for 1 year have mitigated the performance advantage of their former male physiology.

Assumption 1: high testosterone levels provide an all-purpose benefit in sport

It is well recognised that testosterone contributes to physiological factors including body composition, skeletal structure, and the cardiovascular and respiratory systems across the life span, with significant influence during the pubertal period.[19] These physiological factors underpin strength, speed and recovery-[21]with all three elements required to be competitive in almost all sports. An exception is equestrian, and for this reason, elite equestrian competition is not gender-segregated. As testosterone underpins strength, speed and recovery, it follows that testosterone benefits athletic performance.

A complicating factor arises. The extent to which strength, speed and recovery are advantageous changes depending on the sport in question. For example, lawn bowls requires less strength, speed and recovery than many other sports, and so the advantage that testosterone provides is less pronounced. Similarly, the advantage acquired from an individual having high levels of testosterone will apply to team sport, but will be more pronounced in individual sport.

There is little question that strength, speed and recovery influence athletic performance. When considering performance, the parameters that comprise body composition requiring consideration are the percentage of muscle versus fat, and bone strength. On average, from puberty onwards, men have considerably more muscle and less body fat than women.[22 23] Puberty is associated with increased testosterone production in men, and many studies now show testosterone is a key driver of muscle mass. For example, young men administered testosterone in a dose-dependent manner (8.8–82 nmol/L) showed significant increases in muscle mass and strength.[24] Testosterone administration to men has also been shown to protect against loss of muscle mass and strength by increasing muscle protein synthesis and decreasing protein degradation.[25]

Testosterone also has a strong influence on bone structure and strength. From puberty onwards, men have, on average, 10% more bone providing more surface area.[23 26 27] The larger surface area of bone accommodates more skeletal muscle so, for example, men have broader shoulders allowing more muscle to build. This translates into 44% less upper body strength for women, providing men an advantage for sports like boxing, weightlifting and skiing.[28–30] In similar fashion, muscle mass differences lead to decreased trunk and lower body strength by 64% and 72%, respectively in women.[22 28 29 31] These differences in body strength can have a significant impact on athletic performance, and largely underwrite the significant differences in world record times and distances set by men and women.[32]

In contrast, the major female hormones, oestrogens, can have effects that disadvantage female athletic performance. For example, women have a wider pelvis changing the hip structure significantly between the sexes. Pelvis shape is established during puberty and is driven by oestrogen.[33] The different angles resulting from the female pelvis leads to decreased joint rotation and muscle recruitment[34–36] ultimately making them slower.[37 38] Oestrogens also affect body composition by influencing fat deposition. Women, on average, have higher percentage body fat, and this holds true even for highly trained healthy athletes (men 5%–10%, women 8%–15%).[39] Fat is needed in women for normal reproduction and fertility, but it is not performance enhancing. This means men with higher muscle mass and less body fat will normally be stronger kilogram for kilogram than women. In short, higher testosterone levels lead to larger and stronger bones as well as more muscle mass providing a body composition-related performance advantage for men for almost all sports. In contrast, higher oestrogen levels lead to changes in skeletal structure and more fat mass that can disadvantage female athletes, in sports in which speed, strength and recovery are important.

Testosterone also influences the cardiovascular and respiratory systems such that men have a more efficient system for delivering oxygen to active skeletal muscle. Three key components required for oxygen delivery include lungs, heart and blood haemoglobin levels. Inherent sex differences in the lung are apparent from early in life and throughout the life span[40] with lung capacity larger in men because of a lower diaphragm placement due to Y-chromosome genetic determinants.[41 42] The greater lung volume is complemented by testosterone-driven enhanced alveolar multiplication rate during the early years of life.[43]

Oxygen exchange takes place between the air we breathe and the bloodstream at the alveoli, so more alveoli allows more oxygen to pass into the bloodstream. Therefore, the greater lung capacity allows more air to be inhaled with each breath. This is coupled with an improved uptake system allowing men to absorb more oxygen. Once in the blood, oxygen is carried by haemoglobin. Haemoglobin concentrations are directly modulated by testosterone[24 44] so men have higher levels and can carry more oxygen than women.[45] Oxygenated blood is pumped to the active skeletal muscle by the heart. The left ventricle chamber of the heart is the reservoir from which blood is pumped to the body. The larger the left ventricle, the more blood it can hold, and therefore, the more blood can be pumped to the body with each heartbeat, a physiological parameter called 'stroke volume'. The female heart size is, on average, 85% that of a male resulting in the stroke volume of women being around 33% less.[46] Putting all of this together, men have a much more efficient cardiovascular and respiratory system, with testosterone being a major driver of enhanced aerobic capacity.

Combining all of this information, testosterone has profound effects on key physiological parameters that underlie athletic performance in men. There is substantial evidence regarding the effects on muscle gain, bone strength, and the cardiovascular and respiratory system, all of which drive enhanced strength, speed and recovery. Together the scientific data point

to testosterone providing an all-purpose benefit across a range of body systems that contribute to athletic performance for almost all sports. This is exemplified best by the male dominance of sporting world records. Therefore, the first assumption underlying the IOC guidelines is true. Transwomen are allowed to compete with testosterone levels just under 10 nmol/L. This is more than five times the upper testosterone level (1.7 nmol/L) of healthy, premenopausal elite cis-women athletes.[19] Given that testosterone (as well as other elements stemming from Y-chromosome-dependent male physiology) provides an all-purpose benefit in sport, suggests that transwomen have a performance advantage.

Assumption 2: transwomen with testosterone levels under 10 nmol/L for 1 year have mitigated the performance advantage of their former male physiology

We now argue that in addition to higher testosterone levels, transwomen will retain some of the advantages of their former male physiology regardless of 1 year of hormone therapy. Contrary to the second assumption underlying the IOC guidelines, lowering a transwoman's testosterone to under 10 nmol/L does not entirely mitigate the physiology of prior exposure to testosterone and other Y-chromosome genetic determinants. The common hormone therapy for transwomen involves lowering testosterone levels coupled with the administration of the major female hormone, oestradiol. The altered hormonal milieu has pronounced effects on male physiology, including breast development and reduction in body hair. However, there is very little scientific evidence to provide assurance that such hormone therapy will mitigate the advantage transwomen may have in comparison with cis-women athletes. As discussed above, testosterone is central to male physiology and subsequent athletic performance. Lowering testosterone levels to 10 nmol/L could impact on muscle mass and haemoglobin levels because these are positively regulated, in a dose-dependent manner, by testosterone. However, it has been demonstrated that healthy young men did not lose significant muscle mass (or power) when their circulating testosterone levels were reduced to 8.8 nmol/L (lower than the IOC guideline of 10 nmol/L) for 20 weeks.[24] Moreover, retention of muscle mass could be compensated for by training or other ergogenic methods.[47] In addition, the phenomenon of muscle memory means muscle mass and strength can be rebuilt with previous strength exercise making it easier to regain muscle mass later in life even after long intervening periods of inactivity and mass loss.[48 49] Also, indirect effects of testosterone will not be altered by hormone therapy. For example, hormone therapy will not alter bone structure, lung volume or heart size of the transwoman athlete, especially if she transitions postpuberty, so natural advantages including joint articulation, stroke volume and maximal oxygen uptake will be maintained.[50]

While testosterone is the well-recognised stimulator of muscle mass gain, administration of oestradiol has also been shown to activate muscle gain via oestrogen receptor-β activation.[51–55] The combination of oestradiol therapy and a baseline testosterone of 10 nmol/L arguably provides transwomen athletes with an added advantage of increased muscle mass, and therefore power. A recent meta-analysis shows that hormone therapy provided to transwomen over 2 years maintains bone density[56] so bone strength is unlikely to fall to levels of cis-women, especially in an elite athlete competing and training at high intensity. Increased bone strength also translates into protection against trauma, helping with recovery and prevention of injury.[57]

Hence, having transwomen compete in the women's division if their testosterone levels are just under 10 nmol/L will not negate all the performance advantages accrued from having a male physiology prior to transition. A transwoman athlete will retain some of the advantages accrued prior to receiving hormone therapy.[58]

In summary, assumption 1 is true—testosterone provides an all-purpose benefit in almost all sports. Assumption 2 is false. A transwoman athlete with testosterone levels under 10 nmol/L for 1 year will retain at least some of the physiological parameters that underpin athletic performance. This, coupled with the fact that transwomen athletes are allowed to compete with more than five times the testosterone level of a cis-woman, suggests transwomen have a performance advantage.

Proponents of inclusion argue because there are no studies on elite transwomen athletes, and our conclusions are extrapolated from studies on male physiology, there is therefore insufficient evidence to show that transwomen have a performance advantage. However, science does show that having both high testosterone levels and prior male physiology means that transwomen will likely have an inherent advantage in almost all sports.

While the science demonstrating a performance advantage is necessary, normative arguments regarding the inclusion of transwomen in the women's division need to be considered.

ARGUMENTS FOR AND AGAINST INCLUSION

We now present the arguments for inclusion of transwomen in the women's category in elite sport, and each counterargument:

1. Those transwomen athletes who meet the criteria set by the IOC should compete in the women's division.
2. Transwomen identify as women and so should be able to compete in the women's division.
3. The advantage that transwomen have is a tolerable unfairness.
4. As transpeople face discrimination in many spheres of life, they should compete in the gender division in which they identify.
5. The science has been developed using a cis-normative lens.
6. The male/female binary is socially constructed, so athletes ought to be able to switch between men's and women's divisions.

(1) Transwomen athletes who have met accepted IOC criteria should compete in the women's division. These criteria should be followed either because (A) they are fair. Or (B) because they are the established eligibility criteria.

The first point could be countered by showing that the eligibility criteria for the women's category are unfair because, as explained in the science section, transwomen have a significant performance advantage from being able to compete with testosterone levels just under 10 nmol/L and because there has been no thorough exploration of the advantage of having a prior male physiology (for the elite athlete). As Teetzel points out, 'adding a rule or eligibility requirement in a policy or rulebook does not make the rule fair, even if the sports organisation can legally enforce the addition'.[59]

The claim that transwomen have a performance advantage is supported by evidence that testosterone in men is a driving force for the striking dominance of men in elite sport. However, currently women with hyperandrogenic states (having excessive levels of the androgenic hormones including testosterone, such as DSD) compete in the women's division without restrictions. This is despite evidence, although contested, that the high

testosterone levels lead to a performance advantage.[60–62] Both DSD (excess hyperandrogenism) and PCOS (mild hyperandrogenism) are overrepresented among elite women athletes when compared with the wider population.[60 63 64] Recently, based on evidence that hyperandrogenism is associated with performance advantage, the IAAF have lowered the acceptable level of testosterone to 5 nmol/L for some 'restricted' events.[65 66] Hence, the eligibility criteria for transwomen may also be unfair to cis-women, and therefore the guidelines need further attention.

Advocates of inclusion could then argue that transwomen athletes who meet the criteria should be able to compete even if the criteria are unfair and afford transwomen a performance advantage because hyperandrogenism already exists in the women's division. As elite sport has eligibility criteria for transwomen, we agree that until a change is made, transwomen who meet such criteria should compete in the women's division.

(2) Inclusion is an important principle in society. Transwomen identify as women and as such should be included within the women's division in elite sport.

The most common counterargument used by the popular press and cis-gender athletes and their coaches is that the inclusion of transwomen is unfair because they have a physiological advantage. As noted above, we take *fairness* to refer to the idea that all athletes must begin from roughly the same starting point. As Simon explains, 'if one individual has a competitive advantage over another due to differences in initial circumstances that were under the control of neither, the favoured individual can claim no credit for the successes that flow from that undeserved head start' ([18 67]; see also Sher 1979, Loland, 1999). For example, it would be unfair to allow a competitive cyclist to use a hidden electric motor, where others do not. The principle of fairness is the reason why there are sporting regulations prohibiting certain technological advancements (swimsuits using sharkskin technology to reduce drag) or standardising equipment for all (fibreglass pole vaults). The purpose of such standardisation is to level the playing field with regards to that particular concern so the most skilful athlete wins.

The principle of fairness relies on acceptance of the 'skill thesis' which states that the purpose of competitive sport is to identify who is the most skilful.[3 67] Neither Simon nor Bianchi define *skill* with respect to sport. It is incredibly difficult to pin down the meaning of *skill*. However, we take the most skilful athlete to be the one who maximises natural talent via training).[3] We recognise that this definition is problematic because, as discussed below, some natural talents (advantages) are deemed fair while others are not.

Nor do Simon or Bianchi explain how to determine the most skilful athlete. The most skilful athlete is not always the winner of the competition in question. This is because who wins a competition is also influenced by luck,[67 68] such as the bounce of a ball or gust of wind. The eligibility criteria also have a part to play. For example, if a heavyweight boxer was eligible to compete against a bantamweight fighter, the heavyweight would almost invariably win yet not necessarily be the most skilful. If the eligibility criteria are unfair, the winner may not be the most skilful.[59] As stated above, fairness requires that all athletes begin from roughly the same starting point. If a transwoman has a performance advantage over her cis-women competitors then she may win the competition yet not be the most skilful.

(3) Sport is not a level playing field as there are significant physical and psychological differences between people. Furthermore, economic and social factors give rise to differences in funding and resources available to athletes. Both these mean that sport is not fair. If sport is not fair, then the advantage transwomen have should be accepted as, to use Devine's terminology, 'tolerable unfairness'.[69]

It is often claimed that sport is uneven due to socioeconomic factors; that athletes from a wealthy country will have resources available that provide a competitive advantage.[70] This is considered a tolerable unfairness. In addition, many physical differences between people are due to the genetic lottery. Some women are very short and some very tall, some have shorter limbs and some longer. Some of these differences are examples of tolerable unfairness, such as the advantage that tall people have in basketball and the advantage that short-limbed people have in weightlifting. The Finnish skier Eero Mäntyranta had a genetic mutation that enabled him to carry more oxygen in his blood which is known to be performance enhancing.[71] Mäntyranta was able to compete despite his significant advantage—this was a tolerable unfairness. This suggests 'prevailing conceptions of fair play seem unconcerned about the effect of the 'natural lottery''.[72] Equally, those with PCOS with higher than average standard testosterone levels for cis-women compete within the women's division suggesting their advantage is considered a tolerable unfairness. By logical extension, it could be argued that transwomen levels should also be accepted within women's sport—as a further tolerable unfairness.

It is not entirely clear how to distinguish between a tolerable and an intolerable unfairness. Bianchi uses high testosterone levels as an example of something that may be an unfair advantage (ie, an intolerable unfairness). Bianchi says this advantage may be unfair because cis-women cannot attain the advantages afforded by high testosterone levels via endogenous means.[3] Nor are cis-women permitted to take exogenous testosterone to raise their levels to those of transwomen[73] in the women's division. This suggests that an advantage is unfair if no member of the category (eg, cis-women) can attain that advantage.[iii]

Another way to distinguish between tolerable and intolerable advantages is to consider whether the property under consideration provides an all-purpose benefit. High testosterone levels are beneficial in almost all sports (an all-purpose benefit) whereas other advantages are beneficial in only some. For example, a short person is highly unlikely to become an elite basketballer but may excel in gymnastics. Individuals of varying height can self-select into sports that suit their physiological make-up, whereas people with average or low testosterone levels cannot. Hence, it is arguable that height is a tolerable unfairness, whereas high testosterone levels (being an all-purpose benefit) is an intolerable unfairness. High testosterone levels and prior male physiology provide an all-purpose benefit, and a substantial advantage. As the IAAF says, "To the best of our knowledge, there is no other genetic or biological trait encountered in female athletics that confers such a huge performance advantage".[66]

Combined, these three factors (cis-women cannot attain the advantage, all-purpose benefit and magnitude of the advantage) provide a strong argument that transwomen have an intolerable advantage over cis-women. In turn, the impact testosterone has on performance might lead us to reconsider the divisions currently used in competitive sport. In addition, if Mäntyranta's genetic mutation or coming from a wealthy nation are also examples of an intolerable unfairness, these would need to be

[iii]The phrase 'intolerable unfairness' is a philosophical concept, and we do not suggest that gender-variant people are intolerable. Nor do we wish to disregard the discrimination they experience. Thanks to an anonymous reviewer.

accounted for to maintain the skill thesis. Inconsistency may arise if we restrict transwomen from competing in the women's division but allow those from wealthy nations to compete without restriction.[72] This is not to suggest that transwomen (and others such as Mäntyranta or those from wealthy nations) have an easy ride in elite sport. All elite athletes possess a great deal of skill, enhanced by hard work and training. Instead, our point is that some have a head start on their competitors and that the head start ought to be accounted for.

(4) As discussed earlier, transpeople, indeed many groups outside the binary distinctions of male and female, have consistently experienced discrimination and exclusion in all spheres of life. To address these concerns transpeople should compete in the gender division in which they identify.

We fully accept that transpeople have experienced discrimination and transphobia in many aspects of life, and this is unconscionable. While there is a need to ensure there is space for transwomen in elite sport, it does not follow that the space afforded to transwomen should automatically be in women's sport. Cis-women have faced, and still face, difficulties being recognised in top-level sport. The 2020 Olympics will be the first in which all events in all sports will have a women's category.[74] Despite this, some cis-women still struggle for space in elite sport. The first year all 206 Olympic nations sent women to the Olympics was 2012, although many nations do not send women in equal numbers. While there has been dramatic growth in women's sport, 'women of this century (the twentieth century) have occupied only a marginal space in the sports world'.[75] Elite women athletes still face discrimination—they are paid less, receive less media coverage, and when they do, discussion is often about their appearance, rather than athletic ability, and so they contend with 'the ordeal of not being taken seriously as … athlete[s]'[74 75]. In short, women have not achieved gender equity in most sports[76] so care should be taken with any guidelines that may undermine this.

Gender separation provides a dedicated and safe space in which women can compete[77–79] although not all agree that this is desirable for women's sport.[80–82] Cis-women may also be discouraged from direct competition against transwomen. For example, New Zealand weightlifter Tracey Lambrechs reportedly dropped a weight class to avoid competing against Laurel Hubbard because she perceived that Hubbard had a performance advantage because of Hubbard's previous male physiology.[83] [iv] Having transwomen compete against cis-women might erode the space painstakingly built for cis-women. Clearly, transwomen have not had an easy ride in engaging in elite sport, and so we need to negotiate a space in which they can compete fairly. However, that place is not the women's category.

It could be argued that the science has been interpreted using a cis-normative lens meaning that it has been carried out by people who have a (perhaps unconscious) bias towards finding that transwomen have a performance advantage, and so are drawn to results that confirm this bias.[84]

As Ritz explains, employing this binary understanding of sex in experimental systems inherently creates a situation in which we seek difference and which thereby plays into and reinstantiates essentialist and biologically determinist biases. At a fundamental level, experiments are implicitly (and statistically) geared to seeking difference because difference is mechanistically informative in a way that sameness is not. … The upshot is that in treating sex as a unitary dichotomous variable and comparing women and men, we are allowing stereotyped thinking about sex and gender to substitute for the mechanistic understanding that is presumably the goal of experimental research, doing a disservice to people of all genders with respect to our understanding of biology and health.[85]

If a cis-normative lens has affected the scientific approach, and if it is found that transwomen do not have a performance advantage, then this would create a counterargument to the fairness claim. However, there may not be a cis-normative lens. Even if there is such a lens, its existence does not necessarily mean that transwomen do not have a performance advantage. That is, science carried out in the absence of a cis-normative lens might also show that transwomen have a performance advantage. Finally, it might be impossible to do science without a lens, be it cis-normative or something else.[86] If so, and if non-cis-normative science shows that transwomen do not have a performance advantage, then it does not straightforwardly follow that the science that shows 'transwomen do have a performance advantage' is wrong. Instead, it means that claims about truth become much more complex. In short, until there is evidence that the cis-normative lens has generated inaccurate findings, it is justifiable to claim there is strong evidence (although some of it is indirect) to suggest that transwomen have a performance advantage over cis-women. This argument requires us to have solid data on elite transwomen athletes, so we do not have to interpret and extrapolate from male data from which a cis-normative bias could emerge.

(5) It might be argued that the categories of man and woman are socially constructed—that is, they are not natural kinds.[87] [v] Social constructionists might claim that male and female only appear to be natural kinds because the world is interpreted through a cis-normative lens which is entrenched in our everyday lives. If the categories of male and female are socially constructed then this supports an athlete switching from the men's to women's division and vice versa.

There are two possible counterarguments. The first would be to show that the male/female binary does exist in the natural world. This is outside the scope of the paper. The second is to claim that elite athletes should not be able to switch between divisions because the binary is a pragmatic categorisation as it preserves space for women. In other words, the binary might be socially constructed, yet not arbitrary. If the binary serves a pragmatic purpose, then some might claim that it should be retained despite the presence of transgender and intersex people.[88] To reject the binary in sport, it would need to be shown that (A) It is not pragmatic; and/or (B) there are fairer, more respectful ways of categorising athletes.

To summarise the arguments and counterarguments, the issue of whether transwomen should compete in the women's division is particularly difficult because elite sport ought to be both inclusive and fair. The 2015 IOC guidelines allow a transwoman to compete, if, inter alia, her testosterone levels have been under 10 nmol/L for 1 year. The science section showed that the IOC 2015 guidelines allow transwomen to have an advantage over their cis-women competitors. While the IOC 2015 guidelines do not adequately address the principle of fairness, it does not follow that transwomen should not compete in the women's

[iv] See also IAAF's president Sebastian Coe's comments on incentivisation.[65]

[v] *Natural kind* is a term used in the philosophy of science.

division. It could be argued that the advantage that transwomen have is a tolerable unfairness,[69] given the wide range of testosterone levels currently permitted in the woman's division (DSD and to a lesser extent PCOS). We dispute that this is a tolerable unfairness as (A) These high testosterone levels are not available to cis-women (except via doping[3]). (B) Testosterone provides an all-purpose benefit for most sports. (C) Hormone therapy does not mitigate all the benefits from a previous male physiology. Hence, we claim that the IOC guidelines as they stand are poorly drawn. At the same time as being inclusive, the guidelines also need to ensure fairness.

POSSIBLE SOLUTIONS

There is a range of possible responses that could be taken to resolve this conflict between fairness and inclusion.

1. Exclude transwomen from competing in the women's division. This might go some way to levelling the playing field (as evidenced by the science above) but is inconsistent with the principle of inclusion.
2. Transwomen compete in the gender with which they identify. While consistent with the principle of inclusion, this option does not satisfy the principle of fairness given the scientific evidence of a performance advantage. Having unregulated testosterone levels would be an intolerable unfairness. This remains even if testosterone levels are kept below 10 nmol/L.
3. Further restrict the IOC's eligibility criteria for transwomen, by setting the maximum testosterone to a lower level. This approach is problematic because physical harm may result from requiring a transwoman to reduce her testosterone levels further.[89] Further, it would not account for the performance advantage transwomen retain from earlier male physiology.

These three options suggest that the principles of inclusion and fairness are incompatible if the male/female binary is retained. *The following solutions are more radical, challenging the male/female binary in sport:*

4. Have a single category for all athletes regardless of gender.[80 81] This option is problematic as elite women athletes would rarely win against elite male athletes,[79] and so would not satisfy the principle of fairness. Moreover, it may erode the space carved out for women's sport.[77]
   a. Have a single category for all athletes regardless of gender and allow exogenous testosterone to be taken to a single safe level.[11] In addition to the problems noted by Foddy and Savulescu,[11] it would not account for the prior male physiology of transwomen.
5. *Change or increase the categories for participation.*
   a. Create a third division for transwomen and intersex women (who retain their testes).[78]
   b. Create many more categories, such as cis-man, cis-woman, transwoman, transman, transwoman who has had gender affirmation surgery and so on.

Having multiple categories (in 5a and 5b) is not novel as there are weight divisions in sports such as boxing and weightlifting. Weight divisions provide safety to competitors, as well as give athletes in each category a real chance of winning. Approach 5b can also recognise and account for the disadvantage that transwomen without testes have against cis-women competitors, and the disadvantage that transmen have in comparison with cis-men.

A problem with 5a and 5b is that transgender athletes may not get to compete in the category in which they identify.[38 69] Separating out gender-variant people may be perceived as stigmatising or as denying the dignity of transgender people.[vi] This problem might be solved by using a system that does not classify athletes on the basis of whether they are cis-gender or transgender, but on their testosterone levels (as suggested in the following potential solution). A further potential problem with 5a and 5b could be the number of athletes is so small in some categories that it may render the competition meaningless.[38]

(A) Calculate a handicap for transwomen based on their testosterone levels, similar to that used in golf. This approach is suggested by Bianchi.[3 79]

While Bianchi's approach accounts for the disadvantage that transwomen without testes and transmen have, there are two problems with her account. It does not account for the performance advantages transwomen have by virtue of their previous male physiology. Second, Bianchi's system might also be considered offensive as it does not account for gender identity.

Our proposed approach is an extension of Bianchi's.

(B) An algorithm that accounts for (a) Social parameters including gender identity and socioeconomic status. (b) Physiological parameters.

Such an algorithm would be analogous to the divisions in the Paralympics, and may also include paralympians. First, there would be multiple divisions rather than simply male and female. Second, based on the results of the algorithm, athletes would be placed into a division which best mitigates unfair physical and social parameters.

The physiological parameters could include

- Size, for example, height and weight.
- Haemoglobin levels.
- Maximal oxygen uptake (VO^2 max).
- Whether the athlete transitioned before, during or well after puberty.
- Past and present testosterone levels, and the presence of testes.
- Previous characteristics of physiology that are not changed via hormone therapy, that is, bone strength or structure, lung capacity and heart size.

The algorithm would need to be tailored to specific sports as a physiological factor will be more advantageous in some sports than others. For example, weight would need to be accounted for in weightlifting, and perhaps wingspan for swimming.

While it would be difficult to generate such an algorithm, a considered response that adapts to evolving scientific evidence is required. Further discussion regarding this algorithm is beyond the scope of this article but is being developed by the authors.

CONCLUSION

To arrive at conclusion about the adequacy of the IOC (2015) guidelines, both scientific and normative analysis is required. The science section has shown that elite cis-men athletes have a performance advantage in almost all sports due to their higher testosterone levels, and other aspects of male physiology. Extrapolating from this it is plausible that transwomen retain some of that advantage. We acknowledge that these findings rest on extrapolations from male physiology, and encourage further specific scientific investigation regarding elite transwomen athletes. The normative section has shown that while inclusion is an important principle of sport, including elite

[vi] Thanks to an anonymous reviewer.

sport, inclusion does not outweigh the importance of fairness for cis-women athletes—the performance advantage likely held by transwomen is not a 'tolerable unfairness'. This is not an attempt to exclude transwomen from elite sport. Instead, we conclude that it is important to both extend and celebrate diversity, while maintaining fairness for cis-women in sport. To be simultaneously inclusive and fair at the elite level the male/female binary must be discarded in favour of a more nuanced approach. We conclude that the gender binary in sport has perhaps had its day.

Acknowledgements We would like to thank Martin Wilkinson, Silvia Camporesi, Mike King, Cindy Townsend and Louise Pearman-Beres for providing feedback on our earlier drafts of this paper. Additionally, thanks must go to Martin Wilkinson for his phrase 'all-purpose benefit'.

Contributors All three authors contributed to the planning, drafting and editing of the article. AH researched and analysed the scientific evidence, and LA and TK applied this empirical evidence to the ethical issues. All three authors were responsible for the arguments made in the final version of this paper. TK did the majority of the groundwork (eg, literature review) and final editing and referencing of the article. LA is the guarantor and the corresponding author for the article.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 IOC. IOC consensus meeting on sex reassignment and hyperandrogenism. 2015 https://stillmed.olympic.org/Documents/Commissions_PDFfiles/Medical_commission/2015-11_ioc_consensus_meeting_on_sex_reassignment_and_hyperandrogenism-en.pdf (Accessed 24 May 2018).
2 Harper J, Martinez-Patino MJ, Pigozzi F, *et al*. Implications of a third gender for elite sports. *Curr Sports Med Rep* 2018;17:42–4.
3 Bianchi A. Transgender women in sport. *Journal of the Philosophy of Sport* 2017;44:229–42.
4 Buzuvis E. Caster semenya and the myth of a level playing field. *Mod Am* 2010;6:36–42.
5 Jones BA, Arcelus J, Bouman WP, *et al*. Sport and transgender people: a systematic review of the literature relating to sport participation and competitive sport policies. *Sports Med* 2017;47:701–16.
6 Pitsiladis Y, Harper J, Betancurt JO, *et al*. Beyond fairness: the biology of inclusion for transgender and intersex athletes. *Curr Sports Med Rep* 2016;15:311–26.
7 CCES. *Creating inclusive environments for trans participants in canadian sport (trans inclusion in sport expert working group guidelines)*. Ottawa, 2016. https://cces.ca/sites/default/files/content/docs/pdf/cces-transinclusionpolicyguidance-e.pdf. (accessed 24 May 2018).
8 TVNZ. She is completely entitled to compete and anyone who says otherwise is prejudiced or jealous. *tvnz.co.nz* 2017 https://www.tvnz.co.nz/one-news/sport/other/she-completely-entitled-compete-and-anyone-says-otherwise-prejudiced-jealous (Accessed 23 May 2018).
9 Wall T. A level playing field? 2018 https://interactives.stuff.co.nz/2018/03/a-level-playing-field/ (accessed 1 Aug 2018).
10 Martínez-Patiño MJ, Vilain E, Bueno-Guerra N. The unfinished race: 30 years of gender verification in sport. *The Lancet* 2016;388:541–3.
11 Foddy B, Savulescu J. Time to re-evaluate gender segregation in athletics? *Br J Sports Med* 2011;45:1184–8.
12 Karkazis K, Jordan-Young R, Davis G, *et al*. Out of bounds? A critique of the new policies on hyperandrogenism in elite female athletes. *Am J Bioeth* 2012;12:3–16.
13 Schultz J. Caster semenya and the "question of too": sex testing in elite women's sport and the issue of advantage. *Quest* 2011;63:228–43.
14 Schneider F, Kliesch S, Schlatt S, *et al*. Andrology of male-to-female transsexuals: influence of cross-sex hormone therapy on testicular function. *Andrology* 2017;5:873–80.
15 Swann C, Moran A, Piggott D. Defining elite athletes: issues in the study of expert performance in sport psychology. *Psychol Sport Exerc* 2015;16:3–14.
16 Wagman B. Including transitioning and transitioned athletes in sport:issues, facts and perspectives. 2009 http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.513.434&rep=rep1&type=pdf (accessed 24 May 2018).
17 Wamsley KB. Social science literature on sport and transitioning/transitioned athletes. *Prep Promis Pract Work TransitioningTransitioned Athletes Sport Proj Univ West Ont* 2008.
18 IOM. The health of lesbian, gay, bisexual, and transgender people: building a foundation for better understanding. *Inst. Med* 2011 http://nationalacademies.org/hmd/Reports/2011/The-Health-of-Lesbian-Gay-Bisexual-and-Transgender-People.aspx (accessed 7 Aug 2018).
19 Handelsman DJ, Hirschberg AL, Bermon S. Circulating testosterone as the hormonal basis of sex differences in athletic performance. *Endocr Rev* 2018;39:803–29.
20 Harper P. Genetics and public health. *BMJ* 1992;304:721.
21 Williams C. ABC of sports medicine. Assessment of physical performance. *BMJ* 1994;309:180–4.
22 Janssen I, Heymsfield SB, Wang ZM, *et al*. Skeletal muscle mass and distribution in 468 men and women aged 18-88 yr. *J Appl Physiol* 2000;89:81–8.
23 Lang TF. The bone-muscle relationship in men and women. *J Osteoporos* 2011;2011:1–4.
24 Bhasin S, Woodhouse L, Casaburi R, *et al*. Testosterone dose-response relationships in healthy young men. *Am J Physiol Endocrinol Metab* 2001;281:E1172–E1181.
25 Ferrando AA, Sheffield-Moore M, Paddon-Jones D, *et al*. Differential anabolic effects of testosterone and amino acid feeding in older men. *J Clin Endocrinol Metab* 2003;88:358–62.
26 Handelsman DJ. Sex differences in athletic performance emerge coinciding with the onset of male puberty. *Clin Endocrinol* 2017;87:68–72.
27 Vanderschueren D, Laurent MR, Claessens F, *et al*. Sex steroid actions in male bone. *Endocr Rev* 2014;35:906–60.
28 Wilmore JH. Alterations in strength, body composition and anthropometric measurements consequent to a 10-week weight training program. *Med Sci Sports* 1974;6:133???138.
29 Laubach LL. Muscular strength of women and men: a comparative study. *Dayton Univ Oh Research Inst* 1976 http://www.dtic.mil/docs/citations/ADA025793 (accessed 6 Jun 2018).
30 Hegge AM, Myhre K, Welde B, *et al*. Are gender differences in upper-body power generated by elite cross-country skiers augmented by increasing the intensity of exercise? *PLoS One* 2015;10:e0127509.
31 Jones MT, Jagim AR, Haff GG, *et al*. Greater strength drives difference in power between sexes in the conventional deadlift exercise. *Sports* 2016;4:43–53.
32 Devries MC. Do transitioned athletes compete at an advantage or disadvantage as compared with physically born men and women: A review of the scientific literature. *Rep Prep AthletesCAN* 2008 http://www.athletescan.com/Content/Resources/Promising%20Practices.asp
33 Huseynov A, Zollikofer CP, Coudyzer W, *et al*. Developmental evidence for obstetric adaptation of the human female pelvis. *Proc Natl Acad Sci U S A* 2016;113:5227–32.
34 Horton MG, Hall TL. Quadriceps femoris muscle angle: normal values and relationships with gender and selected skeletal measures. *Phys Ther* 1989;69:897–901.
35 Benas D. Special considerations in women's rehabilitation programs. In: Hunter LY, Funk FJ, eds. *Rehabilitation of the Injured Knee*. Princeton, NJ: CV Mosby Company, 1984:393–405.
36 Simoneau GG, Hoenig KJ, Lepley JE, *et al*. Influence of hip position and gender on active hip internal and external rotation. *J Orthop Sports Phys Ther* 1998;28:158–64.
37 Kerrigan DC, Todd MK, Della Croce U. Gender differences in joint biomechanics during walking: normative study in young adults. *Am J Phys Med Rehabil* 1998;77:2–7.
38 Sutherland MA, Wassersug RJ, Rosenberg KR. From transsexuals to transhumans in elite athletics. *Transgender Athletes Compet Sport* 2017;173.
39 Jeukendrup AE. *Sport nutrition: an introduction to energy production and performance*. 2nd edn. Champaign, IL: Human Kinetics, 2010.
40 Townsend EA, Miller VM, Prakash YS. Sex differences and sex steroids in lung health and disease. *Endocr Rev* 2012;33:1–47.
41 Bellemare F, Jeanneret A, Couture J. Sex differences in thoracic dimensions and configuration. *Am J Respir Crit Care Med* 2003;168:305–12.
42 Bellemare JF, Cordeau MP, Leblanc P, *et al*. Thoracic dimensions at maximum lung inflation in normal subjects and in patients with obstructive and restrictive lung diseases. *Chest* 2001;119:376–86.
43 Thurlbeck WM. Postnatal human lung growth. *Thorax* 1982;37:564–71.
44 Bachman E, Travison TG, Basaria S, *et al*. Testosterone induces erythrocytosis via increased erythropoietin and suppressed hepcidin: evidence for a new erythropoietin/hemoglobin set point. *J Gerontol A Biol Sci Med Sci* 2014;69:725–35.
45 Murphy WG. The sex difference in haemoglobin levels in adults - mechanisms, causes, and consequences. *Blood Rev* 2014;28:41–7.
46 Wells CL. The limits of female performance. *Bone* 1985;44:36.0.
47 Bhasin S, Storer TW, Berman N, *et al*. The effects of supraphysiologic doses of testosterone on muscle size and strength in normal men. *N Engl J Med* 1996;335:1–7.
48 Staron RS, Leonardi MJ, Karapondo DL, *et al*. Strength and skeletal muscle adaptations in heavy-resistance-trained women after detraining and retraining. *J Appl Physiol* 1991;70:631–40.
49 Taaffe DR, Robinson TL, Snow CM, *et al*. High-impact exercise promotes bone gain in well-trained female athletes. *J Bone Miner Res* 1997;12:255–60 https://doi.org/.
50 Gooren LJ, Kreukels B, Lapauw B, *et al*. (Patho)physiology of cross-sex hormone administration to transsexual people: the potential impact of male-female genetic differences. *Andrologia* 2015;47:5–19.
51 Lowe DA, Baltgalvis KA, Greising SM. Mechanisms behind estrogen's beneficial effect on muscle strength in females. *Exerc Sport Sci Rev* 2010;38:61–7.
52 Tiidus PM, Lowe DA, Brown M. Estrogen replacement and skeletal muscle: mechanisms and population health. *J Appl Physiol* 2013;115:569–78.
53 Mangan G, Bombardier E, Mitchell AS, *et al*. Oestrogen-dependent satellite cell activation and proliferation following a running exercise occurs via the PI3K signalling pathway and not IGF-1. *Acta Physiol* 2014;212:75–85.

54 Velders M, Schleipen B, Fritzemeier KH, *et al*. Selective estrogen receptor-β activation stimulates skeletal muscle growth and regeneration. *Faseb J* 2012;26:1909–20.
55 Parr MK, Zhao P, Haupt O, *et al*. Estrogen receptor beta is involved in skeletal muscle hypertrophy induced by the phytoecdysteroid ecdysterone. *Mol Nutr Food Res* 2014;58:1861–72.
56 Singh-Ospina N, Maraka S, Rodriguez-Gutierrez R, *et al*. Effect of Sex Steroids on the Bone Health of Transgender Individuals: A Systematic Review and Meta-Analysis. *J Clin Endocrinol Metab* 2017;102:3904–13.
57 Goolsby MA, Boniquit N. Bone Health in athletes: the role of exercise, nutrition, and hormones. *Sports Health* 2017;9:108–17.
58 Van Caenegem E, Wierckx K, Taes Y, *et al*. Preservation of volumetric bone density and geometry in trans women during cross-sex hormonal therapy: a prospective observational study. *Osteoporos Int* 2015;26:35–47.
59 Teetzel S. Athletes' perceptions of transgender eligibility policies applied in high-performance sport in Canada. In: Anderson E, Travers A, eds. *Transgender athletes in competitive sport*. London: Routledge, 2017:78–89.
60 Bermon S, Garnier PY, Hirschberg AL, *et al*. Serum androgen levels in elite female athletes. *J Clin Endocrinol Metab* 2014;99:4328–35.
61 Bermon S, Garnier PY. Serum androgen levels and their relation to performance in track and field: mass spectrometry results from 2127 observations in male and female elite athletes. *Br J Sports Med* 2017;51:1309–14.
62 Bermon S, Hirschberg AL, Kowalski J, *et al*. Serum androgen levels are positively correlated with athletic performance and competition results in elite female athletes. *Br J Sports Med* 2018;52:1531–2.
63 Hagmar M, Berglund B, Brismar K, *et al*. Hyperandrogenism may explain reproductive dysfunction in olympic athletes. *Med Sci Sports Exerc* 2009;41:1241–8.
64 Rickenlund A, Carlström K, Ekblom B, *et al*. Hyperandrogenicity is an alternative mechanism underlying oligomenorrhea or amenorrhea in female athletes and may improve physical performance. *Fertil Steril* 2003;79:947–55.
65 IAAF. IAAF introduces new eligibility regulations for female classification| News | iaaf.org. iaaf.org. 2018 https://www.iaaf.org/news/press-release/eligibility-regulations-for-female-classifica (accessed 23 May 2018).
66 IAAF. Explanatory notes IAAF eligibility regulations for the female classification. 2018 https://www.documentcloud.org/documents/4449931-Explanatory-Notes-IAAF-Eligibility-Regulations.html (accessed 7 Aug 2018).
67 Simon R. Deserving to be lucky: reflections on the role of luck and desert in sports. *Journal of the Philosophy of Sport* 2007;34:13–25.
68 Feinberg J. *Doing & deserving; essays in the theory of responsibility*. Princeton University Press: Princeton, 1970.
69 Devine JW. Gender, steroids, and fairness in sport. *Sport Ethics Philos* 2018:1–9.
70 Bernard AB, Busse MR. Who wins the olympic games: economic resources and medal totals. *Rev Econ Stat* 2004;86:413–7.
71 Enriquez J, Gullans S. Olympics: genetically enhanced olympics are coming. *Nature* 2012;487:297.
72 Davis P, Edwards L. The new IOC and IAAF policies on female eligibility: old Emperor, new clothes? *Sport Ethics Philos* 2014;8:44–56.
73 WADA. The world anti-doping code prohibited list. 2018 https://www.wada-ama.org/en/resources/science-medicine/prohibited-list-documents.
74 IOC. Key dates in the history of women in the olympic movement. *Int. Olymp. Comm* 2018 https://www.olympic.org/women-in-sport/background/key-dates (accessed 23 May 2018).
75 Cahn SK. *Coming on strong : gender and sexuality in women's sport*. Illinois: University of Illinois Press, 2015.
76 Senne JA. Examination of gender equity and female participation in sport. *Sport J* 2016;19:1–9.
77 English J. Sex equality in sports. *Philos Public Aff* 1978;7:269–77.
78 Newbould MJ. What do we do about women athletes with testes? *J Med Ethics* 2016;42:256–9.
79 Dreger A. Sex typing for sport. *Hastings Cent Rep* 2010;40:22–4.
80 Sailors PR. Mixed competition and mixed messages. *Journal of the Philosophy of Sport* 2014;41:65–77.
81 McDonagh E, Pappano L. *Playing with the boys: why separate is not equal in sports*. Oxford, New York: Oxford University Press, 2007.
82 Kane MJ. Resistance/transformation of the oppositional binary: exposing sport as a continuum. *J Sport Soc Issues* 1995;19:191–218.
83 Caldwell O. Human rights commission NZ backs transgender weightlifter laurel hubbard for commonwealth games. stuff.co.nz. 2017 https://www.stuff.co.nz/sport/other-sports/99294375/human-rights-commission-nz-backs-transgender-weightlifter-laurel-hubbard-for-commonwealth-games (accessed 24 May 2018).
84 Longino HE. *Science as social knowledge: values and objectivity in scientific inquiry*. Princeton University Press: Princeton, NJ:, 1990.
85 Ritz SA. Complexities of addressing sex in cell culture research. *Signs* 2017;42:307–27.
86 Imber B, Tuana N. Feminist perspectives on science. *Hypatia* 1988;3:139–55.
87 Cavanagh SL, Sykes H. Transsexual bodies at the olympics: the international olympic committee's policy on transsexual athletes at the 2004 athens summer games. *Body Soc* 2006;12:75–102.
88 Burrow S. Trampled autonomy: women, athleticism, and health. *IJFAB: International Journal of Feminist Approaches to Bioethics* 2016;9:67–91.
89 Peterson MD, Belakovskiy A, McGrath R, *et al*. Testosterone Deficiency, Weakness, and Multimorbidity in Men. *Sci Rep* 2018;8:5897.

TRANSWOMEN COMPETING IN WOMEN'S SPORTS: WHAT WE KNOW, AND WHAT WE DON'T

UNK UNIVERSITY OF NEBRASKA KEARNEY

Gregory A. Brown FACSM
Physical Activity and Wellness Lab, Department of Kinesiology and Sport Sciences
University of Nebraska at Kearney, Kearney, NE, USA.

UNK UNIVERSITY OF NEBRASKA KEARNEY

ABSTRACT

The purpose of this presentation is to summarize the differences between males and females regarding athletic performance, review the current knowledge regarding the effects of gender affirming hormone therapy in transwomen on factors that influence athletic performance, and provide an update on legislation regarding the participation of transwomen in women's sports.

It is well documented that males outperform comparably aged and trained females in most measures of physical fitness and athletic events. Generally speaking, males have 10-40% more body mass, 45% more lean body mass, run 10-15% faster, and have 30-60% greater muscle strength than females. These male athletic advantages originate from sex based physiological differences.

In 2015 the International Olympic Committee adopted a new transgender participation policy with much less stringent requirements for transwomen to compete in women's sports. As a result, the inclusion of transgender athletes in the Tokyo Olympic Games has brought to the forefront the issues of safety, fairness, and inclusion of transwomen in women's sports.

To date research indicates that gender affirming hormone therapy in transwomen eliminates the male advantages in hemoglobin concentrations. Gender affirming hormone therapy in transwomen causes 9-4% reductions in body mass, 4-9% reductions in muscle mass, and concomitant increases in fat mass. Furthermore, there have been only 5 papers published evaluating the effects of gender affirming hormone therapy in transwomen on handgrip strength (showing 3-7% reductions in strength), 1 paper published evaluating knee extensor and flexor strength (showing no reductions in strength), and 1 paper published evaluating push-ups (27% reduction), sit-ups (16% reduction) and 1.5 mile running performance (8% reduction in speed). Otherwise the effects of gender affirming hormone therapy in transwomen on athletic performance and factors influencing athletic performance (e.g. VO_2max, lactate threshold, isotonic 1 repetition maximum) remain unknown. However, on the basis of safety and fairness, in the past year legislation has been introduced in 37 states to limit participation in girl's and women's sports to cisgender women, with the legislation being signed into law in seven states. This information is pertinent to the health of transwomen as their choice to engage in gender affirming hormone therapy may influence their eligibility to participate in sports.

HUMAN SEXUAL DIMORPHISM

- The American Psychological Association defines sex as "physical and biological traits that distinguish between males and females" whereas gender "implies the psychological, behavioral, social, and cultural aspects of being male or female (i.e., masculinity or femininity)" (1, 2)
- Sex chromosomes conferred at conception (3)
  - XX = Female, XY = Male
  - Differences/Disorders of Sexual Development (DSD) occur ~1 in 4,500 – 5,500 births (4)
- Male babies are typical larger than female babies, necessitating different fetal growth charts (5)
- Differences in growth necessitate sex specific infant and child growth charts (6)
- There are also sex specific differences in health outcomes immediately before and after birth, including having higher risk for neurological, metabolic and respiratory complications, and overall poorer health in boys (7, 8)
- During puberty, males gain greater amounts of fat free mass, muscle mass, and bone mass, whereas females acquire significantly more fat mass. There are also differences in the distribution of fat mass in the body (9, 10, 11)
- The average adult male is 7.7% taller, 17.0% heavier, has 45% more lean body mass, 30% less fat mass, 57% higher handgrip strength, 54% higher knee extension torque, 30% higher maximal cardiac output, 25% higher VO_2max (relative to body mass), and 11% higher hemoglobin concentration than the average adult female (12)

SEXUAL DIMORPHISM IN ATHLETIC PERFORMANCE

- School based fitness testing in children ages 3 - 18 indicates that males outperform females in measurements of aerobic fitness, muscular strength, and muscular endurance whereas females outperform males in measures of flexibility (13 – 18)
- Competition performance records from USA Track and Field show that males outperform females in running, throwing, and jumping starting in the 8 and under age group, and continuing through all age groups (19)
  - The performance differences increase at the onset of puberty (~ 11 years old; Table 1)

	8 and under	9 and 10	11 and 12
100 M	1.0%	8.4%	4.0%
200 M	3.3%	1.8%	4.4%
400 m	5.8%	1.1%	7.8%
800 m	6.4%	1.8%	8.7%
1500 m	2.0%	4.4%	5.6%
Mean	3.7%	3.5%	5.7%
	13 and 14	15 and 16	17 and 18
100 M	8.0%	8.1%	10.1%
200 M	7.2%	7.9%	13.4%
400 m	11.4%	13.3%	12.6%
800 m	10.3%	13.0%	14.4%
1500 m	9.8%	13.1%	15.2%
Mean	9.0%	11.5%	13.3%

Table 1. Percent difference between male and female youth track running records

- Competition data in adults indicates that males run and swim faster, lift heavier weights, jump farther and higher, and throw farther and faster than females (figure 1)

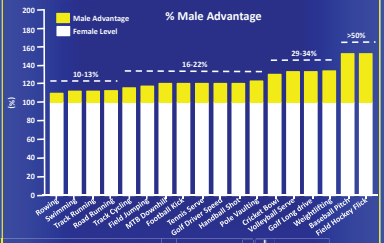


Figure 1. Adapted from Hilton and Lundberg (12). The male performance advantage over females across various selected sporting disciplines. The female levels is set to 100%. In sport events with multiple disciplines, the male value has been averaged across discipline.

- The difference between 1st place and 2nd place in NCAA Outdoor Track Running events = 0.48% and 0.87%, between 2nd and 3rd = 0.46% and 0.57%, for males and females, respectively (20)

TRANSWOMEN IN WOMEN'S SPORTS

- 2015 International Olympic Committee (IOC) policy allows transwomen to compete in women's sports after 1 year of testosterone suppression below 10 nmol/L (2)
  - Many states, countries, and sporting organizations followed this pattern, with the policies varying widely from sport to sport, and state to state within the USA
- June 2019, Soule v. Connecticut Association of Schools, a policy that allows biological males who claim a female identity to compete in girls' athletic events deprived Soule and the other female athletes of honors and opportunities to compete at elite levels
- 2020 Idaho enacts the Fairness in Women's Sports Act, limiting participation in women's sports to biological women
- October 2020 World Rugby recommended that transwomen not play women's contact rugby (21)
  - Input received from experts with the objective of understanding the medical, physiological, psychological, risk, social-ethical and sporting environment. This included input from transgender community representatives, players, medical, research and rugby experts
- January 2021, President Biden signed an executive order that protects against discrimination based on sexual orientation and gender identity and states children should not be "denied access" to "school sports" (22)
- In the 2021 legislative session, bills to limit women's sports to biological women were introduced in 37 states in the USA (23)
  - Laws passed in Montana, Arkansas, Tennessee, Mississippi, Alabama, Florida, and West Virginia
  - Still in legislative debate in some states, many will likely return in 2022
  - Lawsuits abound (e.g. Hecox v. Little; BPJ et al. v. West Virginia State Board Of Education)
- In July 2021 the IOC stated that its current guidelines for transgender athletes are not acceptable in light of recent research (24)
  - Plans to release a new policy after 2022 Winter Games
- In September 2021 the five Sports Councils responsible for supporting and investing in sport across England, Wales, Scotland, and Northern Ireland concluded, after an exhaustive review of research and input from many stakeholders, that there cannot be an equitable balance of inclusion of transwomen in women's sport with the needs for fairness and safety in women's sports due to retained advantages, with or without testosterone suppression, in male-to-female or non-binary athletes (25)

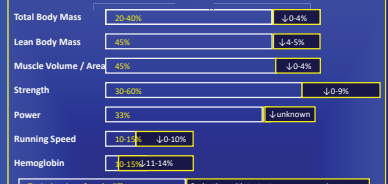


Figure 2. Summary of inherent male vs. female differences and the effects of male-to-female hormones. Adapted from World Rugby (21).

WHAT WE KNOW

- Handgrip strength is reduced by 3-9% due to 1 - 2 years of male-to-female hormone use (26-30)
- Isometric knee extension and flexion, and isokinetic knee extension and flexion are not changed due to 1 year of male-to-female hormone use (31)
- In a small sample of US Air Fore personnel, at baseline transwomen performed 31% more push-ups and 15% more sit-ups in 1 min and ran 1.5 miles 21% faster than age matched females. After 2.5 years of male-to-female hormones, the push-up and sit-up differences disappeared but transwomen were still 12% faster (32)
- Antiandrogenic progestins in 16 year old transgirls attenuates age related increases in handgrip strength, but does not reverse sex based differences in strength, body size, or muscle mass (32)

CONCLUSIONS FROM RECENT REVIEWS

- "Thus, the muscular advantage enjoyed by transgender women is only minimally reduced when testosterone is suppressed." (12)
- "These findings suggest that strength may be well preserved in transwomen during the first 3 years of hormone therapy. Regarding athletic performance, men are bigger, faster, and stronger." (33)
- "While trans athletes have a right to participate in sports, cisgender women have a right to participate in a protected category of sport." (34)

WHAT WE DON'T KNOW

- No controlled training studies with male-to-female hormone use
- No measurements of changes in VO_2max, running economy, lactate threshold, anaerobic power (e.g. Wingate test), vertical jump, 1-Repetetion Maximum (e.g. bench press, leg press, squat, deadlift), or many other common determinants of athletic performance

0021-972X/98/$03.00/0
Journal of Clinical Endocrinology and Metabolism

Vol. 83, No. 12
Printed in U.S.A.

Bone Mass in Prepubertal Children: Gender Differences and the Role of Physical Activity and Sunlight Exposure*

G. JONES AND T. DWYER

Menzies Center for Population Health Research, Hobart, Tasmania 7000, Australia

ABSTRACT

Retrospective studies have suggested that the prepubertal years may be an important window of opportunity to increase bone mass, but there have been few direct studies and little exploration of gender differences in this age group. In this study, we report the associations among physical activity measures, sunlight exposure, body composition, and bone mass in 8-yr-old children. We studied 330 children in 1996 (115 girls and 215 boys; response rate, 60%) who had previously taken part in a cohort study of cot death in 1988. They had measurement of anthropometrics (height, weight, and body composition), sunlight exposure (by questionnaire), and physical activity [questionnaire, muscle strength by dynamometry, and bicycle ergometric physical work capacity at a pulse of 170 beats/min (PWC170)]. Bone mineral density (BMD) was assessed at the spine, femoral neck, and total body by a Hologic QDR2000 densitometer. In females only, PWC170 [hip, 2.4%/quartile (95% confidence interval (CI), 0.3–4.5); spine, 1.7%/quartile (95% CI, 0.0–3.4); size adjusted] was associated with BMD, whereas in males only, BMD was associated with both sports participation (hip, 4.2% (95% CI, 1.1–7.3); spine, 4.3% (95% CI, 0.9–7.7)] and muscle strength [hip, 1.7%/quartile (95% CI, 0.0–3.4); but not spine; size adjusted]. Winter sunlight exposure was associated with BMD in girls [hip, 2.9%/category (95% CI, 0.7–5.0); spine, 3.6%/category (95% CI, 1.4–5.8)], but not in boys [hip, 0.3%/category (95% CI, −1.4 to +2.0); spine, 1.4%/category (95% CI, −0.7 to +3.5)]. Males and females were very similar in body size. However, males had higher size-adjusted BMD at the hip (9.6%; 95% CI, 6.9–14), whereas females had higher size-adjusted BMD at the spine (3.2%; 95% CI, 0.8–5.6%). In conclusion, this study has suggested that physical activity and exposure to sunlight are important in the bone mineralization of prepubertal male and female children. The magnitude of both gender and environmental differences in bone mass in this age group is substantial, suggesting that modification at this stage of life may influence peak bone mass and possibly fracture risk in later life. (*J Clin Endocrinol Metab* **83:** 4274–4279, 1998)

OSTEOPOROTIC fractures are a major public health problem in both males and females (1). Bone density is one of the major predictors of these osteoporotic fractures (2) and is the result of the amount of bone gained in early life (*i.e.* peak bone mass) and subsequent bone loss (3). There is evidence to suggest that physical activity and, to a lesser extent, diet (particularly calcium intake) during adolescence and early adulthood are determinants of peak bone mass (4–7). It seems likely that the vast majority of adult bone mass is attained before age 14 yr (8), but relatively little is known about determinants of bone mass in this group. Recent retrospective studies in ballerinas and gymnasts have suggested that the prepubertal period may be crucial in terms of physical activity and bone mass (6, 7), and calcium supplementation is beneficial in the short term in 8-yr-old children (9). Apart from physical activity and diet, it is possible that other environmental factors are important in the prepubertal age group. One potential candidate is vitamin D. Sunlight is the major source of vitamin D in most age groups (10), and lack of sunlight has been associated with osteopenia in handicapped children (11) and children with cystic fibrosis (12), but there have been no studies to our knowledge in healthy children. Gender differences in bone mass in children are currently controversial. Most studies report higher femoral bone mass in boys (13–17) and higher spinal bone mass in girls (14, 18, 19), whereas others report no difference (20–23). It has been hypothesized that this may be explained by size differences, as areal bone mass only partially estimates true bone mass (24). However, studies of volumetric bone mass have reported gender differences at the hip while eliminating an apparent age or size effect (16), suggesting that other factors, especially the effect of physical activity in different populations, may contribute to gender differences. This hypothesis would appear more attractive given the site-specific gender differences and generally stronger associations between physical activity and femoral bone mass compared to spinal mass (13, 15). Furthermore, it suggests that other factors, perhaps body fat, may be a stronger determinant of bone mass at the spine, hence the higher levels in girls, but there are limited data in prepubertal children. In this study, therefore, we examined the associations among physical activity measures, sunlight exposure, body composition, and bone mass and their association with gender differences in 8-yr-old male and female children from Southern Tasmania.

Received June 1, 1998. Revision received September 3, 1998. Accepted September 14, 1998.

Address all correspondence and requests for reprints to: Dr. Graeme Jones, Menzies Center for Population Health Research, GPO Box 252–23, Hobart, Tasmania 7000, Australia. E-mail: g.jones@utas.edu.au.

* This work was supported in part by the National Health and Medical Research Council of Australia, the Royal Hobart Hospital Acute Care Program, the Lions Club of Australia, Coca-Cola Amatil, the Tasmanian Dairy Authority, and Talays.

Subjects and Methods

In 1988, there were 6779 live births in Tasmania (latitude 42°S). Of these, 1380 were identified as being at high risk of sudden infant death syndrome by previously published criteria (25) and were invited to take part in a longitudinal study. In Southern Tasmania, there were 735 births who met these criteria. Of these, the parents of 696 (95%) agreed to an

in-hospital interview, and the parents of 581 (80%) agreed to the 1 month follow-up.

The 696 subjects who agreed to the in-hospital interview were approached during 1996 to take part in a bone mass study. After 8 yr, we were able, through the use of school lists, to definitely identify 551 of these subjects (or 80%, which is in close agreement to the Australian Bureau of Statistics data on annual outward migration rates from Tasmania of 2.5%). Subjects who provided informed consent to take part underwent an extensive protocol involving measurement of bone mass, anthropometrics, physical activity and fitness, diet, and sunlight exposure. Ethical approval for this study was obtained from the University of Tasmania ethics committee (human experimentation).

Bone mass was assessed using the technique of dual energy x-ray absorptiometry at the total body, lumbar spine, and right femoral neck (Hologic QDR2000 densitometer, Waltham, MA). Bone mass was examined in two separate ways: bone mineral content (BMC) and bone mineral density (BMD). Precision estimates *in vivo* are not available in our subjects for ethical reasons. The longitudinal coefficient of variation for our machine during 1996 using daily measurements of a spine phantom was 0.54%. Body composition estimates were also available for these subjects.

Height was measured using a stadiometer with the subject in bare feet. Weight was measured using bathroom scales that were calibrated daily using known weights.

Physical activity measures included questionnaire items regarding sports participation (defined as taking part in organized sport for at least 3 months of the last 12), lunchtime activities (five-point ordinal scale), and recess school activities (three-point ordinal scale). Objective measures included measurement of muscle strength by dynamometry at three sites: lower limb (involving both legs simultaneously), upper limb push, and upper limb pull. The child was instructed in each technique before testing, and each measurement was performed twice. Repeatability estimates (Cronbach's α) were as follows; lower limb, 0.91; upper limb pull, 0.89; and upper limb push, 0.83. The devices were calibrated by suspending known weights at regular intervals. Physical work capacity was also assessed by a bicycle ergometer (26). Subjects were asked to cycle at a constant 60 rpm for 3 min at each of three successively increasing, but submaximal, workloads. Heart rate was recorded at 1-min intervals at each workload using an electric heart rate monitor. Work capacity at 170 beats/minute (PWC170) was then assessed by linear regression with extrapolation of the line of best fit to a heart rate of 170 beats/min. The PWC170 was not considered a technically adequate measure unless subjects had spent a minimum of 2 minutes at each workload, and the pulse rate increased by at least 5 beats/min with increasing workloads. We have previously found this measure to correlate well (r = 0.8) with treadmill assessment of maximal oxygen intake in Australian school children, aged 9–15 yr (27). Repeatability was not assessed in our subjects, but has previously been reported as 0.92 (26). The usual diet in the last 12 months was assessed by food frequency questionnaire completed by the parent/guardian under supervision by the research assistant.

Sunlight exposure was assessed by questionnaire relating to the amount of daily exposure during school days, weekends, and on school holidays in both summer and winter. Categories were as follows: 1, less than 2 h; 2, 2–3 h; 3, 3–4 h; and 4, more than 4 h. We have previously validated this measure of exposure against actual exposure with polysulfone badges in teenage children and found them to correlate well in summer (r = 0.62) (28). Parents were also asked to record their child's sleep duration (time of falling asleep and time of waking).

Statistics

Linear modelling techniques were used to examine the relationship between bone mass and study factors of causal interest. For BMC, the approach taken was initially univariate. Any factor was then adjusted for bone area at that site. Any factor with an area adjusted $P < 0.15$ was then placed in a multivariate model with height and weight. The exception to this approach was if a factor was associated with BMC at one site and not others; then the univariate coefficient and confidence limits at other sites are reported even if they include the null point. A similar approach was adopted for BMD, except that any univariate association was adjusted for height and weight in step 2 before further modelling. To determine whether associations were mediated by lean body mass, associations were then separately performed by the addition of lean body mass to the model. Physical activity and sunlight variables were examined separately in the modelling procedure. A statistically significant result was regarded as $P < 0.05$ (two-tailed) or a 95% confidence limit not including the null point. All statistical calculations were carried out using SPSS version 6.1 for Windows.

Results

A total of 330 subjects took part (males, n = 215; females, n = 115), representing an overall response rate of 60% of those available in 1996. Table 1 documents study factors and BMD broken down by sex. Males and females were very similar in terms of age, weight, height, and calcium intake. However, there were marked differences in other factors. Males had higher levels of sports participation, sunlight exposure, muscle strength, PWC170, and lean body mass but lower fat mass compared to females.

Physical activity

Associations between the various measures of physical activity and fitness are reported in Table 2 and Fig. 1. In males, sports participation had the strongest association with BMD, with those participating in organized sport having a

TABLE 1. Demographic details of subjects

	Males (n = 215)[a]	Females (n = 115)[a]	*P* value for difference
Age (yr)	8.17 (7.32–8.82)	8.26 (7.34–8.92)	0.02
Wt (kg)	27.9 (5.4)	28.0 (5.9)	0.92
Ht (cm)	127.9 (5.9)	127.3 (5.5)	0.36
% Sports participation	66	54	0.04
PWC170	1.16 (0.35)	1.02 (0.27)	0.001
Lower limb muscle strength	35.7 (11.5)	29.9 (9.6)	<0.0001
Sunlight in winter (mean score)[b]	2.87 (1–4)	2.50 (1–4)	0.009
Sunlight in summer (mean score)[b]	3.74 (1–4)	3.60 (1–4)	0.08
Bone free lean body mass (kg)	21.8 (2.8)	20.0 (3.9)	<0.0001
Fat mass (kg)	5.8 (3.4)	7.3 (4.1)	0.001
Calcium intake (mg/day)	1374 (681)	1414 (632)	0.60
Areal bone mass			
Femoral neck (g/cm^2)	0.66 (0.07)	0.59 (0.07)	<0.0001
Lumbar spine (g/cm^2)	0.60 (0.07)	0.61 (0.07)	0.20

[a] All are the mean ± SD, except for age and sunlight exposure, where it is the mean score (range), which approximately equates to mean hours of exposure.

[b] Tests of significance are based on Mann-Whitney U test. All others are unpaired *t* tests.

TABLE 2. Physical measures and bone mass in prepubertal children

Variable	Univariate [change in g/cm^2 (95% CI)]	Adjusted for size[a] [change in g/cm^2 (95% CI)]	Adjusted for lean body mass [change in g/cm^2 (95% CI)]
a. Femoral neck			
Males			
Sports participation (yes/no)	+0.027 (+0.007–+0.047)	+0.018 (−0.001–+0.037)	+0.014 (−0.005–+0.033)
Muscle strength (per quartile)[b]	+0.020 (+0.012–+0.028)	+0.011 (+0.003–+0.020)	+0.008 (0.000–+0.016)
PWC170 (per quartile)[b]	+0.014 (+0.004–+0.024)	+0.004 (−0.009–+0.010)	−0.001 (−0.095–+0.093)
Females			
Sports participation (yes/no)	+0.009 (−0.016–+0.034)	+0.005 (−0.015–+0.025)	+0.008 (−0.016–+0.032)
Muscle strength (per quartile)[b]	+0.005 (−0.005–+0.015)	+0.003 (−0.008–+0.014)	+0.000 (−0.010–+0.010)
PWC170 (per quartile)[b]	+0.025 (+0.013–+0.037)	+0.014 (+0.002–+0.026)	+0.021 (+0.008–+0.034)
b. Lumbar spine			
Males			
Sports participation (yes/no)	+0.025 (+0.005–+0.045)	+0.013 (−0.004–+0.041)	+0.001 (−0.016–+0.018)
Muscle strength (per quartile)[b]	+0.010 (+0.002–+0.018)	−0.004 (−0.012–+0.004)[c]	−0.008 (−0.016–−0.000)[c]
PWC170 (per quartile)[b]	+0.013 (+0.004–+0.022)	+0.001 (−0.007–+0.009)	−0.004 (−0.012–+0.006)
Females			
Sports participation (yes/no)	+0.005 (−0.021–+0.031)	+0.004 (−0.018–+0.026)	+0.004 (−0.020–+0.024)
Muscle strength (per quartile)[b]	+0.008 (−0.004–+0.020)	+0.003 (−0.007–+0.013)	+0.005 (−0.006–+0.016)
PWC170 (per quartile)[b]	+0.021 (+0.009–+0.033)	+0.010 (−0.001–+0.021)	+0.017 (+0.005–+0.029)

Underlined results denote statistical significance.
[a] Height and weight.
[b] Quartiles are as follows: muscle strength: males, 10–26, 27–34, 35–40, 41–70; females, 10–23, 24–29, 30–35, 36–60; PWC170; males, 0.38–0.92, 0.93–1.11, 1.12–1.35, 1.36–3.07; females, 0.48–0.82, 0.83–1.00, 1.01–1.16, 1.18–1.91).
[c] Significant collinearity.

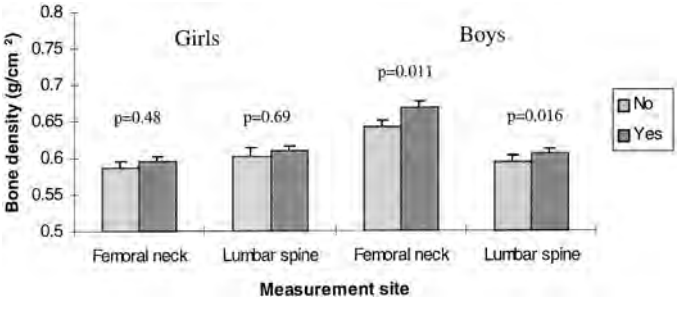


FIG. 1. Sports participation and bone mass in 8-yr-old children. Sports participation in the past year is associated with bone mass that is 4% higher in boys and 1% higher in girls. Results are expressed as the mean ± SEM.

4.2% higher BMD at the femoral neck and a 4.3% BMD higher at the spine. Adjustment for bone free lean body mass negated these associations, suggesting that they may be mediated by changes in lean mass. Lower limb muscle strength was also important at the femoral neck in males. This association, although decreased in magnitude, persisted after adjustment for both body size and lean body mass, suggesting a local loading mechanism. A boy in the highest quartile of muscle strength had a size-adjusted BMD 5.1% higher than those in the lowest quartile. This association was not present at the spine.

In contrast, in females, it was PWC170 that had the strongest association with BMD at both sites. This adjustment persisted after adjustment for both size and lean body mass. A girl in the highest quartile of PWC170 had size-adjusted BMD 7.2% higher at the femoral neck and 5.1% higher at the spine compared to those of girls in the lowest quartile. Neither sports participation nor muscle strength was associated with BMD in girls.

Lunchtime and playtime activities and hours of sleep were not associated with BMD at any site (data not shown).

Sunlight

Associations between winter sunlight exposure and BMD are presented in Table 3 and Fig. 2. In males, winter sunlight exposure is weakly associated with BMD. Compared to those in the lowest category of sunlight exposure, males in the highest category had 0.001 g/cm^2 higher BMD at the femoral neck [95% confidence interval (CI), −0.039 to 0.041] and 0.020 g/cm^2 higher BMD at the lumbar spine (95% CI, −0.014 to 0.062). This equates to 0.9% at the femoral neck and 4.2% at the spine, respectively.

In females, however, sunlight exposure was significantly associated with BMD at all sites. This was present both for a continuous association (Table 3 and Fig. 2) and when categorized. Compared to those in the lowest category of sunlight exposure, females in the highest category had 0.057 g/cm^2 higher BMD at the femoral neck (95% CI, 0.016–0.099) and 0.058 g/cm^2 higher BMD at the lumbar spine (95% CI, 0.014–0.102). This equates to 8.9% at the femoral neck and 10.5% at the spine, respectively. These associations persisted after adjustment for body size and composition (Table 3). The associations were not altered by adjustment for PWC170 (data not shown).

Sunlight exposure during winter weekdays was more weakly associated with BMD in both sexes than weekend exposure. No category of sunlight exposure during the summer was associated with BMD (data not shown).

Body composition and bone mass

In multivariate analysis, BMD was strongly predicted by bone free lean mass at both sites and weakly by fat mass at the spine in boys. In contrast, in girls, BMD was strongly predicted by fat mass at both sites and weakly by bone-free lean mass at the spine (Table 4).

TABLE 3. Sunlight exposure in winter and bone mass in prepubertal children expressed as change in bone mass for each increasing category of exposure

Variable	Univariate [change in g/cm^2 (95% CI)]	Adjusted for size[a] [change in g/cm^2 (95% CI)]	Adjusted for lean body mass [change in g/cm^2 (95% CI)]
Males			
Femoral neck	+0.002 (−0.009–+0.013)	+0.001 (−0.009–+0.011)	+0.001 (−0.009–+0.011)
Lumbar spine	+0.008 (−0.004–+0.020)	+0.006 (−0.002–+0.014)	+0.005 (−0.003–+0.013)
Females			
Femoral neck	+0.016 (+0.004–+0.028)	+0.015 (+0.004–+0.026)	+0.014 (+0.002–+0.026)
Lumbar spine	+0.021 (+0.008–+0.034)	+0.018 (+0.006–+0.030)	+0.019 (+0.007–+0.031)

Underlined results denote statistical significance. Numbers in each category are as follows: males, 1–19, 2–58, 3–73, 4–57; females, 1–15, 2–45, 3–30, 4–22.

[a] Height and weight.

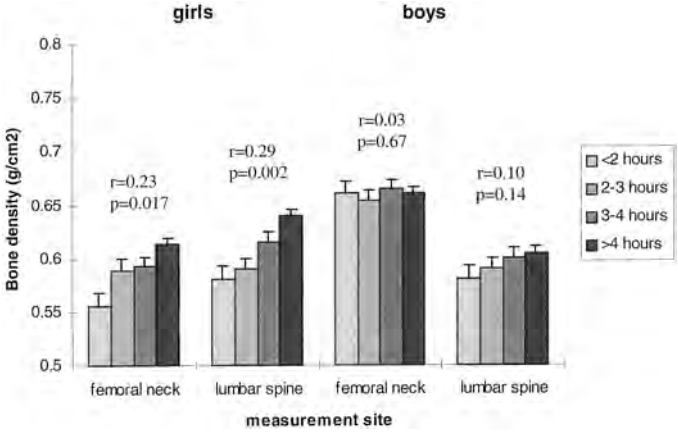


FIG. 2. Sunlight exposure on winter weekends and bone mass in 8-yr-old children. There is a dose-response relationship between increasing levels of sun exposure and bone mass in girls. The relationship is weaker and not statistically significant in boys. Results are expressed as the mean ± SEM.

Gender differences in bone mass

Despite being similar in terms of body size, males had significantly higher BMD at the femoral neck but not at the lumbar spine (Table 1). The magnitude of the differences at the femoral neck was reduced by adjustment for body size, composition, physical activity, and sunlight exposure (adjusted difference, 9.6%; 95% CI, 6.9–14). However, at the lumbar spine, after the same adjustment, girls had higher BMD than boys (adjusted difference, 3.2%; 95% CI, 0.8–5.6%).

The analysis for BMC was virtually identical to that for BMD, and the associations with total body BMD were very similar to those with femoral neck for physical activity, sunlight exposure, body composition, and gender differences, so the results of these analyses are not presented here.

Discussion

This cross-sectional study has found that both physical activity and winter sunlight exposure are significantly associated with bone mass in prepubertal children. There are gender differences in both bone mass and the strength of exposure associations. These differences are partially explicable by the higher levels of sun exposure and physical activity in males. The magnitude of the associations are substantial even after adjustment for size, and if maintained until the attainment of peak bone mass would suggest that optimizing these exposures may decrease fracture risk in later life.

TABLE 4. Gender differences in multivariate body composition bone mass associations (both factors simultaneously)

	Femoral neck coefficient (g/cm^2 · kg)	Lumbar spine coefficient (g/cm^2 · kg)
Males		
Bone free lean mass (kg)	0.012 ($P < 0.00001$)	0.013 ($P < 0.00001$)
Fat mass (kg)	0.007 ($P = 0.64$)	0.030 ($P = 0.03$)
Females		
Bone free lean mass (kg)	0.002 ($P = 0.34$)	0.004 ($P = 0.03$)
Fat mass (kg)	0.077 ($P < 0.00001$)	0.053 ($P = 0.0009$)

This study has a number of potential limitations. The children who took part in this study are not representative of Tasmanian children. They were originally selected on the basis of having a higher risk of sudden infant death syndrome (25). As a result, there was a higher proportion of males, premature babies, teenage mothers, and smoking during pregnancy. These findings suggest that this group is of a lower socio-economic status than the Tasmanian population as a whole. According to Mietinen (29), an analytical cohort study to be generalizable to other populations does not have to be representative of the community from which it was selected provided it meets the following key criteria with regard to definition of eligible participants, sample size, and a proper distribution of determinants, modifiers, and confounders. This study fulfills all three criteria. The study population was explicitly defined and is of adequate sample size; furthermore, it has considerable heterogeneity of exposure to factors of causal interest. Furthermore, there was no association between maternal education and household income levels and bone mass in these children, suggesting that this bias may not be of major concern. Overall, these observations would suggest that the exposure-bone mass associations reported in this sample may be generalizable to other prepubertal populations, particularly with regard to physical activity, but less so for sunlight exposure for other reasons (see below).

We found that a variety of physical activity measures were associated with bone mass, particularly at the femoral neck. In boys, sports participation was the most strongly associated at both sites and for lower limb muscle strength at the femoral neck. The association with sports participation appears to be mediated by increases in lean body mass, as adjustment for this negated the association. The association with muscle

strength was only partially reduced by such adjustment and suggests an additional role for local biomechanical loading. In contrast, in girls, neither of these two was associated, but physical work capacity was, and this association also persisted after adjustment for body size and lean body mass. This suggest that physical fitness may have an effect on bone mass independent of these factors that may be due to common genetic factors or, alternatively, may be due to physical activity independently leading to both higher bone mass and higher fitness. However, it is widely recognized that the assessment of physical activity in children is difficult (13), and it may be that measurement of fitness is an indirect measure of physical activity in girls that has not been adequately captured by our other measurements. It is also possible that there is a threshold effect with physical activity, by which levels above a certain cut-off point may confer no further benefit. There are some data to support this in our sample, as males in the lowest quartile of PWC170 had BMD identical to females in the highest quartile of PWC170 at the femoral neck, but not at the spine. However, the gender differences between physical activity measures and bone mass that we report here in prepubertal children are both unexpected and largely unexplained at present. The types of sports played by both are similar, so this is not an adequate explanation. The magnitude of the increase in bone mass with physical activity is substantial, with increments ranging from 4–7% (size adjusted). These findings contrast with those of Bass *et al.* (6), who found differences in the order of 10–15% in gymnasts, but, as stated by the researchers, it is likely that these are close to the maximum attainable, and our reported figures are more likely to be representative of what is possible in normal children. The variations we report indicate the need for further research in prepubertal children and strongly suggest that more objective measures that accurately quantify actual and habitual physical activity, such as pedometers, need to be further developed and validated so that measurement error can be minimized.

A novel finding of this study was an association between the amount of sunlight exposure on winter weekends (but not summer) and bone mass at all sites. The sunlight association is most likely to be through photosynthesis of vitamin D as well as physical activity, as relatively little vitamin D comes from dietary sources in Australia, and dietary supplementation is not considered necessary due to the abundant sunlight. Ambient UV light in the environment depends on many factors, including latitude, elevation, stratospheric ozone, sunspot activity, and atmospheric pollution (30). Furthermore, vitamin D photosynthesis depends on the action of UV radiation on skin and thus can also be affected by subject activity during peak exposure periods, amount of clothing worn, sun angle on skin, skin color, humidity, temperature, wind speed, sunscreens, and sweating (30). Hobart has levels of UV exposure comparable to those of other regions within Australia during summer at 20 minimal erythermal doses. In winter, however, the levels drop markedly to 1.6 minimal erythermal doses (31). Although comparative data exist, it is difficult to directly compare Northern and Southern latitudes. In New Zealand, it has been estimated that UV levels are 13–50% higher than equivalent Northern latitudes due primarily to lower ozone levels (32). This may explain the observation that sunlight intensity in Boston (latitude 42°N) during the winter months is insufficient to lead to photosynthesis of vitamin D (10), whereas our data suggest that 4 h or more of sunlight each day on winter weekends is required in Hobart (latitude 42°S) to attain vitamin D stores that are optimal for bone health. This relationship appears plausible, in that our questionnaire assessment of sun exposure correlates well with actual sun exposure measured by polysulfone badges in teenagers (28). It is possible that this association may be confounded by physical activity as winter organized sport is generally played on the weekend in Tasmania. However, adjustment for lean body mass or PWC170 did not alter the strength of the association. Furthermore, the association was not present for summer exposure, which would also be expected to be confounded, suggesting that this is not the case. The association with sunlight was stronger in girls than in boys. A possible explanation for this is that overall boys had higher levels of sunlight exposure than girls with fewer in the lower exposure categories, suggesting that more boys achieved sufficient exposure for optimal bone mineralization, making it more difficult to show statistically significant results even with the higher number of boys compared to girls. Furthermore, anecdotal evidence would suggest that boys may expose more skin than girls in Tasmania and, thus, may synthesize more vitamin D for the same amount of exposure. We do not have measures of vitamin D stores for these children, but our results indicate the need to directly measure serum 25-hydroxyvitamin D_3 in both sexes to validate this finding, as vitamin D supplementation, at least in winter, may be required to achieve optimal bone mineralization in Tasmanian children (and possibly those in higher latitudes in other areas of the world where routine supplementation is not practiced), as the amount of sunlight required closely parallels the average daily sunlight hours, which are 3.9 h in winter compared to 7.4 h in summer (M. Nunez, personal communication). An optimum level of vitamin D remains controversial, but recent studies have suggested that minimum levels of 50–80 nmol/L are required in both adults and children (33, 34), which is well above levels indicative of overt deficiency.

The area of gender differences in bone mass is controversial. Areal bone mass is a two-dimensional approximation of volumetric bone density that is strongly correlated with actual breaking strength of bone (35). It appears likely that a major explanation for the reported gender differences in bone density and fracture thresholds (2) are due to artifactual differences in areal bone mass related to bone size. Although a number of methods have been developed to deal with this bias, including bone mineral apparent density (16) or adjusting for body size in the analysis as we have done, none is free of problems. Although we did not have a direct measure of bone size for ethical reasons, male and female children were virtually identical in terms of height and weight, suggesting that this explanation is unlikely in this case. Recent American studies have reported racial, but not gender, differences in prepubertal children (20). However, this study did not report on hip BMD and found a trend for higher spinal BMD in females, as reported in the present sample. A possible explanation for the gender differences is the obser-

vation that male and female children in this study differed markedly in terms of physical activity measures, sunlight exposure, and body composition, all of which were associated with areal bone mass, suggesting that the gender differences documented are not size related, but are related to environmental and/or constitutional differences in the prepubertal years. Variations in physical activity and body composition would appear the most plausible explanations for our observation. The fact that multivariate analysis did not completely remove the gender difference is consistent with residual confounding introduced by measurement error (36) in the assessment of physical activity and sunlight exposure. Better assessment of these factors may negate apparent gender differences.

In conclusion, this study has suggested that physical activity and exposure to sunlight are important in the bone mineralization of prepubertal male and female children. The magnitude of both gender and environmental differences in bone mass in this age group is substantial and suggests that modification at this stage of life may influence peak bone mass and possibly fracture risk in later life. However, these findings need to be confirmed in further studies of a longitudinal nature with the development of more objective measures of exposure, particularly physical activity and vitamin D.

Acknowledgment

Special thanks to Jack Allan the bus driver, Carole Goff and the staff of the Medical Imaging Department, Royal Herbert Hospital.

References

1. **Jones G, Nguyen T, Sambrook PN, Kelly PJ, Gilbert C, Eisman JA.** 1994 Symptomatic fracture incidence in elderly men and women: the Dubbo Osteoporosis Epidemiology Study (DOES). Osteoporosis Int. 4:277–282.
2. **Nguyen T, Sambrook P, Kelly P, et al.** 1993 Prediction of osteoporotic fractures by postural instability and bone density. Br Med J. 307:1111–1115.
3. **Hansen MA, Kirsten O, Riis BJ, Christiansen C.** 1991 Role of peak bone mass and bone loss in postmenopausal osteoporosis: 12 years study. Br Med J. 303:961–964.
4. **Valimaki M, Karkkainen M, Allardt C, et al.** 1994 Exercise, smoking and calcium intake during adolescence and early adulthood as determinants of peak bone mass. Br Med J. 309:230–235.
5. **Welten DC, Kemper HC, Post GB, et al.** 1994 Weight-bearing activity during youth is a more important factor for peak bone mass than calcium intake. J Bone Miner Res. 9:1089–1096.
6. **Bass S, Pearce G, Bradney M, et al.** 1998 Exercise before puberty may confer residual benefits in bone density in adulthood: studies in active prepubertal and retired gymnasts. J Bone Miner Res. 13:500–508.
7. **Khan K, Bennell KL, Hopper JL, et al.** 1998 Self-reported ballet classes undertaken at age 10–12 years and hip bone mineral density in later life. Osteoporosis Int. 8:165–173.
8. **Sabatier J-P, Guaydier-Souquieres, Laroche D, et al.** 1996 Bone mineral acquisition during adolescence and early adulthood: a study in 574 healthy females 10–24 years of age. Osteoporosis Int. 6:141–148.
9. **Bonjour JP, Carrie AL, Ferrari S, et al.** 1997 Calcium-enriched foods and bone mass growth in prepubertal girls: a randomized, double-blind, placebo-controlled trial. J Clin Invest. 99:1287–1294.
10. **Holick MF.** 1994 Vitamin D: new horizons for the 21st century. Am J Clin Nutr. 60:619–630.
11. **Lee JJK, Lyne ED, Kleerekoper M, Logan MS, Belfi RA.** 1989 Disorders of bone metabolism in severely handicapped children and young adults. Clin Orthop Rel Res. 245:297–302.
12. **Reiter EO, Brugman SM, Pike JW, et al.** 1985 Vitamin D metabolites in adolescents and young adults with cystic fibrosis: effect of sun and season. J Pediatr. 106:21–26.
13. **Bailey DA, Faulkner RA, McKay HA.** 1996 Growth, physical activity and bone mineral acquisition. Exer Sports Sc Rev. 24:234–266.
14. **Kroger H, Kotaniemi A, Vainio P, Alhava E.** 1992 Bone densitometry of the spine and femur in children by dual energy x-ray absorptiometry. Bone Miner. 17:75–85.
15. **Slemenda CW, Miller JZ, Hui SL, Reister TK, Johnston CC.** 1991 Role of physical activity in the development of skeletal mass in children. J Bone Miner Res. 6:1227–1233.
16. **Lu PW, Cowell CT, Lloyd-Jones SA, Briody JN, Howman-Giles R.** 1996 Volumetric bone mineral density in normal subjects aged 5–27 years. J Clin Endocrinol Metab. 81:1586–1590.
17. **Specker BL, Brazerol W, Tsang RC, Levin R, Searcy J, Steichen J.** 1987 Bone mineral content in children 1 to 6 years of age. Am J Dis Child. 141:343–344.
18. **McCormick DP, Ponder SW, Fawcett HD, Palmer JL.** 1991 Spinal bone mineral density in 335 normal and obese children and adolescents: evidence for ethnic and sex differences. J Bone Miner Res. 6:507–513.
19. **Glastre C, Braillon P, David L, Cochat P, Meunier PJ, Delmas PD.** 1990 Measurement of bone mineral content of the lumbar spine by dual energy x-ray absorptiometry in normal children: correlations with growth parameters. J Clin Endocrinol Metab. 70:1330–1333.
20. **Nelson DA, Simpson PM, Johnson CC, Barondess DA, Kleerekoper M.** 1997 The accumulation of whole body skeletal mass in third- and fourth-grade children: effects of age, gender, ethnicity, and body composition. Bone. 20:73–78.
21. **Gilsanz V, Gibbens DT, Roe TF, et al.** 1988 Vertebral bone density in children: effect of puberty. Radiology. 166:847–850.
22. **Geusens P, Cantatore F, Nijs J, Proesmans W, Emma F, Dequeker J.** 1991 Heterogeneity of growth of bone in children at the spine radius and total skeleton. Growth Dev Aging. 55:249–256.
23. **Southard RN, Morris JD, Mahan JD, et al.** 1991 Bone mass in healthy children: measurement with quantitative DXA. Radiology. 179:735–738.
24. **Seeman E.** 1997 From density to structure: growing up and growing old on the surfaces of bone. J Bone Miner Res. 12:1–13.
25. **Dwyer T, Ponsonby AL, Newman NM, Gibbons LE.** 1991 Prospective cohort study of prone sleeping position and sudden infant death syndrome. Lancet. 337:1244–1247.
26. **Withers RT, Davies GJ, Crouch RG.** 1977 A comparison of 3 W170 protocols. Eur J Appl Physiol. 37:123–128.
27. **Pyke JE.** 1985 Australian Health and Fitness survey. Edwardstown, Australia: KB Printing Services.
28. **Dwyer T, Blizzard CL, Gies PH, Ashbolt R, Roy C.** 1996 Assessment of habitual sun exposure in adolescents via questionnaire: a comparison with objective measurement using polysulphone badges. Melanoma Res. 6:231–239.
29. **Mietinen OS.** 1985 Theoretical epidemiology: principles of occurrence research in medicine. New York: Wiley and Sons.
30. **Elwood JM, Diffey BL.** 1993 A consideration of ambient solar ultraviolet radiation in the interpretation of studies of the aetiology of melanoma. Melanoma Res. 3:113–122.
31. **Gies P.** 1994 Ambient ultraviolet radiation. Washington, DC: Society of Photo-Optical Instrumentation Engineers. Vol 2282. 282.
32. **McKenzie RL, Elwood JM.** 1990 Intensity of ultraviolet radiation and its implications for skin cancer. NZ Med J. 103:152–154.
33. **Docio S, Riancho JA, Perez A, Olmos JM, Amado JA, Gonzalez-Macias J.** 1998 Seasonal deficiency of vitamin D in children: a potential target for osteoporosis-preventing strategies. J Bone Miner Res. 13:544–548.
34. **Chapuy M-C, Preziosi P, Maamer M, et al.** 1997 Prevalence of vitamin D deficiency in an adult normal population. Osteoporosis Int. 7:439–443.
35. **Carter DR, Hayes WC.** 1977 The compressive behavior of bone as a two phase porous structure. J Bone Joint Surg. 59:954–962.
36. **Dwyer T, Blizzard L.** 1997 Inferring cause in observational studies. Aust Epidemiologist. 4:18–22.