Colin Proksel (034133)
OSBORN MALEDON, P.A.
2929 North Central Avenue, 21st Floor
Phoenix, Arizona 85012-2793
State Bar No. 034133
Telephone: (602) 640-9000
Facsimile: (602) 640-9050
Email: cproksel@omlaw.com
*Attorney for Plaintiffs*
*Additional counsel listed in signature block*

**UNITED STATES DISTRICT COURT**
**FOR THE DISTRICT OF ARIZONA**
**TUCSON DIVISION**

Jane Doe, by her next friend and parents Helen Doe and James Doe; and Megan Roe, by her next friend and parents, Kate Roe and Robert Roe,

Plaintiffs,

v.

Thomas C. Horne in his official capacity as State Superintendent of Public Instruction; Laura Toenjes, in her official capacity as Superintendent of the Kyrene School District; Kyrene School District; The Gregory School; and Arizona Interscholastic Association Inc.,

Defendants.

Case No. 4:23-cv-00185-JGZ

**PLAINTIFFS' EXHIBIT LIST FOR PLAINTIFFS' MOTION FOR PRELIMINARY INJUNCTION**

Plaintiffs submit the following list of exhibits, along with copies of exhibits not already filed on the docket, pursuant to the Court's June 14, 2023 Order (ECF No. 80). Plaintiffs respectfully reserve the right to amend this exhibit list in advance of the hearing.

| Ex. No. | Description | Location |
|---|---|---|
| 1 | Declaration of Jane Doe | ECF No. 6 |
| 2 | Declaration of Helen Doe | ECF No. 7 |
| 3 | Second Declaration of Helen Doe | ECF No. 78 |
| 4 | Declaration of Megan Roe | ECF No. 8 |
| 5 | Declaration of Kate Roe | ECF No. 9 |
| 6 | Declaration of Stephanie Budge, Ph.D. | ECF No. 4 |
| 7 | Rebuttal Declaration of Stephanie Budge, Ph.D. | ECF No. 65-1 |
| 8 | Declaration of Daniel Shumer, M.D., MPH | ECF No. 5 |

| | | |
|---|---|---|
| 9 | Rebuttal Declaration of Daniel Shumer, M.D., MPH | ECF No. 65-2 |
| 10 | AIA's Constitution, Bylaws, Policies, and Procedures 2022-2023, Transgender Policy | ECF No. 51-1 |
| 11 | Photographs of the Doe Family (*filed under seal*)[1] | Attached |
| 12 | Photographs of the Roe Family (*filed under seal*) | Attached |
| 13 | Jane Doe's Name Change Court Order (*filed under seal*) | Attached |
| 14 | Megan Roe's Name and Gender Change Court Order (*filed under seal*) | Attached |
| 15 | Jane Doe's Passport (*filed under seal*) | Attached |
| 16 | Megan Roe's Passport (*filed under seal*) | Attached |
| 17 | Consideration of Bills: Hearing on S.B. 1165 Before S. Comm. on Judiciary, Jan. 20, 2022, 55th Leg., 2nd Reg. Sess., 00:08:08–01:30:05 (*filed as a non-electronic exhibit*)[2] | Cover page attached |
| 18 | David Handelsman, et al., *Circulating Testosterone as the Hormonal Basis of Sex Differences in Athletic Performance*, 39 Endocrine Revs. 803 (2018) | Attached |
| 19 | David Handelsman, *Sex Differences in Athletic Performance Emerge Coinciding with the Onset of Male Puberty*, 87 Clinical Endocrinology 68 (2017) | Attached |
| 20 | Jonathon W. Senefeld et al., *Sex Differences in Youth Elite Swimming*, 14 PLOS ONE 1 (2019) | Attached |
| 21 | Joanna Harper, *Race Times for Transgender Athletes*, 6 J. Sporting Cultures & Identities 1 (2015) | Attached |
| 22 | Marnee McKay & Joshua Burns, *When it Comes to Sport, Boys "Play Like a Girl,"* The Conversation (Aug. 3, 2017), https://theconversation.com/when-it-comes-to-sport-boys-play-like-a-girl-80328 | Attached |
| 23 | Marnee McKay, et al., *Normative Reference Values for Strength and Flexibility of 1,000 Children and Adults*, Neurology, 88 (1) (2017) | Attached |
| 24 | World Rugby Transgender Women's Guidelines (2020), https://www.world.rugby/the-game/player-welfare/guidelines/transgender/women | Attached |
| 25 | Governor Douglas A. Ducey's Letter to Arizona Secretary of State re: Senate Bill 1138 and 1165 | Attached |

[1] Plaintiffs respectfully request by forthcoming separate motion that Exhibits 11-16 be filed under seal because these documents provide sensitive and private information about the Plaintiffs and their families, who the Court has ordered may proceed in this case under pseudonyms.

[2] Plaintiffs respectfully request by forthcoming separate motion that Exhibit 17 be filed as a non-electronic exhibit with the Court because the exhibit is a multimedia file that cannot be filed using the CM/ECF system.

Respectfully submitted this 29th day of June, 2023.

*/s/ Colin M. Proksel*

Colin M. Proksel (034133)
OSBORN MALEDON, P.A.
2929 North Central Avenue, 21st Floor
Phoenix, Arizona 85012-2793
Telephone: (602) 640-9000
Facsimile: (602) 640-9050
Email: cproksel@omlaw.com

Jyotin Hamid*
Justin R. Rassi*
Amy C. Zimmerman*
DEBEVOISE & PLIMPTON LLP
66 Hudson Boulevard
New York, New York 10001
Telephone: (212) 909-6000
Facsimile: (212) 909-6836
Email: jhamid@debevoise.com
Email: jrassi@debevoise.com
Email: azimmerman@debevoise.com

Amy Whelan*
Rachel Berg*
NATIONAL CENTER FOR LESBIAN RIGHTS
870 Market Street, Suite 370
San Francisco, California 94102
Telephone: (415) 343-7679
Facsimile: (415) 392-8442
Email: awhelan@nclrights.org
Email: rberg@nclrights.org

**Admitted pro hac vice.*

# Exhibit 11
# Filed Under Seal

# Exhibit 12
# Filed Under Seal

# Exhibit 13
# Filed Under Seal

# Exhibit 14
# Filed Under Seal

# Exhibit 15
# Filed Under Seal

# Exhibit 16
# Filed Under Seal

# Exhibit 17

# Filed as a Non-Electronic Exhibit

# Exhibit 18

# Circulating Testosterone as the Hormonal Basis of Sex Differences in Athletic Performance

David J. Handelsman,[1,2] Angelica L. Hirschberg,[3,4] and Stephane Bermon[5,6]

[1]ANZAC Research Institute, University of Sydney, Concord, New South Wales 2139, Australia; [2]Department of Andrology, Concord Hospital, Sydney, New South Wales 2139, Australia; [3]Department of Women's and Children's Health, Karolinska Institutet, 171 76 Stockholm, Sweden; [4]Department of Gynecology and Reproductive Medicine, Karolinska University Hospital, 171 76 Stockholm, Sweden; [5]Laboratoire Motricité Humaine, Education, Sport, Santé, Université Côte d'Azur, 06000 Nice, France; and [6]Health and Science Department, International Association of Athletics Federations, 98000 Monaco

**ABSTRACT** Elite athletic competitions have separate male and female events due to men's physical advantages in strength, speed, and endurance so that a protected female category with objective entry criteria is required. Prior to puberty, there is no sex difference in circulating testosterone concentrations or athletic performance, but from puberty onward a clear sex difference in athletic performance emerges as circulating testosterone concentrations rise in men because testes produce 30 times more testosterone than before puberty with circulating testosterone exceeding 15-fold that of women at any age. There is a wide sex difference in circulating testosterone concentrations and a reproducible dose-response relationship between circulating testosterone and muscle mass and strength as well as circulating hemoglobin in both men and women. These dichotomies largely account for the sex differences in muscle mass and strength and circulating hemoglobin levels that result in at least an 8% to 12% ergogenic advantage in men. Suppression of elevated circulating testosterone of hyperandrogenic athletes results in negative effects on performance, which are reversed when suppression ceases. Based on the nonoverlapping, bimodal distribution of circulating testosterone concentration (measured by liquid chromatography–mass spectrometry)—and making an allowance for women with mild hyperandrogenism, notably women with polycystic ovary syndrome (who are overrepresented in elite athletics)—the appropriate eligibility criterion for female athletic events should be a circulating testosterone of <5.0 nmol/L. This would include all women other than those with untreated hyperandrogenic disorders of sexual development and noncompliant male-to-female transgender as well as testosterone-treated female-to-male transgender or androgen dopers. ***(Endocrine Reviews 39: 803 – 829, 2018)***

Virtually all elite sports are segregated into male and female competitions. The main justification is to allow women a chance to win, as women have major disadvantages against men who are, on average, taller, stronger, and faster and have greater endurance due to their larger, stronger muscles and bones as well as a higher circulating hemoglobin level. Hence, elite female competition forms a protected category with entry that must be restricted by an objective eligibility criterion related, by necessity, to the relevant sex-specific physical advantages. The practical need to establish an eligibility criterion for elite female athletic competition led the International Association of Athletic Federations (IAAF) to establish a rule in 2011, endorsed by the International Olympic Committee (IOC) in 2012, for hyperandrogenic women. That IAAF regulation stated that for athletes to be eligible to compete in female events, the athlete must be legally recognized as a female and, unless she has complete androgen insensitivity, maintain serum testosterone <10 nmol/L. That IAAF eligibility rule was challenged by an athlete to the Court for Arbitration in Sports, which ruled in 2015 that, although an eligibility criterion was justified, there was insufficient evidence of the extent of the competitive advantage enjoyed by hyperandrogenic athletes who had circulating testosterone >10 nmol/L over female athletes with circulating testosterone in the normal female range. The Court for Arbitration in Sports suspended the rule pending receipt of such evidence. In that context, the present review presents the available evidence on the hormonal basis for the sex difference

ISSN Print: 0163-769X
ISSN Online: 1945-7189
Printed: in USA

Received: 28 January 2018
Accepted: 18 June 2018
First Published Online: 13 July 2018

**ESSENTIAL POINTS**

- It is widely accepted that elite athletic competitions should have separate male and female events
- The main justification is that men's physical advantages in strength, speed, and endurance mean that a protected female category, with objective entry criteria, is required
- Prior to puberty, there is no sex difference in circulating testosterone concentrations and athletic performance
- From male puberty onward, the sex difference in athletic performance emerges as circulating testosterone concentrations rise as the testes produce 30 times more testosterone than before puberty, resulting in men having 15- to 20-fold greater circulating testosterone than children or women at any age
- This wide, bimodal sex difference in circulating testosterone concentrations and the clear dose-response relationships between circulating testosterone and muscle mass and strength, as well as the hemoglobin level, largely account for the sex differences in athletic performance
- Based on the nonoverlapping, bimodal distribution of circulating testosterone concentration (measured by liquid chromatography–mass spectrometry) with 95% references ranges of 7.7 to 29.4 nmol/L in healthy men and 0 to 1.7 nmol/L in healthy premenopausal women—making an allowance for women with the mild hyperandrogenism of polycystic ovary syndrome, who are overrepresented in elite athletics—the eligibility criterion for female athletic events should be a circulating testosterone concentration of $<5.0$ nmol/L

in athletic performance. It concludes that the evidence justifies a revised eligibility criterion of a threshold circulating testosterone concentration of 5 nmol/L (measured by a mass spectrometry method).

## Sex, Fairness, and Segregation in Sport

If sports are defined as the organized playing of competitive games according to rules (1), fixed rules are fundamental in representing the boundaries of fair sporting competition. Rule breaking, whether by breaching eligibility or competition rules, such as use of banned drugs, illegal equipment, or match fixing, creates unfair competitive advantages that violate fair play. Cheating constitutes a fraud against not just competitors but also spectators, sponsors, the sport, and the public. In the absence of genuine fair competition, elite sports would lose their wide popular appeal and ability to captivate and inspire with the authentic attraction of genuine contest between highly trained athletes.

Nevertheless, fairness is an elusive, subjective concept with malleable boundaries that may change over time as social concepts of fairness evolve. For example, until the late 19th century when organized sports trainers emerged, training itself was considered a breach of fairness because competition was envisaged at that time as a contest based solely on natural endowments. Similarly, sports once distinguished between amateurs and professionals. The concept of fairness has deep and complex philosophical roots mainly focused on notions of distributive justice. These considerations affect sports through the universal application of antidiscrimination and human rights legislation. Less attention is given to the philosophical basis of fair competition in elite sports, where the objectives are not egalitarian but aim to discover a hierarchy of achievement derived from a mixture of unequal natural talent and individual training effort. Excellent, insightful discussion of the legal and moral complexities of sex and fair competition in elite sports from a legal scholar and former elite female athlete is available (2).

The terms *sex* and *gender* are often confused and used as if interchangeable. *Sex* is an objective, specific biological state, a term with distinct, fixed facets, notably genetic, chromosomal, gonadal, hormonal, and phenotypic (including genital) sex, each of which has a characteristic defined binary form. Whereas all facets of biological sex are almost always aligned so that assignment of sex at birth is straightforward, rare instances in which two or more facets of biological sex conflict constitute an intersex state, now referred to as disorders (or differences) of sex development (DSDs) (3). In contrast, *gender* is a subjective, malleable, self-identified social construct that defines a person's individual gender role and orientation. Prompted by biological, personal, and societal factors, volitional expression of gender can take on virtually any form limited only by the imagination, with some individuals asserting they have not just a single natal gender but two genders, none, a distinct third gender, or gender that varies (fluidly) from time to time. Hence, whereas gender is usually consistent with biological sex as assigned at birth, in a few it can differ during life. For example, if gender were the basis for eligibility for female sports, an athlete could conceivably be eligible to compete at the same Olympics in both female and male events. These features render the unassailable personal assertion of gender identity incapable of forming a fair, consistent sex classification in elite sports.

The strongest justification for sex classification in elite sports is that after puberty men produce 20 times more testosterone than women (4–7), resulting in circulating testosterone concentrations 15-fold higher than in children or women of any age. Age-grade competitive sporting records show no sex differences prior to puberty, whereas from the age of male puberty onward there is a strong and ongoing male advantage (8). The striking male postpubertal increase in circulating testosterone provides a major, ongoing, cumulative, and durable physical advantage in sporting contests by creating larger and stronger bones, greater muscle mass and strength, and higher circulating hemoglobin as well as possible psychological (behavioral) differences. In concert, these render women, on average, unable to compete effectively against men in power-based or endurance-based sports.

Sex classification in sports therefore requires proof of eligibility to compete in the protected (female) category. This deceptively simple requirement for fairness is taken for granted by peer female competitors who regard participation by males, or athletes with physical features closely resembling males, as unfair. This makes policing of eligibility inescapable for sports, to avoid unfair male participation in female events. However, such policing inevitably intrudes into highly personal matters so that it must be achieved with respect for dignity and privacy, demanding use of the least invasive, scientifically reliable means. Unsurprisingly, this dilemma has always been highly contentious since it first entered international elite sports in the early 20th century, and it has become increasingly prominent and contentious in recent decades; nevertheless, the requirement to maintain fair play in female events will not disappear as long as separate female competitions exist. During recent decades, there has been progressively better understanding of the complex biology of genetic sex determination and the impact of pubertal sexual maturation in establishing phenotypic sexual dichotomy in physical capabilities. These sex-dichotomous physical features form the basis of, but remain quite distinct from, adult gender roles and identity. During the last century, as knowledge grew, the attempts to formalize a scientific basis for the unavoidable necessity of policing eligibility for the female category have been continually challenged. Most recently, the increasing assertion of gender self-identification as a social criterion has further challenged the hegemony of biology for determining "sports sex," Coleman's apt term (2). Allowing subjective gender self-identification to become the sole criterion of sports sex would allow for gaming and perceptions of systematic unfairness to grow. The case for women's sports being defined by sex rather than gender, including the consequences of acceding to gender-based classification, has been outlined (9) in arguing the importance of proper medical management of athletes intending to compete in female events.

Separate male and female events in sports is a dominant form of classification that is superimposed on other graduated age group and weight classifications (*e.g.*, in weightlifting, power lifting, wrestling, boxing, rowing), which reflect differences in strength, power, and speed to ensure fairness in terms of opportunity to win and, additionally, safety in contact sports. Age and weight classifications rely on objective criteria (birth date, weigh-in weight) for eligibility, and so should sex classification. Nevertheless, some power sports dependent on explosive strength and power (*e.g.*, throwing events, sprinting) do not segregate weight classes, whereas other sports where height is an advantage (*e.g.*, basketball, jockeys) do not have height classifications. These sports disproportionately attract athletes with greater weight and/or power-to-weight ratio or advantageous stature, respectively. If sex classification were eliminated, such open or mixed competitions would be dominated almost exclusively by men. It therefore seems highly unlikely that sex classification would ever be discarded, despite calls on philosophical or sociological grounds to end "gender" classification in sport (10).

## Sex Difference in Circulating Testosterone Levels

### Testosterone biosynthesis, secretion, and regulation in men and women

An androgen is a hormone capable of developing and maintaining masculine characteristics in reproductive tissues (notably the genital tract, as well as in other tissues and organs associated with secondary sexual characteristics and fertility) and contributing to the anabolic status of nonreproductive body tissues (11). The two dominant bioactive androgens circulating in mature mammals, including humans—testosterone and its more potent metabolite DHT—account for the development and maintenance of all androgen-dependent characteristics, and their circulating levels in men and nonpregnant women arise from steroids synthesized *de novo* in the testes, ovary, or adrenals (12).

The sexually undifferentiated gonads in the embryo develop into either ovaries or testes according to whether a Y chromosome (or at least the *sry* gene) is present. After birth and until puberty commences, circulating testosterone concentrations are essentially the same in boys and girls, other than briefly in the neonatal period of boys when higher levels prevail. The onset of male puberty, a brain-driven process triggered by a still mysterious hypothalamic or higher cerebral mechanism (13), initiates a hormonal cascade. In males, this leads to enhanced pituitary LH secretion that stimulates the 500 million Leydig cells in the testes

to secrete 3 to 10 mg (mean, 7 mg) of testosterone daily (4, 6, 7, 14, 15). This creates a very high local concentration of testosterone within the testis as well as a steep downhill concentration gradient into the bloodstream that maintains circulating testosterone levels at adult male levels, which are tightly regulated by strong negative hypothalamic feedback of circulating testosterone. In the absence of testes, these mechanisms do not function in females. In girls, serum testosterone increases during puberty (16), peaking at age 20 to 25 years before declining gradually with age (17, 18), but it remains <2 nmol/L at all ages, as determined by a reliable method (see below).

In adult women, circulating testosterone is derived from three roughly equal sources: direct secretion from the adrenal gland or the ovary and indirect extraglandular conversion (in liver, kidney, muscle, fat, skin) from testosterone precursors secreted by the adrenal and ovary. Only when circulating testosterone concentrations rise in male adolescents above the prepubertal concentrations does the virilization characteristic of men commence, progress, and endure throughout adult life, at least until old age (18). In combination, these different sources produce ~0.25 mg of testosterone daily so that throughout life women maintain circulating testosterone levels of <2 nmol/L. Circulating testosterone concentrations in women are subject to little dynamic physiological regulation. As a result, circulating testosterone concentrations in healthy premenopausal women are stable (nonfluctuating) and not subject to strong negative feedback by exogenous testosterone (as happens in men). Even the small rise (50%) at the time of the mid-cycle LH surge triggering ovulation (19) remains within the physiological range for premenopausal females.

### Male and female reference ranges for circulating testosterone

A reliable threshold for circulating testosterone must be set using measurement by the reference method of liquid chromatography–mass spectrometry (LC-MS) rather than using one of the various available commercial testosterone immunoassays. The necessary reliance on steroid mass spectrometry for clinical applications in endocrinology, reproductive medicine, and sports medicine is widely recognized. It has been standard for decades in antidoping science (20), and the growing consensus is that it is required for high-quality clinical research and practice recognized by cognate professional societies (21, 22) and editorials in leading clinical endocrinology (23) and reproductive medicine (24) journals. The inherently limited specificity of testosterone immunoassays arises from antibody cross-reactivity with structurally related steroids (such as precursors and metabolites) other than the intended target. As a result, all steroid immunoassays, including for testosterone, display method-specific bias whereby, for example, the lower limit of a testosterone reference range in healthy young men varies from 7.3 to 12.6 nmol/L according to the immunoassay used, so that no consensus definition of a lower limit could be obtained independent of the commercial immunoassay method used (25). Furthermore, testosterone immunoassays are optimized for circulating levels in men but display increasing inaccuracy at the lower, by an order of magnitude, circulating testosterone concentrations in women or children. In contrast to immunoassays, LC-MS–based methods are highly specific and do not depend on proprietary antibodies. Using LC-MS–based measurements, method-specific bias can be avoided and a fixed consensus lower reference limit defined (Table 1). Hence, for the precision required in sports medicine, whether for eligibility criteria or antidoping applications, testosterone in serum must be measured by LC-MS methods.

Prior to puberty, levels of circulating testosterone as determined by LC-MS are the same in boys and girls (16). They remain lower than 2 nmol/L in women of all ages. However, from the onset of male puberty the testes secrete 20 times more testosterone resulting in circulating testosterone levels that are 15 times greater in healthy young men than in age-similar women. Using LC-MS measurement, circulating testosterone in adults has a strikingly nonoverlapping bimodal distribution with wide and complete separation between men and women. Table 1 (25–36) summarizes data from appropriate reported studies using mass spectrometry–based methods to measure serum testosterone in healthy men and women. Based on a number-weighted pooling with conventional 95% two-sided confidence limits of the eight available studies using LC-MS measurements of serum testosterone, the reference range for healthy young men (18 to 40 years) is 7.7 nmol/L to 29.4 nmol/L. Similarly, summarizing the nine available studies for healthy menstruating women under 40 years, the 95% (two-sided) reference range is 0 to 1.7 nmol/L. These reference limits do not control for factors such as oral contraceptive use (35, 36), menstrual phase (19), SHBG (37, 38), overweight (39, 40), fasting and smoking (41), diet (40), and physical activity (42, 43) in women and men, all of which have small effects on circulating testosterone but without materially influencing the divergence between the nonoverlapping bimodal distribution of male and female reference ranges of circulating testosterone.

In creating a threshold for eligibility for female events it is also necessary to make allowance for women with polycystic ovary syndrome (PCOS) and nonclassical adrenal hyperplasia. PCOS is a relatively common disorder among women of reproductive ages with a prevalence of 6% to 10%, depending on the diagnostic criteria used (44), in which mild hyperandrogenism is a key clinical feature and has higher than expected prevalence among elite female athletes

**Table 1.** Serum Testosterone Measurements by LC-MS Methods in Studies of Healthy Men and Women

| Study | Sample (Age 18–40 y) | N | Lower 95% CL (nmol/L) | Upper 95% CL (nmol/L) |
|---|---|---|---|---|
| Men | | | | |
| Sikaris *et al.*, 2005 (25) | Elite, eugonadal | 124 | 10.4 | 30.1 |
| Turpeinen *et al.*, 2008 (26) | Convenience | 30 | 10.1 | 31.2 |
| Kushnir *et al.*, 2010 (27) | Convenience | 132 | 7.2 | 24.2 |
| Salameh *et al.*, 2010 (28) | Convenience | 264 | 7.1 | 39.0 |
| Neale *et al.*, 2013 (29) | Convenience | 67 | 10.6 | 31.9 |
| Kelsey *et al.*, 2014 (30) | Secondary pooled analysis | 1058 | 7.2 | 25.3 |
| Hart *et al.*, 2015 (31) | Birth cohort | 423 | 7.4 | 28.0 |
| Travison *et al.*, 2017 (32) | Pooled two cohorts | 1656 | 7.9 | 31.1 |
| Number-weighted mean | | | 7.7 | 29.4 |
| Women | | | | |
| Turpeinen *et al.*, 2008 (26) | Convenience | 32 | 0.8 | 2.8 |
| Kushnir *et al.*, 2010 (27) | Convenience | 104 | 0.3 | 2.0 |
| Salameh *et al.*, 2010 (28) | Convenience | 235 | 0.03 | 1.5 |
| Haring *et al.*, 2012 (33) | Population-based | 263 | 0.04 | 2.0 |
| Neale *et al.*, 2013 (29) | Convenience | 90 | 0 | 1.7 |
| Bui *et al.*, 2013 (34) | Convenience | 25 | 0.30 | 1.69 |
| Rothman *et al.*, 2013 (19) | Convenience | 31 | 0.4 | 0.92 |
| Bermon and Garnier, 2017 (35) | Elite athletes | 1652 | 0 | 1.62 |
| Eklund *et al.*, 2017 (36) | Elite athletes and controls | 223 | 0.26 | 1.73 |
| Number-weighted mean | | | 0.06 | 1.68 |

Abbreviation: CL, confidence limit.

(36, 45–47). Nonclassical adrenal hyperplasia is a milder and later (adult) onset variant of classical congenital adrenal hyperplasia (48) with a much higher but still rare population prevalence (1:1000 vs 1:16,000 for the classical variant) (49). Table 2 (50–64) summarizes clinical studies (n = 16, ≥40 women) reporting serum testosterone concentrations measured by LC-MS in samples from women with PCOS.

The pooled data reveal that the upper limit of serum testosterone in women with PCOS is 3.1 nmol/L (95% CI, one-sided) or 4.8 nmol/L (using a 99.99% CI, one-sided) (Table 3). Hence, a conservative threshold for circulating testosterone of 5 nmol/L measured by LC-MS would identify <1:10,000 women with PCOS as false positives, based on circulating testosterone measurement alone. Circulating testosterone higher than this threshold is likely to be due to testosterone-secreting adrenal or ovarian tumors, intersex/DSD, badly controlled or noncompliant male-to-female (M2F) transgender athletes, or testosterone doping.

### The physiological effects of testosterone depend on the circulating testosterone, not its source (endogenous or exogenous)

Testosterone, whether of a natural endogenous or manufactured exogenous source, has an identical chemical structure and biological effects, aside from minor differences in isotopic composition, which are biologically insignificant. At equivalent doses and circulating levels, exogenous testosterone exerts the same biological and clinical effects on every known androgen-responsive tissue or organ as endogenous testosterone, apart from effects on spermatogenesis, which as discussed below is only a matter of degree. Consequently, exogenous testosterone is a fully effective substitute for endogenous testosterone in therapeutic use, countering the effects of testosterone deficiency due to hypogonadism (reproductive system disorders). Any purported differences between endogenous and exogenous testosterone are due to corresponding differences in the endogenous production rate or exogenous dose. Such differences in

**Table 2.** Summary of Serum Testosterone (nmol/L) by LC-MS in Women With PCOS From 16 Studies

| Study | N | Mean | SD |
|---|---|---|---|
| Moran *et al.*, 2017 (50) | 92 | *0.24* | *0.08* |
| Münzker *et al.*, 2017 (51) | 274 | *0.93* | *0.19* |
| O'Reilly *et al.*, 2017 (52) | 114 | *0.55* | *0.19* |
| Handelsman *et al.*, 2017 (53) | 152 | 0.38 | 0.25 |
| Pasquali *et al.*, 2016 (54) | 156 | 1.17 | 0.47 |
| Yang *et al.*, 2016 (55) | 1159 | 2.2 | 1.44 |
| Tosi *et al.*, 2016 (56) | 116 | 1.33 | 0.55 |
| Daan *et al.*, 2015 (57) | 170 | *1.64* | *0.53* |
| Bui *et al.*, 2015 (58) | 44 | *0.85* | *0.3* |
| Keefe *et al.*, 2014 (59) | 52 | *1.7* | *0.97* |
| Yasmin *et al.*, 2013 (60) | 165 | 1.99 | 1.02 |
| Janse *et al.*, 2011 (61) | 200 | *1.12* | *0.47* |
| Jedel *et al.*, 2011 (62) | 72 | *0.23* | *0.08* |
| Legro *et al.*, 2010 (Mayo) (63) | 596 | *2.12* | *0.89* |
| Legro *et al.*, 2010 (Quest) (63) | 596 | *1.98* | *0.97* |
| Stener-Victorin *et al.*, 2010 (64) | 74 | 1.53 | 0.62 |
| Sum | 4032 | | |
| Number-weighted mean | | 1.69 | 0.87 |

Data taken directly from paper or interpolated from other data (*e.g.*, median, quartiles, ranges, sample size) supplied as described by Wan *et al.*, 2014 (Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. BMC Med Res Methodol 14: 135) are shown in italics.

effective exposure lead to corresponding differences in circulating testosterone levels and its effects according to the dose-response curves for testosterone.

Similar to all hormones and drugs, over their effective range of biological activity the dose-response relationship for testosterone is usually a sigmoidal curve with lower and upper plateaus joined by a monotonically rising middle region, which may be linear in the natural scale but more often log-linear (linear on the log or similar transformed scale). In the middle portion of the typical sigmoidal dose-response curve for the same increase in testosterone dose (or concentration), the response would be increased in simple proportional (*i.e.*, linear) but more often on a logarithmic scale. In contrast, at the lower and upper plateaus of dose or concentrations, changes in testosterone exposure may evoke minimal or no response on the endpoint. For example, in women of any age circulating testosterone concentrations are along the lower plateau of the dose-response curve, so that increases in circulating testosterone concentrations within that lower plateau may have minimal or no effect. In female athletes with the mild hyperandrogenism of PCOS, higher performance has been shown (47), with their muscle mass and power performance correlating with androgen levels (36). However, beyond these effects where endogenous testosterone concentrations are in the high-normal adult female range, it is only when the increases in circulating testosterone concentrations substantially and consistently exceed those prevailing in childhood (<2 nmol/L) and among women including those with PCOS (<5 nmol/L) that the effects would replicate the effects of rising testosterone concentrations of boys in middle to late puberty (typically >8 nmol/L), that is, the masculinizing effects of increased muscle, bone, and hemoglobin characteristics of men. As shown above, the circulating testosterone of most women never reaches consistently >5 nmol/L, a level that boys must sustain for some time to exhibit the masculinizing effects of male puberty.

In addition, the effects of testosterone are modulated in a form of fine tuning by the patterns of exposure, such as whether the circulating testosterone is delivered in the unphysiological steady-state format (*e.g.*, quasi–steady-state delivery by implant or transdermal products) or by the peak-and-trough delivery of injections, as opposed to the natural state of endogenous fluctuations in serum testosterone around the average adult male levels. However, these latter pattern effects are subtle and the dominant effect remains that of dose and average testosterone

concentrations in blood, however they arise. Furthermore, there is evidence that the androgen sensitivity of responsive tissues differs and may be optimal at different circulating testosterone concentrations (65).

Male sexual function is maintained by endogenous testosterone at adult male circulating concentrations. These effects can be replicated by exogenous testosterone if and only if it achieves comparable circulating testosterone concentrations. For example, in a well-controlled prospective study of older men with prostate cancer (66), androgen deprivation achieving castrate levels of circulating testosterone sustained during 12 months markedly suppressed sexual desire and function, whereas those effects did not occur in age-matched men having nonhormonal treatment of prostate cancer or those without prostate cancer. In healthy younger men whose endogenous testosterone was fully suppressed, sexual function completely recovered when circulating testosterone was restored to the physiological male range by administration of exogenous testosterone (67). Similar effects were also observed in healthy, middle-aged men in whom male sexual function was fully maintained (compared with placebo) during 2 years of treatment with an exogenous androgen (DHT) despite that treatment causing sustained, complete suppression of endogenous testosterone (68). This further supports the key interpretation that the biological effects of exogenous or endogenous testosterone are the same at comparable circulating levels.

Clinically, exogenous testosterone replicates fully all effects of endogenous testosterone on every reproductive and nonreproductive organ or tissue, with the sole exception of the testis. Sperm production in the testis requires a very high concentration of testosterone (typically 100-fold greater than in the general bloodstream), which is produced in nature only by the action of the pituitary hormone LH. LH stimulates the Leydig cells in the interstitial space of the testis between seminiferous tubules to produce high intratesticular concentrations of testosterone, which are necessary and sufficient to initiate and maintain sperm production in the adjacent seminiferous tubules. This high concentration of testosterone also provides a downhill gradient to supply the rest of the body, where circulating testosterone acts on androgen-responsive tissues to produce and maintain masculine patterns of androgenization. When exogenous testosterone (or any other androgen) is administered to men, pituitary LH is suppressed by negative feedback and the sperm production halts for as long as exogenous testosterone or androgen exposure continues, after which it recovers (69). However, even the reduction in spermatogenesis and testis size when men are treated with exogenous testosterone is only a matter of degree. It is well established in rodents (70, 71) that spermatogenesis is induced by exogenous testosterone when the testosterone concentrations in the testis are high enough to replicate what occurs naturally via LH stimulation (72). However, direct replication that high-dose testosterone also initiates and maintains spermatogenesis in humans is not feasible, as these testosterone doses are 10- to 100-fold higher than could be safely given to humans. Nevertheless, confirmatory evidence in humans is available from rare cases of men with an activating mutation of the chorionic gonadotropin/LH receptor (73, 74). This mutation causes autonomous testicular testosterone secretion leading to precocious puberty arising from the premature adult male circulating testosterone concentrations that lead to complete suppression of circulating gonadotropin (LH, FSH) secretion. In this illustrative case the testis was exposed to nonphysiologically high testosterone concentrations (but without any gonadotropin stimulation) that induced sperm production and allowed for natural paternity (73). This indicates that even for spermatogenesis, exogenous testosterone can replicate all biological effects of endogenous testosterone in accordance with the relevant dose-response characteristics.

The most realistic view is that increasing circulating testosterone from the childhood or female range to the adult male range will have the same physiological effects whether the source of the additional testosterone is endogenous or exogenous. This is strongly supported by well-established knowledge about the relationship of circulating testosterone concentrations

**Table 3.** Upper Confidence Limits on Serum Testosterone in Women With PCOS

| Confidence Interval | Likelihood[a] | SD[b] | One-Sided[c] | Two-Sided[c] |
|---|---|---|---|---|
| 95% | 1:20 | 1.96 | 3.13 | 3.39 |
| 99% | 1:100 | 2.35 | 3.47 | 3.73 |
| 99.9% | 1:1000 | 3.10 | 4.21 | 4.39 |
| 99.99% | 1:10,000 | 3.72 | 4.77 | 4.95 |

[a]Likelihood that a woman with PCOS would exceed that limit by chance.
[b]Number of SDs for each confidence limit.
[c]Two-sided CIs are conventional for a result that could exceed or fall below confidence limits, but here we focus only on values exceeding the upper limit, so that one-sided confidence limits are appropriate.

with the timing and manifestations of male puberty. The characteristic clinical features of masculinization (*e.g.*, muscle growth, increased height, increased hemoglobin, body hair distribution, voice change) appear only if and when circulating testosterone concentrations rise into the range of males at mid-puberty, which are higher than in women at any age even after the rise in circulating testosterone in female puberty. If and only if the pubertal rise in circulating testosterone fails will the males affected be clinically considered hypogonadal. Such a failure of male puberty may occur for genetic reasons (arising from mutations that inactivate any of the cascade of proteins whose activity is critical in the hypothalamus to trigger male puberty) or as a result of acquired conditions, caused by pathological disorders of the hypothalamus or pituitary or functional defects arising from severe deficits of energy or nutrition (*e.g.*, extreme overtraining, undernutrition), with the latter being comparable with hypothalamic amenorrhea or anorexia nervosa in female athletes/ballet dancers. If male puberty fails, testosterone replacement therapy is fully effective in replicating all of the distinctive masculine features apart from spermatogenesis.

#### Elevated circulating testosterone concentration caused by DSDs

Rare genetic intersex conditions known as DSDs can lead to markedly increased circulating testosterone in women. When coupled with ambiguous genitalia at birth, they may appear as undervirilized males or virilized females. This can cause athletes who were raised and identify as women to have circulating testosterone levels comparable to those of men and greatly exceeding those of non-DSD (and nondoped) women, including those with PCOS. Key congenital disorders in this category are 46,XY DSDs, namely 5$\alpha$ reductase deficiency (75), 17$\beta$-hydroxysteroid dehydrogenase type 3 deficiency (76), and androgen insensitivity (77, 78), as well as congenital adrenal hyperplasia (79), which is a 46,XX DSD. There is evidence that the first three conditions, components of 46,XY DSDs, are 140-fold more prevalent among elite female athletes than expected in the general population (80).

Genetic 5$\alpha$ reductase deficiency is due to an inactivating mutation in the 5$\alpha$ reductase type II enzyme (75). This leads to a deficit of DHT during fetal life when DHT is required for converting the sex-undifferentiated embryonic and fetal tissue to form the sex-differentiated masculine form external genitalia. Although genetic males (46,XY) with 5$\alpha$ reductase deficiency will develop testes, they usually remain undescended and labial fusion to form a scrotum and phallic growth does not occur. Hence, at birth the external genitalia may appear feminine, leading to a female assigned natal sex. Thus, individuals with 5$\alpha$ reductase deficiency may have male chromosomal sex (46,XY), gonadal sex (testes), and hormonal sex (adult male testosterone concentrations), but such severely undervirilized genitalia that affected individuals may be raised from birth as females rather than as undervirilized males. However, from the onset of male puberty, testicular Leydig cells start producing large amounts of testosterone, and the steep rise in circulating testosterone to adult male levels (with the permissive role of 5$\alpha$ reductase activity) leads to masculine virilization, including male patterns of muscle and bone growth, hemoglobin levels, and other masculine body habitus features (hair growth pattern, voice change), as well as phallic growth (80). Such changes of male puberty prompt around half affected individuals who had female sex assigned at birth and developed as girls prior to puberty to adopt a male gender identity and role at puberty (81). Sperm are formed in the testes so that, using *in vitro* fertilization, these individuals may father children (82).

17$\beta$-Hydroxysteroid dehydrogenase type 3 deficiency (76) has a natural history similar to that of 5$\alpha$ reductase deficiency. This disorder is due to inactivating mutations in a steroidogenic enzyme expressed only in the testis and that is essential for testosterone formation in the fetus. In the absence of a functional enzyme, the testis makes little testosterone but instead secretes large amounts of androstenedione, the steroid immediately prior to the enzymatic block. In the circulation, the excess of androstenedione is converted to testosterone (mainly by the enzyme AKR1C3) (12). Although the circulating testosterone is then converted to circulating DHT, insufficient DHT is formed locally within the urogenital sinus to virilize genitalia at birth. This causes the same severe undervirilization of the external genitalia of genetically male individuals, leading to ambiguous genitalia at birth despite male chromosomal, gonadal, and hormonal sex. When puberty arrives, the testes start producing the adult male testosterone output. Again, this leads to marked virilization and subsequent assumption of a male gender identity by some affected individuals, conflicting with a female assigned natal sex and childhood upbringing.

Androgen insensitivity, which arises from mutation in the androgen receptor (AR), poses different but complex challenges for eligibility for female athletic events. As the AR is located on the X chromosome, genetic males (46,XY) are hemizygous, so that an inactivating mutation in the AR can be partially or fully insensitive to androgen action. Affected individuals have male internal genitalia (testes in the inguinal canal or abdomen with Wolffian ducts) and consequently adult male circulating testosterone concentrations after puberty. These nonlethal mutations have a wide spectrum of functional effects, ranging from full resistance to all androgen action in complete androgen insensitivity syndrome (CAIS) where individuals have a full female phenotype with

normal female external genitalia, to partial androgen insensitivity syndrome (PAIS) where some androgen action is still exerted, leading to various degrees of ambiguous genitalia, or to mild androgen insensitivity, which produces a very mild, undervirilized male phenotype (normal male genital and somatic development but with little body hair and no male pattern balding) (77). Testosterone (and dihydrotestosterone) have no consistent effect of inducing normal nitrogen retention (anabolic) responses in patients with CAIS (83–86), although some reduced androgen responsiveness is retained by patients with PAIS (84, 87–90). Athletes with CAIS can compete fairly as females because the circulating testosterone, although at adult male levels, has no physiological effect so that, in terms of androgen action and the ensuing physical somatic advantages of male sex, affected individuals are indistinguishable from females and gain no benefits of the sex difference arising from unimpeded testosterone action. A more complex issue arises with athletes having PAIS reflecting the degree of incomplete impairment of AR function. Residual androgen action in such AR mutations is harder to characterize quantitatively, as there is no standardized, objective *in vitro* test to quantify AR functionality. Hence, individuals with PAIS may have adult male circulating testosterone concentrations but variable androgen sensitivity. At present, determination of eligibility to compete in the female category requires a case-by-case evaluation, primarily based on the degree of virilization. The current best available clinical approach to determining the functional impact (degree of functionality/sensitivity) of an AR mutation is based on the degree of somatic, primarily genital, virilization assessed according to the Quigley classification of grade of androgen sensitivity (91).

Congenital adrenal hyperplasia (CAH) is a relatively common defect in adrenal steroidogenesis in the enzymatic pathway, leading to synthesis of cortisol, aldosterone, and sex steroid precursors. The disease varies in severity from life-threatening (adrenal failure) to mild (hirsutism and menstrual irregularity), or even asymptomatic and undiagnosed. The most common mutations causing CAH occur in the 21-hydroxylase enzyme, accounting for 95% of cases (79). The defect leads to a bottleneck, creating a major backing up of precursor steroids that then overflow into other steroid pathways, leading to diagnostic high levels of 17-hydroxyprogesterone and, in female patients, excessive circulating testosterone or other adrenal-source androgen precursors (*e.g.*, androstenedione, dehydroepiandrosterone) that may be converted to testosterone in tissues. A common clinical problem with management of CAH is that glucocorticoid/mineralocorticoid treatment is not always fully effective partly due to variable compliance, which may leave high circulating testosterone, including well into or even above the normal male range (92). It is unlikely that mild nonclassical congenital adrenal hyperplasia is a major contributor to the mild hyperandrogenism prevalent among elite female athletes. The prevalence of PCOS (6% to 16%) is about 100-fold higher than mild nonclassical congenital adrenal hyperplasia (0.1%) (49), whereas a disproportionately high number of elite female athletes (especially in power sports) have PCOS (45). In one study of hyperandrogenic female athletes, even mild nonclassic adrenal hyperplasia was ruled out by normal 17-hydroxyprogesterone (36) and, in another (47), reported serum androstenedione and cortisol did not differ from controls, ruling out significant congenital adrenal hyperplasia.

## Sex Difference in Muscle, Hemoglobin, Bone, and Athletic Performance Relating to Adult Circulating Testosterone Concentrations

Following puberty, testosterone production increases (16) but remains <2 nmol/L in women, whereas in men testosterone production increases 20-fold (from 0.3 mg/d to 7 mg/d), leading to 15-fold higher circulating testosterone concentrations (15 vs 1 nmol/L). The greater magnitude of sex difference in testosterone production (20-fold) compared with circulating levels (15-fold) is due to women's higher circulating SHBG, which retards testosterone clearance, creating a slower circulating half-time of testosterone. This order-of-magnitude difference in circulating testosterone concentrations is the key factor in the sex difference in athletic performance due to androgen effects principally on muscle, bone, and hemoglobin.

*"Sex differences in height, where they exist, are largely dependent on postpubertal differences in circulating testosterone."*

### Muscle

#### *Biology*

It has been known since ancient times that castration influences muscle function. Modern knowledge of the molecular and cellular basis for androgen effects on skeletal muscle involves effects due to androgen (testosterone, DHT) binding to the AR that then releases chaperone proteins, dimerizes, and translocates into the nucleus to bind to androgen response elements in the promoter DNA of androgen-sensitive genes. This leads to increases in (1) muscle fiber numbers and size, (2) muscle satellite cell numbers, (3) numbers of myonuclei, and (4) size of motor neurons (93). Additionally, there is experimental evidence that testosterone increases skeletal muscle myostatin expression (94), mitochondrial biogenesis (95), myoglobin expression (96), and IGF-1 content (97), which may augment energetic and power generation of skeletal muscular activity.

Customized genetic mouse models can provide unique experimental insight into mammalian physiology that is unobtainable by human experimentation.

The tight evolutionary conservation of the mammalian reproductive system explains why genetic mouse models have provided consistent, high-fidelity replication of the human reproductive system (98, 99). Genetic males (46,XY) with androgen insensitivity displaying similar features occur through the spontaneous production of inactivating AR mutations in all mammalian species studied, including humans, where they are known as women with CAIS. The converse, genetic females (46,XX) resistant to all androgen action cannot occur naturally in humans or other mammals. This is because fully androgen-resistant females must have both X chromosomes carrying an inactivated AR. In turn, this requires acquiring one X chromosome from their father, and hemizygous males bearing a single X chromosome with an inactive AR produce no sperm, as a functional AR is biologically indispensable for making sperm in any mammal. However, androgen-resistant females can be bred by genetic engineering using the Cre-Lox system (100). An important finding from such studies is that androgen-resistant female mice have essentially the same muscle mass and function as wild-type androgen-sensitive females bearing normal AR, whereas androgen-resistant male mice have smaller and weaker muscle mass and function than do wild-type males and comparable instead with wild-type females (101). This indicates that androgen action, represented by circulating testosterone, is the key determinant of the higher muscle mass and strength characteristic of males compared with females. Furthermore, endogenous circulating testosterone has minimal effects on skeletal muscle mass and strength in female mice because of its low levels. Although these experiments cannot be replicated in humans, their key insight is that the higher circulating testosterone in males is the determinant of the male's greater muscle mass and function compared with females. Nevertheless, there is also evidence that hyperandrogenic women, mostly with PCOS, have increased muscle mass and strength that correlates with mildly increased circulating testosterone in the high-normal female range (36, 47).

### *Observational data*

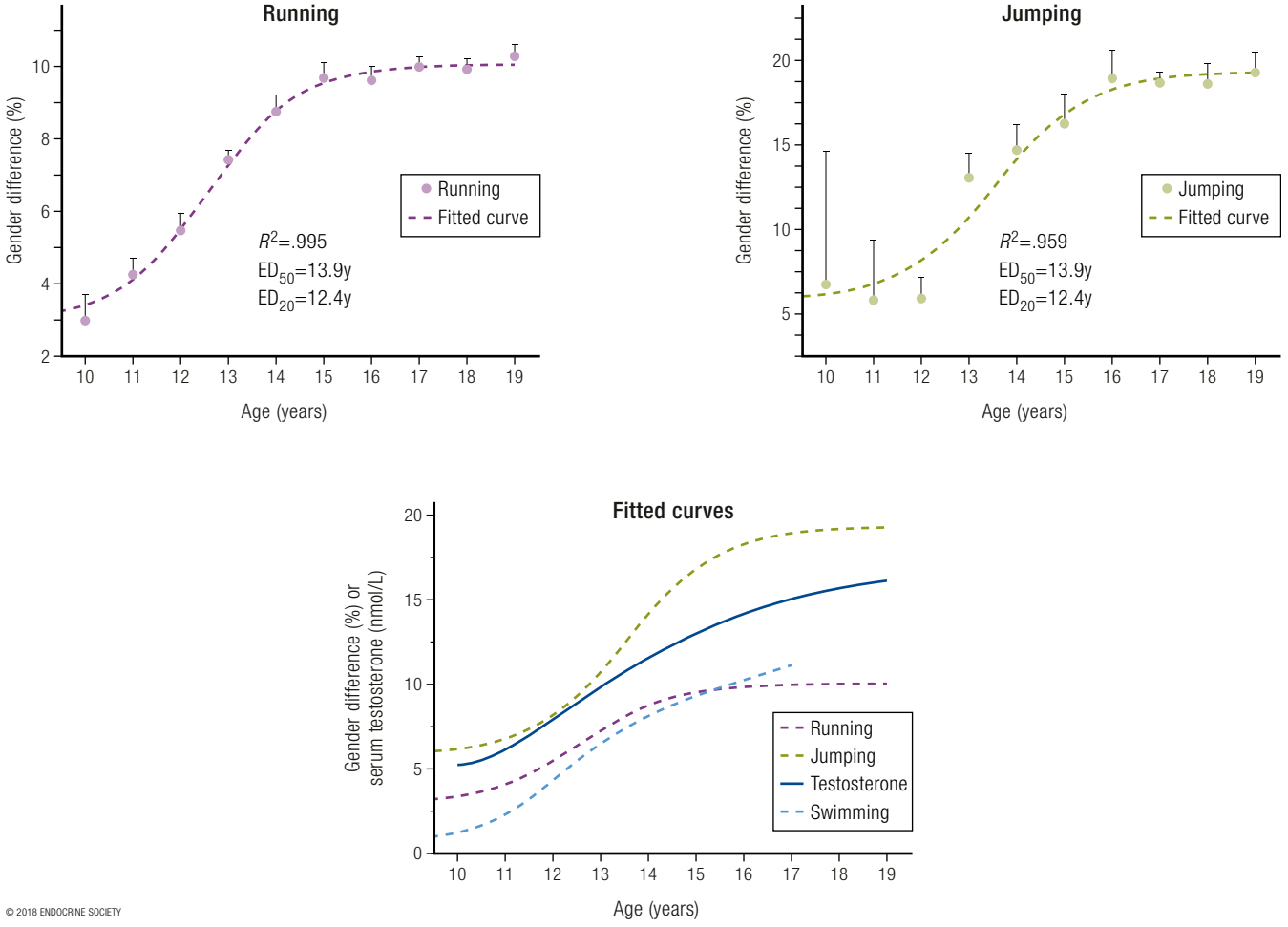
There is a clear sex difference in both muscle mass and strength (102–104) even adjusting for sex differences in height and weight (104, 105). On average, women have 50% to 60% of men's upper arm muscle cross-sectional area and 65% to 70% of men's thigh muscle cross-sectional area, and women have 50% to 60% of men's upper limb strength and 60% to 80% of men's leg strength (106). Young men have on average a skeletal muscle mass of >12 kg greater than age-matched women at any given body weight (104, 105). Whereas numerous genes and environmental factors (including genetics, physical activity, and diet) may contribute to muscle mass, the major cause of the sex difference in muscle mass and strength is the sex difference in circulating testosterone.

Age-grade competitive sports records show minimal or no female disadvantage prior to puberty, whereas from the age of male puberty onwards there is a strong and ongoing male advantage. Corresponding to the endogenous circulating testosterone increasing in males after puberty to 15 to 20 nmol/L (sharply diverging from the circulating levels that remain <2 nmol/L in females), male athletic performances go from being equal on average to those of age-matched females to 10% to 12% better in running and swimming events, and 20% better in jumping events (8) (Fig. 1). Corroborative findings are provided by a Norwegian study that examined performance of adolescents in certain athletic events but without reference to contemporaneous circulating testosterone concentrations (107). The striking postpubertal increase in male circulating testosterone provides a major, ongoing, cumulative, and durable advantage in sporting contests by creating greater muscle mass and strength. These sex differences render women unable to compete effectively against men, especially (but not only) in power sports.

These findings are supported by studies of non-athletic women showing that muscle mass is increased in proportion to circulating testosterone in women with mildly elevated testosterone levels due to PCOS (108, 109), a condition that is more prevalent among elite female athletes who exhibit these features (36, 45, 47), often undiagnosed (46), but that may provide an ergogenic advantage (47), consistent with the graded effects of circulating testosterone on explosive performance in men and women (110).

Studies of elite female athletes further corroborate these findings. One study demonstrates dose-response effects of better performance in some (400 m running, 400 m hurdles, 800 m running, hammer throw, pole vault) but not all athletic events correlated with significantly higher endogenous testosterone in female, but not male, athletes. Even within the low circulating testosterone levels prevailing within the normal female range, in these events there was a significant advantage of 1.8% to 4.5% among those in the highest tertile compared with the lowest tertile of endogenous testosterone (35). A further study of elite female athletes corroborates and extends these observations in that endogenous androgens are associated with a more anabolic body composition as well as enhanced muscular performance (36). In this study, 106 Swedish Olympic female athletes were compared with 117 age- and weight (body mass index)-matched sedentary control women for their muscle and bone mass (by dual-energy X-ray absorptiometry), their muscular strength (squat and countermovement jumps), and testosterone and DHT, as well as androgen precursors (dehydroepiandrosterone, androstenedione) and urinary androgen glucuronide metabolites (androsterone,

**Figure 1.** Sex differences in performance (in percentage) according to age (in years) in running events, including 50 m to 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) [for details, see Ref. (8)]. The lower panel is a fitted sigmoidal curve plot of sex differences in performance (in percentage) according to age (in years) in running, jumping, and swimming events, as well as the rising serum testosterone concentrations from a large dataset of serum testosterone of males. Note that in the same dataset, female serum testosterone concentrations did not change over those ages, remaining the same as in prepubertal boys and girls. Data are shown as mean and SEM of the pooled sex differences by age. Reproduced with permission from Handelsman DJ. Sex differences in athletic performance emerge coinciding with the onset of male puberty. *Clin Endocrinol (Oxf)*. 2017;**87**:68–72.






etiocholanolone, 3 and 17 3$\alpha$-diols) measured by LC-MS (36). The athletes displayed higher muscle (and bone) mass than did the sedentary control women, with strength tests correlating strongly with muscle mass whether in total or just in the legs. In turn, muscle mass and strength were correlated with androgens and androgen precursors. Considering that such studies may be confounded by factors such as menstrual phase and dysfunction, as well as heterogeneous sports disciplines, which weaken the power of the study, these findings can be regarded as quite robust.

### *Interventional data*

Dose-response studies show that in men whose endogenous testosterone is fully suppressed, add-back administration of increasing doses of testosterone that produce graded increases in circulating testosterone causes a dose-dependent (whether expressed according to testosterone dose or circulating levels) increase in muscle mass (measured as lean body mass) and strength (65, 111). Taken together, these studies prove that testosterone doses leading to circulating concentrations from well below to well above the normal male range have unequivocal dose-dependent effects on muscle mass and strength. These data strongly and consistently suggest that the sex difference in lean body mass (muscle) is largely, if not exclusively, due to the differences in circulating testosterone between men and women. These findings have strong implications for power-dependent sport performance and largely explain the potent efficacy of androgen doping in sports.

The key findings providing conclusive evidence that testosterone has prominent dose-response effects in men are reported in studies by Bhasin and colleagues that proved a monotonic dose response,

extending from subphysiological to supraphysiological range for men for testosterone effects on muscle mass, size, and strength in healthy young men, findings that have been replicated and confirmed by an independent group (65). Both sets of studies used a common design of fully suppressing all endogenous testosterone (to castrate levels) for the full duration of the experiment by administering a GnRH analog. In the Bhasin and colleagues studies, participants were then randomized to five groups and each received weekly injections of 25 mg, 50 mg, 125 mg, 300 mg, or 600 mg of testosterone enanthate for 20 weeks. In effect, this was two subphysiological and two supraphysiological testosterone doses. In these studies, the lowest testosterone dose produced a mean serum testosterone of 253 ng/dL (8.8 nmol/L) in younger men and 176 ng/dL (6.1 nmol/L) in older men. The studies showed a consistent dose response for muscle mass and strength that was clearly related to testosterone dose and consequential blood testosterone concentrations (Fig. 2, upper panel).

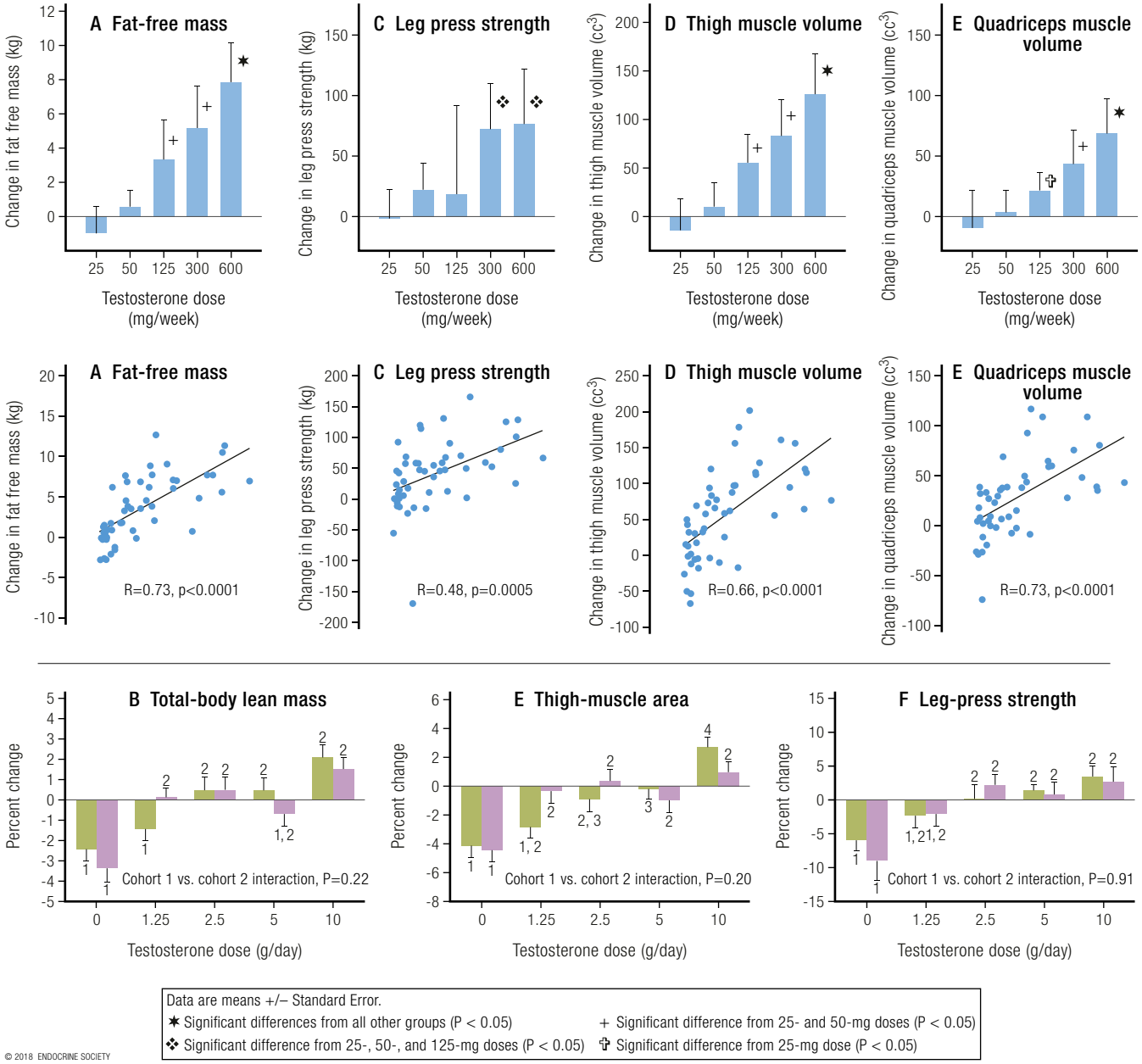
The study of Finkelstein *et al.* (65) involved the same design and involved 400 healthy men aged 20 to 50 years who had complete suppression of endogenous testosterone for the 16 weeks of the study, with testosterone added back using daily doses of 0, 1.25 g, 2.5 g, 5 g, or 10 g of a topical 1% testosterone gel. This again created a graded dose-response curve for serum testosterone and for muscle mass and strength. The inclusion of a 0 (placebo) dose allowed differentiation between the 0 and lowest testosterone dose. The placebo (0) dose produced a serum testosterone of 0.7 nmol/L (the typical mean for castrated men, childhood, and women of any age). Meanwhile, the lowest testosterone dose (1.25 g of gel per day) produced a serum testosterone of 6.9 nmol/L, which is equivalent to that of a male in early to middle puberty. A key finding for this review is that, from this study of men, the increase in serum testosterone from mean of normal female concentration (0.9 nmol/L) to supraphysiological female concentrations (6.9 nmol/L) produced significant increases of 2.3% for total body lean (muscle) mass, 3.0% for thigh muscle area, and 5.5% increase in leg press strength (digitized data pooling of both cohorts from lower panel, Fig. 2).

Studies of the ergogenic effects of supraphysiological concentrations of circulating testosterone require studies administering graded doses of exogenous testosterone for months. Owing to ethical concerns regarding risks of unwanted virilization and hormone-dependent cancers, however, few studies have administered supraphysiological testosterone doses to healthy women. One well-designed, randomized placebo-controlled study of postmenopausal women investigated the effects of different testosterone doses on muscle mass and performance and physical function (112). Sixty-two women (mean age, 53 years) all had a standard estrogen-replacement dose administered during a 12-week run-in period (to eliminate any hypothetical confounding effects of estrogen deficiency), after which they were randomized to one of five groups receiving weekly injections of testosterone enanthate (doses: 0, 3 mg, 6.25 mg, 12.5 mg, and 25 mg, respectively) for 24 weeks. The increasing doses of testosterone produced an expected dose response in serum testosterone concentrations (by LC-MS), with the highest testosterone dose (25 mg/wk) producing a mean nadir concentration of 7.3 nmol/L. The women whose testosterone concentrations were increased to 7.3 nmol/L achieved significant increases in muscle mass and strength (Table 4), ranging from 4.4% for muscle (lean) mass to between 12% and 26% for measures of muscle strength (chest and leg press, loaded stair climb). As muscle strength measurement is effort-dependent, the placebo-controlled design of the Huang *et al.* (112) study supports the further interpretation that the highest dose of testosterone also had prominent mental motivational effects in the effort-dependent tests of muscle strength. These findings provide salient direct evidence of the ergogenic effects of hyperandrogenism in female athletes confirming that at least up to average circulating testosterone concentrations of 7.3 nmol/L, women display a dose-response relationship similar to that of men, with supraphysiological doses of testosterone leading to significant gains in muscle mass and power.

These effects of testosterone administration on circulating testosterone concentrations and muscle mass and strength in females may be compared with the effects in males from the Finkelstein *et al.* (65) and Bhasin and colleagues studies. In men, the lowest testosterone dose (1.25 g/d) increased mean serum testosterone to 6.9 nmol/L (equivalent to levels seen in early to middle male puberty), resulting in significant increases of total body lean (muscle) mass (2.3%), thigh muscle area (3.0%), and leg press strength (5.5%) compared with the placebo dose that resulted in a serum testosterone of 0.7 nmol/L. In the Huang *et al.* (112) study (Fig. 3), muscle mass and strength in postmenopausal women displayed a flat response at the three lower doses, when circulating testosterone concentrations remain $<5$ nmol/L, and displayed a significant increase only when the mean circulating testosterone concentration produced by the highest testosterone dose first increased circulating testosterone concentrations $>5$ nmol/L. This pattern, flat at lower doses and rising at the highest dose, represents the lower plateau and the earliest rising portion, respectively, of the sigmoidal dose-response curve of testosterone for muscle.

Data corroborating the Huang *et al.* study results comes from another well-controlled study in which postmenopausal women who were administered methyl testosterone following a run-in period of estrogen replacement displayed a significant increase in lean (muscle) mass as well as upper and lower limb

**Figure 2.** Strong dose-response relationship between testosterone dose and circulating concentration with muscle mass and strength in men. The upper panels [from Bhasin *et al.* (111)] display the strong dose-response relationships of muscle mass shown as (A) "lean" or "fat-free" mass or volume of (D) thigh and (E) quadriceps muscle and (C) of leg muscle strength with increasing testosterone dose (upper row) or circulating concentration (middle row). Serum testosterone concentrations are in US units (ng/dL; divide by 28.8 to get nmol/L). Adapted with permission from 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. The lower panels [from Finkelstein *et al.* (65)] show the strong dose-response relationships of (B) whole-body muscle mass, (E) thigh muscle mass, and (F) leg press strength with increasing testosterone dose. Cohorts 1 and 2 were treated with the same increasing doses of testosterone but either without (green fill, cohort 1) or with (purple fill, cohort 2) an aromatase inhibitor (anastrozole), which prevents conversion of testosterone to estradiol. The differences between cohorts (*i.e.*, use of anastrozole) was not significant for muscle mass and strength and can be ignored with results of the two cohorts being pooled. Reproduced with permission from Finkelstein JS, Lee H, Burnett-Bowie SA, Pallais JC, *et al.* Gonadal steroids and body composition, strength, and sexual function in men. N Engl J Med 2013;369:1011–1022.






power during a 16-week double-blind, parallel group study (113).

Similarly, two prospective studies of the first 12 months of treatment of transmen [female-to-male (F2M) transgender] shows a consistent major increase in muscle mass and strength due to testosterone administration. In one study testosterone treatment of 17 transmen achieving adult male circulating testosterone levels

**Table 4. Effects of Testosterone on Muscle Mass and Strength in Women**

| Androgen-Sensitive Variable | Baseline | Increase | % Increase |
|---|---|---|---|
| Lean muscle mass, kg | 43 ± 6 | 1.9 ± 0.5 | 4.4 |
| Chest press, W | 100 ± 26 | 26 ± 7 | 26 |
| Leg press, N | 744 ± 172 | 90 ± 30 | 12 |
| Loaded stair-climb power, W | 406 ± 77 | 56 ± 13 | 14 |

With data from 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. Data are shown as mean and SEM derived from Table 1 and digitized from Figure 4 from Huang *et al.* (112) showing the effects of testosterone (mean circulating concentration, 7.3 nmol/L) on muscle mass and strength in women treated with the highest testosterone dose (n = 11; 25 mg of testosterone enanthate per week).

(mean, 31 nmol/L) increased muscle mass by 19.2% (114). In a second study, 23 transmen administered adult male testosterone doses also produced striking increases in total body muscle size and limb muscle size (by 6.5% to 16.6%) and grip strength (by 18%) compared with age-matched untreated control women (115). Conversely, testosterone suppression (using an estrogen-based treatment regimen) in 20 transwomen (M2F transgender) that reduced circulating testosterone levels from adult male range to adult female range led to a 9.4% reduction in muscle mass (measured as cross-sectional area).

#### *Effects on athletic performance*

Muscle growth, as well as the increase in strength and power it brings, has an obvious performance-enhancing effect, in particular in sports that depend on strength and (explosive) power, such as track and field events (107, 110). There is convincing evidence that the sex differences in muscle mass and strength are sufficient to account for the increased strength and aerobic performance of men compared with women and is in keeping with the differences in world records between the sexes (116). The basis for the sex difference in muscle mass and strength is the sex difference in circulating testosterone as clearly shown (for example) by (1) the enhanced athletic performance of men compared with prepubertal boys and women (8); (2) the close correspondence of muscle growth (muscle size) with muscle strength in ascending dose studies in men by Bhasin *et al.* (111, 117–119) and Finkelstein *et al.* (65) and in postmenopausal women by Huang *et al.* (112); (3) the effect of male castration in reducing muscle size and strength, effects that are fully rectified by testosterone replacement; and (4) the striking efficacy of androgen doping on the sports performances of German Democratic Republic female athletes (120).

### Hemoglobin

#### *Biology*

It is well known that levels of circulating hemoglobin are androgen-dependent and consequently higher in men than in women by 12% on average; however, the physiological mechanism by which androgens such as testosterone boosts circulating hemoglobin is not fully understood (121). Testosterone increases secretion of and sensitivity to erythropoietin, the main trophic hormone for erythrocyte production and thereby hemoglobin synthesis, as well as suppressing hepcidin (122), a crucial iron regulatory protein that governs the body's iron economy. Hepcidin has to balance the need for iron absorption from foods (the only source of iron required for the body's iron-containing proteins) against the fact that the body has no mechanism to shed excess iron, which can be toxic. Adequate iron availability is essential for normal erythropoiesis and synthesis of key heme, iron-containing oxygen-transporting proteins such as hemoglobin and myoglobin (123) as well as other iron-dependent proteins such as cytochromes and DNA synthesis and repair enzymes. Experimental evidence in mice shows that testosterone increases myoglobin content of muscle with potential for augmenting aerobic exercise performance (96), but this has not been evaluated in humans.

Increasing the amount of hemoglobin in the blood has the biological effect of increasing oxygen transport from lungs to tissues, where the increased availability of oxygen enhances aerobic energy expenditure. This is exploited to its greatest effect in endurance sports (1). The experiments of Ekblom *et al.* (124) in 1972 (Fig. 4) demonstrated strong linear relationships between changes in hemoglobin [due to withdrawal or retransfusion of 1, 2 or 3 U (400 mL) of blood] and aerobic capacity, established by repeated testing of maximal exercise-induced oxygen consumption before and after each procedure (124). As already noted, circulating hemoglobin levels are on average 12% higher in men than women (125). It may be estimated that as a result the average maximal oxygen transfer will be ~10% greater in men than in women, which has a direct impact on their respective athletic capacities.

#### *Observational data*

The proposition that the sex difference in circulating hemoglobin levels is likely to be due to the sex difference in average circulating testosterone concentrations is supported by the fact that male castration (*e.g.*, for advanced prostate cancer) (126) and androgen deficiency due to reproductive system disorders (127) reduce circulating

hemoglobin in men, eliminating the sex difference, whereas testosterone replacement therapy restores circulating hemoglobin to adult male levels (121, 127, 128).

An unusually informative observational study of women with CAH provides unique insight into testosterone effects on circulating hemoglobin in otherwise healthy women (92). Women with CAH require glucocorticoid replacement therapy but exhibit widely varying levels of hormonal control (79). The degree of poor control is associated with increasing levels of circulating testosterone ranging from normal female concentrations up to 36 nmol/L, and these levels correlate closely ($r$ = 0.56) with levels of circulating hemoglobin (Fig. 5). Interpolating from the dose-response regression, increases in circulating testosterone measured by LC-MS from 0.9 nmol/L to 5 nmol/L, 7 nmol/L, 10 nmol/L, and 19 nmol/L were associated with increases in circulating hemoglobin of 6.5%, 7.8%, 8.9%, and 11%, respectively, establishing a strong dose-response relationship. An 11% increase in circulating hemoglobin translates to a 10% difference in maximal oxygen transfer (124), which may account for virtually all the 12% sex difference in male and female circulating hemoglobin (125). To put this into context, any drug that achieved such increases in hemoglobin would be prohibited in sports for blood doping, as this difference is sufficient to have ergogenic effects, even without taking into account any testosterone effects on muscle mass or strength (for which data were not available in that study). Conversely, among elite female athletes with circulating testosterone in the healthy premenopausal female range, circulating hemoglobin does not correlate with athletic performance (35). In women with the mild hyperandrogenism of PCOS, circulating hemoglobin and hematocrit are reported as not (129) or marginally increased (130), findings that may be influenced by the fact that PCOS is associated with reduced or absent menstruation, thereby reducing the iron loss of regular menstruation.

### *Interventional data*

In the Bhasin *et al.* (111) studies, in both young and older men the highest testosterone dose produced a 12% increase in blood hemoglobin compared with the lowest dose, reflecting a strong dose-response relationship (Fig. 6) (131). Analogous findings were reported for testosterone treatment effects in postmenopausal women where the highest dose (25 mg weekly) of testosterone, which increased mean serum testosterone to 7.3 nmol/L, had the largest increase (3%) in blood hemoglobin and hematocrit (112).

Corroborative findings are available from studies of transmen (F2M transgender), that is, natal females who subsequently receive testosterone treatment at replacement doses to create adult male circulating testosterone concentrations, who exhibit increases in circulating hemoglobin to male levels [reviewed in (132–134)]. Testosterone treatment in 17 (F2M) transmen that created mean circulating testosterone levels of 31 nmol/L also increased hemoglobin levels by 15% (114). Conversely, one prospective 12-month study of transgender (nonathlete) individuals reported that testosterone suppression (by an estrogen-based regimen) to normal female levels in 20 (M2F) transwomen reduced hemoglobin by 14%.

If such an increase in hemoglobin were produced by any chemical substance, it would be considered doping, according to the World Anti-Doping Code.

## Bone

### *Biology*

There is extensive experimental evidence from genetic mouse models showing that the sex differences in bone

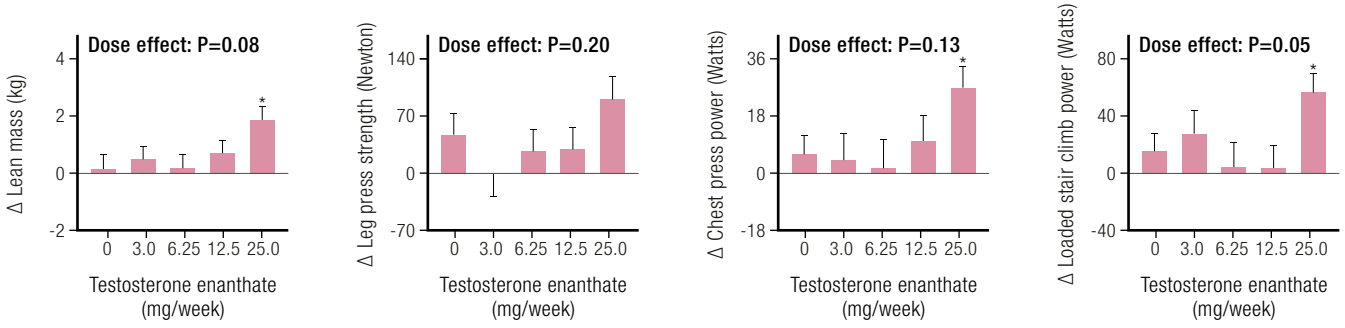
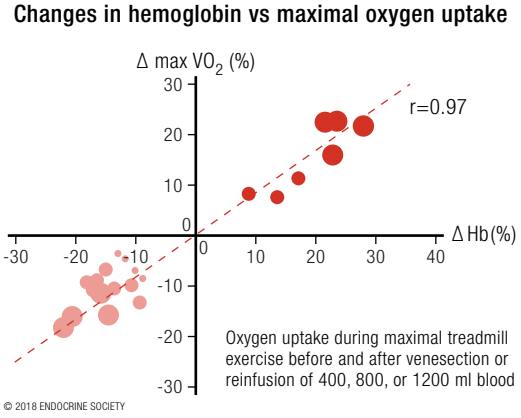





**Figure 3.** From Huang *et al.* (112): Dose-response effects on lean (muscle) mass and three measures of muscle strength as a result of increasing doses of weekly testosterone enanthate injections in women. Note the effects on all four parameters (three statistically significant) of the highest testosterone dose, the only one that produced circulating testosterone levels exceeding the normal female range. Reproduced with permission from 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.

**Figure 4.** Redrawn results from Ekblom *et al.* (124). Results from the transfusion of additional blood are shown in dark red circles and those after blood withdrawal in light red circles. Adapted with permission from Ekblom B, Goldbarg AN, Gullbring B. Response to exercise after blood loss and reinfusion. J Appl Physiol 1972;33:175–180.




size, mass, and function are due to the sex difference in circulating testosterone. These effects have been reported from studies of global and tissue or cell-selective inactivation of ARs or estrogen receptors that show that androgen effects are mediated by both direct effects on the AR as well as indirect effects mediated via aromatization of testosterone to estradiol to act on estrogen receptors [reviewed in (135)]. Bone grows in length due to epiphyseal chondral growth plates that provide cartilage, forming the matrix for lengthening of long bone, which is terminated by an estrogen-dependent mechanism that depends on aromatization of testosterone to estradiol. Similarly, bone width and density are increased through appositional growth from periosteal and endosteal expansion that depend on bone loading and androgen exposure together with other factors. An important difference between androgen effects on bone compared with effects on muscle or hemoglobin is that developmental bone effects of androgens are likely to be irreversible.

### *Observational data*

Men have distinctively greater bone size, strength, and density than do women of the same age. As with muscle, sex differences in bone are absent prior to puberty but then accrue progressively from the onset of male puberty due to the sex difference in exposure to adult male circulating testosterone concentrations [reviewed in (135)]. The earlier onset of puberty and the related growth spurt in girls as well as earlier estrogen-dependent epiphyseal fusion explains shorter stature of girls than boys. As a result, on average men are 7% to 8% taller with longer, denser, and stronger bones, whereas women have shorter humerus and femur cross-sectional areas being 65% to 75% and 85%, respectively, those of men (106). These changes create an advantage of greater bone strength and stronger fulcrum power from longer bones. Additionally, whereas passing through puberty enhances male physical performance, the widening of the female pelvis during puberty, balancing the evolutionary demands of obstetrics and locomotion (136, 137), retards the improvement in female physical performance, possibly driven by ovarian hormones rather than the absence of testosterone (138, 139).

Sex differences in height have been the most thoroughly investigated measure of bone size, as adult height is a stable, easily quantified measure in large population samples. Extensive twin studies show that adult height is highly heritable with predominantly additive genetic effects (140) that diverge in a sex-specific manner from the age of puberty onwards (141, 142), the effects of which are likely to be due to sex differences in adult circulating testosterone concentrations.

Bone density (total and medullary cross-sectional area) is increased in women with CAH with variably elevated serum testosterone (including into the male range) when it is only partially suppressed by glucocorticoid treatment (143), although more effective glucocorticoid suppression lowers bone density (144).

### *Interventional data*

Well-designed, placebo-controlled direct interventional studies of supraphysiological androgen effects on bone in females are few, rarely feasible, and unlikely to be performed for ethical and practical reasons. Unlike muscle, which responds relatively rapidly to androgen effects so that muscle studies in humans can be completed within 3 to 4 months (65, 111, 112, 119, 145), comparable bone studies would typically take a year or more to reach plateau effects. Hence, such direct investigational studies in otherwise healthy women would risk side effects of virilization that may be only slowly and partly reversible, if at all, as well as potential promotion of hormone-dependent cancers making such studies ethically and practically not feasible.

### *Effects on athletic performance*

The major effects of men's larger and stronger bones would be manifest via their taller stature as well as the larger fulcrum with greater leverage for muscular limb power exerted in jumping, throwing, or other explosive power activities. The greater cortical bone density and thereby resistance to long bone fractures is unlikely to be relevant to the athletic performance of young athletes, in whom fractures during competition are extremely rare and not expected to be linked to sex. Alternatively, stress fractures in athletes, mostly involving the legs, are more frequent in females with the male protection attributable to their larger and thicker bones (146).

### Other androgen-sensitive sex dichotomous effects

#### *Biology and observational data*

Many if not most other aspects of physiology exhibit sex differences and may therefore enhance the impact of the male advantage in sports performance of the dominant determinants (muscle and hemoglobin). Examples include sex differences in exercise-induced cardiac (147, 148) and lung (149) function and mitochondrial biogenesis and energetics (95). However, the limited knowledge of the magnitude and hormonal mechanisms involved, specifically the degree of androgen dependence of these mechanisms, means that it is difficult to estimate their contribution, if any, toward the sex difference in athletic performance. The sex difference in pulmonary function may be largely explained by the androgen-sensitive sex difference in height, which is a strong predictor of lung capacity and function (149). Further physiological studies of the androgen dependence of other physiological sex differences are awaited with interest.

Psychological differences between men and women on mental function (*e.g.*, rotational orientation) (150) as well as mood, motivation, and behavioral effects may involve androgen-sensitive effects during prenatal and perinatal as well as postpubertal effects (151, 152).

#### *Interventional data*

There is some limited direct evidence from well-designed, placebo-controlled trials that administration of testosterone or other androgens at supraphysiological doses directly affect mood and behavior, notably inducing hypomania (153). In a randomized placebo-controlled study of testosterone administration in postmenopausal women (112), in case of those receiving the highest dose (the only one causing circulating testosterone levels to exceed the normal female range), there was not only an increase in muscle mass (4.4%) but a strikingly greater increase in muscle strength (12% to 26%), suggesting an enhanced mental motivational effect of testosterone on the effort-dependent tests of muscle strength.

## Alternative Mechanisms Proposed to Explain Sex Differences in Athletic Performance

Alternative explanations for the sex difference in athletic performance, other than it being due to the sex difference in postpubertal circulating testosterone, have been proposed, including (1) sex differences in height because height is a predictor of muscle mass (116), (2) genetic sex differences due to the influence of unspecified Y chromosome genes (154), and (3) sex differences in GH secretion (116).

### Effects of height

One proposal has been that, as men are taller than women, height differences may explain the sex differences in muscle mass and function, which explains some athletic success (116). Numerous factors contribute to the regulation of adult muscle mass, including genetics, race, adiposity, hormones, physical activity (exercise/training), diet, birth order, and bone size (including height) [reviewed in (155)]. Among the nonhormonal factors, genetics explains a large proportion [~50% to 60% from pooled twin studies (156)] of the variability in muscle mass and strength (157, 158) and may be explained in turn by the equally high genetic contributions to circulating testosterone (37, 38). Some factors influencing muscle mass and strength such as physical activity, adiposity, and bone size are also partly androgen-dependent. Prior to puberty there is no sex difference in skeletal features, including height (159, 160). However, with the onset of puberty, girls aged 11 and 12 years are transiently taller than peer-aged boys due to their earlier onset of the female pubertal growth spurt, but from the age of 14 years onward the taller stature in males emerges and stabilizes (141). Hence, similar to muscle mass, sex differences in bone size (including length, density, and height) arise after male puberty establishes the marked dichotomy between men and women in adult circulating testosterone concentrations. Taller height is

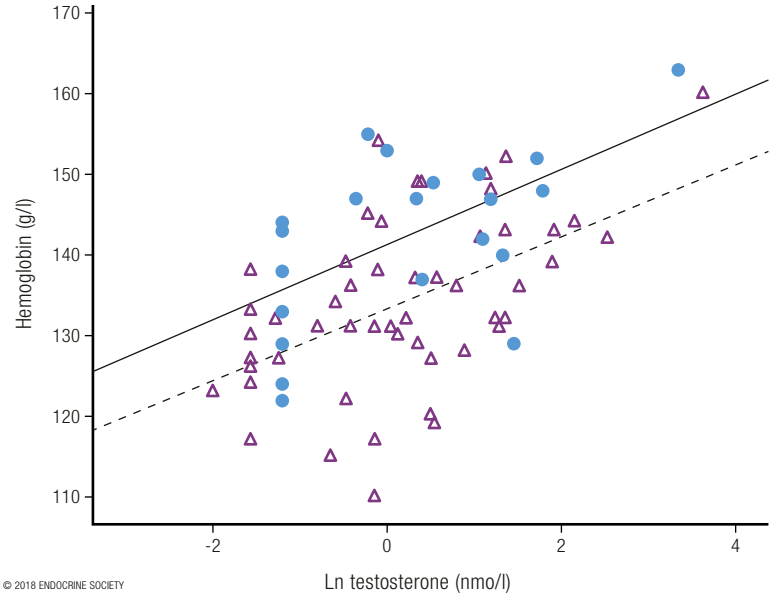



**Figure 5.** Plot of circulating hemoglobin against the natural logarithm of serum testosterone in women with congenital adrenal hyperplasia [from Karunasena *et al.* (92)]. The filled circles represent a cohort where serum testosterone was measured by immunoassay. The open triangles denote a second cohort, where serum testosterone was measured by LC-MS. Note the systematic overestimation of testosterone by the immunoassay used in cohort 1 vs LC-MS measurement in cohort 2. Despite that overestimation, however, the correlations were similar in both cohorts. Reproduced under a Creative Commons BY-NC-ND 4.0 license from Karunasena N, Han TS, Mallappa A, *et al.* Androgens correlate with increased erythropoiesis in women with congenital adrenal hyperplasia. Clin Endocrinol (Oxf) 2017;86:19–25.

advantageous in some sports (basketball, some football codes, combat sports), but in others (horse racing jockeys, cycling, gymnastics, weightlifting, bodybuilding) short stature provides a greater power/strength-to-weight ratio as well as superior rotational balance, speed, and agility. However, the male advantages in speed, strength, and endurance apply regardless of whether height is advantageous. Hence, the sex differences in height, where they exist, are largely dependent on postpubertal differences in circulating testosterone when sex differences in height are first expressed.

### Genetic effects of Y chromosome

It has also been proposed that the sex difference in athletic performance may be due to genetic effects of an unspecified Y chromosome gene that may dictate taller stature (154), as height is correlated with men's greater muscle mass. The small human Y chromosome has few functional genes and none with a known effect on height other than the short stature homeobox (SHOX) gene, located in the pseudoautosomal regions of the tip of the short arms of X and Y chromosomes (161). Adult height displays an apparent dose dependency on SHOX gene copy number that is a major factor contributing to explaining both the short stature of 45,XO females (Turner syndrome), who have a single copy of the SHOX gene, as well as the tall stature of 47,XXY males (Klinefelter syndrome), who have three copies (161). However, when SHOX copy number is the same, men with additional supernumerary Y chromosomes (*e.g.*, 47,XYY) are the same height as 47,XXY men (162). Hence, there is no evidence supporting dosage-dependent Y chromosomal gene effects on height independent of SHOX gene copy number, nor does men's possession of a Y chromosome explain the height difference between adult men and women. On the contrary, the tall stature of 47,XXY men is at least partly due to the concomitant androgen deficiency leading to pubertal delay. Pubertal delay prolongs long bone growth due to delayed epiphyseal closure, an estrogen-dependent effect that requires adequate production of testosterone as a substrate for aromatization to estradiol, resulting in tall stature. Similar eunuchoidal features and taller stature are evident in 46,XY men with congenital hypogonadotropic hypogonadism (Kallmann syndrome and its variants) with comparable congenital onset of androgen deficiency, also manifest as pubertal delay and long bone overgrowth. Hence, taller height is better explained by impaired testicular function with delayed puberty and epiphyseal closure rather than unspecified Y chromosome dosage effects. In any case, rare aneuploidies in themselves do not explain the sex difference in height in the general population of individuals with normal sex chromosomes.

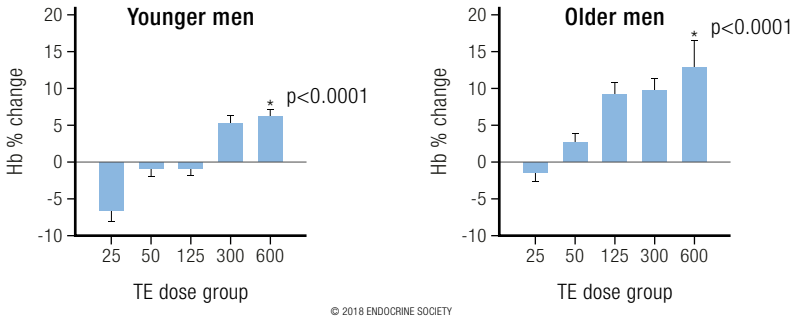



**Figure 6.** From Coviello *et al.* (131): Depicts the strong dose-response relationship between increasing testosterone dose with resulting change in blood hemoglobin in young and older men. Reproduced with permission from Coviello AD, Kaplan B, Lakshman KM, *et al.* Effects of graded doses of testosterone on erythropoiesis in healthy young and older men. J Clin Endocrinol Metab 2008;93:914–919.

### Growth hormone

The proposal that the sex difference in muscle mass and function might be due to sex differences in endogenous GH secretion (116) is refuted by the extensive and conclusive clinical evidence that endogenous GH secretion in young women is consistently higher (typically twice as high) as in young men of similar age (163–170). Those findings cannot explain the male advantage in muscle mass and strength unless GH retards muscle growth/function, for which there is no evidence. Furthermore, estrogens inhibit GH-dependent, hepatic IGF-1 production, the major pathway of GH action (171, 172). The weak observational association between low circulating IGF-1 and some, but not other, measures of weak muscle strength and limited mobility among older women may reflect general age-associated debility rather than any specific hormonal effects (173). Finally, the evidence that endogenous GH plays no role in sex differences in muscle mass and function is supported by evidence from the most extensive interventional study of GH treatment to non-GH–deficient adults, daily GH administration for 8 weeks to healthy recreational athletes produced only marginally significant improvement in exercise performance of men and none in women (174). These findings are consistent with the speculation that GH (or IGF-1) may be an amplifier of testosterone effects and therefore be a consequence of the sex difference in circulating testosterone rather than its cause.

## The Impact of Adult Male Circulating Testosterone Concentrations on Sports Performance

Plausible estimates of the magnitude of the ergogenic advantage of adult male circulating testosterone concentrations are feasible from the limited available observational and interventional studies.

Population data on the ontogeny of puberty show that prior to puberty boys and girls have comparable athletic performance, whereas sex differences in athletic performance emerge coinciding with the rise in circulating testosterone from the onset of male puberty. Male puberty results in circulating testosterone concentrations rising from the prepubertal and female postpubertal range (<2 nmol/L) to adult male circulating testosterone concentrations (18). This is associated with a 10% to 12% better performance in running and swimming events and 20% enhancement in jumping events (8).

A minimal estimate of the impact of adult male testosterone concentrations on muscle size and strength in females is provided by the Huang *et al.* (112) study of postmenopausal women. In this study the highest testosterone dose (weekly injections of 25 mg of testosterone enanthate) increased mean circulating testosterone from 0.9 nmol/L to 7.3 nmol/L, which is equivalent to the circulating testosterone of boys in early to middle puberty. After 24 weeks of testosterone treatment, the increase in circulating testosterone concentrations led to significant increases in muscle size of 4.4% and in muscle strength of 12% to 26%. Given the limited testosterone dose (and concentration) as well as study duration, it is likely that these findings underestimate the magnitude of the impact that sex difference in circulating testosterone has on muscle mass and strength, and therefore on athletic performance.

Converse effects of reduced athletic performance in athletes who undergo suppression of circulating testosterone concentrations from those in the male into the female range have been reported. Among recreational (nonelite) athletes, an observational study showed a consistent deterioration in athletic performance of transwomen (M2F transgender) athletes corresponding closely to the suppression of circulating testosterone concentrations (175). Similarly, among elite athletes with circulating testosterone in the male range due to DSDs, comparable findings of athletic performance reduced by an average of 5.7% when circulating testosterone was suppressed from the male range to <10 nmol/L (176). Subsequently, when the IAAF hyperandrogenism rule was suspended in 2015, and so these elite athletes could train and compete with unsuppressed serum testosterone levels, their athletic performances increased by a similar amount. Additionally, circulating hemoglobin levels in these untreated DSD athletes were comparable with male athletes or with female athletes doping with erythropoietin (Fig. 7). However, when circulating testosterone was suppressed to <10 nmol/L the levels of circulating hemoglobin were 12% lower and again comparable with nondoped, non-DSD females, corresponding to the 12% magnitude of the sex difference in hemoglobin between men and women (125).

Congruent findings are also known for an elite female athlete whose serial athletic performance based on publicly available best annual times between 2008 and 2016 for the 800-m running event are depicted in relationship to the original 2011 IAAF hyperandrogenism regulation (Fig. 8).

Based on the established dose-response relationships, suppression of circulating testosterone to <10 nmol/L would not eliminate all ergogenic benefits of testosterone for athletes competing in female events. For example, according to the Huang *et al.* (112) study, reducing circulating testosterone to a mean of 7.3 nmol/L would still deliver a 4.4% increase in muscle size and a 12% to 26% increase in muscle strength compared with circulating testosterone at the normal female mean value of 0.9 nmol/L. Similarly, according to the Karunasena *et al.* (92) study, reducing circulating testosterone concentration to 7 nmol/L would still deliver 7.8% more circulating hemoglobin than the normal female mean value. Hence, the magnitude of the athletic performance advantage in DSD athletes, which depends on the magnitude of elevated circulating testosterone concentrations, is considerably greater than the 5% to 9% difference observed in reducing levels to <10 nmol/L.

The physiological mechanism underlying these observations is further strengthened by prospective controlled studies of initiation of cross-sex hormone treatment in transgender individuals (114, 177). These show that during the first 12 months muscle mass (area) was decreased by 9.4% and hemoglobin levels by 14% in 20 transwomen (M2F transgender) treated with an estrogen-based regimen that reduced circulating testosterone concentrations from the male range to the female range. Conversely, in 17 transmen (F2M transgender) treated for the first time with testosterone for 12 months (which increased circulating testosterone levels to a mean of 31 nmol/L), muscle mass increased by 19.2% and hemoglobin by 15% (114). The muscle mass findings remained stable between 1 and 3 years after initiation of treatment, although fat mass continued to change between 1 and 3 years of testosterone treatment (177). These studies did not report muscle strength, but other studies of testosterone dose-response relationships for muscle mass and strength show consistently positively correlation (65, 93, 117, 119), although with disproportionately greater effect on muscle strength than on muscle mass. Hence, the muscle mass estimates in these prospective treatment initiation studies in transgender individuals likely underestimate the muscle strength gains from elevated testosterone levels where the circulating testosterone markedly exceeds female range to be within the male range as occurs in severe hyperandrogenism of DSD females, poorly controlled transwomen (M2F transgender), or transmen (F2M transgender). These effects are also the biological

basis of the ergogenic efficacy of androgen doping in women.

Finally, to put these competitive advantages into context, the winning margin (the difference in performance by which a competitor misses a gold medal, any medal, or making the final) in elite athletic or swimming events during the last three Olympics is <1% equally for both male and female events (Table 5).

## Gaps in Knowledge and Research Limitations

The major limitations on scientific knowledge of the impact of adult male circulating testosterone concentrations on the sex difference in athletic performance is the lack of well-designed studies. Ideally, these would need to replicate adult male circulating testosterone concentrations for sufficient time in women to investigate the effects on muscle, hemoglobin, bone, and other androgen-sensitive measures that display consistent sex dichotomy in the population. However, the ethical and safety concerns preventing such studies hitherto are likely to remain formidable obstacles due to the risk of unacceptable and potentially irreversible virilization as well as of promoting hormone-dependent cancers in women.

With the exception of one interventional study administering a relatively low testosterone dose (*i.e.*, low for males) to women (112), the available evidence comprises observational studies that can only examine the effects of serum testosterone within physiological female limits or sparse and mostly uncontrolled data from intersex/DSD athletes. Although the available observational findings in healthy females are informative, the key question is the magnitude and dose response of effects at still higher circulating testosterone concentrations on the performances of women. Whereas a testosterone dose-response relationship has been established in women at relatively low (for men) testosterone dose and circulating concentrations, it remains unproven (even if clearly plausible) that the testosterone dose-response relationships established in men for muscle, hemoglobin, and bone can be extrapolated to women when they are exposed to higher circulating testosterone concentrations (*i.e.*, comparable with male levels). It is theoretically possible there could be differences between men and women in muscle responses to testosterone, as muscle cell populations might express genetic differences in androgen sensitivity (for which there are no data), or alternatively the long-term prior pattern of testosterone exposure from conception to adulthood might lead to differences in testosterone dose responsiveness after maturity. Although the dose-response relationship in women may be similar to what is seen in men, there is also anecdotal evidence that the dose-response curves may be left shifted so that testosterone has greater potency in women than in men at comparable doses and circulating levels. The prediction is supported by the anecdotal evidence from the surreptitious East German national doping program in which the supervising doctors asserted from their experience of illicit cheating that androgens had more potent ergogenic effects in women than in men (120), a speculative opinion shared by many experienced sports medicine physicians.

There is no known means of increasing endogenous testosterone in women to anything like the requisite degree to attempt to answer these questions. In healthy men, circulating testosterone originates almost exclusively from a single source (testicular Leydig cells) and is subject to tight hypothalamic negative feedback control, so that either direct stimulation (by human chorionic gonadotropin) or indirect reflex effects (*e.g.*, from estrogen blockers operating via negative feedback) to enhance Leydig cell testosterone secretion are feasible. However, similar mechanisms do not operate in women, in whom circulating testosterone originates from three different sources (adrenal, ovary, extraglandular conversion of androgen precursors), none of which is subject to tight testosterone negative feedback control. As a result, it is not feasible to produce a sufficient increase in circulating testosterone in women either by direct ovarian stimulation or indirect reflex effects to test this hypothesis even if doing so were deemed ethical and safe. Alternatively, carefully controlled, graded-dose studies in F2M transgender individuals might be informative but are largely lacking at this time.

Hence, the only feasible design of such studies would be testosterone (or another androgen) administration to healthy young women. The only well-designed, placebo-controlled study of testosterone in

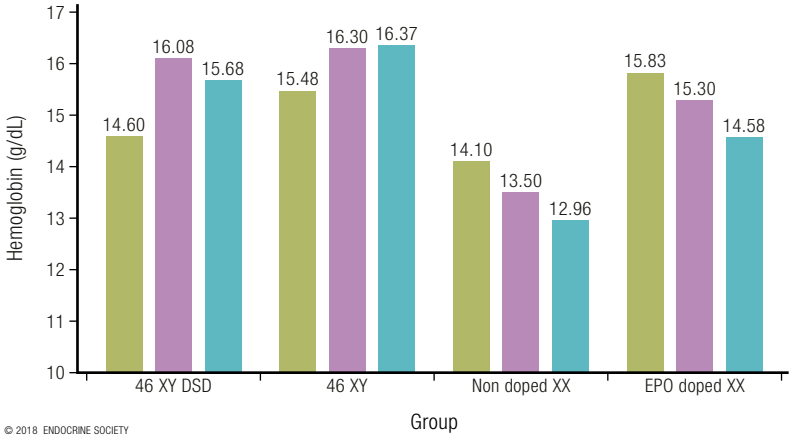



**Figure 7.** Mean hemoglobin concentrations (g/dL) of 12 elite athletes in 4 groups of 3 XY or XX middle-distance runners. The hemoglobin concentrations were collected as a part of the Athlete Biological Passport and analyzed according to the World Anti-Doping Agency standard methods. Each bar (athlete) is the mean of a minimum of three blood samples. In the 46,XY DSD group, blood was collected in a period when the athlete was not undergoing hormonal suppressive treatment.

otherwise healthy postmenopausal women was restricted to relatively low testosterone doses that, although clearly supraphysiological for women, were only 20% to 25% of male testosterone replacement doses (112). We are currently performing a double-blind, randomized, placebo-controlled study of the effects of moderately increased testosterone concentration on physical performance and behavior in young healthy women (ClinicalTrials.gov no. NCT03210558). However, obtaining ethical approval to administer supraphysiological testosterone doses that maintain circulating testosterone in the male range for sufficiently prolonged periods, as well as the practical difficulties in recruitment, are likely to remain obstacles to definitive resolution of this question.

In men, analogous ethical concerns over short- and long-term adverse effects delayed the definitive studies of supraphysiological testosterone doses to healthy young and older men but were eventually overcome. This was despite the fact that, uniquely among hormones, there is no known disease state in men due to pathologically excessive testosterone secretion. In contrast, in women, supraphysiological testosterone effects are known to produce virilization side effects that may be only slowly and partially, if at all, reversible. However, maintaining clearly supraphysiological testosterone concentrations would require treatment of months (muscle) or years (bone) and would replicate not only a known hyperandrogenic disease state (PCOS) but also potentially increasing risk of hormone-dependent cancers. In these circumstances, it could only be justifiable to replicate in women the salient testosterone dose-response studies available from men if the available evidence of dose-response relationship in men was not sufficiently convincing and/or there was reason to think that these dose-response characteristics would be substantially different in women. Overall, the unequivocal dose-response evidence in men together with the available overlap evidence in women appears sufficiently persuasive, so that it is doubtful that women would respond differently from men if their circulating testosterone levels were raised to the male range. More broadly, there is no more reason to require separate studies in women vs men than there is for every different ethnic subgroup of people. An aesthetic preference for splitting categories is not a sound reason to require the virtually impossible standard of establishing fresh and comprehensive empirical evidence in women of testosterone dose-response effects ranging into male circulating testosterone concentrations.

An analogy can be drawn to the World Anti-Doping Agency's practice of accepting salient surrogate evidence for banning the plethora of existing and new drugs with potential but individually unproven ergogenic effects where it is not feasible or ethical to require direct proof of the ergogenic effects. In that context, the firmly established ergogenic efficacy of androgens (on muscle mass and strength) and increased hemoglobin (on endurance) [evidence reviewed in (1)] mean that chemical substances or methods that increase endogenous testosterone, erythropoietin, or hemoglobin are also considered ergogenic (178). By parity of reasoning, if a condition causes a female athlete's circulating testosterone levels to be in the male range, well exceeding normal female levels, with consequential increases in muscle, hemoglobin, and bone effects (at least), an ergogenic effect may reasonably be assumed.

## Conclusions

The available, albeit incomplete, evidence makes it highly likely that the sex difference in circulating testosterone of adults explains most, if not all, the sex differences in sporting performance. This is based on the dose-response effects of circulating testosterone to increase muscle mass and strength, bone size and strength (density), and circulating hemoglobin, each of which alone increases athletic capacity, as well as other possible sex dichotomous, androgen-sensitive contributors such as mental effects (mood, motivation, aggression) and muscle myoglobin content. These facts explain the clear sex difference in athletic performance in most sports, on which basis it is commonly accepted that competition has to be divided into male and female categories.

The first IAAF hyperandrogenism regulation specified a hormonal eligibility criterion of a serum testosterone of <10 nmol/L for an androgen-sensitive athlete's participation in the protected category of female athletic events. This threshold was based on serum testosterone measurements by immunoassays.

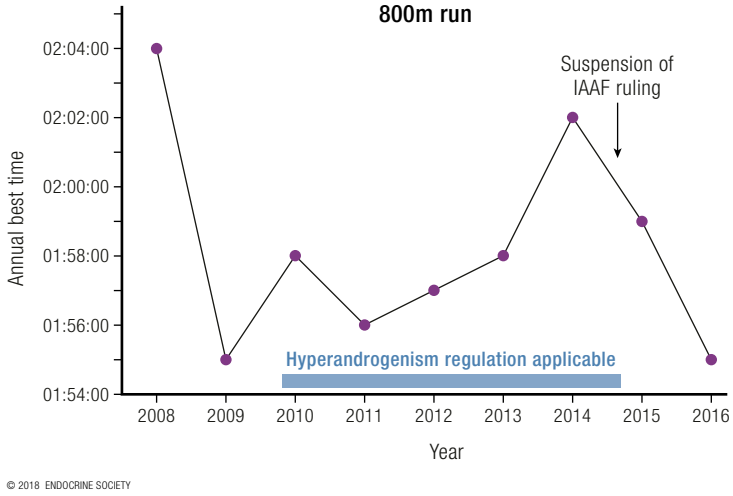



**Figure 8.** Best annual 800-m times of an elite female athlete between 2008 and 2016. Data provided by Dr. Richard Auchus, University of Michigan, Ann Arbor, Michigan.

**Table 5.** The Winning Margin in Elite Athletic or Swimming Events During the Last Three Olympics

| Median Margin (%)[a] | n | Win Gold | Win Medal | Make Final |
|---|---|---|---|---|
| Athletics[b] | | | | |
| Running | 81 | 0.62 | 0.31 | 0.22 |
| Jumping | 24 | 0.92 | 0.42 | 0.92 |
| Throwing | 24 | 1.93 | 0.70 | 0.75 |
| Swimming[c] | | | | |
| Backstroke | 12 | 0.56 | 0.28 | 0.16 |
| Breaststroke | 12 | 0.84 | 0.14 | 0.17 |
| Butterfly | 12 | 0.52 | 0.48 | 0.12 |
| Freestyle | 30 | 0.49 | 0.23 | 0.14 |
| Relay | 18 | 0.37 | 0.35 | 0.12 |

[a]Winning margin is defined as the difference (expressed as a percentage of the faster time) between first and second place (Win Gold), between third and fourth place (Win Medal), and between the last into the final and the first that missed out (Make Final). Years (2008, 2012, 2016) and sexes were combined as there were no significant differences in winning margin between them.

[b]Running includes 100 m, 200 m, 400 m, 800 m, 1500 m, 5000 m, 10,000 m, marathon, and 3000-m steeplechase, 110-m (male)/100-m (female) and 400-m hurdles, 4 × 100-m and 4 × 400-m relays, and 20-km and 50-km walk events. Jumping includes high jump, long jump, triple jump, and pole vault events. Throwing includes javelin, shot put, discus, and hammer events. Heptathlon and decathlon were not included as their final results are in points, not times.

[c]Events comprise 100 m and 200 m for the form strokes and 50 m, 100 m, 200 m, 400 m, 800 m (female)/1500 m (male) and marathon 10 km, with the relays being the 4 × 100-m medley and 4 × 100-m and 4 × 200-m freestyle relays.

However, no reliable method-independent consensus threshold could be established using commercial testosterone immunoassays, as these assays differ systematically due to method-specific bias arising unavoidably from the specificity of the different proprietary antibodies employed (25). Based on measurements using the more accurate and specific mass spectrometry methods, if the objective is to require female athletes with congenital conditions that cause them to have serum testosterone concentrations in the normal male range to bring those levels down to the same range as other female athletes, then (allowing for PCOS athletes) the threshold used should not be >5.0 nmol/L. This represents a conservative criterion that includes all healthy young (<40 years) women, including those with PCOS. Conversely, this criterion is generous to intersex/DSD females in allowing them to maintain a higher serum testosterone (2 to 5 nmol/L) than most non-PCOS competitors in female events even though increases in muscle mass and strength and hemoglobin would be expected in this range. This is so even though the range remains below the circulating testosterone levels of middle male puberty when the major biological effects of men's higher circulating testosterone begin to be fully expressed. Ongoing compliance with the eligibility criterion is also an important variable because the estrogen-based suppression of circulating testosterone, typically using daily administered estrogen products, has a rapid onset and offset. Adequate monitoring to prevent gaming of eligibility criteria would require regular random rather than announced blood sampling.

A related matter is how long such a threshold of circulating testosterone should be maintained prior to competition. In both intersex/DSD and transgender individuals, the developmental effects of adult male circulating testosterone concentrations will have established the sex difference in muscle, hemoglobin, and bone, some of which is fixed and irreversible (bone size) and some of which is maintained by the male circulating testosterone concentrations (muscle, hemoglobin). The limited available prospective evidence from initiation of transgender cross-sex hormone treatment suggests that the advantageous increases in muscle and hemoglobin due to male circulating testosterone concentrations are induced or reversed during the first 12 months and the androgenic effects may plateau after time. This time course is much faster than the somatic effects of male puberty, which evolve over years and for some variables (*e.g.*, peak bone mass) are not complete for up to a decade after the start of puberty. However, the abrupt hormonal changes induced by medical treatment in intersex/DSD or transgender individuals may be telescoped compared with male puberty where circulating testosterone concentrations increase irregularly and incompletely for some years. Additional data are available from the unique investigative model of men undergoing castration for prostate cancer. Just as androgen sensitivity to testosterone may differ between tissues (65), the time course of offset of

androgen effects following withdrawal of male testosterone concentrations may also differ between the major androgen-responsive tissues. For example, circulating hemoglobin shows a progressive fall for 6 months reaching a nadir and plateau at 12 to 16 months in six studies involving 534 men undergoing medical castration for prostate cancer (179–184). Although these studies of older men with prostate cancer must be extrapolated with caution, age, stage of disease, race, and baseline circulating testosterone concentration did not affect the rate or extent of decline in hemoglobin (179, 181). Comparable longitudinal studies of muscle loss, strength, and performance following castration for prostate cancer are well summarized (185), showing progressive loss for 24 months (see Fig. 4). Further clinical studies to define the time course of changes, mainly offset, in testosterone-dependent effects, notably on muscle and hemoglobin, are badly needed to determine the optimal duration for cross-sex hormone effects in sports.

## References


1. Handelsman DJ. Performance enhancing hormones in sports doping. In: DeGroot LJ, Jameson JL, eds. *Endocrinology*. 7th ed. Philadelphia, PA: Elsevier Saunders; 2015:441–454.
2. Coleman DL. Sex in sport. Available at: ssrn.com/abstract=2928106. Accessed 22 October 2017.
3. Lee PA, Nordenström A, Houk CP, Ahmed SF, Auchus R, Baratz A, Baratz Dalke K, Liao LM, Lin-Su K, Looijenga LH III, Mazur T, Meyer-Bahlburg HF, Mouriquand P, Quigley CA, Sandberg DE, Vilain E, Witchel S; Global DSD Update Consortium. Global disorders of sex development update since 2006: perceptions, approach and care [published correction appears in *Horm Res Paediatr*. 2016;85(3):180]. *Horm Res Paediatr*. 2016;**85**(3):158–180.
4. Southren AL, Tochimoto S, Carmody NC, Isurugi K. Plasma production rates of testosterone in normal adult men and women and in patients with the syndrome of feminizing testes. *J Clin Endocrinol Metab*. 1965;**25**(11):1441–1450.
5. Horton R, Tait JF. Androstenedione production and interconversion rates measured in peripheral blood and studies on the possible site of its conversion to testosterone. *J Clin Invest*. 1966;**45**(3):301–313.
6. Southren AL, Gordon GG, Tochimoto S. Further study of factors affecting the metabolic clearance rate of testosterone in man. *J Clin Endocrinol Metab*. 1968;**28**(8):1105–1112.
7. Saez JM, Forest MG, Morera AM, Bertrand J. Metabolic clearance rate and blood production rate of testosterone and dihydrotestosterone in normal subjects, during pregnancy, and in hyperthyroidism. *J Clin Invest*. 1972;**51**(5):1226–1234.
8. Handelsman DJ. Sex differences in athletic performance emerge coinciding with the onset of male puberty. *Clin Endocrinol (Oxf)*. 2017;**87**(1):68–72.
9. Auchus RJ. Endocrinology and women's sports: the diagnosis matters. *Law Contemp Probl*. 2017;**80**:127–138.
10. Foddy B, Savulescu J. Time to re-evaluate gender segregation in athletics? *Br J Sports Med*. 2011;**45**(15):1184–1188.
11. Handelsman DJ. Androgen physiology, pharmacology and abuse. In: DeGroot LJ, Jameson JL, eds. *Endocrinology*. 7th ed. Philadelphia, PA: Elsevier Saunders; 2015:2368–2393.
12. Miller WL, Auchus RJ. The molecular biology, biochemistry, and physiology of human steroidogenesis and its disorders. *Endocr Rev*. 2011;**32**(1):81–151.
13. Abreu AP, Kaiser UB. Pubertal development and regulation. *Lancet Diabetes Endocrinol*. 2016;**4**(3):254–264.
14. Horton R, Shinsako J, Forsham PH. Testosterone production and metabolic clearance rates with volumes of distribution in normal adult men and women. *Acta Endocrinol (Copenh)*. 1965;**48**:446–458.
15. Rivarola MA, Saez JM, Meyer WJ, Jenkins ME, Migeon CJ. Metabolic clearance rate and blood production rate of testosterone and androst-4-ene-3,17-dione under basal conditions, ACTH and HCG stimulation. Comparison with urinary production rate of testosterone. *J Clin Endocrinol Metab*. 1966;**26**(11):1208–1218.
16. Courant F, Aksglaede L, Antignac JP, Monteau F, Sorensen K, Andersson AM, Skakkebaek NE, Juul A, Bizec BL. Assessment of circulating sex steroid levels in prepubertal and pubertal boys and girls by a novel ultrasensitive gas chromatography-tandem mass spectrometry method. *J Clin Endocrinol Metab*. 2010;**95**(1):82–92.
17. Davison SL, Bell R, Donath S, Montalto JG, Davis SR. Androgen levels in adult females: changes with age, menopause, and oophorectomy. *J Clin Endocrinol Metab*. 2005;**90**(7):3847–3853.
18. 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. *Ann Clin Biochem*. 2016;**53**(Pt 3):377–384.
19. Rothman MS, Carlson NE, Xu M, Wang C, Swerdloff R, Lee P, Goh VH, Ridgway EC, Wierman ME. Reexamination of testosterone, dihydrotestosterone, estradiol and estrone levels across the menstrual cycle and in postmenopausal women measured by liquid chromatography–tandem mass spectrometry. *Steroids*. 2011;**76**(1-2):177–182.
20. Müller RK. History of doping and doping control. *Handb Exp Pharmacol*. 2010;(195):1–23.
21. Rosner W, Hankinson SE, Sluss PM, Vesper HW, Wierman ME. Challenges to the measurement of estradiol: an Endocrine Society position statement. *J Clin Endocrinol Metab*. 2013;**98**(4):1376–1387.
22. Rosner W, Auchus RJ, Azziz R, Sluss PM, Raff H. Position statement: utility, limitations, and pitfalls in measuring testosterone: an Endocrine Society position statement. *J Clin Endocrinol Metab*. 2007;**92**(2):405–413.
23. Handelsman DJ, Wartofsky L. Requirement for mass spectrometry sex steroid assays in the *Journal of Clinical Endocrinology and Metabolism*. *J Clin Endocrinol Metab*. 2013;**98**(10):3971–3973.
24. Handelsman DJ. Mass spectrometry, immunoassay and valid steroid measurements in reproductive medicine and science. *Hum Reprod*. 2017;**32**(6):1147–1150.
25. Sikaris K, McLachlan RI, Kazlauskas R, de Kretser D, Holden CA, Handelsman DJ. Reproductive hormone reference intervals for healthy fertile young men: evaluation of automated platform assays. *J Clin Endocrinol Metab*. 2005;**90**(11):5928–5936.
26. Turpeinen U, Linko S, Itkonen O, Hämäläinen E. Determination of testosterone in serum by liquid chromatography-tandem mass spectrometry. *Scand J Clin Lab Invest*. 2008;**68**(1):50–57.
27. Kushnir MM, Blamires T, Rockwood AL, Roberts WL, Yue B, Erdogan E, Bunker AM, Meikle AW. Liquid chromatography–tandem mass spectrometry assay for androstenedione, dehydroepiandrosterone, and testosterone with pediatric and adult reference intervals. *Clin Chem*. 2010;**56**(7):1138–1147.
28. Salameh WA, Redor-Goldman MM, Clarke NJ, Reitz RE, Caulfield MP. Validation of a total testosterone assay using high-turbulence liquid chromatography tandem mass spectrometry: total and free testosterone reference ranges. *Steroids*. 2010;**75**(2):169–175.
29. Neale SM, Hocking R, Biswas M, Turkes A, Rees D, Rees DA, Evans C. Adult testosterone and calculated free testosterone reference ranges by tandem mass spectrometry. *Ann Clin Biochem*. 2013;**50**(Pt 2):159–161.
30. Kelsey TW, Li LQ, Mitchell RT, Whelan A, Anderson RA, Wallace WH. A validated age-related normative model for male total testosterone shows increasing variance but no decline after age 40 years [published correction appears in *PLoS One*. 2015;10(2):e0117674]. *PLoS One*. 2014;**9**(10):e109346.
31. Hart RJ, Doherty DA, McLachlan RI, Walls ML, Keelan JA, Dickinson JE, Skakkebaek NE, Norman RJ, Handelsman DJ. Testicular function in a birth cohort of young men. *Hum Reprod*. 2015;**30**(12):2713–2724.
32. Travison TG, Vesper HW, Orwoll E, Wu F, Kaufman JM, Wang Y, Lapauw B, Fiers T, Matsumoto AM, Bhasin S. Harmonized reference ranges for circulating testosterone levels in men of four cohort studies in the United States and Europe. *J Clin Endocrinol Metab*. 2017;**102**(4):1161–1173.
33. Haring R, Hannemann A, John U, Radke D, Nauck M, Wallaschofski H, Owen L, Adaway J, Keevil BG, Brabant G. Age-specific reference ranges for serum testosterone and androstenedione concentrations in women measured by liquid chromatography-tandem mass spectrometry. *J Clin Endocrinol Metab*. 2012;**97**(2):408–415.
34. Bui HN, Sluss PM, Blincko S, Knol DL, Blankenstein MA, Heijboer AC. Dynamics of serum testosterone during the menstrual cycle evaluated by daily measurements with an ID-LC-MS/MS method and a 2nd generation automated immunoassay. *Steroids*. 2013;**78**(1):96–101.

35. 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**(17):1309–1314.
36. Eklund E, Berglund B, Labrie F, Carlström K, Ekström L, Hirschberg AL. Serum androgen profile and physical performance in women Olympic athletes. *Br J Sports Med*. 2017;**51**(17):1301–1308.
37. Travison TG, Zhuang WV, Lunetta KL, Karasik D, Bhasin S, Kiel DP, Coviello AD, Murabito JM. The heritability of circulating testosterone, oestradiol, oestrone and sex hormone binding globulin concentrations in men: the Framingham Heart Study. *Clin Endocrinol (Oxf)*. 2014;**80**(2):277–282.
38. Coviello AD, Zhuang WV, Lunetta KL, Bhasin S, Ulloor J, Zhang A, Karasik D, Kiel DP, Vasan RS, Murabito JM. Circulating testosterone and SHBG concentrations are heritable in women: the Framingham Heart Study. *J Clin Endocrinol Metab*. 2011;**96**(9):E1491–E1495.
39. Fui MN, Dupuis P, Grossmann M. Lowered testosterone in male obesity: mechanisms, morbidity and management. *Asian J Androl*. 2014;**16**(2):223–231.
40. Corona G, Rastrelli G, Monami M, Saad F, Luconi M, Lucchese M, Facchiano E, Sforza A, Forti G, Mannucci E, Maggi M. Body weight loss reverts obesity-associated hypogonadotropic hypogonadism: a systematic review and meta-analysis. *Eur J Endocrinol*. 2013;**168**(6):829–843.
41. Sartorius G, Spasevska S, Idan A, Turner L, Forbes E, Zamojska A, Allan CA, Ly LP, Conway AJ, McLachlan RI, Handelsman DJ. Serum testosterone, dihydrotestosterone and estradiol concentrations in older men self-reporting very good health: the healthy man study. *Clin Endocrinol (Oxf)*. 2012;**77**(5):755–763.
42. Webb ML, Wallace JP, Hamill C, Hodgson JL, Mashaly MM. Serum testosterone concentration during two hours of moderate intensity treadmill running in trained men and women. *Endocr Res*. 1984;**10**(1):27–38.
43. Cano Sokoloff N, Misra M, Ackerman KE. Exercise, training, and the hypothalamic-pituitary-gonadal axis in men and women. *Front Horm Res*. 2016;**47**:27–43.
44. Bozdag G, Mumusoglu S, Zengin D, Karabulut E, Yildiz BO. The prevalence and phenotypic features of polycystic ovary syndrome: a systematic review and meta-analysis. *Hum Reprod*. 2016;**31**(12):2841–2855.
45. Hagmar M, Berglund B, Brismar K, Hirschberg AL. Hyperandrogenism may explain reproductive dysfunction in Olympic athletes. *Med Sci Sports Exerc*. 2009;**41**(6):1241–1248.
46. Eliakim A, Marom N, Galitskaya L, Nemet D. Hyperandrogenism among elite adolescent female athletes. *J Pediatr Endocrinol Metab*. 2010;**23**(8):755–758.
47. Rickenlund A, Carlström K, Ekblom B, Brismar TB, von Schoultz B, Hirschberg AL. Hyperandrogenicity is an alternative mechanism underlying oligomenorrhea or amenorrhea in female athletes and may improve physical performance. *Fertil Steril*. 2003;**79**(4):947–955.
48. Falhammar H, Nordenström A. Nonclassic congenital adrenal hyperplasia due to 21-hydroxylase deficiency: clinical presentation, diagnosis, treatment, and outcome. *Endocrine*. 2015;**50**(1):32–50.
49. Auchus RJ. The classic and nonclassic concenital adrenal hyperplasias. *Endocr Pract*. 2015;**21**(4):383–389.
50. Moran LJ, Mundra PA, Teede HJ, Meikle PJ. The association of the lipidomic profile with features of polycystic ovary syndrome. *J Mol Endocrinol*. 2017;**59**(1):93–104.
51. Münzker J, Lindheim L, Adaway J, Trummer C, Lerchbaum E, Pieber TR, Keevil B, Obermayer-Pietsch B. High salivary testosterone-to-androstenedione ratio and adverse metabolic phenotypes in women with polycystic ovary syndrome. *Clin Endocrinol (Oxf)*. 2017;**86**(4):567–575.
52. O'Reilly MW, Kempegowda P, Jenkinson C, Taylor AE, Quanson JL, Storbeck KH, Arlt W. 11-Oxygenated C19 steroids are the predominant androgens in polycystic ovary syndrome. *J Clin Endocrinol Metab*. 2017;**102**(3):840–848.
53. Handelsman DJ, Teede HJ, Desai R, Norman RJ, Moran LJ. Performance of mass spectrometry steroid profiling for diagnosis of polycystic ovary syndrome. *Hum Reprod*. 2017;**32**(2):418–422.
54. Pasquali R, Zanotti L, Fanelli F, Mezzullo M, Fazzini A, Morselli Labate AM, Repaci A, Ribichini D, Gambineri A. Defining hyperandrogenism in women with polycystic ovary syndrome: a challenging perspective. *J Clin Endocrinol Metab*. 2016;**101**(5):2013–2022.
55. Yang Y, Han Y, Wang W, Du T, Li Y, Zhang J, Yang D, Zhao X. Assessing new terminal body and facial hair growth during pregnancy: toward developing a simplified visual scoring system for hirsutism. *Fertil Steril*. 2016;**105**(2):494–500.
56. Tosi F, Fiers T, Kaufman JM, Dall'Alda M, Moretta R, Giagulli VA, Bonora E, Moghetti P. Implications of androgen assay accuracy in the phenotyping of women with polycystic ovary syndrome. *J Clin Endocrinol Metab*. 2016;**101**(2):610–618.
57. Daan NM, Jaspers L, Koster MP, Broekmans FJ, de Rijke YB, Franco OH, Laven JS, Kavousi M, Fauser BC. Androgen levels in women with various forms of ovarian dysfunction: associations with cardiometabolic features. *Hum Reprod*. 2015;**30**(10):2376–2386.
58. Bui HN, Sluss PM, Hayes FJ, Blincko S, Knol DL, Blankenstein MA, Heijboer AC. Testosterone, free testosterone, and free androgen index in women: reference intervals, biological variation, and diagnostic value in polycystic ovary syndrome. *Clin Chim Acta*. 2015;**450**:227–232.
59. Keefe CC, Goldman MM, Zhang K, Clarke N, Reitz RE, Welt CK. Simultaneous measurement of thirteen steroid hormones in women with polycystic ovary syndrome and control women using liquid chromatography-tandem mass spectrometry. *PLoS One*. 2014;**9**(4):e93805.
60. Yasmin E, Balen AH, Barth JH. The association of body mass index and biochemical hyperandrogenaemia in women with and without polycystic ovary syndrome. *Eur J Obstet Gynecol Reprod Biol*. 2013;**166**(2):173–177.
61. Janse F, Eijkemans MJ, Goverde AJ, Lentjes EG, Hoek A, Lambalk CB, Hickey TE, Fauser BC, Norman RJ. Assessment of androgen concentration in women: liquid chromatography–tandem mass spectrometry and extraction RIA show comparable results. *Eur J Endocrinol*. 2011;**165**(6):925–933.
62. Jedel E, Gustafson D, Waern M, Sverrisdottir YB, Landén M, Janson PO, Labrie F, Ohlsson C, Stener-Victorin E. Sex steroids, insulin sensitivity and sympathetic nerve activity in relation to affective symptoms in women with polycystic ovary syndrome. *Psychoneuroendocrinology*. 2011;**36**(10):1470–1479.
63. Legro RS, Schlaff WD, Diamond MP, Coutifaris C, Casson PR, Brzyski RG, Christman GM, Trussell JC, Krawetz SA, Snyder PJ, Ohl D, Carson SA, Steinkampf MP, Carr BR, McGovern PG, Cataldo NA, Gosman GG, Nestler JE, Myers ER, Santoro N, Eisenberg E, Zhang M, Zhang H; Reproductive Medicine Network. Total testosterone assays in women with polycystic ovary syndrome: precision and correlation with hirsutism. *J Clin Endocrinol Metab*. 2010;**95**(12):5305–5313.
64. Stener-Victorin E, Holm G, Labrie F, Nilsson L, Janson PO, Ohlsson C. Are there any sensitive and specific sex steroid markers for polycystic ovary syndrome? *J Clin Endocrinol Metab*. 2010;**95**(2):810–819.
65. Finkelstein JS, Lee H, Burnett-Bowie SA, Pallais JC, Yu EW, Borges LF, Jones BF, Barry CV, Wulczyn KE, Thomas BJ, Leder BZ. Gonadal steroids and body composition, strength, and sexual function in men. *N Engl J Med*. 2013;**369**(11):1011–1022.
66. Donovan KA, Gonzalez BD, Nelson AM, Fishman MN, Zachariah B, Jacobsen PB. Effect of androgen deprivation therapy on sexual function and bother in men with prostate cancer: a controlled comparison. *Psychooncology*. 2018;**27**(1):316–324.
67. Buena F, Swerdloff RS, Steiner BS, Lutchmansingh P, Peterson MA, Pandian MR, Galmarini M, Bhasin S. Sexual function does not change when serum testosterone levels are pharmacologically varied within the normal male range. *Fertil Steril*. 1993;**59**(5):1118–1123.
68. Sartorius GA, Ly LP, Handelsman DJ. Male sexual function can be maintained without aromatization: randomized placebo-controlled trial of dihydrotestosterone (DHT) in healthy, older men for 24 months. *J Sex Med*. 2014;**11**(10):2562–2570.
69. Liu PY, Swerdloff RS, Christenson PD, Handelsman DJ, Wang C; Hormonal Male Contraception Summit Group. Rate, extent, and modifiers of spermatogenic recovery after hormonal male contraception: an integrated analysis. *Lancet*. 2006;**367**(9520):1412–1420.
70. Walsh PC, Swerdloff RS. Biphasic effect of testosterone on spermatogenesis in the rat. *Invest Urol*. 1973;**11**(3):190–193.
71. Singh J, O'Neill C, Handelsman DJ. Induction of spermatogenesis by androgens in gonadotropin-deficient (*hpg*) mice. *Endocrinology*. 1995;**136**(12):5311–5321.
72. Handelsman DJ, Spaliviero JA, Simpson JM, Allan CM, Singh J. Spermatogenesis without gonadotropins: maintenance has a lower testosterone threshold than initiation. *Endocrinology*. 1999;**140**(9):3938–3946.
73. Juel Mortensen L, Blomberg Jensen M, Christiansen P, Rønholt AM, Jørgensen A, Frederiksen H, Nielsen JE, Loya AC, Grønkær Toft B, Skakkebæk NE, Rajpert-De Meyts E, Juul A. Germ cell neoplasia in situ and preserved fertility despite suppressed gonadotropins in a patient with testotoxicosis. *J Clin Endocrinol Metab*. 2017;**102**(12):4411–4416.
74. Cunha-Silva M, Brito VN, Macedo DB, Bessa DS, Ramos CO, Lima LG, Barroso PS, Arnhold IJP, Segaloff DL, Mendonca BB, Latronico AC. Spontaneous fertility in a male patient with testotoxicosis despite suppression of FSH levels. *Hum Reprod*. 2018;**33**(5):914–918.
75. Mendonca BB, Batista RL, Domenice S, Costa EM, Arnhold IJ, Russell DW, Wilson JD. Steroid 5α-reductase 2 deficiency. *J Steroid Biochem Mol Biol*. 2016;**163**:206–211.
76. Mendonca BB, Gomes NL, Costa EM, Inacio M, Martin RM, Nishi MY, Carvalho FM, Tibor FD, Domenice S. 46,XY disorder of sex development (DSD) due to 17β-hydroxysteroid dehydrogenase type 3 deficiency. *J Steroid Biochem Mol Biol*. 2017;**165**(Pt A):79–85.

77. Quigley CA, De Bellis A, Marschke KB, el-Awady MK, Wilson EM, French FS. Androgen receptor defects: historical, clinical, and molecular perspectives. *Endocr Rev*. 1995;**16**(3):271–321.
78. Lucas-Herald A, Bertelloni S, Juul A, Bryce J, Jiang J, Rodie M, Sinnott R, Boroujerdi M, Lindhardt Johansen M, Hiort O, Holterhus PM, Cools M, Guaragna-Filho G, Guerra-Junior G, Weintrob N, Hannema S, Drop S, Guran T, Darendeliler F, Nordenstrom A, Hughes IA, Acerini C, Tadokoro-Cuccaro R, Ahmed SF. The long-term outcome of boys with partial androgen insensitivity syndrome and a mutation in the androgen receptor gene. *J Clin Endocrinol Metab*. 2016;**101**(11):3959–3967.
79. El-Maouche D, Arlt W, Merke DP. Congenital adrenal hyperplasia. *Lancet*. 2017;**390**(10108):2194–2210.
80. Bermon S, Garnier PY, Hirschberg AL, Robinson N, Giraud S, Nicoli R, Baume N, Saugy M, Fénichel P, Bruce SJ, Henry H, Dollé G, Ritzen M. Serum androgen levels in elite female athletes. *J Clin Endocrinol Metab*. 2014;**99**(11):4328–4335.
81. Imperato-McGinley J, Peterson RE, Gautier T, Sturla E. Androgens and the evolution of male-gender identity among male pseudohermaphrodites with 5α-reductase deficiency. *N Engl J Med*. 1979;**300**(22):1233–1237.
82. Kang HJ, Imperato-McGinley J, Zhu YS, Rosenwaks Z. The effect of 5α-reductase-2 deficiency on human fertility. *Fertil Steril*. 2014;**101**(2):310–316.
83. Strickland AL, French FS. Absence of response to dihydrotestosterone in the syndrome of testicular feminization. *J Clin Endocrinol Metab*. 1969;**29**(9):1284–1286.
84. Rosenfield RL, Lawrence AM, Liao S, Landau RL. Androgens and androgen responsiveness in the feminizing testis syndrome. Comparison of complete and "incomplete" forms. *J Clin Endocrinol Metab*. 1971;**32**(5):625–632.
85. Hamilton CR Jr, Kliman B. Anabolic effect of dihydrotestosterone in testicular feminization syndrome. *Metabolism*. 1971;**20**(9):870–877.
86. Zachmann M, Zagalak M, Völlmin JA, Gitzelmann RP, Prader A. Influence of testosterone on urinary $^{15}$N-balance in normal subjects and patients with testicular feminization. *Clin Chim Acta*. 1977;**77**(2):147–157.
87. Tincello DG, Saunders PT, Hodgins MB, Simpson NB, Edwards CR, Hargreaves TB, Wu FC. Correlation of clinical, endocrine and molecular abnormalities with in vivo responses to high-dose testosterone in patients with partial androgen insensitivity syndrome. *Clin Endocrinol (Oxf)*. 1997;**46**(4):497–506.
88. Grino PB, Isidro-Gutierrez RF, Griffin JE, Wilson JD. Androgen resistance associated with a qualitative abnormality of the androgen receptor and responsive to high dose androgen therapy. *J Clin Endocrinol Metab*. 1989;**68**(3):578–584.
89. Lundberg Giwercman Y, Nikoshkov A, Lindsten K, Byström B, Pousette A, Knudtzon J, Alm J, Wedell A. Response to treatment in patients with partial androgen insensitivity due to mutations in the DNA-binding domain of the androgen receptor. *Horm Res*. 2000;**53**(2):83–88.
90. Holterhus PM, Sinnecker GH, Hiort O. Phenotypic diversity and testosterone-induced normalization of mutant L712F androgen receptor function in a kindred with androgen insensitivity. *J Clin Endocrinol Metab*. 2000;**85**(9):3245–3250.
91. Quigley CA. *The androgen receptor: physiology and pathophysiology*. In: Nieschlag E, Behre HM, eds. *Testosterone: Action, Deficiency, Substitution*. 2nd ed. Berlin, Germany: Springer-Verlag; 1998:33–106.
92. Karunasena N, Han TS, Mallappa A, Elman M, Merke DP, Ross RJ, Daniel E. Androgens correlate with increased erythropoiesis in women with congenital adrenal hyperplasia. *Clin Endocrinol (Oxf)*. 2017;**86**(1):19–25.
93. Herbst KL, Bhasin S. Testosterone action on skeletal muscle. *Curr Opin Clin Nutr Metab Care*. 2004;**7**(3):271–277.
94. Dubois V, Laurent MR, Sinnesael M, Cielen N, Helsen C, Clinckemalie L, Spans L, Gayan-Ramirez G, Deldicque L, Hespel P, Carmeliet G, Vanderschueren D, Claessens F. A satellite cell-specific knockout of the androgen receptor reveals myostatin as a direct androgen target in skeletal muscle. *FASEB J*. 2014;**28**(7):2979–2994.
95. Usui T, Kajita K, Kajita T, Mori I, Hanamoto T, Ikeda T, Okada H, Taguchi K, Kitada Y, Morita H, Sasaki T, Kitamura T, Sato T, Kojima I, Ishizuka T. Elevated mitochondrial biogenesis in skeletal muscle is associated with testosterone-induced body weight loss in male mice. *FEBS Lett*. 2014;**588**(10):1935–1941.
96. Mänttäri S, Anttila K, Järvilehto M. Testosterone stimulates myoglobin expression in different muscles of the mouse. *J Comp Physiol B*. 2008;**178**(7):899–907.
97. Ferrando AA, Sheffield-Moore M, Yeckel CW, Gilkison C, Jiang J, Achacosa A, Lieberman SA, Tipton K, Wolfe RR, Urban RJ. Testosterone administration to older men improves muscle function: molecular and physiological mechanisms. *Am J Physiol Endocrinol Metab*. 2002;**282**(3):E601–E607.
98. Matzuk MM, Lamb DJ. The biology of infertility: research advances and clinical challenges. *Nat Med*. 2008;**14**(11):1197–1213.
99. Matzuk MM, Lamb DJ. Genetic dissection of mammalian fertility pathways. *Nat Cell Biol*. 2002;**4**(Suppl):S41–S49.
100. Walters KA, Simanainen U, Handelsman DJ. Molecular insights into androgen actions in male and female reproductive function from androgen receptor knockout models. *Hum Reprod Update*. 2010;**16**(5):543–558.
101. 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*. 2008;**22**(8):2676–2689.
102. Morrow JR Jr, Hosler WW. Strength comparisons in untrained men and trained women athletes. *Med Sci Sports Exerc*. 1981;**13**(3):194–197.
103. Miller AE, MacDougall JD, Tarnopolsky MA, Sale DG. Gender differences in strength and muscle fiber characteristics. *Eur J Appl Physiol Occup Physiol*. 1993;**66**(3):254–262.
104. Janssen I, Heymsfield SB, Wang ZM, Ross R. Skeletal muscle mass and distribution in 468 men and women aged 18–88 yr. *J Appl Physiol*. 2000;**89**(1):81–88.
105. Hosler WW, Morrow JR Jr. Arm and leg strength compared between young women and men after allowing for differences in body size and composition. *Ergonomics*. 1982;**25**(4):309–313.
106. Sale DG. Neuromuscular function. In: Tarnopolsky M, ed. *Gender Differences in Metabolism: Practical and Nutritional Implications*. Boca Raton, FL: CRC Press; 1999:61–86.
107. Tønnessen 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.
108. Carmina E, Guastella E, Longo RA, Rini GB, Lobo RA. Correlates of increased lean muscle mass in women with polycystic ovary syndrome. *Eur J Endocrinol*. 2009;**161**(4):583–589.
109. Douchi T, Oki T, Yamasaki H, Kuwahata R, Nakae M, Nagata Y. Relationship of androgens to muscle size and bone mineral density in women with polycystic ovary syndrome. *Obstet Gynecol*. 2001;**98**(3):445–449.
110. Cardinale M, Stone MH. Is testosterone influencing explosive performance? *J Strength Cond Res*. 2006;**20**(1):103–107.
111. Bhasin S, Woodhouse L, Casaburi R, Singh AB, Bhasin D, Berman N, Chen X, Yarasheski KE, Magliano L, Dzekov C, Dzekov J, Bross R, Phillips J, Sinha-Hikim I, Shen R, Storer TW. Testosterone dose-response relationships in healthy young men. *Am J Physiol Endocrinol Metab*. 2001;**281**(6):E1172–E1181.
112. Huang G, Basaria S, Travison TG, Ho MH, Davda M, Mazer NA, Miciek R, Knapp PE, Zhang A, Collins L, Ursino M, Appleman E, Dzekov C, Stroh H, Ouellette M, Rundell T, Baby M, Bhatia NN, Khorram O, Friedman T, Storer TW, Bhasin S. 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**(6):612–623.
113. Dobs AS, Nguyen T, Pace C, Roberts CP. Differential effects of oral estrogen versus oral estrogen-androgen replacement therapy on body composition in postmenopausal women. *J Clin Endocrinol Metab*. 2002;**87**(4):1509–1516.
114. Elbers JM, Asscheman H, Seidell JC, Gooren LJ. Effects of sex steroid hormones on regional fat depots as assessed by magnetic resonance imaging in transsexuals. *Am J Physiol*. 1999;**276**(2 Pt 1):E317–E325.
115. 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). *Eur J Endocrinol*. 2015;**172**(2):163–171.
116. Sonksen P. Determination and regulation of body composition in elite athletes. *Br J Sports Med*. 2018;**52**(4):219–229.
117. Storer TW, Woodhouse L, Magliano L, Singh AB, Dzekov C, Dzekov J, Bhasin S. Changes in muscle mass, muscle strength, and power but not physical function are related to testosterone dose in healthy older men. *J Am Geriatr Soc*. 2008;**56**(11):1991–1999.
118. Bhasin S, Parker RA, Sattler F, Haubrich R, Alston B, Umbleja T, Shikuma CM; AIDS Clinical Trials Group Protocol A5079 Study Team. Effects of testosterone supplementation on whole body and regional fat mass and distribution in human immunodeficiency virus-infected men with abdominal obesity. *J Clin Endocrinol Metab*. 2007;**92**(3):1049–1057.
119. Bhasin S, Woodhouse L, Casaburi R, Singh AB, Mac RP, Lee M, Yarasheski KE, Sinha-Hikim I, Dzekov C, Dzekov J, Magliano L, Storer TW. 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**(2):678–688.
120. Franke WW, Berendonk B. Hormonal doping and androgenization of athletes: a secret program of the German Democratic Republic government. *Clin Chem*. 1997;**43**(7):1262–1279.
121. Shahani S, Braga-Basaria M, Maggio M, Basaria S. Androgens and erythropoiesis: past and present. *J Endocrinol Invest*. 2009;**32**(8):704–716.
122. Bachman E, Travison TG, Basaria S, Davda MN, Guo W, Li M, Connor Westfall J, Bae H, Gordeuk V,

Bhasin S. 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**(6):725–735.

123. Ordway GA, Garry DJ. Myoglobin: an essential hemoprotein in striated muscle. *J Exp Biol*. 2004;**207**(Pt 20):3441–3446.
124. Ekblom B, Goldbarg AN, Gullbring B. Response to exercise after blood loss and reinfusion. *J Appl Physiol*. 1972;**33**(2):175–180.
125. Murphy WG. The sex difference in haemoglobin levels in adults—mechanisms, causes, and consequences. *Blood Rev*. 2014;**28**(2):41–47.
126. Grossmann M, Zajac JD. Hematological changes during androgen deprivation therapy. *Asian J Androl*. 2012;**14**(2):187–192.
127. Snyder PJ, Peachey H, Berlin JA, Hannoush P, Haddad G, Dlewati A, Santanna J, Loh L, Lenrow DA, Holmes JH, Kapoor SC, Atkinson LE, Strom BL. Effects of testosterone replacement in hypogonadal men. *J Clin Endocrinol Metab*. 2000;**85**(8):2670–2677.
128. Roy CN, Snyder PJ, Stephens-Shields AJ, Artz AS, Bhasin S, Cohen HJ, Farrar JT, Gill TM, Zeldow B, Cella D, Barrett-Connor E, Cauley JA, Crandall JP, Cunningham GR, Ensrud KE, Lewis CE, Matsumoto AM, Molitch ME, Pahor M, Swerdloff RS, Cifelli D, Hou X, Resnick SM, Walston JD, Anton S, Basaria S, Diem SJ, Wang C, Schrier SL, Ellenberg SS. Association of testosterone levels with anemia in older men: a controlled clinical trial. *JAMA Intern Med*. 2017;**177**(4):480–490.
129. Berria R, Gastaldelli A, Lucidi S, Belfort R, De Filippis E, Easton C, Brytzki R, Cusi K, Jovanovic L, DeFronzo R. Reduction in hematocrit level after pioglitazone treatment is correlated with decreased plasma free testosterone level, not hemodilution, in women with polycystic ovary syndrome. *Clin Pharmacol Ther*. 2006;**80**(2):105–114.
130. Han Y, Kim HS, Lee HJ, Oh JY, Sung YA. Metabolic effects of polycystic ovary syndrome in adolescents. *Ann Pediatr Endocrinol Metab*. 2015;**20**(3):136–142.
131. 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**(3):914–919.
132. Irwig MS. Testosterone therapy for transgender men. *Lancet Diabetes Endocrinol*. 2017;**5**(4):301–311.
133. Velho I, Fighera TM, Ziegelmann PK, Spritzer PM. Effects of testosterone therapy on BMI, blood pressure, and laboratory profile of transgender men: a systematic review. *Andrology*. 2017;**5**(5):881–888.
134. Jacobeit JW, Gooren LJ, Schulte HM. Safety aspects of 36 months of administration of long-acting intramuscular testosterone undecanoate for treatment of female-to-male transgender individuals. *Eur J Endocrinol*. 2009;**161**(5):795–798.
135. Almeida M, Laurent MR, Dubois V, Claessens F, O'Brien CA, Bouillon R, Vanderschueren D, Manolagas SC. Estrogens and androgens in skeletal physiology and pathophysiology. *Physiol Rev*. 2017;**97**(1):135–187.
136. Sharma K, Gupta P, Shandilya S. Age related changes in pelvis size among adolescent and adult females with reference to parturition from Naraingarh, Haryana (India). *Homo*. 2016;**67**(4):273–293.
137. Fischer B, Mitteroecker P. Allometry and sexual dimorphism in the human pelvis. *Anat Rec (Hoboken)*. 2017;**300**(4):698–705.
138. Riesenfeld A. Functional and hormonal control of pelvic width in the rat. *Acta Anat (Basel)*. 1978;**102**(4):427–432.
139. Berdnikovs S, Bernstein M, Metzler A, German RZ. Pelvic growth: ontogeny of size and shape sexual dimorphism in rat pelves. *J Morphol*. 2007;**268**(1):12–22.
140. Polderman TJ, Benyamin B, de Leeuw CA, Sullivan PF, van Bochoven A, Visscher PM, Posthuma D. Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nat Genet*. 2015;**47**(7):702–709.
141. Jelenkovic A, Sund R, Hur YM, Yokoyama Y, Hjelmborg JV, Möller S, Honda C, Magnusson PK, Pedersen NL, Ooki S, Aaltonen S, Stazi MA, Fagnani C, D'Ippolito C, Freitas DL, Maia JA, Ji F, Ning F, Pang Z, Rebato E, Busjahn A, Kandler C, Saudino KJ, Jang KL, Cozen W, Hwang AE, Mack TM, Gao W, Yu C, Li L, Corley RP, Huibregtse BM, Derom CA, Vlietinck RF, Loos RJ, Heikkilä K, Wardle J, Llewellyn CH, Fisher A, McAdams TA, Eley TC, Gregory AM, He M, Ding X, Bjerregaard-Andersen M, Beck-Nielsen H, Sodemann M, Tarnoki AD, Tarnoki DL, Knafo-Noam A, Mankuta D, Abramson L, Burt SA, Klump KL, Silberg JL, Eaves LJ, Maes HH, Krueger RF, McGue M, Pahlen S, Gatz M, Butler DA, Bartels M, van Beijsterveldt TC, Craig JM, Saffery R, Dubois L, Boivin M, Brendgen M, Dionne G, Vitaro F, Martin NG, Medland SE, Montgomery GW, Swan GE, Krasnow R, Tynelius P, Lichtenstein P, Haworth CM, Plomin R, Bayasgalan G, Narandalai D, Harden KP, Tucker-Drob EM, Spector T, Mangino M, Lachance G, Baker LA, Tuvblad C, Duncan GE, Buchwald D, Willemsen G, Skytthe A, Kyvik KO, Christensen K, Öncel SY, Aliev F, Rasmussen F, Goldberg JH, Sørensen TI, Boomsma DI, Kaprio J, Silventoinen K. Genetic and environmental influences on height from infancy to early adulthood: an individual-based pooled analysis of 45 twin cohorts. *Sci Rep*. 2016;**6**(1):28496.
142. Jelenkovic A, Hur YM, Sund R, Yokoyama Y, Siribaddana SH, Hotopf M, Sumathipala A, Rijsdijk F, Tan Q, Zhang D, Pang Z, Aaltonen S, Heikkilä K, Öncel SY, Aliev F, Rebato E, Tarnoki AD, Tarnoki DL, Christensen K, Skytthe A, Kyvik KO, Silberg JL, Eaves LJ, Maes HH, Cutler TL, Hopper JL, Ordoñana JR, Sánchez-Romera JF, Colodro-Conde L, Cozen W, Hwang AE, Mack TM, Sung J, Song YM, Yang S, Lee K, Franz CE, Kremen WS, Lyons MJ, Busjahn A, Nelson TL, Whitfield KE, Kandler C, Jang KL, Gatz M, Butler DA, Stazi MA, Fagnani C, D'Ippolito C, Duncan GE, Buchwald D, Derom CA, Vlietinck RF, Loos RJ, Martin NG, Medland SE, Montgomery GW, Jeong HU, Swan GE, Krasnow R, Magnusson PK, Pedersen NL, Dahl-Aslan AK, McAdams TA, Eley TC, Gregory AM, Tynelius P, Baker LA, Tuvblad C, Bayasgalan G, Narandalai D, Lichtenstein P, Spector TD, Mangino M, Lachance G, Bartels M, van Beijsterveldt TC, Willemsen G, Burt SA, Klump KL, Harris JR, Brandt I, Nilsen TS, Krueger RF, McGue M, Pahlen S, Corley RP, Hjelmborg JV, Goldberg JH, Iwatani Y, Watanabe M, Honda C, Inui F, Rasmussen F, Huibregtse BM, Boomsma DI, Sørensen TI, Kaprio J, Silventoinen K. Genetic and environmental influences on adult human height across birth cohorts from 1886 to 1994. *eLife*. 2016;**5**:e20320.
143. Bechtold S, Beyerlein A, Bonfig W, Dalla Pozza R, Putzker S, Otto R, Schmidt H, Schwarz HP. Sexual difference in bone geometry of adult patients with classical congenital adrenal hyperplasia: data using peripheral quantitative computed tomography. *Horm Res Paediatr*. 2014;**82**(3):171–178.
144. Falhammar H, Filipsson H, Holmdahl G, Janson PO, Nordenskjöld A, Hagenfeldt K, Thorén M. Fractures and bone mineral density in adult women with 21-hydroxylase deficiency. *J Clin Endocrinol Metab*. 2007;**92**(12):4643–4649.
145. Bhasin S, Storer TW, Berman N, Callegari C, Clevenger B, Phillips J, Bunnell TJ, Tricker R, Shirazi A, Casaburi R. The effects of supraphysiologic doses of testosterone on muscle size and strength in normal men. *N Engl J Med*. 1996;**335**(1):1–7.
146. Moreira CA, Bilezikian JP. Stress fractures: concepts and therapeutics. *J Clin Endocrinol Metab*. 2017;**102**(2):525–534.
147. Foryst-Ludwig A, Kintscher U. Sex differences in exercise-induced cardiac hypertrophy. *Pflugers Arch*. 2013;**465**(5):731–737.
148. Gibala MJ, Gillen JB, Percival ME. Physiological and health-related adaptations to low-volume interval training: influences of nutrition and sex. *Sports Med*. 2014;**44**(Suppl 2):S127–S137.
149. Townsend EA, Miller VM, Prakash YS. Sex differences and sex steroids in lung health and disease. *Endocr Rev*. 2012;**33**(1):1–47.
150. Levine SC, Foley A, Lourenco S, Ehrlich S, Ratliff K. Sex differences in spatial cognition: advancing the conversation. *Wiley Interdiscip Rev Cogn Sci*. 2016;**7**(2):127–155.
151. Hines M. Prenatal testosterone and gender-related behaviour. *Eur J Endocrinol*. 2006;**155**(Suppl 1):S115–S121.
152. Hines M, Spencer D, Kung KT, Browne WV, Constantinescu M, Noorderhaven RM. The early postnatal period, mini-puberty, provides a window on the role of testosterone in human neurobehavioural development. *Curr Opin Neurobiol*. 2016;**38**:69–73.
153. Pope HG Jr., Kouri EM, Hudson JI. Effects of supraphysiologic doses of testosterone on mood and aggression in normal men: a randomized controlled trial. *Arch Gen Psychiatry*. 2000;**57**(2):133–140.
154. Ferguson-Smith MA, Bavington LD. Natural selection for genetic variants in sport: the role of Y chromosome genes in elite female athletes with 46, XY DSD. *Sports Med*. 2014;**44**(12):1629–1634.
155. Heymsfield SB, Gonzalez MC, Lu J, Jia G, Zheng J. Skeletal muscle mass and quality: evolution of modern measurement concepts in the context of sarcopenia. *Proc Nutr Soc*. 2015;**74**(4):355–366.
156. Silventoinen K, Sammalisto S, Perola M, Boomsma DI, Cornes BK, Davis C, Dunkel L, De Lange M, Harris JR, Hjelmborg JV, Luciano M, Martin NG, Mortensen J, Nistico L, Pedersen NL, Skytthe A, Spector TD, Stazi MA, Willemsen G, Kaprio J. Heritability of adult body height: a comparative study of twin cohorts in eight countries. *Twin Res*. 2003;**6**(5):399–408.
157. Beunen G, Thomis M. Gene powered? Where to go from heritability (h2) in muscle strength and power? *Exerc Sport Sci Rev*. 2004;**32**(4):148–154.
158. Silventoinen K, Magnusson PK, Tynelius P, Kaprio J, Rasmussen F. Heritability of body size and muscle strength in young adulthood: a study of one million Swedish men. *Genet Epidemiol*. 2008;**32**(4):341–349.
159. Seeman E. Pathogenesis of bone fragility in women and men. *Lancet*. 2002;**359**(9320):1841–1850.
160. Nishiyama KK, Macdonald HM, Moore SA, Fung T, Boyd SK, McKay HA. Cortical porosity is higher in boys compared with girls at the distal radius and distal tibia during pubertal growth: an HR-pQCT study. *J Bone Miner Res*. 2012;**27**(2):273–282.
161. Oliveira CS, Alves C. The role of the SHOX gene in the pathophysiology of Turner syndrome. *Endocrinol Nutr*. 2011;**58**(8):433–442.
162. Ottesen AM, Aksglaede L, Garn I, Tartaglia N, Tassone F, Gravholt CH, Bojesen A, Sørensen K, Jørgensen N, Rajpert-De Meyts E, Gerdes T, Lind AM, Kjaergaard S, Juul A. Increased number of sex

chromosomes affects height in a nonlinear fashion: a study of 305 patients with sex chromosome aneuploidy. *Am J Med Genet A*. 2010;**152A**(5):1206–1212.
163. Wideman L, Weltman JY, Shah N, Story S, Veldhuis JD, Weltman A. Effects of gender on exercise-induced growth hormone release. *J Appl Physiol*. 1999;**87**(3):1154–1162.
164. Veldhuis JD, Roemmich JN, Rogol AD. Gender and sexual maturation-dependent contrasts in the neuroregulation of growth hormone secretion in prepubertal and late adolescent males and females—a general clinical research center-based study. *J Clin Endocrinol Metab*. 2000;**85**(7):2385–2394.
165. Veldhuis JD. Gender differences in secretory activity of the human somatotropic (growth hormone) axis. *Eur J Endocrinol*. 1996;**134**(3):287–295.
166. Ho KY, Evans WS, Blizzard RM, Veldhuis JD, Merriam GR, Samojlik E, Furlanetto R, Rogol AD, Kaiser DL, Thorner MO. Effects of sex and age on the 24-hour profile of growth hormone secretion in man: importance of endogenous estradiol concentrations. *J Clin Endocrinol Metab*. 1987;**64**(1):51–58.
167. Veldhuis JD, Roelfsema F, Keenan DM, Pincus S. Gender, age, body mass index, and IGF-I individually and jointly determine distinct GH dynamics: analyses in one hundred healthy adults. *J Clin Endocrinol Metab*. 2011;**96**(1):115–121.
168. Veldhuis JD, Patrie JT, Brill KT, Weltman JY, Mueller EE, Bowers CY, Weltman A. Contributions of gender and systemic estradiol and testosterone concentrations to maximal secretagogue drive of burst-like growth hormone secretion in healthy middle-aged and older adults. *J Clin Endocrinol Metab*. 2004;**89**(12):6291–6296.
169. Roelfsema F, Veldhuis JD. Growth hormone dynamics in healthy adults are related to age and sex and strongly dependent on body mass index. *Neuroendocrinology*. 2016;**103**(3-4):335–344.
170. Pritzlaff-Roy CJ, Widemen L, Weltman JY, Abbott R, Gutgesell M, Hartman ML, Veldhuis JD, Weltman A. Gender governs the relationship between exercise intensity and growth hormone release in young adults. *J Appl Physiol*. 2002;**92**(5):2053–2060.
171. Leung KC, Doyle N, Ballesteros M, Sjogren K, Watts CK, Low TH, Leong GM, Ross RJ, Ho KK. Estrogen inhibits GH signaling by suppressing GH-induced JAK2 phosphorylation, an effect mediated by SOCS-2. *Proc Natl Acad Sci USA*. 2003;**100**(3):1016–1021.
172. Ho KK, O'Sullivan AJ, Wolthers T, Leung KC. Metabolic effects of oestrogens: impact of the route of administration. *Ann Endocrinol (Paris)*. 2003;**64**(2):170–177.
173. Cappola AR, Bandeen-Roche K, Wand GS, Volpato S, Fried LP. Association of IGF-I levels with muscle strength and mobility in older women. *J Clin Endocrinol Metab*. 2001;**86**(9):4139–4146.
174. Meinhardt U, Nelson AE, Hansen JL, Birzniece V, Clifford D, Leung KC, Graham K, Ho KK. The effects of growth hormone on body composition and physical performance in recreational athletes: a randomized trial. *Ann Intern Med*. 2010;**152**(9):568–577.
175. Harper J. Race times for transgender athletes. *Journal of Sporting Cultures and Identities*. 2015;**6**(1):1–9.
176. Bermon S. Androgens and athletic performance of elite female athletes. *Curr Opin Endocrinol Diabetes Obes*. 2017;**24**(3):246–251.
177. Elbers JM, Asscheman H, Seidell JC, Megens JA, Gooren LJ. Long-term testosterone administration increases visceral fat in female to male transsexuals. *J Clin Endocrinol Metab*. 1997;**82**(7):2044–2047.
178. Handelsman DJ. Clinical review: the rationale for banning human chorionic gonadotropin and estrogen blockers in sport. *J Clin Endocrinol Metab*. 2006;**91**(5):1646–1653.
179. Asbell SO, Leon SA, Tester WJ, Brereton HD, Ago CT, Rotman M. Development of anemia and recovery in prostate cancer patients treated with combined androgen blockade and radiotherapy. *Prostate*. 1996;**29**(4):243–248.
180. Strum SB, McDermed JE, Scholz MC, Johnson H, Tisman G. Anaemia associated with androgen deprivation in patients with prostate cancer receiving combined hormone blockade. *Br J Urol*. 1997;**79**(6):933–941.
181. Bogdanos J, Karamanolakis D, Milathianakis C, Repousis P, Tsintavis A, Koutsilieris M. Combined androgen blockade-induced anemia in prostate cancer patients without bone involvement. *Anticancer Res*. 2003;**23**(2C):1757–1762.
182. Choo R, Chander S, Danjoux C, Morton G, Pearce A, Deboer G, Szumacher E, Loblaw A, Cheung P, Woo T. How are hemoglobin levels affected by androgen deprivation in non-metastatic prostate cancer patients? *Can J Urol*. 2005;**12**(1):2547–2552.
183. Chander S, Choo R, Danjoux C, Morton G, Pearse A, Deboer G, Szumacher E, Loblaw A, Cheung P, Woo T. Effect of androgen suppression on hemoglobin in prostate cancer patients undergoing salvage radiotherapy plus 2-year buserelin acetate for rising PSA after surgery. *Int J Radiat Oncol Biol Phys*. 2005;**62**(3):719–724.
184. Golfam M, Samant R, Eapen L, Malone S. Effects of radiation and total androgen blockade on serum hemoglobin, testosterone, and erythropoietin in patients with localized prostate cancer. *Curr Oncol*. 2012;**19**(4):e258–e263.
185. Storer TW, Miciek R, Travison TG. Muscle function, physical performance and body composition changes in men with prostate cancer undergoing androgen deprivation therapy. *Asian J Androl*. 2012;**14**(2):204–221.

## Acknowledgments

The authors are grateful for helpful insights and comments from Alan Vernec and Osquel Barroso (World Anti-Doping Agency), Peter Harcourt (Australian Football League, Federation of International Basketball Associations), and Richard Budgett (IOC).

***Correspondence and Reprint Requests:*** David J. Handelsman, PhD, ANZAC Research Institute, University of Sydney, Hospital Road, Concord Hospital, Sydney, New South Wales 2139, Australia. E-mail: djh@anzac.edu.au.

***Disclosure Summary:*** D.J.H. is a medical and scientific consultant for the IAAF and to the Australian Sports Anti-Doping Agency. He is a member of the World Anti-Doping Agency's Health, Medicine and Research Committee and of the IOC working group on hyperandrogenic female and transgender athletes. He has received institutional grant support from Besins Healthcare and Lawley for investigator-initiated clinical studies in testosterone pharmacology and has provided expert testimony in testosterone litigation. A.L.H. is a medical and scientific consultant for the Swedish Olympic Committee and a member of the IAAF and IOC working groups on hyperandrogenic female athletes and transgender athletes. She has received grant support from the IAAF for a study on testosterone and physical performance in women. S.B. is a medical and scientific consultant for the IAAF and a member of the IAAF and IOC working groups on hyperandrogenic female athletes and transgender athletes. The authors have no other involvement with any entity having a financial interest in the material discussed in the manuscript. Opinions expressed in this review are the personal views of the authors and do not represent those of the IAAF, IOC, World Anti-Doping Agency, or Swedish Olympic Committee.

## Abbreviations

AR, androgen receptor; CAH, congenital adrenal hyperplasia; CAIS, complete androgen insensitivity syndrome; DSD, disorder (or difference) of sex development; F2M, female-to-male; IAAF, International Association of Athletic Federations; IOC, International Olympic Committee; LC-MS, liquid chromatography–mass spectrometry; M2F, male-to-female; PAIS, partial androgen insensitivity syndrome; PCOS, polycystic ovary syndrome; SHOX, short stature homeobox.

# Exhibit 19

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

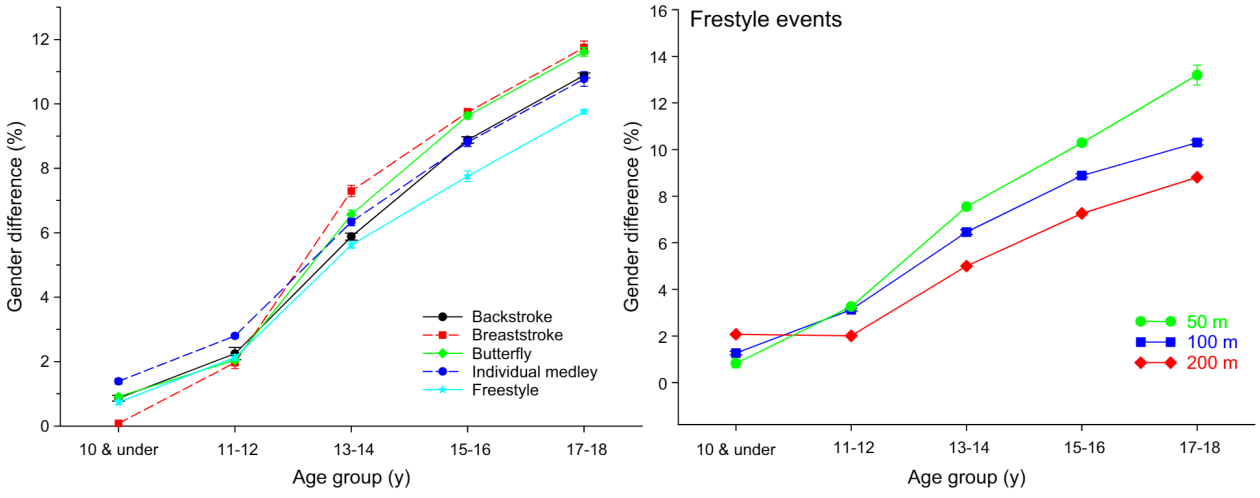



**FIGURE 1** Gender differences in performance (in percentage) according to age group and stroke (left panel) or distance in freestyle events (right) in swimming events. Data shown as mean and standard error of the mean. Note greatest increase after the age of 12 years by age in breaststroke and least 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

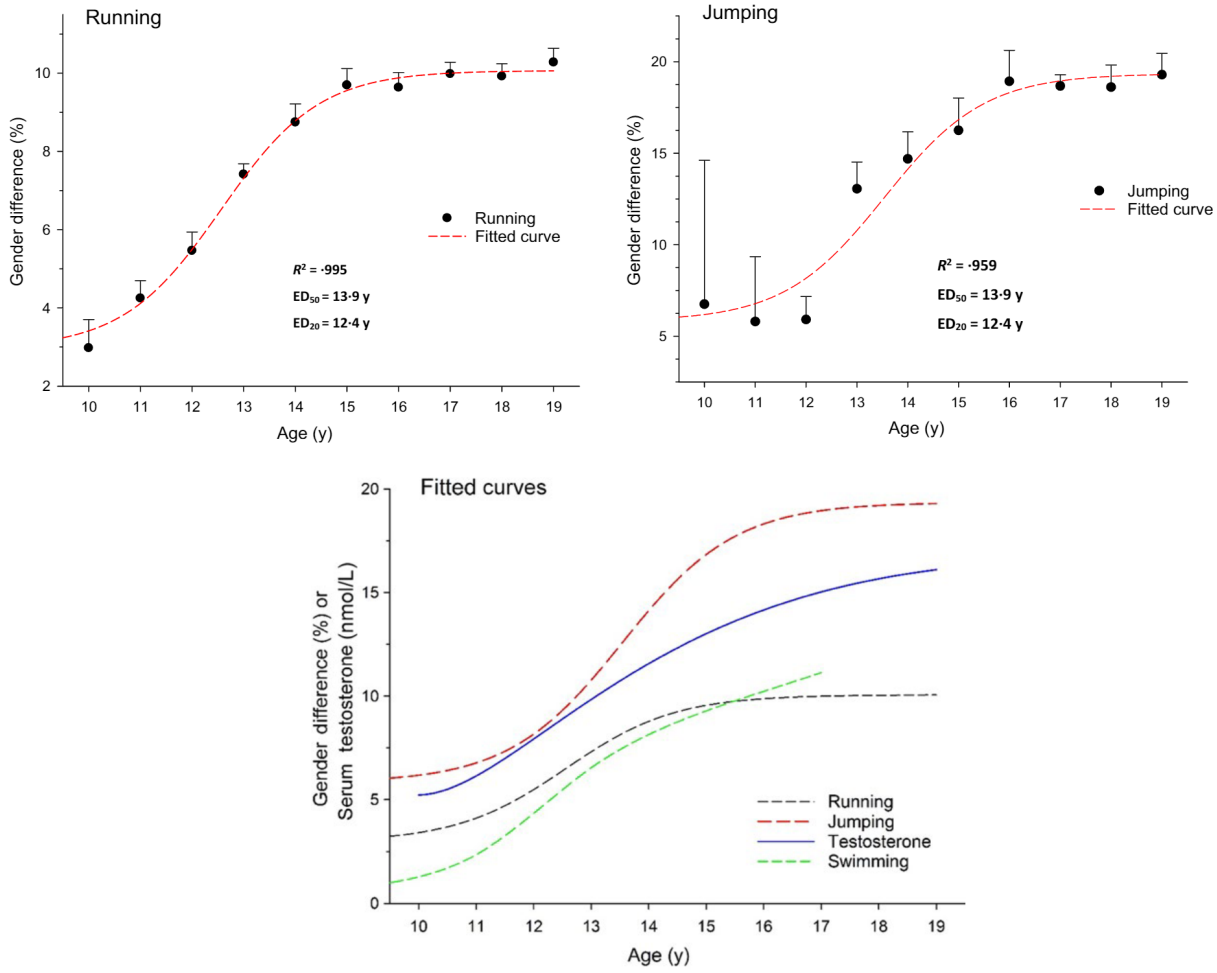



**FIGURE 2** Gender differences in performance (in percentage) according to age (in years) in running events including 50 m, 60 m, 100 m, 200 m, 300 m, 400 m, 500 m, 600 m, 800 m, 1000 m, 1500 m, 1 mile, 2000 m, 3000 m and 2 miles (upper left panel) and in jumping events including high jump, pole vault, 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]

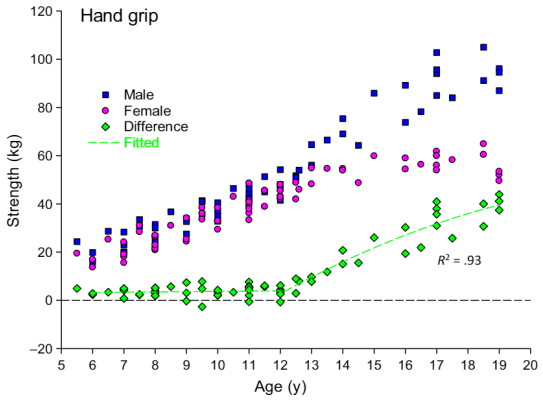



**FIGURE 3** Hand-grip strength in children and adolescents from 19 studies including 5676 males (square) and 5489 females (circles) and the differences between male and females (diamonds) conducted between 1966 and 2009. The dotted line represents the fitted curve using a piecewise linear-quadratic 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

# Exhibit 20

PLOS ONE

RESEARCH ARTICLE

# Sex differences in youth elite swimming

**Jonathon W. Senefeld[1]*, Andrew J. Clayburn[1], Sarah E. Baker[1], Rickey E. Carter[2], Patrick W. Johnson[2], Michael J. Joyner[1]**

**1** Department of Anesthesiology and Perioperative Medicine, Mayo Clinic, Rochester, Minnesota, United States of America, **2** Department of Health Sciences Research, Mayo Clinic, Jacksonville, Florida, United States of America

* senefeld.jonathon@mayo.edu



OPEN ACCESS

**Citation:** Senefeld JW, Clayburn AJ, Baker SE, Carter RE, Johnson PW, Joyner MJ (2019) Sex differences in youth elite swimming. PLoS ONE 14 (11): e0225724. https://doi.org/10.1371/journal.pone.0225724

**Editor:** Daniel Boullosa, Universidade Federal de Mato Grosso do Sul, BRAZIL

**Received:** September 18, 2019

**Accepted:** November 10, 2019

**Published:** November 22, 2019



**Data Availability Statement:** All data are available from the publicly accessible online database found at the USA Swimming Data Hub website (https://www.usaswimming.org/Home/times/data-hub).

**Funding:** The authors received no specific funding for this work.

**Competing interests:** The authors have declared that no competing interests exist.

## Abstract

### Background

The timing and magnitude of sex differences in athletic performance during early human development, prior to adulthood, is unknown.

### Objective

To compare swimming velocity of boys and girls for all Olympic-length freestyle swimming events to determine the age of divergence in swimming performance.

### Methods

We collected the all-time top 100 U.S. freestyle swimming performance times of boys and girls age 5 to 18 years for the 50m to 1500m events.

### Results

Swimming performance improved with increasing age for boys and girls ($p<0.001$) until reaching a plateau, which initiated at a younger age for girls (15 years) than boys (17 years; sex×age; $p<0.001$). Prior to age 10, the top 5 swimming records for girls were 3% faster than the top boys ($p<0.001$). For the $10^{th}$-$50^{th}$ places, however, there were no sex-related differences in swimming performance prior to age 10 ($p = 0.227$). For both the top 5 and $10^{th}$-$50^{th}$ places, the sex difference in performance increased from age 10 (top 5, 2.5%; $10^{th}$-$50^{th}$ places, 1.0%) until age 17 (top 5, 7.6%; $10^{th}$-$50^{th}$ places, 8.0%). For all places, the sex difference in performance at age 18 was larger for sprint events (9.6%; 50-200m) than endurance events (7.1%; 400-1500m; $p<0.001$). Additionally, the sex-related difference in performance increased across age and US ranking from 2.4% for $1^{st}$ place to 4.3% for $100^{th}$ place ($p<0.001$), indicating less depth of performance in girls than boys. However, annual participation was ~20% higher in girls than boys for all ages ($p<0.001$).

### Conclusion

The top 5 girls demonstrated faster swimming velocities and the $10^{th}$-$50^{th}$ place girls demonstrated similar swimming velocities than boys (until ~10 years). After age 10, however, boys demonstrated increasingly faster swimming velocities than girls until 17 years. Collectively,

these data suggest girls are faster, or at least not slower, than boys prior to the performance-enhancing effects of puberty.

## Introduction

Recent high profile cases have raised controversy about whether transgender athletes and XY (intersex) women with differences in sexual development should be allowed to participate in competitions restricted to women, for example Caster v. IAAF [1]. Stemming from the recognized performance-enhancing effects of androgens [2–6], regulation of endogenous androgen levels is now required to be eligible for participation in many women's sports competitions. To better understand the timing and magnitude of sex differences in athletic performance during human development we examined elite swimming performances in youth as a proxy to estimate the expected divergence of athletic performance between girls and boys during periods of low androgen concentrations (pre-puberty) and increasing concentrations of androgens throughout puberty. As reviewed previously [6], human androgens—primarily testosterone and its related metabolite dihydrotestosterone (DHT)—are key factors in the development of muscle, bone and hemoglobin. Androgens largely contribute to bigger and more powerful muscle mass, higher hemoglobin concentrations (and subsequent oxygen carrying capacity), and greater bone strength in men than women [6]. In combination, these androgen-driven and sex-based differences in muscle, bone and hemoglobin contribute to a ~10% higher maximal oxygen consumption capacity ($\dot{V}O_{2max}$) in men compared with women [7, 8].

Androgen levels are not different between the sexes prior to puberty, however, after completion of puberty, circulating testosterone levels are on average ~10–20 times greater in men than children or women at any age [9, 10]. This sex-based difference in circulating testosterone is the basic premise to explain why men have faster performance times than women in many time-based sports including running, cycling, swimming, rowing, etc. [11–17]. Thus, prior to puberty, it would be theorized that sex differences in performance between boys and girls would be negligible—which has been observed previously in athletes ~10–12 years of age [5, 18]. However, the sex-based differences in performance prior to age 10 are unknown and there is no previous data on long distance swimming which likely has the smallest influence from sociocultural biases [13]. Historically, women have had less opportunity to participate in most sports than men, and these differences in opportunity are suspected to contribute to the larger sex differences in performance than would be predicted based on physiological differences between the men and women [14, 19]. Women have been permitted to participate in swimming at the highest levels for many years (since ~1912) and currently more girls typically participate than boys, thus, swimming is an ideal 'experiment of nature' for this question [15]. Elite swimmers are also generally homogenous for high socio-economic status, meaning that sex differences in nutrition or access to medical care are unlikely [20]. There is intensive training from a young age, and practices and competitions are inclusive of both sexes. Additionally, standardized environmental conditions along with state of the art facilities are widely available during championship competitions.

Accordingly, the objective of our study was to determine the age of the divergence of swimming performance between elite boys and girls. To our knowledge, our study is the first to analytically investigate the role of normal human hormonal changes on sex-related differences in sprint and endurance performance in elite youth swimming. We hypothesized that: 1) there would be no sex-differences in swimming performance of girls and boys with similar and low

testosterone concentrations (pre-pubescent years), 2) boys would be faster than girls after the initiation of puberty, and 3) the faster performance of boys would plateau after age 16, as androgen concentrations plateau [21].

## Materials and methods

### Methods

Finishing times of the top 100 All-Time Freestyle Swimming Records for Long Course Meters for boys and girls between 5 and 18 years of age in one-year age brackets were analyzed for all distances with full datasets ($n$ = 100). Swimming times were downloaded from the USA Swimming Database (https://www.usaswimming.org /Home/times/data-hub) for six freestyle swimming distances from 50 to 1500 meters (50, 100, 200, 400, 800 and 1500 m) on April 4, 2019. Average swimming velocity ($m \cdot min^{-1}$) was calculated from the finishing time as: (race distance) × (finishing time)$^{-1}$. Sex differences in swimming velocity were calculated for each place and event distance as: [(boy's velocity)–(girl's velocity)] × (boy's velocity)$^{-1}$ × 100%. The reduction in swimming velocities of boys and girls across world record place (between 1$^{st}$ and 100$^{th}$ place) was calculated as: (velocity of $n^{th}$ place) × (velocity of 1$^{st}$ place)$^{-1}$ × 100%, for $n$ = 1 to 100. Participation data was accessed via publicly-available membership demographics reports prepared by the USA Swimming Member Services staff for 2015 to 2018 (https://www.usaswimming.org). Additionally, circulating testosterone concentrations of a nationally representative sample of the United States population were downloaded from the National Health and Nutrition Examination Survey (NHANES) coordinated and conducted by the Centers for Disease Control and Prevention (CDC) (https://wwwn.cdc.gov/nchs/nhanes/Search/DataPage.aspx?Compon ent = Laboratory). As described previously [22], testosterone was quantified via isotope dilution liquid chromatography tandem mass spectrometry (ID-LC-MS/MS) based on the National Institute for Standards and Technology's reference method, optimized by the CDC. This analytical quantification method initiated in 2013–2014 testing cycle, and data were analyzed for two consecutive testing cycles (2013–2014 and 2015–2016). These data are representative of the national population in demographic characteristics, and notably, are not specific to an elite-athletic population. All procedures accessed public information and did not require ethical review as determined by the Mayo Clinic Institutional Review Board in accordance with the Code of Federal Regulations, 45 CFR 46.102, and the *Declaration of Helsinki*.

### Statistical analysis

Data were reported as means ± SD within the text. Separate full factorial univariate analyses of variance (ANOVAs) were used to compare the dependent variables (swimming velocity and relative performance (%1$^{st}$ place) of boys and girls, and sex differences in swimming velocity) between three independent variables [age (5–18 years), US ranking (1$^{st}$-100$^{th}$) and event distance (50 m– 1500 m)]. *Post hoc* analyses (Tukey's HSD multiple comparisons) were used to test for differences between pairs within a data set when significant main effects or interactions were identified for age, US ranking or event distance. Recognizing early puberty may exhibit high statistical leverage on observed sex effects; a sensitivity analysis was conducted by filtering the data to only consist of the top 10$^{th}$ through 50$^{th}$ performance times by age, sex and distance. *Post hoc* Student's t-tests were used to test for differences between boys and girls when a significant interaction of sex was identified. Bonferroni corrected $p$-values for multiple comparisons ($p < 0.025$) were used for all *post hoc* analyses. Pearson correlation coefficients (r) were used to determine associations between the sex difference in swimming performance and average circulating testosterone concentrations. For all other analyses, significance was

determined at $p < 0.05$. All analyses were performed with IBM Statistical Package for Social Sciences version 25 statistical package (IBM, Armonk, NY, USA) and R version 3.4.2 (Vienna, Austria).

## Results

For boys and girls, swimming velocity improved with advancing age according to a quadratic growth curve that was reproducible for each swimming distance (Fig 1). The quadratic growth curve demonstrates rapid improvements in swimming velocity up to 10 years of age after which the age-related improvement in performance slows and approaches a plateau (horizontal asymptote). There were many distinct differences between the age-related performance enhancement curves between girls and boys however. The plateau of swimming velocity was 8.4% lower for girls than boys ($p<0.001$), and the age at which the plateau in performance initiated was younger for girls (15 years) than boys (17 years) for all swimming distances aggregated ($p<0.001$). These data indicate that boys had faster swimming performances than girls particularly at older ages, thus, there was a sex-related difference in swimming performance that increased with age ($p<0.001$).

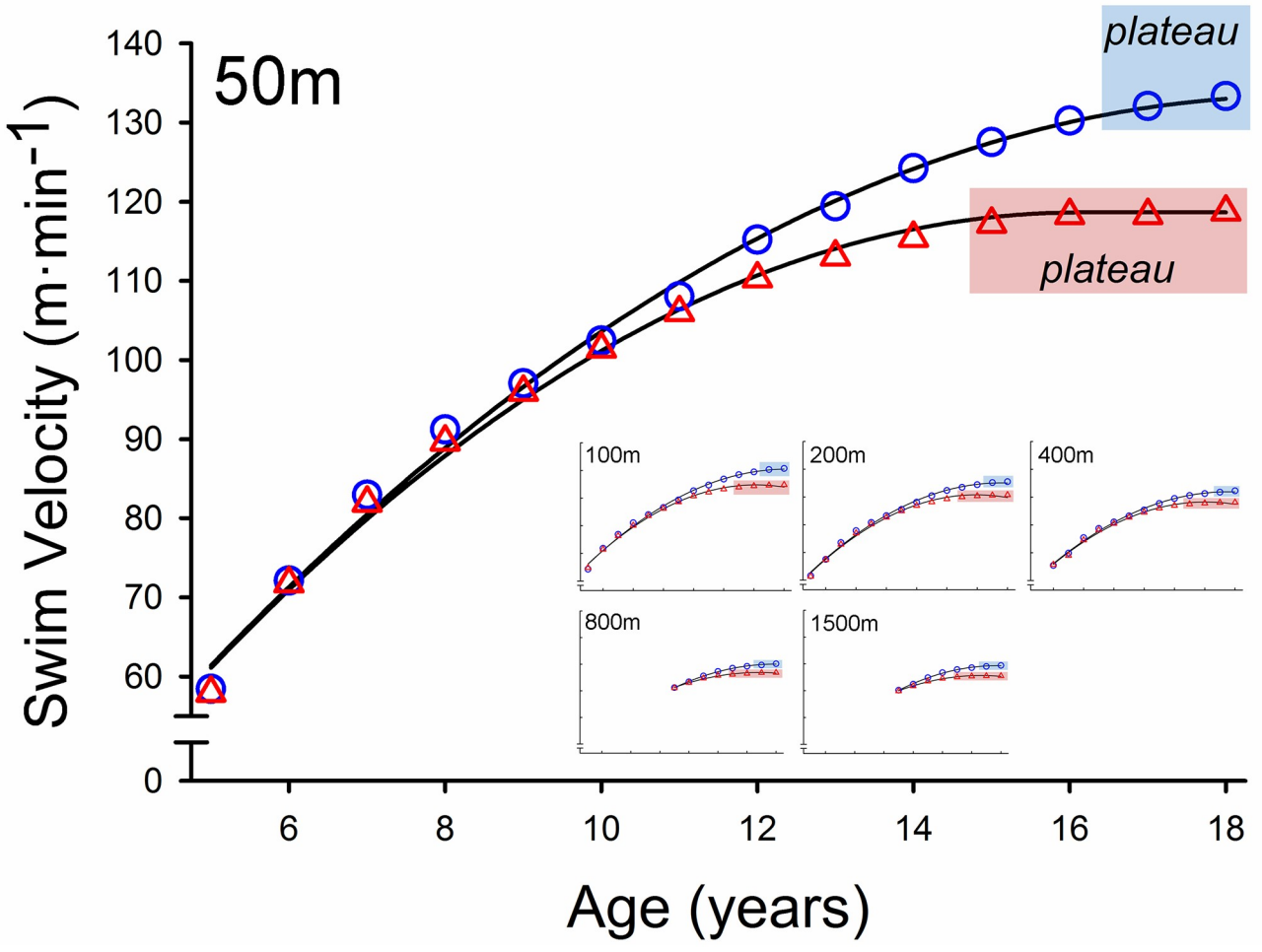



**Fig 1. Elite swimming performance.** Mean swimming velocity from 5 to 18 years of the top 100 fastest US boys (blue circles) and girls (red triangles) for the 50, 100, 200, 400, 800 and 1500m freestyle swimming distances.

https://doi.org/10.1371/journal.pone.0225724.g001

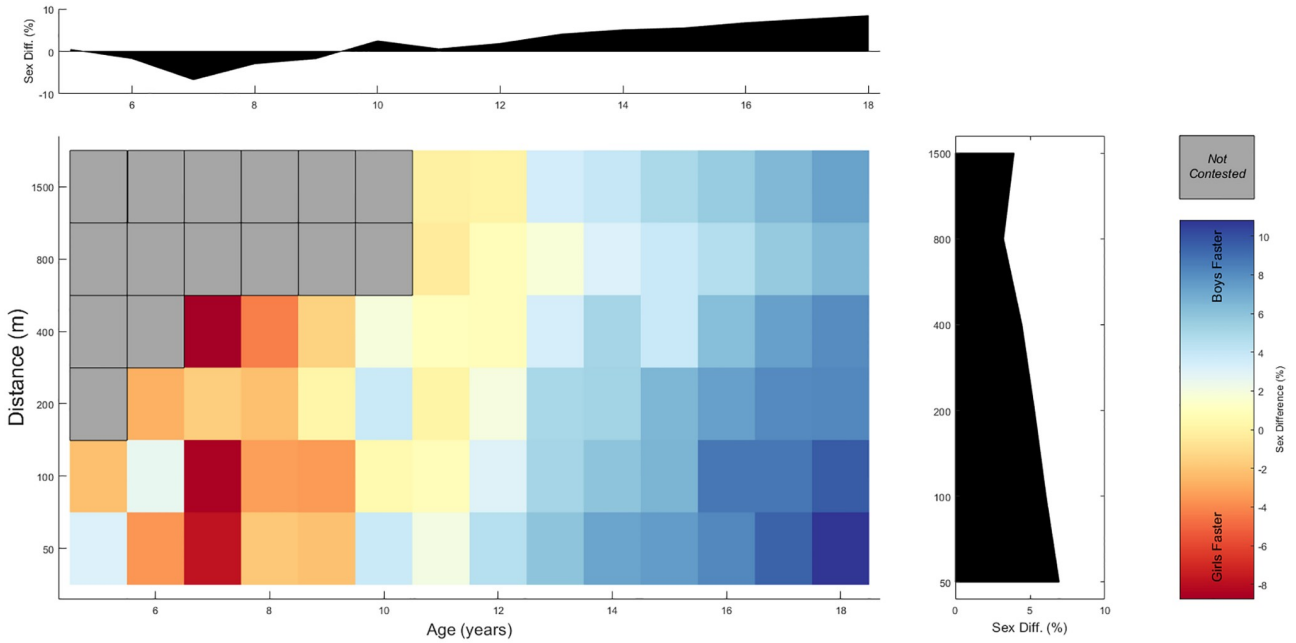



**Fig 2. Sex differences in performance of the top 5 places.** The primary plot (heat map) displays the sex differences in swimming velocity (% boy's swimming velocity) of the *top 5 US rankings* in each freestyle event distance and age, negative values (red) indicate faster performance of girls. The top displays the mean sex difference across age, and the right plot displays the mean sex difference across swimming event distance.

https://doi.org/10.1371/journal.pone.0225724.g002

Considering the most elite competitors in the top 5 places, girls had 3% faster swimming performance than boys prior to age 10 ($p<0.001$; Fig 2). However, considering the 10th-50th places, there were no sex-related differences in swimming performance between boys and girls prior to age 10 ($p>0.05$; Fig 3D). In the 50m for ages 5–9 years, for example, the average sex-related difference in performance was -2.5% for top 5 (indicating faster performance for girls)

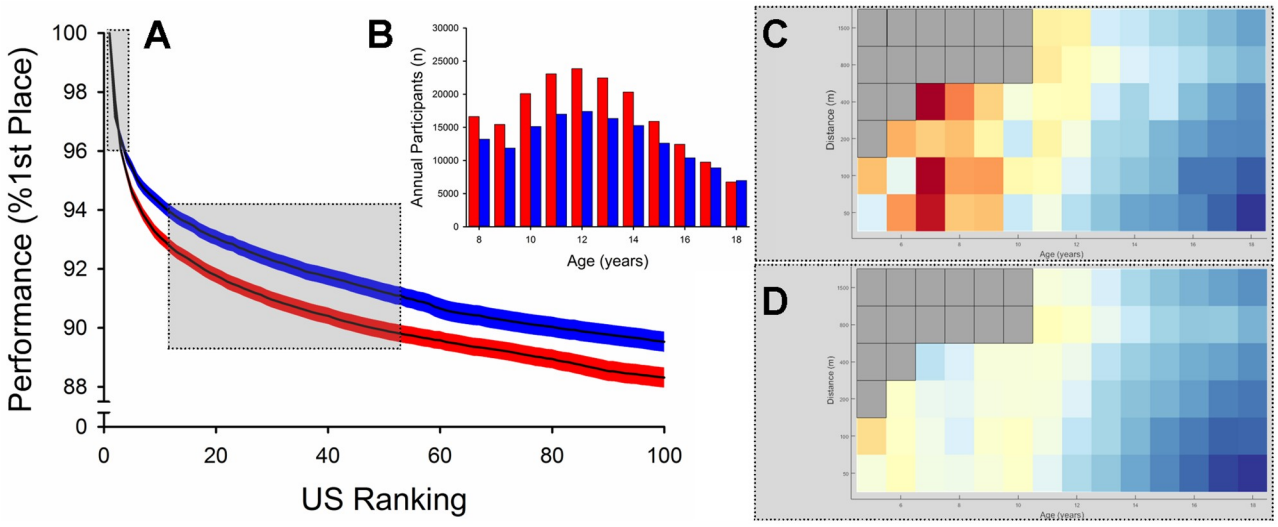



**Fig 3. Relative performance decline across US ranking.** The decline in swimming performance (% 1st Place) across US ranking for boys (blue) and girls (red; mean ± 95% confidence interval; Panel A). The average annual membership numbers (Panel B) for boys (blue) and girls (red) of USA Swimming. The heat maps (Panels C and D) display the sex differences in swimming velocity of the top 5 US Rankings (Panel C) and the 10th-50th US Rankings (Panel D) using the same color values displayed in Fig 2.

https://doi.org/10.1371/journal.pone.0225724.g003

and 1.2% (indicating faster performance for boys) for the 10th-50th places. Importantly, for both analyses, boys do not exhibit statistically faster swimming velocities than girls at young ages (<10 years). For both the top 5 and the 10th-50th places, pairwise comparisons indicated that the sex difference in performance increased from age 10 (top 5, boys 2.5% faster; 10th-50th places, boys 1.0% faster) incrementally increased for each age until the sex difference plateaued at age 17 (top 5, boys 7.6% faster; 10th-50th places, boys 8.0% faster). Thus, beginning at age 10, boys had faster swimming performance than girls and the sex-difference in performance plateaued at age 17 (Fig 3C & 3D). For top 100 places aggregated, the sex-related difference in performance of 17–18 year olds was larger for the sprint distance events (9.6%; 50-200m) compared with the endurance distance events (7.1%; 400-1500m; $p<0.001$). The larger sex-related differences in performance for sprint distance events were observed for all ages (Fig 2).

Comparison of the relative reductions in velocity between the 1st and 100th place (age groups and event distances aggregated) demonstrated that the girls had greater reductions in relative velocity across place than boys (Fig 3A; sex × place, $p<0.001$). The average 100th place US record holder swam at 90.7 ± 8.8% the velocity of the first place US record holder for the boys and 89.3 ± 9.0% for the girls (age groups and distances pooled). Thus, the sex difference in swimming performance progressively increased with US record place between first place (boys 2.4% faster) to 100th place (boys 4.3% faster) across all ages and distances ($p<0.001$). These data indicate that there was less depth of performance in girls than boys (Fig 3). Despite the lesser depth of performance in girls, annual participation was higher in girls compared to boys ($p<0.001$; Fig 3B).

Circulating testosterone concentration data comprised results from 2,085 measurements. Boys had a greater than 100-fold increase in serum testosterone from ages 6 to 18 (3.6 ± 16.4 to 482.0 ± 232.0 ng·dL$^{-1}$, $p<0.001$), and this testosterone level began to plateau at 16 years. During the years of low testosterone for boys (<10 ng·dL$^{-1}$)—6 to 10 years—there was no association between testosterone ($p = 0.500$). However, during the years of rapidly increasing testosterone levels for boys—11 to 17 years—mean testosterone was strongly, linearly correlated with the mean sex difference in swimming performance (pooled for all race distances and places; $p<0.001$, $r = 0.990$). See Fig 4.

## Discussion

Using 'big data' as a proxy to estimate the ergogenic advantage of androgens in boys compared to girls, we determined the age of the sex-related divergence in elite swimming performance. As expected, boys had faster swimming performance than girls at 18 years in sprint and endurance distances. Participation data provides evidence that there were equal opportunities to participate in swimming between boys and girls (Fig 3B), thus, we propose that the observed mean sex difference in performance across all freestyle events (8.4%) is *solely* due to physiological differences between the sexes. In support of this, the sex difference in world record swimming performances is similar (8.5%) until ~50 years [13]. Between the ages of 11 and 17 years, the sex difference in performance was strongly associated with circulating testosterone concentrations of boys from a nationally-representative sample ($r = 0.990$, $r^2 = 0.980$). These data suggest that endogenous testosterone explains 98% of the variance of the sex difference in performance, and support the previous assertion that the sex difference in circulating testosterone of adults explains most of the sex difference in sporting performance [6].

However, prior to the ergogenic effects of puberty/androgen hormones, there are no sex-related differences in performance for the 10th-50th places and the top 5 girls have faster performances than the top 5 boys. Importantly, the faster performance of the top girls is clearly not due to earlier initiation of puberty because girls are faster at 5 years of age, well before the age

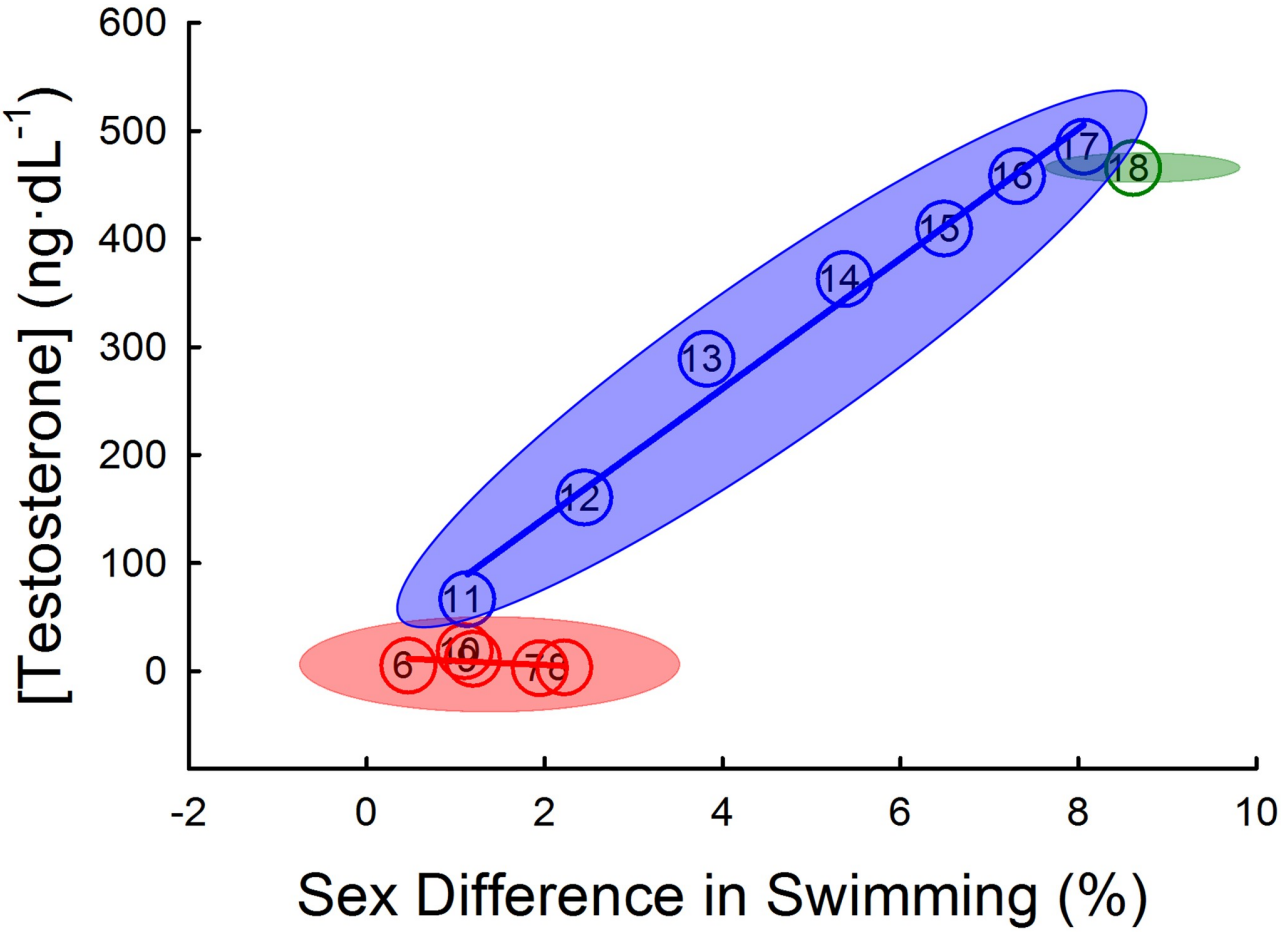



**Fig 4. Correlation of boy's serum testosterone and sex differences in swimming performance across age.** The mean circulating testosterone concentrations from NHANES database were strongly, linearly correlated with the mean sex difference in swimming performance during the years of rapidly increasing testosterone levels for boys (11 to 17 years; $p<0.001$, $r = 0.990$), but not during the years of low testosterone for boys (6 to 10 years; $p = 0.500$). Each circle represents the mean sex difference in swimming for each age group pooled for all race distances and places (x-axis) and mean circulating testosterone level for boys (y-axis) with the age group denoted using corresponding Arabic numeral within each circle. The colored ellipses represent the standard error of three separate groupings for the correlation analysis, low-testosterone group (6–10 years, red), increasing testosterone group (11–17 years, blue) and plateaued testosterone group (18 years, green).

https://doi.org/10.1371/journal.pone.0225724.g004

of puberty. Although the precise mechanisms are unclear, these data suggest that girls are inherently faster swimmers than boys (or at least not slower) than boys *sans* the performance-enhancing effects of androgens and puberty. As expected before puberty, there are minimal sex differences between boys and girls in stature [23], hand grip strength [5] and hemoglobin content [24]. Thus, if girls are inherently faster than boys prior to puberty, these sex-based differences would likely be due to optimized composition of the genes encoded on the X chromosome (e.g. genes associated with regulation of blood pressure, angiotensin-related enzymes) in girls with two copies of the X chromosome than boys with only 1 copy of the X chromosome [25]. After puberty however, our data and previous data [5, 6] suggest that ~15× greater concentrations of androgens [21] and subsequent physiological changes in boys compared to girls account for the ~8.5% sex-related difference in performance [6].

These data also demonstrate greater participation in swimming for girls than boys, however, this sex-difference in participation narrows with advancing age (Fig 3B). Limited experimental data exist to explain the greater participation among girls, however, the leading

hypothesis is that more girls participate because of longstanding opportunity to participate and compete with (and often outperform) boys. Interestingly, these data also show that despite greater participation for girls than boys, girls have less depth in performance. A prevailing hypothesis (*'sociocultural conditions hypothesis'* [17]) suggests that decreased opportunities and participation contribute to sex differences in sports performance. Indeed, in a previous study examining collegiate rowing, a sport sanctioned by the US National Collegiate Athletic Association (NCAA) for women but not men, greater participation for women was associated with greater depth of performance for women in the heavyweight class [17]. Thus it is unclear why in this study; girls have greater participation and less depth of performance. However it is clear that these data provide one of the only examples of faster (or at least not slower) sports performance for girls than boys.

## Conclusion

We conclude that prior to the performance-enhancing effects of puberty; the best girls outperform the best boys at sprint and endurance swimming events. Our findings are in direct opposition to nearly universal findings in elite adult athletes that boys are faster than girls. These data provide evidence that the Y chromosome *per se* does not provide an advantage in sports performance. Rather, our data are consistent with 'doping' ideology and findings that sustained and augmented levels of endogenous androgens induce performance-enhancing adaptations regardless of genotype of the sex chromosomes. This information may be of use to governing bodies of athletic competitions as eligibility regulations for participation in female events are refined.

## Acknowledgments

We thank USA Swimming for publicly available archives of swimming performances and demographics reports prepared by the USA Swimming Member Services staff.

## Author Contributions

**Conceptualization:** Jonathon W. Senefeld, Andrew J. Clayburn, Sarah E. Baker, Rickey E. Carter, Michael J. Joyner.

**Data curation:** Jonathon W. Senefeld, Andrew J. Clayburn.

**Formal analysis:** Jonathon W. Senefeld, Rickey E. Carter, Patrick W. Johnson.

**Writing – original draft:** Jonathon W. Senefeld, Andrew J. Clayburn, Sarah E. Baker, Rickey E. Carter, Michael J. Joyner.

**Writing – review & editing:** Jonathon W. Senefeld, Andrew J. Clayburn, Sarah E. Baker, Rickey E. Carter, Patrick W. Johnson, Michael J. Joyner.

## References

1. CAS 2018/O/5794 Mokgadi Caster Semenya v. International Association of Athletics Federations: Hearing before the Court of Arbitration for Sport(2018).
2. Saudan C, Baume N, Robinson N, Avois L, Mangin P, Saugy M. Testosterone and doping control. British journal of sports medicine. 2006; 40 Suppl 1:i21–4.
3. Wood RI, Stanton SJ. Testosterone and sport: current perspectives. Hormones and behavior. 2012; 61 (1):147–55. https://doi.org/10.1016/j.yhbeh.2011.09.010 PMID: 21983229
4. Franke WW, Berendonk B. Hormonal doping and androgenization of athletes: a secret program of the German Democratic Republic government. Clin Chem. 1997; 43(7):1262–79. PMID: 9216474

5. Handelsman DJ. Sex differences in athletic performance emerge coinciding with the onset of male puberty. Clinical endocrinology. 2017; 87(1):68–72. https://doi.org/10.1111/cen.13350 PMID: 28397355
6. 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 PMID: 30010735
7. Joyner MJ. Physiological limiting factors and distance running: influence of gender and age on record performances. Exercise and sport sciences reviews. 1993; 21:103–33. PMID: 8504840
8. Cheuvront SN, Carter R, Deruisseau KC, Moffatt RJ. Running performance differences between men and women:an update. Sports medicine. 2005; 35(12):1017–24. https://doi.org/10.2165/00007256-200535120-00002 PMID: 16336006
9. 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. Ann Clin Biochem. 2016; 53(Pt 3):377–84. https://doi.org/10.1177/0004563215610589 PMID: 26438522
10. Vesper HW, Wang Y, Vidal M, Botelho JC, Caudill SP. Serum Total Testosterone Concentrations in the US Household Population from the NHANES 2011–2012 Study Population. Clin Chem. 2015; 61 (12):1495–504. https://doi.org/10.1373/clinchem.2015.245969 PMID: 26510959
11. Joyner MJ, Coyle EF. Endurance exercise performance: the physiology of champions. The Journal of physiology. 2008; 586(1):35–44. https://doi.org/10.1113/jphysiol.2007.143834 PMID: 17901124
12. Senefeld J, Smith C, Hunter SK. Sex Differences in Participation, Performance, and Age of Ultramarathon Runners. International journal of sports physiology and performance. 2016; 11(7):635–42. https://doi.org/10.1123/ijspp.2015-0418 PMID: 26561864
13. Senefeld J, Joyner MJ, Stevens A, Hunter SK. Sex differences in elite swimming with advanced age are less than marathon running. Scandinavian journal of medicine & science in sports. 2016; 26(1):17–28.
14. Hunter SK, Stevens AA. Sex differences in marathon running with advanced age: physiology or participation? Medicine and science in sports and exercise. 2013; 45(1):148–56. https://doi.org/10.1249/MSS.0b013e31826900f6 PMID: 22843112
15. Tanaka H, Seals DR. Age and gender interactions in physiological functional capacity: insight from swimming performance. Journal of applied physiology. 1997; 82(3):846–51. https://doi.org/10.1152/jappl.1997.82.3.846 PMID: 9074973
16. Tanaka H, Seals DR. Endurance exercise performance in Masters athletes: age-associated changes and underlying physiological mechanisms. The Journal of physiology. 2008; 586(1):55–63. https://doi.org/10.1113/jphysiol.2007.141879 PMID: 17717011
17. Keenan KG, Senefeld JW, Hunter SK. Girls in the boat: Sex differences in rowing performance and participation. PloS one. 2018; 13(1):e0191504. https://doi.org/10.1371/journal.pone.0191504 PMID: 29352279
18. 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
19. Hunter SK, Stevens AA, Magennis K, Skelton KW, Fauth M. Is there a sex difference in the age of elite marathon runners? Medicine and science in sports and exercise. 2011; 43(4):656–64. https://doi.org/10.1249/MSS.0b013e3181fb4e00 PMID: 20881885
20. Dukes RL. Parental commitment to competitive swimming. Free inquiry in creative sociology. 2002; 30 (2):185–98.
21. Buttler RM, Peper JS, Crone EA, Lentjes EGW, Blankenstein MA, Heijboer AC. Reference values for salivary testosterone in adolescent boys and girls determined using Isotope-Dilution Liquid-Chromatography Tandem Mass Spectrometry (ID-LC-MS/MS). Clin Chim Acta. 2016; 456:15–8. https://doi.org/10.1016/j.cca.2016.02.015 PMID: 26920638
22. Zhou H, Wang Y, Gatcombe M, Farris J, Botelho JC, Caudill SP, 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):5943–54. https://doi.org/10.1007/s00216-017-0529-x PMID: 28801832
23. Abbassi V. Growth and normal puberty. Pediatrics. 1998; 102(2 Pt 3):507–11. PMID: 9685454
24. Hawkins WW, Kline DK. Hemoglobin levels among 7 to 14 year old children in Saskatoon, Canada. Blood. 1950; 5(3):278–85. PMID: 15404069
25. Araujo FC, Milsted A, Watanabe IK, Del Puerto HL, Santos RA, Lazar J, et al. Similarities and differences of X and Y chromosome homologous genes, SRY and SOX3, in regulating the renin-angiotensin system promoters. Physiol Genomics. 2015; 47(5):177–86. https://doi.org/10.1152/physiolgenomics.00138.2014 PMID: 25759379

# Exhibit 21

# Race Times for Transgender Athletes

**Article** *in* Journal of Sporting Cultures and Identities · January 2015

DOI: 10.18848/2381-6678/CGP/v06i01/54079

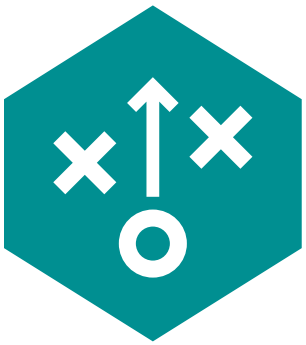
CITATIONS
51

READS
10,275

**1 author:**

Joanna Marie Harper
Loughborough University
**10** PUBLICATIONS **207** CITATIONS

SEE PROFILE



VOLUME 6 ISSUE 1

Journal of

# Sporting Cultures and Identities

## Race Times for Transgender Athletes



JOANNA HARPER

COMMON GROUND

SPORTANDSOCIETY.COM

**JOURNAL OF SPORTING CULTURES AND IDENTITIES**
http://sportandsociety.com
ISSN: 2381-6678 (Print)
ISSN: 2381-6694 (Online)
http://doi.org/10.18848/2381-6678/CGP (Journal)

First published by Common Ground Research Networks in 2015
University of Illinois Research Park
2001 South First Street, Suite 202
Champaign, IL 61820 USA
Ph: +1-217-328-0405
http://cgnetworks.org

*The International Journal of Sport and Society*
is a peer-reviewed, scholarly journal.

# Race Times for Transgender Athletes

Joanna Harper, Providence Portland Medical Center, USA

*Abstract: In recent years, organizations such as the International Olympic Committee have created regulations to allow those athletes who have undergone gender reassignment to compete in their chosen gender. Despite these rules, there is still a widespread belief that transgender female athletes have an inherent advantage over 46,XX female competitors. Until now, there has not been any published data, based on performances of transgender athletes, to either support or refute this belief. There are two main stumbling blocks to creating such a study: the first is to determine an appropriate metric to examine and the second is to find participants for the study. This study analyzed race times for eight transgender female runners, who have competed in distance races as both male and female, using a mathematical model called age grading. Collectively, the age graded scores for these eight runners are the same in both genders.*

*Keywords: Transgender, Athletes, Distance Running, Gender, Research*

## Introduction

Athletes have historically been divided into male and female for the purpose of most sporting competitions. Two components of biological sex, first external genitalia, and later chromosomes were used to make the determination of who was allowed to compete in women's sport. Chromosome testing was initiated for the 1968 Olympics (Elsas et al 2000, 249-254) and thereafter, only those people with XX sex chromosomes among their 23 chromosome pairs, or 46,XX females, were allowed into women's sports. Human biology, however, does not neatly divide into two categories. For instance, some people have neither a 46,XY nor a 46,XX karyotype. Additionally, some people are born with a 46,XY pattern, but with mutations which cause them to be assigned female gender at birth. Chromosome based requirements for participation in female athletics were discontinued in the 1990s (Elsas et al, 249-254), but controversy surrounding athletes with karyotypes other than 46,XX competing in women's sport continues. (Karkazis et al 2012, 3-16).

Transgender people are those whose innate sense of gender, or gender identity, does not match their biological sex. Some transgender people seek gender reassignment. Such people have been termed transsexual, and although the term is descriptive, it is now often viewed unfavorably within the transgender community. While transgender surgery can alter external and internal genitalia, and hormone therapy changes many secondary sex characteristics, neither can alter karyotype; hence it is questionable whether one could claim a change in sex as a result of any intervention. Unambiguous reassignment of gender is, however, possible.

Those who are satisfied with the gender assigned to them at birth can be described as cisgender.

Transgender athletes have sought to compete against other athletes on the basis of their reassigned gender, rather than on their biological sex. While there has been little resistance to the presence of transgender male athletes, sporting organizations were unwilling to allow transgender women to compete against 46,XX women prior to the 21st century. It is notable that in the 1970s, Rene Richards, probably the best-known transgender athlete in history, sued in the United States court system in order to be allowed to play women's tennis (Abrams 2010).

In 2004, the International Olympic Committee (IOC) enacted the Stockholm Consensus (Ljungqvist et al., 2003), that allows transgender women to compete in women's sport once a) gender reassignment surgery had been completed, b) the athlete was legally recognized as female, and c) they had undergone two years of hormone replacement therapy. Transgender men were permitted to compete against cisgender men, although transgender men must file a therapeutic use exception (TUE) form to cover their use of testosterone injections.

*Journal of Sporting Cultures and Identities*
Volume 6, Issue 1, 2015, www.sportandsociety.com, ISSN: 2381-6678

https://doi.org/10.18848/2381-6678/CGP/v06i01/54079

At the time of the Stockholm Consensus, there was no published scientific literature that would justify the inclusion of transgender women. The committee that created the Stockholm consensus relied heavily on information from Dr. Louis Gooren from Amsterdam (Ljungqvist 2104). Dr. Gooren was an expert in transgender studies and would go on to co-author an important paper which studied nineteen transgender women after commencement of hormone therapy (Gooren and Bunck, 2004, 425-429). After one year of testosterone suppression, the subjects had testosterone levels below those of 46,XX women, and hemoglobin levels equal to those of 46,XX women (red blood cell content is very important in endurance sports). Muscle mass differences between the two groups were cut in half. The height of the individuals did not change. There were no additional changes noted at three years. This study was not undertaken on athletes, nor did the researchers directly measure any physical component of athleticism, such as strength, speed, explosiveness, or endurance. The authors concluded that it was reasonable to allow transgender women to compete against cisgender women after appropriate hormone therapy.

It is notable that the Stockholm consensus required two years of hormone therapy, while the published study noted that there were no physical changes in the subjects after one year. This discrepancy was due to conservative estimates given to the committee by Dr Gooren prior to the publication of his study (Ljungqvist 2014).

Many sports followed the lead of the IOC, and in subsequent years there have been transgender women competing in sports such as golf (Mianne Bagger and Lana Lawless), cycling (Natalie Van Gogh, Michelle Dumaresq, and Kristin Worley), martial arts (Nong Toom, and Fallon Fox), and basketball (Gabrielle Ludwig). None of these women has been particularly successful at the highest levels of sport after gender reassignment, and one could argue that this lack of success over ten years would be a strong indication of the fairness of permitting transgender women to compete against cisgender women.

Instead of acceptance, however, there has been a substantial amount of controversy over the presence of transgender women in female athletics. Most people contend that transgender women have an unfair advantage in women's competition (Cavanagh and Sykes, 2006). Opponents of transgender inclusion often point to physical characteristics such as height and hand size, which are not changed by gender reassignment, and suggest that transgender women will always maintain an unfair advantage over cisgender women. These arguments continue today and are not confined to competition at the highest levels. Recently, there were 10,000 emails sent in to protest the decision by the State of Minnesota to allow high school transgender athletes to compete in their chosen gender (Minnesota Star Tribune 2014).

Those in favor of allowing transgender athletic participation inevitably point to the fact that every major sporting organization to look at the issue since 2004 has agreed to allow transgender women to compete against other women. Proponents also will often suggest that science is on their side. However, the only existing published study related to transgender women in sport is the original one by Gooren and Bunk. The science supporting transgender inclusion is very thin indeed.

A thorough literature review of studies applicable to transgender athletes was undertaken for the Canadian Government (Devries, 2008). This review found that "To date no study has conducted any sort of exercise test to assess athletic performance" and concluded that there were no data indicating any sporting advantage or disadvantage for transgender women as compared to over 46,XX women.

The lack of such a study should not come as a surprise. There are two major obstacles involved in compiling any study involving transgender athletes. The first problem is how to formulate a study to create a meaningful measurement of athletic performance, both before and after testosterone suppression. No methodology has been previously devised to make meaningful measurements.

The second problem is to find study participants. There are few transgender athletes, and even fewer who will want to be identified. In order to create a study, a small cohort of competitive transgender athletes must be found in one given sport. Fortunately, there is mass participation in distance running races throughout much of world. All major cities hold road races with many thousands of runners, giving the sport a large base of adult competitors. Thus, the sport of distance running is an obvious choice to try to find suitable candidates.

In 2011, the international governing body for track and field, the IAAF, amended its rules to allow anyone who was legally and hormonally female to compete in the women's category (IAAF, 2011). The portion of the ruling applicable to transgender women lists no requirement for surgical intervention, or specific duration of hormone therapy. It does require an endocrine evaluation prior to any declaration of eligibility. In many parts of the world, legal gender reassignment is not allowed, and this will be a barrier to participation for many transgender athletes.

In 2012, the IOC also adopted a testosterone-based rule for eligibility for women's sport (IOC, 2012); however, the IOC maintained their previous rules pertaining to transgender women. Hence, it would be possible for a transgender woman to compete against other women in the IAAF sponsored 2015 world track championships, but not be eligible to do so in the IOC-sponsored 2016 Olympics.

## Methods

Race times from eight transgender women runners were collected over a period of seven years and, when possible, verified. The collection process consisted of seeking out female transgender distance runners, mostly online, and then asking them to submit race times. Even in 2014 few people are open about being transgender, so the submission of race times represented a large leap of faith for the participants. When possible, race times were then verified using online services listing race results. For six of the eight runners, online checking made it possible to verify approximately half of the submitted times. Two of the subjects, runners three and four, would only participate anonymously, creating an ethical dilemma over the use of their times, versus respect their privacy.

Seven of the eight subjects experienced a substantial reduction in running speed upon transition. There are a few methods of comparing men's and women's race times. The simplest involves the well-known approximation that men will, on average, run 10% faster than women (Berman et al. 2013 63–65). There are a couple of other comparison methods as well, but there is only one method that also factors in age. Correcting for age is important because most of the runners in the study were more than 30 years old, and would be faced with declining performance as they grew older, following their gender transition.

Age grading (Grubb, 1998, 509-521) is a method of comparing the performance of athletes of all ages and both sexes. For running events, the athlete's race time (RT) is compared to the fastest time ever run by a person of that age and sex, or the age standard (AS). The resultant age grade (AG) percentage is obtained by the following formula:

$$AG\ (\%) = (AS \times 100) / RT$$

All times are measured in seconds.

In order to understand how age grading works, let's examine two forty-year-old runners who run a 5-kilometer race (5k). The male runner runs 19:30 (1170 seconds). In order to determine his age grade, one compares his time to the fastest time ever run by a forty-year-old male 5k runner, i.e. 13:39 (819 seconds). The equation becomes

$$AG = (819 \text{ seconds} \times 100)/1170 \text{ seconds} = 70$$

and our male runner gets a score of 70.

The female runner has a time of 21:51 (1311 seconds) and her time is compared to the fastest ever time by a forty-year-old woman, i.e. 15:18 (918 seconds). The equation for her AG is

AG = (918 seconds x 100) /1311 seconds = 70

Thus, our male and female runners score the same age grade despite the fact that the male ran more than two minutes faster than the female did. This is fair. Men run faster than women. The two runners are both well above average runners for their age and sex, and deserve to receive equal accolades.

Age grading has become the standard way of comparing performances by older track and field athletes of both sexes. The method has also been rigorously evaluated and improved, specifically with regard to the curve fitting that is needed to connect the age standards associated with different ages. Mathematician Alan Jones (Jones 2010) has made significant improvements to the age-graded tables that Howard Grubb developed in the 1990s.

## Results

Collectively, the eight runners had much slower race times in the female gender than as males. Time differences were, in fact, so great, that age graded performances stayed virtually constant for the group. Tables one through four summarize the data from all eight runners over four frequently run race distances varying from 5k to the marathon (42 kilometers). Not all eight women submitted times for all four of these distances.

Table 1: 5k Race Times

| | Male | Races | | Female | Races | |
|---|---|---|---|---|---|---|
| **Runner No.** | **Age** | **Time** | **AG** | **Age** | **Time** | **AG** |
| **One** | 48 | 18:27 | 78.7 | 52 | 22:43 | 75.7 |
| **Two** | 30 | 15:56 | 81.4 | 36 | 17:51 | 82 |
| **Four (a)** | 30 | 17:35 | 73.6 | 33 | 21:04 | 70.6 |
| **Five** | 34 | 19:39 | 66.7 | 35 | 23:43 | 63 |
| **Six (b)** | 24 | 15:07 | 83.5 | 53 | 20:22 | 85.5 |
| **Eight** | 27 | 20:29 | 62.2 | 30 | 22:51 | 64.8 |

Table 2: 10k Race Times

| | Male | Races | | Female | Races | |
|---|---|---|---|---|---|---|
| **Runner No.** | **Age** | **Time** | **AG** | **Age** | **Time** | **AG** |
| **One** | 49 | 0:39:05 | 77.9 | 56 | 0:48:45 | 76.1 |
| **Two (b)** | 22 | 0:32:37 | 82.4 | 36 | 0:36:58 | 83.1 |
| **Five** | 34 | 0:45:33 | 60.1 | 36 | 0:57:40 | 53.3 |
| **Six (a)** | 46 | 0:37:10 | 80 | 48 | 0:42:01 | 80.5 |

Table 3: Half-marathon Race Times

| | Male | Races | | Female | Races | |
|---|---|---|---|---|---|---|
| **Runner No.** | **Age** | **Time** | **AG** | **Age** | **Time** | **AG** |
| **Five** | 33 | 1:53:06 | 52.4 | 37 | 2:05:38 | 53.3 |
| **Six (b) (d)** | 26 | 1:08:38 | 86.3 | 53 | 1:32:27 | 83.8 |
| **Six (a) (d)** | 46 | 1:23:11 | 77.8 | 48 | 1:34:01 | 77.5 |
| **Seven (c)** | 19 | 1:48:47 | 55.7 | 28 | 1:48:45 | 60.5 |

Table 4: Marathon Race Times

| | Male | Races | | Female | Races | |
|---|---|---|---|---|---|---|
| **Runner No.** | **Age** | **Time** | **AG** | **Age** | **Time** | **AG** |
| **Three** | 49 | 3:18:58 | 69.5 | 54 | 4:12:31 | 67.2 |
| **Five** | 34 | 3:16:59 | 63.4 | 35 | 4:08:33 | 55.3 |
| **Seven (c)** | 19 | 3:49:55 | 55.7 | 31 | 2:59:10 | 75.7 |
| **Eight** | 29 | 3:08:53 | 66.1 | 30 | 3:44:55 | 60.2 |

## Notes

(a) These races were run over the same course within three years' time and represent the best comparison points.

(b) Races compared over a long time period have more uncertainty associated with them, but both runner two and runner six reported stable training patterns over this time range. These races also help to confirm the age-grading methodology for tracking progress of a runner over the course of a multi-year time frame.

(c) Runner seven represented the biggest evaluation challenge. She raced as a 19 year-old male recreational runner and then resumed running years later as a female. She got serious about the sport after she resumed, doubled her training load and dropped 10 kg of weight. Not surprisingly, she got faster. This improvement can be seen in the fact that her AG went from 60.5 at age 28 to 75.7 at age 31 (both in female gender). This 15 point change in age grade was much larger than the 5-point change she experienced after transition from male to female.

(d) It is useful to compare times for the same runner over different race courses and at different time periods. The two lower scores occurred on a hilly course at a period of average fitness for runner six. The two higher scores were on flat courses at times of peak fitness.

Table five indicates the average AGs from all eight runners in each gender and the overall averages of all eight.

Table six shows the highest AGs from each runner and the average of these highest AGs. Two tailed t tests were run on both the mean and peak AGs. The p values were p=0.84 for the average AGs and p=0.68 for the highest AG. A p value of less than 0.05 is needed for the values to be considered significantly different, and these p values are very much higher.

Table 5: Average Age Grades

| | Average male AG | Average female AG |
|---|---|---|
| Runner 1 | 75.2 | 77.1 |
| Runner 2 | 81.8 | 82.8 |
| Runner 3 | 69.5 | 70.8 |
| Runner 4 | 71.4 | 64.8 |
| Runner 5 | 57.7 | 49.3 |
| Runner 6 | 83.8 | 81.9 |
| Runner 7 | 55.7 | 61.9 (e) |
| Runner 8 | 54.3 | 59.1 |
| **Average** | 68.7 | 68.5 |

Table 6: Highest Age Grade

| | Highest male AG | Highest female AG |
|---|---|---|
| Runner 1 | 78.7 | 79.2 |
| Runner 2 | 82.9 | 83.2 |
| Runner 3 | 69.5 | 74.3 |
| Runner 4 | 74.1 | 74.1 |
| Runner 5 | 66.7 | 63.0 |
| Runner 6 | 87.5 | 85.6 |
| Runner 7 | 55.7 | 63.4 (e) |
| Runner 8 | 66.1 | 64.8 |
| **Average** | 72.7 | 73.4 |

(e) The 2:59 marathon time by runner seven was considered an outlier, the result of her substantially altered training and was not used in these tables.

## Discussion

The majority of scientists believe that testosterone is primarily responsible for the difference in athletic results between the sexes (Bermon et al. 2014, 4328–4335), although there are dissenters (Healy et al. 2014, 294-303). There have been multiple studies on men's and women's testosterone levels with some variation in results, but a typical set of values would be as follows: Men's range — 10 to 35 nmol/l; female range — 0.35 to 2.0 nmol/l (Haring et al. 2012, 408–415).

Transgender women who have undertaken testosterone suppression change from normal male testosterone levels to normal female levels, in fact, after surgery their testosterone levels are below the mean for 46,XX women (Gooren and Bunck, 425–429). Largely as a result of their vastly reduced testosterone levels, transgender women lose strength, speed, and virtually every other component of athletic ability.

Since this study looks at endurance capabilities of athletes both pre and post testosterone suppression, it is also of significant interest to look at hematocrit or hemoglobin levels of transgender women. One year after testosterone suppression, hemoglobin levels in transgender women fell from 9.3 mmol/l to 8.0 mmol/l. This latter number is statistically identical to the mean hemoglobin level for cisgender women (Gooren and Bunck 425–429).

The reduction of testosterone and hemoglobin levels of transgender women after transition would suggest that endurance capabilities of transgender women athletes should be similar to those of 46,XX women.

The difficulty of finding suitable subjects is underscored by the fact that it took seven years to amass data from eight participants.

The times submitted by the eight runners were self-selected and self-reported. The self-reporting by the subjects certainly affects the strength of the findings. As mentioned previously, almost half of the race times were double checked by the author for accuracy. None of the subjects incorrectly reported any result.

Collectively, the eight runners were much slower in the female gender; slow enough, in fact, that their age graded performances were almost identical to their male AGs. Two of the runners had higher average AGs in male gender than in female gender, while one runner had higher female AGs than male ones. The changes in the age grades of these runners mirrored changes in their training habits.

After transition, runner four began to experience a significant number of injuries which prevented her from training as rigorously as she previously had. It is not surprising that her results got worse as time went on. Runner five experienced both weight gain and a loss

motivation in the years after her transition. In fact her motivation declined to the point that she gave up racing not long after the submission of her results.

On the other hand, runner seven blossomed as a runner after transition. Eventually, she doubled her weekly training distance. She also lost approximately 10 kg of body mass after she started to train harder. It is not surprising that her times and age grade scores showed a subsequent improvement.

The other runners in the study reported relatively stable training loads in both male and female mode, and this is reflected in their more stable age grades in both genders.

Since training loads vary over time for all runners, the author believes that highest age grade might be the best comparison of male versus female athletic potential. But, whether one uses average or highest age grades, the subjects scored statistically identical age grades both as male and as female.

It is significant to note that none of the eight subjects was a truly elite runner. An optimal study would use world-class runners and the results could be used to justify the presence of transgender women in events such as the Olympic Games. Unfortunately, there simply are no world-class transgender distance runners. Three of the eight runners have achieved notable success at the national level, and two of the other runners could be described as sub-elite. Resistance to the presence of transgender women occurs at all levels of sport, and so there is still much merit to the study.

One interesting trend was noted in runners five, six and eight, who age graded higher in shorter events as women than they did in longer events. Runners six and eight scored higher age grades in 5k races than they did as males but lower AGs in longer races – half marathon and up. Runner five scored lower across the board as female than as male but her 5k AGs were much closer to her male ones, than her marathon AGs were. Transgender women carry more muscle mass than 46,XX women (Gooren and Bunck 2004, 425–429). This extra muscle mass might cause increased speed when compared to cisgender women, and hence faster times and higher AGs at shorter distances. Increased muscle mass and heavier bones are not conducive to long distance running, and would actually be a disadvantage when running distances of a half marathon and higher, causing slower times and lower AGs. This effect is small in the three mentioned runners, and none of the other five runners submitted data over a wide enough range of distances to determine whether or not this pattern held true for them; more research would be needed to confirm or refute the hypothesis of distance related variations in age grade scores for transgender women.

It should be noted that these results are only valid for distance running. Transgender women are taller and larger, on average, than 46,XX women (Gooren and Bunck, 2004, 425-429), and these differences probably would result in performance advantages in events in which height and strength are obvious precursors to success - events such as the shot put and the high jump. Conversely, transgender women will probably have a notable disadvantage in sports such as gymnastics, where greater size is an impediment to optimal performance.

The Grubb and Jones age-grading methodology applies only to track-and-field and distance running, but, it should be possible to create a similar analytic method to compare results for other sports, such as swimming, weightlifting, or ski-jumping, which also measure results in times, distances or weights – the so called CGS (centimeter, grams, and seconds) sports. It would be very difficult, however, to devise such a method to analyze performances in most other types of sports.

## Conclusions

Despite the fact that transgender women have been allowed to compete against cisgender ones since 2004, there has been no study used to justify this decision beyond the original work of Gooren and Bunck. It bears repeating that this original study was not undertaken on athletes, nor

did it directly measure any aspect of athleticism. In fact, this is the first time a study has been developed to measure the performance of transgender athletes. The author overcame two significant barriers which have prevented any previous study from being performed, i.e. the difficulty in determining an appropriate metric to measure athletic performance both before and after testosterone suppression, and the difficulty in finding enough willing study participants in any given sport.

The author chose to use the standard age-grading methodology which is commonly used in master's (over forty) track meets worldwide, to evaluate the performance of eight distance runners who had undergone gender transition from male to female. As a group, the eight study participants had remarkably similar age grade scores in both male and female gender, making it possible to state that transgender women run distance races at approximately the same level, for their respective gender, both before and after gender transition.

It should be noted that this conclusion only applies to distance running and the author makes no claims as to the equality of performances, pre and post gender transition, in any other sport. As such, the study cannot, unequivocally, state that it is fair to allow to transgender women to compete against 46,XX women in all sports, although the study does make a powerful statement in favor of such a position.

It should also probably be noted that the publication of this study will likely not appreciably change the resistance faced by transgender women who compete against cisgender ones. There will continue to be strong opposition by athletes, parents and fans to the inclusion of transgender women. It will take many more years before the average sports enthusiast understands that transgender women who have undergone testosterone suppression will not dominate women's sports.

## REFERENCES


Abrams, Roger I. "*Sports justice: the law and the business of sports* Boston, Mass".: Northeastern University Press, 2010

Bermon, S., Garnier, P.Y., Hirschberg, A.L., et al. 2014. "Serum Androgen Levels in Elite Female Athletes." J Clin Endocrinol Metab. Nov;99(11):4328-35.

Cavanagh, S.L., & Sykes, H. 2006. "Transsexual bodies at the Olympics: The International Olympic Committee's policy on transsexual athletes at the 2004 Athens Summer Games". *Body & Society, 12*, 75-102.

Devries, M. 2008. "Do Transitioned Athletes Compete at an Advantage or Disadvantage as compared with Physically Born Men and Women": A review of the Scientific Literature

Elsas LJ, Ljungqvist A, Ferguson-Smith MA, et al. 2000. "Gender verification of female athletes." Genet Med;2:249–54

Gooren L**,** and Bunck M. 2004. "Transsexuals and competitive sports." *European Journal of Endocrinology* 151: 425-429.

Grubb, H.J. 1998. "Models for Comparing Athletic Performances". The Statistician*,* 47, 509-521.

Haring, R., A. Hannemann, U. John, et al. 2012. "Age-specific reference ranges for serum testosterone and androstenedione concentrations in women measured by liquid chromatography–tandem mass spectrometry." *Journal of Clinical Endocrinology and Metabolism* 97(2): 408–415.

Healy, M.L., Gibney, R., Pentecost, C., Wheeler, M.J., and Sonksen, P. H. 2014 "Endocrine profiles in 693 elite athletes in the post competition setting." Clin Endocrinol (Oxf). Aug;81(2):294-305.

International Association of Athletics Federations 2011. "IAAF Regulations Governing Eligibility of Athletes who have Undergone Sex Reassignment to Compete in Women's Competition"

International Olympic Committee 2012. "IOC Regulations on Female Hyperandrogenism". www.olympic.org/Documents/Commissions_PDFfiles/Medical_commission/2012-06-22-IOC-Regulations-on-Female-Hyperandrogenism-eng.pdf

Jones, Alan. 2010. "Age Grading Running Races". http://home.roadrunner.com/~alanjones/AgeGrade.html

Karkazis, K., R. Jordan-Young, G. Davis, and S. Camporesi. 2012. "Out of bounds? A critique of the new policies on hyperandrogenism in elite female athletes." American Journal of Bioethics 12(7): 3–16.

Ljungqvist, A., Cohen-Haguenauer , O., Genel, M., Simpson, J., Ritzen, M., Fellous, M., Schamasch 2003. "Statement of the Stockholm consensus on sex reassignment in sports." www.olympic.org/Documents/Reports/EN/en_report_905.pdf

Ljungqvist, A., 2014. Private Communication. *Minnesota Star Tribune* http://www.startribune.com/opinion/commentaries/284678481.html

## ABOUT THE AUTHOR

***Joanna Harper:*** Medical Physicist, Radiation Oncology, Providence Portland Medical Center, Portland, Oregon, USA

***Journal of Sporting Culture and Identities*** is one of the four thematically focused journals that support the Sport and Society knowledge community—its journal collection, book series, and online community.

The journal focuses on sport's motivations, meanings and purposes. It publishes articles that examine the psychology of sport and the interplay between individual identities and access, equity, and participation in sports.

*Journal of Sporting Culture and Identities* is a peer-reviewed scholarly journal.

ISSN 2381-6678

# Exhibit 22

https://theconversation.com/when-it-comes-to-sport-boys-play-like-a-girl-80328

Marnee McKay

9 min read

# When it comes to sport, boys 'play like a girl'

Primary school-aged boys and girls can play in mixed teams until they reach high school, our research suggests. Clappstar/Flickr, CC BY-SA

Published: August 3, 2017 4.16pm EDT

## Authors

1. Marnee McKay

   Lecturer of Musculoskeletal Physiotherapy, University of Sydney

2. Joshua Burns

   Professor, Sydney Children's Hospitals Network, University of Sydney

## Disclosure statement

Joshua Burns receives funding from NIH (National Institute of Neurological Diseases and Stroke and National Center for Advancing Translational Sciences, Inherited Neuropathies Consortium, Rare Disease Clinical Research Network #2U54NS065712), Charcot-Marie Tooth Association of Australia, Charcot-Marie Tooth Association (USA), Diabetes Australia, Multiple Sclerosis Research Australia, Sydney Southeast Asia Centre, New Zealand Neuromuscular Research Foundation Trust, Elizabeth Lottie May Rosenthal Bone Bequest and Perpetual Limited.

Marnee McKay does not work for, consult, own shares in or receive funding from any company or organization that would benefit from this article, and has disclosed no relevant

affiliations beyond their academic appointment.

## Partners


University of Sydney provides funding as a member of The Conversation AU.

View all partners

Girls in primary school are just as physically capable as their male classmates, according to our research, taking the sting out of the insult “you play like a girl”.

When we compared primary school children’s physical capabilities, differences between girls and boys were not as important as people think. So, they should be happily playing with and competing against each other in the backyard, playground and sporting fields.

---

***Read more:*** ***It's not just the toy aisles that teach children about gender stereotypes***

---

As part of wider research to assess people’s physical capabilities across the lifespan, we tested 300 children and adolescents between the ages of 3 and 19.

We tested each child for over two hours, taking more than 100 measurements. These included measuring the strength of 14 muscle groups, the flexibility of 13 joints and 10 different types of balance. We looked at factors including hand dexterity, reaction times, how far kids could walk, how high and how long they could jump, as well as their gait.

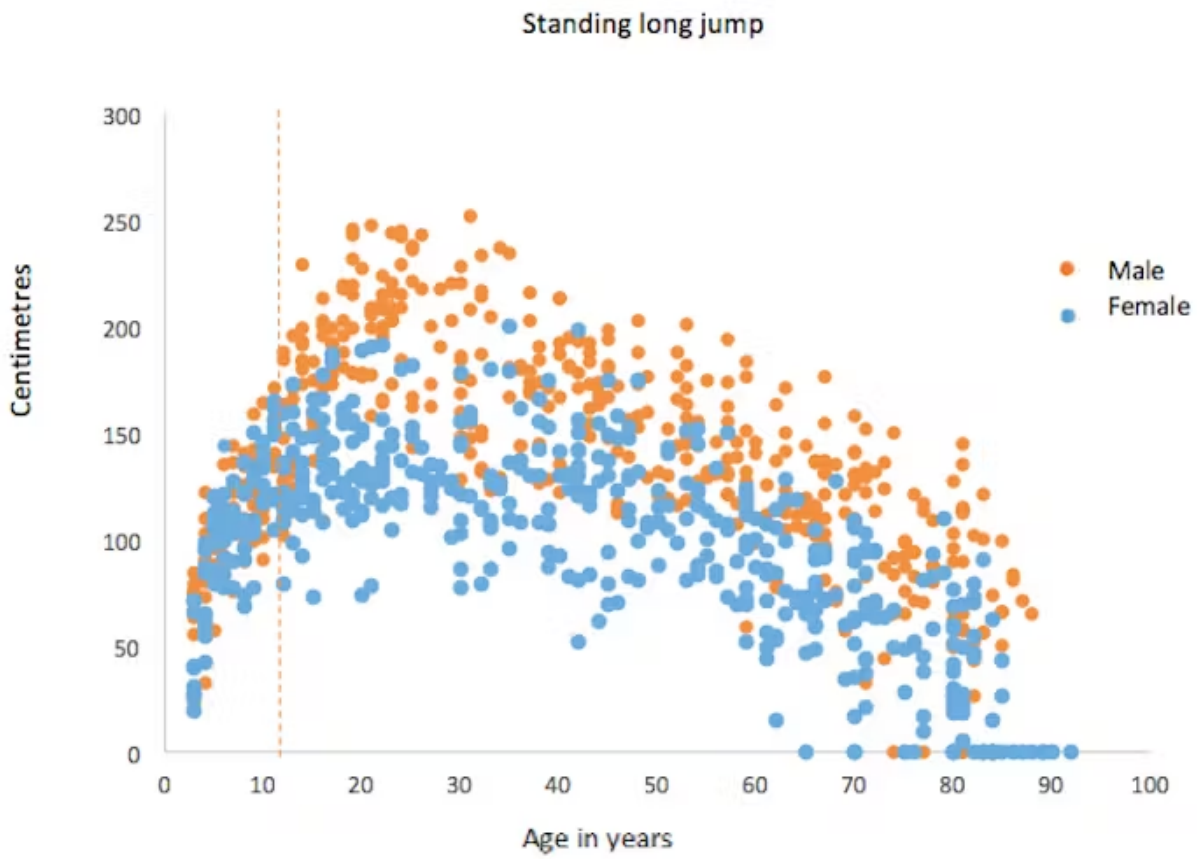
## What did we find?

Across all measures of physical performance, there was one consistent finding. There was no statistical difference in the capabilities of girls and boys until high-school age (commonly age 12).

Let's use standing long jump (also known as a broad jump test) as an example. This provides a measure of your legs' explosive power. It needs minimal equipment and the results compare well with the type of information you get from strength testing using expensive equipment. It's also one of the tests would-be American NFL (National Football League) players take to impress talent scouts.

The standing long jump is a test football scouts use to assess explosive power.

We found no difference between boys and girls before they turn 12 (see graph below). Every physical measure followed this pattern.




Before the age of 12, boys and girls do just as well as each other in the standing long jump. Author provided

## How do our findings compare?

Other studies have had similar results. These have included ones testing muscle strength, walking, jumping and balancing.

However, it's difficult to directly compare data from one study to another, as different studies have different sample sizes, include children of different age ranges, and assess different measures. For example, we were the first to use the timed stairs test and stepping reaction time to find what regular children were capable of.

Some studies found differences in physical capabilities between primary school-aged boys and girls using the same types of tests we used. And others reported small differences in the jump height of boys and girls aged 6-17 years but not with the long jump.

These differences can in part be attributed to sampling methods that were limited to specific age ranges or locations and socioeconomic backgrounds, the latter potentially having a significant impact on physical health and activity.

By contrast, the children in our research were generally representative of the Australian population, using data from the Australian Bureau of Statistics about socioeconomic status, ethnicity and body mass index.

## What do our findings mean for kids, coaches and parents?

There is no consensus across schools or among different sports about mixed-gender sports for primary school children.

For instance, boys and girls compete separately in most local Little Athletics after age five but field hockey can have mixed gender teams until age 17.

And in tennis, primary school-aged girls and boys play separately in singles matches but can play against each other in mixed doubles.

Our findings support the push for boys and girls to compete in mixed sporting teams until the end of primary school, after which the hormonal changes of puberty mean boys tend to perform better in sports and tasks requiring strength and speed.

---

***Read more:*** ***Our 'sporting nation' is a myth, so how do we get youngsters back on the field?***

---

There are also some practical advantages to mixed sport in primary school and in weekend competitions:

- fewer scheduling conflicts for councils (allowing school and sport administrations to fit games more conveniently into busy sporting venues)

- fewer clubs or organisations to share already stretched government and private sector funding
- consolidation of coaching and manager talent, and most importantly
- fewer parent-taxi drop offs.

Perhaps perceived differences in physical capability between boys and girls are based on outdated gender stereotypes that appear at birth, when some boys are given their first footy and some girls their first doll.

But whatever the origin of the idea young boys are physically more capable than young girls, the evidence is clear. Boys “play like a girl”, and that’s certainly no insult.

# Exhibit 23

# Normative reference values for strength and flexibility of 1,000 children and adults

Marnee J. McKay, MHlthSc
Jennifer N. Baldwin, BAppSc (Hons)
Paulo Ferreira, PhD
Milena Simic, PhD
Natalie Vanicek, PhD
Joshua Burns, PhD
For the 1000 Norms Project Consortium

Correspondence to M.J. McKay: marnee.mckay@sydney.edu.au

## ABSTRACT

**Objective:** To establish reference values for isometric strength of 12 muscle groups and flexibility of 13 joint movements in 1,000 children and adults and investigate the influence of demographic and anthropometric factors.

**Methods:** A standardized reliable protocol of hand-held and fixed dynamometry for isometric strength of ankle, knee, hip, elbow, and shoulder musculature as well as goniometry for flexibility of the ankle, knee, hip, elbow, shoulder, and cervical spine was performed in an observational study investigating 1,000 healthy male and female participants aged 3–101 years. Correlation and multiple regression analyses were performed to identify factors independently associated with strength and flexibility of children, adolescents, adults, and older adults.

**Results:** Normative reference values of 25 strength and flexibility measures were generated. Strong linear correlations between age and strength were identified in the first 2 decades of life. Muscle strength significantly decreased with age in older adults. Regression modeling identified increasing height as the most significant predictor of strength in children, higher body mass in adolescents, and male sex in adults and older adults. Joint flexibility gradually decreased with age, with little sex difference. Waist circumference was a significant predictor of variability in joint flexibility in adolescents, adults, and older adults.

**Conclusions:** Reference values and associated age- and sex-stratified *z* scores generated from this study can be used to determine the presence and extent of impairments associated with neuromuscular and other neurologic disorders, monitor disease progression over time in natural history studies, and evaluate the effect of new treatments in clinical trials. ***Neurology*®** **2017;88:36–43**

Meaningful, reliable, and sensitive outcome measures are required to monitor treatment and progression of neuromuscular and other neurologic disorders. While there have been substantial advances in the understanding of the pathogenesis and natural history of many neuromuscular disorders, the identification and development of new outcome measures that best reflect the efficacy of specific treatments have not advanced at the same rate.[1] Establishing valid and responsive outcome measures is a priority for the field.[2] To assist in the development of new outcome measures, normative reference values generated from large populations across the lifespan using standardized methods are required. Normative reference values can be utilized to generate *z* scores, which can be used in multicenter studies to improve outcome measure precision and responsiveness.

Muscle weakness and joint contractures predispose to numerous pathologies requiring intervention. Reference data play an important role in identifying and quantifying these impairments and evaluating the effectiveness of interventions. Currently, few comprehensive datasets detail the normal variation of active range of motion in healthy individuals and are limited by the number of joints assessed,[3] the age range of participants,[4,5] or insufficient sex representation.[6,7]

**Supplemental data at Neurology.org**

From the Faculty of Health Sciences (M.J.M., J.N.B., P.F., M.S., J.B.), University of Sydney, New South Wales, Australia; Department of Sport, Health and Exercise Science (N.V.), University of Hull, UK; and Sydney Children's Hospitals Network (Randwick and Westmead) and Paediatric Gait Analysis Service of New South Wales (J.B.), Children's Hospital at Westmead, Sydney, Australia.

Coinvestigators are listed at Neurology.org.

Go to Neurology.org for full disclosures. Funding information and disclosures deemed relevant by the authors, if any, are provided at the end of the article.

Similar limitations exist in strength reference datasets, relevant only to children[8–11] or adult populations,[12–16] or strength measured using equipment not readily available in clinic.[17–19]

The purpose of this study was to generate a reference dataset of normative values across the lifespan for an extensive set of isometric muscle strength and joint flexibility items, stratified for age and sex, and to investigate the influence of demographic and anthropometric factors.

## METHODS

**Study design and participants.** Data were collected as part of the 1,000 Norms Project, an observational study investigating physical function and self-reported health in 1,000 people across the lifespan (see full protocol[20]). One thousand people aged between 3 and 101 years from the Greater Sydney metropolitan area in Australia participated in the project. Participants were recruited from January 2014 to September 2015 using highly structured convenience and snowball sampling techniques, including advertising via social media, e-newsletters, and community flyers. Presentations were held at social and volunteer groups, aged care organizations, playgroups, and schools. Eligible participants were aged ≥3 years, considered themselves healthy for their age, and could participate in age-appropriate activities of daily living. People with significant health conditions affecting physical performance or an inability to follow age-appropriate instructions were excluded. Potential participants with the following conditions were also excluded: diagnosed diabetes mellitus; malignant cancers; demyelinating, inflammatory, or degenerative neurologic conditions; pregnancy; class 3 obesity; severe cardiac or pulmonary disease; joint replacement; infectious or inflammatory arthropathies; or severe mobility impairment necessitating dependence on mobility aids for all ambulation. Equal numbers of male and female participants were recruited and were stratified into 9 age categories. One hundred people per decade were recruited in the age groups of 20–29, 30–39, 40–49, 50–59, 60–69, 70–79, and 80+ years. In order to represent the rapid periods of growth and maturation and to distinguish between young children and adolescents, 20 children per year from 3 to 9 years of age and 16 per year from 10 to 19 years of age were recruited.

**Standard protocol approvals, registrations, and patient consents.** Ethical approval was granted by the institutional Human Research Ethics Committee (HREC 2013/640) and written informed consent was obtained from all participants or parents/guardians of children.

**Procedure.** Participants attended the University of Sydney Performance Laboratory once for a 2-hour assessment. Participants had their height, body mass, waist circumference, and lower limb alignment measured. Foot structure was assessed using the Foot Posture Index, a 6-item summed scale from −12 (supinated) to +12 (pronated).[21] Age, sex, current work status, and self-reported ethnicity were collected from all participants or parents/guardians. Work status was classified as working (full-time, part-time, or unpaid) or not currently working (unemployed, student, or retired). Ethnicity was classified into 5 categories: British/European, American, Asian, African, and Aboriginal/Torres Strait Islander.

Two experienced clinical evaluators (physiotherapists) assessed isometric strength and joint flexibility using standardized methodology, including instructions, positioning, and scoring.[20] The dominant limb was assessed and determined as the hand used to write with and the foot used to kick a ball. The strength of 12 muscle groups—hand grip, ankle dorsiflexors and plantarflexors, knee flexors and extensors, hip abductors, internal and external rotators, elbow flexors and extensors, and shoulder internal and external rotators—were assessed by maximal voluntary isometric contraction using a portable hand-held dynamometer (Citec dynamometer CT 3001; CIT Technics, Groningen, Netherlands). The dynamometer was calibrated 0–500 N with certified weights monthly throughout data collection. The strength of knee musculature in participants ≥12 years of age was assessed by fixed dynamometry (CSMi; HUMAC NORM, Stoughton, MA). For unit of measure consistency, knee flexor and extensor strength in children aged 3–11 years were converted to Newton-meters (Nm) using anthropometric tables.[22] Rather than using the fixation device outlined in the protocol, ankle plantarflexion strength was assessed using hand-held dynamometry in long sitting, heel over plinth edge.

Joint flexibility was assessed using a universal goniometer, digital inclinometer, or bubble inclinometer (Baseline; Fabrication Enterprises Inc., White Plains, NY) depending on the joint assessed. Thirteen active joint range movements were assessed: ankle dorsiflexion and plantarflexion, knee flexion and extension, hip flexion, internal and external rotation, elbow flexion and extension, shoulder internal and external rotation, and cervical flexion and extension. Interrater reliability of the clinical evaluators demonstrated satisfactory repeatability of all strength and flexibility measures (intraclass correlation coefficient$_{2,1}$ 0.80–0.99) in a pilot study of 10 participants aged 6–67 years.

**Data analysis.** Data were collected and managed using REDCap electronic data capture and manually checked for transcription errors. Reference values were generated for each age group and sex in SPSS v22 Statistics for Windows (IBM SPSS; Armonk, NY). Normality of the data was assessed using Kolmogorov-Smirnov test. For analysis, age categories were children (3–9 years), adolescents (10–19 years), adults (20–59 years), and older adults (60+ years). To determine if strength and flexibility differed between male and female participants, independent $t$ tests were conducted. A series of multiple regression models was constructed to determine the extent to which muscle strength and joint flexibility were influenced by participant demographic (age and sex) and anthropometric factors (height, body mass, waist circumference, foot posture, and lower limb alignment). First, Pearson product-moment correlation coefficients ($r$) were generated to explore the bivariate relationships between strength and anthropometric and demographic factors. The same correlations were explored for each joint flexibility measure. Second, factors identified to have an association ($r \geq 0.3$, $p < 0.05$) with strength or joint flexibility were entered simultaneously into a stepwise multiple regression model, which was reduced to a set of factors that best predicted and could be regarded as independent determinants of each strength and joint flexibility measure. To avoid multicollinearity, only one variable from highly correlated ($r \geq 0.7$) variables was included. Standardized β weights were calculated to provide an indication of the relative importance of the contribution of the various factors entered into the model to explain the variance in joint flexibility

**Table 1** Isometric strength reference values of children (3–9 years), adolescents (10–19 years), adults (20–59 years), and older adults (60+ years)[a]

| Muscle group | Entire sample | 3–9 years | | 10–19 years | | 20–59 years | | 60+ years | |
|---|---|---|---|---|---|---|---|---|---|
| | | Male | Female | Male | Female | Male | Female | Male | Female |
| Grip, N | 187.1 (94.9) | 55.8 (29.1) | 50.6 (25.3) | 195.7 (83.0)[b] | 153.8 (46.0) | 305.0 (73.0)[b] | 190.2 (50.4) | 221.6 (49.5)[b] | 128.7 (35.4) |
| Ankle dorsiflexors, N | 164.7 (61.2) | 87.1 (38.2) | 81.6 (29.2) | 197.2 (56.6)[b] | 166.0 (37.8) | 224.6 (48.9)[b] | 166.5 (41.6) | 173.3 (44.0)[b] | 131.5 (38.9) |
| Ankle plantarflexors, N | 257.0 (85.6) | 151.7 (52.3) | 142.7 (45.9) | 309.9 (74.9)[b] | 261.2 (52.7) | 338.8 (66.8)[b] | 243.9 (59.2) | 281.4 (62.7)[b] | 216.3 (60.3) |
| Knee flexors, Nm[c] | 68.6 (33.8) | 27.0 (13.9) | 25.2 (11.5) | 89.8 (34.6)[b] | 65.9 (19.7) | 106.3 (28.6)[b] | 64.4 (18.9) | 76.3 (20.1)[b] | 45.8 (13.3) |
| Knee extensors, Nm[d] | 124.0 (66.4) | 34.9 (18.1) | 34.2 (14.9) | 152.8 (71.1)[b] | 116.9 (36.6) | 202.1 (56.1)[b] | 122.6 (33.6) | 136.2 (35.6)[b] | 81.9 (26.8) |
| Hip internal rotators, N | 146.8 (69.5) | 63.3 (31.7) | 61.1 (25.8) | 178.5 (67.2)[b] | 143.2 (45.0) | 217.7 (62.4)[b] | 136.1 (44.6) | 169.7 (55.0)[b] | 108.4 (33.8) |
| Hip external rotators, N | 110.4 (52.6) | 49.7 (22.7) | 43.8 (16.8) | 141.7 (53.7)[b] | 104.0 (28.7) | 169.4 (45.8)[b] | 100.7 (29.1) | 125.5 (33.9)[b] | 76.3 (23.7) |
| Hip abductors, N | 116.1 (50.6) | 52.3 (23.1) | 52.4 (21.9) | 143.4 (47.2)[b] | 116.6 (31.9) | 170.7 (43.9)[b] | 113.1 (32.4) | 124.8 (32.8)[b] | 83.8 (23.5) |
| Elbow flexors, N | 176.3 (80.0) | 71.7 (29.1) | 66.0 (26.4) | 213.8 (81.1)[b] | 148.5 (36.8) | 270.2 (59.6)[b] | 164.4 (42.3) | 209.4 (48.4)[b] | 129.7 (33.9) |
| Elbow extensors, N | 135.8 (57.4) | 66.8 (24.4) | 62.0 (19.7) | 159.3 (56.8)[b] | 118.3 (30.0) | 203.2 (46.1)[b] | 121.2 (30.2) | 162.1 (36.8)[b] | 102.8 (25.3) |
| Shoulder internal rotators, N | 126.7 (64.7) | 56.1 (27.1)[e] | 47.7 (17.4) | 151.5 (63.2)[b] | 101.6 (27.7) | 202.4 (55.9)[b] | 109.7 (33.6) | 159.7 (42.9)[b] | 86.0 (27.5) |
| Shoulder external rotators, N | 86.4 (41.0) | 38.7 (19.5) | 34.7 (13.0) | 100.6 (38.8)[b] | 73.4 (19.1) | 134.7 (39.6)[b] | 82.2 (20.9) | 96.7 (25.3)[b] | 63.3 (19.2) |

[a] Mean values (SD).
[b] Significant ($p < 0.01$).
[c] Participants aged 3–11 years measured with hand-held dynamometry (mean 94.4, SD 40.6) was converted to Nm.
[d] Participants aged 3–11 years measured with -held dynamometry (mean 126.4, SD 53.5) was converted to Nm.
[e] Significant sex differences ($p < 05$).

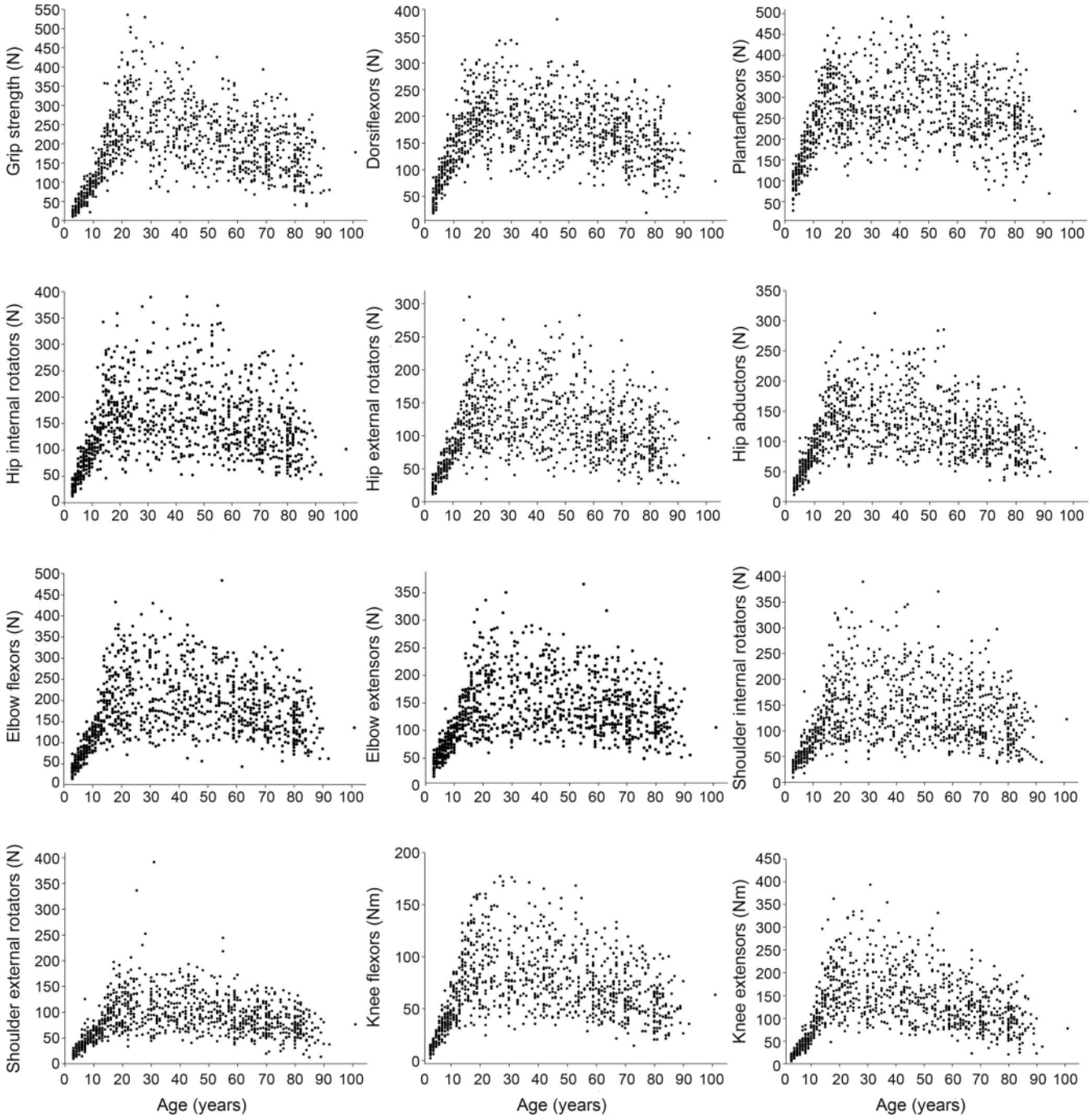
or muscle strength. Variables were retained in the multiple regression model if $p < 0.05$.

**RESULTS** To recruit 1,000 participants, 2,972 e-mails and 240 phone calls were logged. Ninety-one potential participants were excluded in accordance with the inclusion and exclusion criteria. Among adults aged over 18 years, 56% were currently working and 44% were not (31% retired and 13% students or unemployed). Participants were of diverse geographic ancestry, although the majority of participants were British/European ethnicity (74.4%), followed by Asian (16.6%), North or South American (5.1%), African (2.4%), and Aboriginal or Torres Strait Islander (1.5%). The sample mean (SD) age was 40.9 years (26.1), body mass 62.9 kg (21.1), height 1.61 m (0.02), waist circumference 78.6 cm (15.4), Foot Posture Index 3.5 (2.4), and lower limb alignment 1.8° (2.7). Ninety-three percent were right-footed and 91% were right-handed.

All missing data were accounted for. Four children declined to perform ankle dorsiflexion strength and 12 children were unable to perform cervical flexion and extension joint movements in accordance with the protocol. Ankle plantarflexors for 7 male adults and ankle dorsiflexors for 1 male adult were not assessed with hand-held dynamometry during periods of offsite calibration and servicing; 6 adults and 8 older adults were not assessed using fixed dynamometry due to safety concerns.

Normative reference values for the strength of 13 muscle groups per age category (children, adolescents, adults, and older adults) and sex are presented in table 1 and per decade in table e-1 at Neurology.org. From adolescence, male participants were significantly stronger in all muscle groups across all ages. There were no significant ($p < 0.05$) differences between the strength measures of boys or girls aged 3–9 years, except for shoulder internal rotators ($p = 0.031$), where boys were stronger. Correlations between strength and participant demographics and anthropometrics for children, adolescents, adults, and older adults are presented in table e-2. In children and adolescents, strength and age were highly correlated ($p < 0.05$), confirming that children become significantly stronger as they age from childhood and through adolescence. From 20 years of age, the relationship between strength and age changed. In adults aged 20–59 years, reduced strength with age was evidenced by significant, although weak, correlations with hand grip, ankle dorsiflexors, knee flexors and extensors, and shoulder external rotators. In older adults, decreased strength with increasing age was evidenced in all muscle groups. All muscle groups across all age categories demonstrated that greater height, body mass, and waist circumference were significantly associated with greater strength. The changes in strength measures with advancing age are shown in figure 1. Table e-3 shows the results of the multiple analyses. In children, height, followed

**Figure 1** Scatterplots of muscle strength vs age for 1,000 children and adults




by waist circumference, was the most significant predictor of strength. In adolescents, a combination of body mass, sex, and age were shown to be the strongest independent predictors of strength. Sex (male) was the most significant predictor of strength in adults, followed by height and body mass with lower predictive values. In older adults, sex (male) was the most significant predictor, with body mass, height, and age demonstrating lower predictive values.

Normative reference values for active range of motion per age category (children, adolescents, adults, and older adults) and sex are presented in table 2 and per decade in table e-4. There was no significant difference ($p < 0.05$) in joint flexibility between boys and girls aged 3–9 years, except for hip internal rotation ($p = 0.017$), where girls had greater flexibility. Active range of motion was greatest in children compared to older adults. Figure 2 illustrates the inverse relationship between joint flexibility in all joints with age. Pearson correlations (table e-2) demonstrate that a decrease in flexibility with aging occurred in 8 of 13 joints of both adolescents and

**Table 2** Joint flexibility reference values of children (3–9 years), adolescents (10–19 years), adults (20–59 years), and older adults (60+ years)[a]

| | | 3–9 years | | 10–19 years | | 20–59 years | | 60+ years | |
|---|---|---|---|---|---|---|---|---|---|
| Movement, degrees | Entire sample | Male | Female | Male | Female | Male | Female | Male | Female |
| Ankle dorsiflexion | 30 (6.7)[b] | 33 (7.2) | 31 (5.7) | 32 (5.7) | 31 (7.1) | 32 (6.1)[b] | 29 (6.4) | 31 (6.1)[b] | 26 (6.3) |
| Ankle plantarflexion | 59 (8.6)[b] | 63 (7.3) | 63 (9.2) | 58 (7.6)[b] | 63 (7.3) | 56 (7.5)[b] | 62 (8.9) | 53 (6.8)[b] | 57 (7.2) |
| Knee flexion | 137 (7.8) | 145 (5.5) | 144 (5.7) | 140 (6.7)[c] | 142 (6.6) | 136 (6.1) | 137 (6.2) | 133 (7.2) | 131 (8.1) |
| Knee extension | 1 (2.9) | 4 (3.3) | 4 (3.9) | 2 (2.6) | 2 (2.6) | 1 (2.3) | 2 (2.7) | −1 (2.4)[c] | 1 (1.6) |
| Hip flexion | 121 (11.8)[c] | 133 (9.1) | 133 (9.8) | 120 (9.9)[b] | 124 (10.2) | 120 (8.7)[b] | 123 (10.0) | 115 (10.7) | 114 (12.6) |
| Hip internal rotation | 37 (9.0)[b] | 40 (8.4) | 43 (9.1) | 37 (9.3) | 39 (7.7) | 36 (7.9)[b] | 40 (8.8) | 33 (8.0)[c] | 35 (8.4) |
| Hip external rotation | 28 (8.5)[b] | 32 (8.1) | 32 (9.2) | 31 (6.4) | 31 (9.1) | 30 (8.3)[b] | 27 (8.3) | 26 (7.0)[b] | 22 (6.7) |
| Elbow flexion | 148 (5.4)[b] | 146 (5.4) | 147 (5.8) | 148 (5.4)[c] | 150 (4.5) | 147 (4.9)[b] | 149 (5.4) | 146 (6.0)[b] | 149 (4.7) |
| Elbow extension | 3 (5.9)[b] | 7 (4.6) | 7 (5.1) | 4 (5.4)[b] | 7 (5.6) | 2 (5.0)[b] | 4 (5.1) | −1 (5.0)[c] | 0 (5.1) |
| Shoulder internal rotation | 62 (12.9)[b] | 67 (14.2) | 67 (13.2) | 62 (12.2) | 66 (12.3) | 58 (12.0)[b] | 63 (14.0) | 57 (11.0)[b] | 63 (11.6) |
| Shoulder external rotation | 83 (16.6) | 98 (12.2) | 99 (12.9) | 93 (12.4) | 93 (13.2) | 83 (13.2) | 83 (15.8) | 71 (12.2) | 72 (13.9) |
| Cervical flexion | 60 (12.6)[b] | 72 (13.5) | 68 (12.4) | 66 (12.3) | 64 (10.5) | 59 (10.8)[b] | 56 (10.2) | 55 (12.0) | 53 (10.6) |
| Cervical extension | 59 (19.5)[c] | 82 (16.0) | 80 (19.2) | 67 (13.8)[c] | 73 (15.3) | 58 (13.1)[c] | 61 (14.7) | 40 (12.3)[c] | 43 (13.3) |

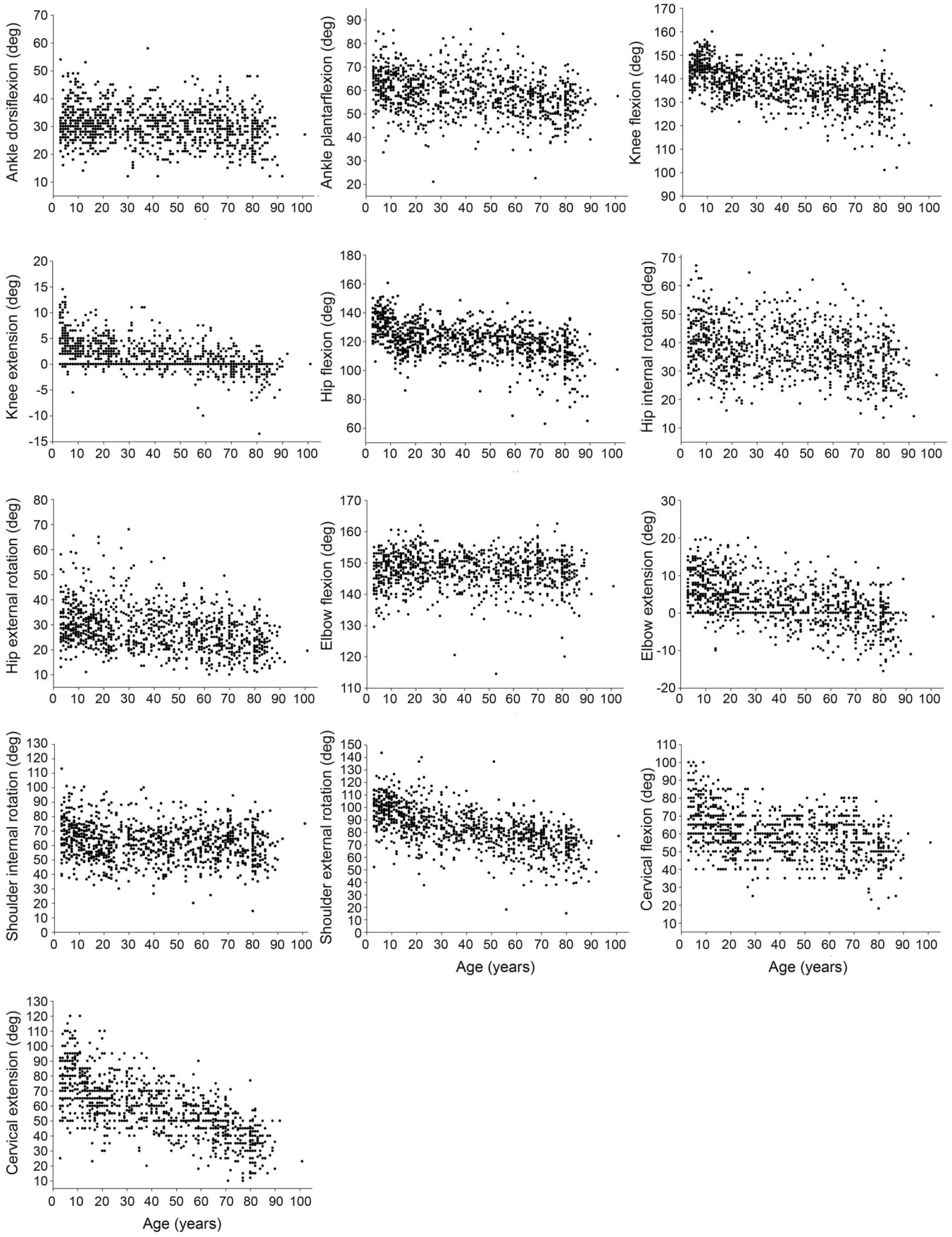
[a] Mean values (SD).
[b] Significant ($p < 0.01$).
[c] Significant sex differences ($p < 0.05$).

adults and in 12 joints of older adults ($p < 0.05$). Greater body mass and waist circumference were associated with a decrease in joint flexibility from 10 years of age. The correlation between height and joint flexibility was strongest in adolescents, where taller individuals demonstrated less joint range of motion. From adolescence to older adulthood, a more pronated foot posture was associated with a greater range of ankle dorsiflexion. Lower limb alignment did not demonstrate any significant correlations, beyond very weak associations, with measures of flexibility in any age category. In children, 2 multiple regression models (see table e-4) reached significance (knee extension and neck extension) and revealed height as the most significant predictor. Age, waist circumference, and height were the strongest independent predictors of flexibility in adolescents. For all adults older than 20 years, age, sex, and waist circumference were the strongest predictors of joint flexibility.

**DISCUSSION** This study has established a comprehensive reference dataset of isometric muscle strength and joint flexibility in 1,000 healthy people aged 3–101 years. The associations between strength and flexibility measures with demographic and anthropometric variables within different age categories identified some important relationships. As expected, there is a highly significant increase in strength of all muscle groups as children rapidly develop through to early adulthood. From adulthood, this relationship changes and a decrease in muscle strength with aging starts to occur; by older adulthood, all muscle groups demonstrate loss of strength with aging. From 10 years of age, a time of life coinciding with rapid growth and maturation, males are significantly stronger in all measures. In contrast, joint flexibility demonstrates a steady decline with age and no meaningful difference between males and females.

Our results are consistent with previous studies investigating isometric muscle strength in children[8,9] and adults.[12–14] However, direct comparison is limited by differences in age range, sample size, and muscle groups evaluated. Some studies report body mass as the strongest correlate with muscle strength in children,[8–11] while others demonstrate as we did that height showed the strongest relationship and was the most significant predictor of strength.[23,24] In adults, a decline in strength was most strongly associated with aging in 5 muscle groups (hand grip, ankle dorsiflexors, shoulder external rotators, knee flexors, and extensors), while in older adulthood all muscle groups demonstrated a significant decline in strength associated with aging. These results suggest a muscle-specific response to aging during adulthood and that generalized weakness does not occur until older adulthood. This highlights the importance of using age- and sex-matched reference data for specific muscle groups to avoid overrepresentation or underrepresentation of the force capabilities of a particular muscle group. Similar relationships between aging and muscle weakness have been reported in a limited number of adult studies.[12–14]

**Figure 2** Scatterplots of joint flexibility vs age for 1,000 children and adults



The association between waist circumference and strength and flexibility has not been reported previously. Waist circumference was identified as a significant predictor of flexibility in adolescents, adults, and older adults and of strength in children. Epidemiologic studies have identified an association between waist circumference and tendon pathology,[25] with preliminary evidence supporting either a mechanical effect (due to increased load) or systemic effect (due to circulating lipids).[26] The influence of adiposity on localized musculo-tendinous tissues in neuromuscular disorders will be an important factor to evaluate with the increasing rates of obesity in society.

Few studies report normative reference values for flexibility in children. Our normative reference values for adults are consistent with the literature.[3,4,7] We identified only one sex difference in the flexibility of children (namely hip internal rotation), and only small differences (2°–6°) between men and women from adolescence through to older adulthood. As such, sex does not seem to have a clinically important effect on the joint flexibility of healthy adolescents and adults. There is no consensus in the literature regarding sex differences and flexibility; some studies report, as we have, that there is no clinically relevant difference,[4] while others report sex differences.[5] We identified a linear decrease in joint flexibility associated with advancing age, consistent with the adult literature.[3] It is likely that in healthy individuals, joint flexibility declines gradually and steadily with age and a substantial or sudden decline should be considered indicative of an underlying pathology.

Studies characterizing the functional decline and rate of progression of neuromuscular disorders such as amyotrophic lateral sclerosis,[27] Duchenne muscular dystrophy,[28] and Charcot-Marie-Tooth disease[29] depend on hand-held dynamometry to capture and track relevant changes in muscle strength. Access to reliable and expansive normative reference values and associated age- and sex-matched *z* scores are necessary to accurately and precisely quantify response to new interventions and to establish minimum clinically important differences.

This study is not without limitation. Participants were recruited through convenience sampling methods and with the exclusion criteria of conditions affecting physical performance may have resulted in a population that were particularly healthy and physically capable for their age. While the mixed ethnicity of our sample is reflective of the Australian population, the ethno-geographic variation in strength and flexibility measures could not be established. The cross-sectional study design was effective in achieving our study aim of generating a reference dataset of strength and flexibility across the lifespan; however, the direction of some of the cause and effect relationships can only be identified in longitudinal studies that track the changes in these measures over time. Ankle plantarflexion strength in healthy adolescents and adults can only ever be estimated with hand-held dynamometry due to the very high force capability (often exceeding 1,000 N).[30] The reported reference values for ankle plantarflexors are likely to underestimate the force capabilities of this muscle group, and values should be used as a lower threshold for weakness in patients with neuromuscular and other neurologic disorders. Finally, the strength and flexibility reference values are specific to the Citec hand-held dynamometer and Baseline goniometer and inclinometer and may not be interchangeable with data obtained from other devices.

The normative reference data generated from this study can be used to determine the presence and extent of impairments associated with neuromuscular disorders and to monitor disease progression over time. The reference values and associated age- and sex-matched *z* scores can be used to develop outcome measures with enhanced precision and responsiveness to be used in clinical trials for neuromuscular and other neurologic disorders.

## AUTHOR CONTRIBUTIONS

Marnee J. McKay: study design, data collection, analysis and interpretation, drafting and revising the manuscript. Jennifer N. Baldwin: study design and data collection. Paulo Ferreira: study design and revising the manuscript. Milena Simic: study design and revising the manuscript. Natalie Vanicek: study design and revising the manuscript. Joshua Burns: study conceptualization and design, data interpretation, drafting and revising the manuscript.

## ACKNOWLEDGMENT

The authors thank John Eisenhuth for technical assistance, Ray Patton for calibration of the hand-held dynamometer, and the 1,000 volunteers.

## STUDY FUNDING

Study funded by National Health and Medical Research Council of Australia Centre for Research Excellence in Neuromuscular Disorders (NHMRC 1031893) and the Australian Podiatry Education and Research Foundation.

## DISCLOSURE

The authors report no disclosures relevant to the manuscript. Go to Neurology.org for full disclosures.

*Received July 14, 2016. Accepted in final form September 22, 2016.*

## REFERENCES

1. Lunn MP, Van den Bergh PY. Outcome measures in neuromuscular disease: is the world still flat? J Peripher Nerv Syst 2015;20:255–259.
2. Mercuri E, Messina S, Pane M, Bertini E. Current methodological issues in the study of children with inherited neuromuscular disorders. Dev Med Child Neurol 2008;50:417–421.
3. Roach KE, Miles TP. Normal hip and knee active range of motion: the relationship to age. Phys Ther 1991;71:656–665.
4. Hallaceli H, Uruc V, Uysal HH, et al. Normal hip, knee and ankle range of motion in the Turkish population. Acta Orthop Traumatol Turc 2014;48:37–42.

5. Sengupta P, De S, Pal A, Maity P, Banerjee M, Dhara PC. Variation of range of joint motion in Bengalee (Indian) healthy adult subjects. J Life Sci 2012;4:123–133.
6. Boone DC, Azen SP. Normal range of motion of joints in male subjects. J Bone Joint Surg Am 1979;61:756–759.
7. Gunal I, Kose N, Erdogan O, Gokturk E, Seber S. Normal range of motion of the joints of the upper extremity in male subjects, with special reference to side. J Bone Joint Surg Am 1996;78:1401–1404.
8. Backman E, Odenrick P, Hendriksson KG, Ledin T. Isometric muscle force and anthropometric values in normal children aged between 3.5 and 15 years. Scand J Rehab Med 1989;21:105–114.
9. Beenakker E, Van der Hoeven J, Fock J, Maurits N. Reference values of maximum isometric muscle force obtained in 270 children aged 4–16 years by hand-held dynamometry. Neuromuscul Disord 2001;11:441–446.
10. Eek MN, Kroksmark A-K, Beckung E. Isometric muscle torque in children 5 to 15 years of age: normative data. Arch Phys Med Rehabil 2006;87:1091–1099.
11. Hébert LJ, Maltais DB, Lepage C, Saulnier J, Crête M. Hand-held dynamometry isometric torque reference values for children and adolescents. Pediatr Phys Ther 2015;27:414.
12. Andrews AW, Thomas MW, Bohannon RW. Normative values for isometric muscle force measurements obtained with hand-held dynamometers. Phys Ther 1996;76:248–259.
13. Bäckman E, Johansson V, Häger B, Sjöblom P, Henriksson K. Isometric muscle strength and muscular endurance in normal persons aged between 17 and 70 years. Scand J Rehab Med 1995;27:109–117.
14. Bohannon RW. Reference values for extremity muscle strength obtained by hand-held dynamometry from adults aged 20 to 79 years. Arch Phys Med Rehabil 1997;78:26–32.
15. Meldrum D, Cahalane E, Conroy R, Fitzgerald D, Hardiman O. Maximum voluntary isometric contraction: reference values and clinical application. Amyotroph Lateral Scler 2007;8:47–55.
16. Van der Ploeg R, Fidler V, Oosterhuis H. Hand-held myometry: reference values. J Neurol Neurosurg Psychiatry 1991;54:244–247.
17. Danneskiold-Samsøe B, Bartels E, Bülow P, et al. Isokinetic and isometric muscle strength in a healthy population with special reference to age and gender. Acta Physiol 2009;197:1–68.
18. Hogrel J-Y, Payan CA, Ollivier G, et al. Development of a French isometric strength normative database for adults using quantitative muscle testing. Arch Phys Med Rehabil 2007;88:1289–1297.
19. The National Isometric Muscle Strength (NIMS) Database Consortium. Muscular weakness assessment: use of normal isometric strength data. Arch Phys Med Rehabil 1996;77:1251–1255.
20. McKay MJ, Baldwin JN, Ferreira P, et al. 1000 Norms Project: protocol of a cross-sectional study cataloguing human variation. Physiotherapy 2016;102:50–56.
21. Redmond AC, Crosbie J, Ouvrier RA. Development and validation of a novel rating system for scoring standing foot posture: the Foot Posture Index. Clin Biomech 2006;21:89–98.
22. Winter DA. Biomechanics and Motor Control of Human Movement. New York: John Wiley & Sons; 2009.
23. Hogrel JY, Decostre V, Alberti C, et al. Stature is an essential predictor of muscle strength in children. BMC Musculoskel Disord 2012;13:176.
24. Macfarlane TS, Larson CA, Stiller C. Lower extremity muscle strength in 6-to 8-year-old children using hand-held dynamometry. Pediatr Phys Ther 2008;20:128–136.
25. Gaida JE, Ashe MC, Bass SL, Cook JL. Is adiposity an under-recognized risk factor for tendinopathy? A systematic review. Arthritis Care Res 2009;61:840–849.
26. Scott A, Zwerver J, Grewal N, et al. Lipids, adiposity and tendinopathy: is there a mechanistic link? Crit Rev Br J Sports Med 2015;49:984–988.
27. Shefner JM, Liu D, Leitner ML, et al. Quantitative strength testing in ALS clinical trials. Neurology 2016:87:617–624.
28. Kirschner J, Schessl J, Schara U, et al. Treatment of Duchenne muscular dystrophy with cyclosporine A: a randomised, double-blind, placebo-controlled multicentre trial. Lancet Neurol 2010;9:1053–1059.
29. Cornett KM, Menezes MP, Bray P, et al. Phenotypic variability of childhood Charcot-Marie-Tooth disease. JAMA Neurol 2016;73:645–651.
30. Kelln BM, McKeon PO, Gontkof LM, Hertel J. Hand-held dynamometry: reliability of lower extremity muscle testing in healthy, physically active, young adults. J Sport Rehabil 2008;17:160.

# Exhibit 24

WORLD RUGBY

# Transgender Women Guidelines

## Can transgender women play rugby?

- Transgender women who transitioned pre-puberty and have not experienced the biological effects of testosterone during puberty and adolescence can play women's rugby (subject to confirmation of medical treatment and the timing thereof)
- Transgender women who transitioned post-puberty and have experienced the biological effects of testosterone during puberty and adolescence cannot currently play women's rugby
- Transgender women can play mixed-gender non-contact rugby
- World rugby are committed to ongoing evaluation of the guidelines and will remain current on all published research that pertains to the biological and physiological implications of testosterone suppression, with a formal review of the Guideline every three years. In support of this, World Rugby will prioritise support for high quality research projects on transgender rugby players, as part of this commitment to evidence-based guidelines.

## Why can't transgender women play women's rugby?

## Effects of testosterone

re reference is made to "females" and "males" to explain the effects of testosterone, the references are used to diffe

Testosterone is an androgenic-anabolic hormone whose functions include reproductive maturation, along with the genesis of male secondary sex characteristics. From puberty onwards, testosterone levels increase 20-fold in males, but remain low in females, resulting in circulating testosterone concentrations at least 15 times higher in males than in females of any age [1,2]. Among the biological changes initiated by testosterone and its derivatives are:

- Larger and denser lean muscle mass [3,4];
- Greater force-producing capacity of skeletal muscle [5,6];
- Stiffer connective tissue [7];
- Reduced fat mass and different distribution of body fat and lean muscle mass [3];
- Longer, larger and denser skeletal structure [8,9];
- Changes to cardiovascular and respiratory function that include higher haemoglobin concentration, greater cross-sectional area of the trachea and lower oxygen cost of respiration (as described in [1,10-12]).

Collectively, these biological differences account for large sporting performance differences between males and females. These include gaps between 9% and 15% for running, swimming and jumping events [13], between 15% and 35% for functional tasks like kicking, throwing, bowling and weightlifting, and in excess of 50% for tasks that involve upper body force production [10], since the biological effects of testosterone creates disproportionately greater strength on their upper compared to lower body, while females show the inverse [14,15]. In weight-lifting events, for instance, even when matched for mass and stature, males lift approximately 30% more weight than females. Evaluated differently, males are able to lift weights similar to females who weigh 30% to 40% more than them [10]. Functional movements such as explosive jumping are similarly larger in elite males than females, with approximately 30% more power generated during a counter movement jump [10].

The result of these biological differences is that males outperform females in all sporting activities where speed, size, power, strength, cardiorespiratory and anthropometric characteristics are crucial determinants of performance. This is true for many thousands of boys and men who have undergone a testosterone-

performance differences varies depending on the contributions made by each of the biological variables to performance, and indeed, some may be detrimental to performance in some events (mass during endurance running or cycling events, for example). Generally, however, there is no overlap in performance between males compared to females at all matched levels of competition from high school to the elite level. The performance disparity is illustrated by the observation that thousands of teenage boys and adult males are able to outperform the very best biological females every year [13].

Similar performance differences between males and females have been described in non athletically trained individuals. Males have muscle mass 30% to 40% greater than females [4], maximal cardiorespiratory capacities (VO2max) 25% to 50% greater than in females [17], cardiovascular parameters between 11% and 43% greater than in females, lower limb strength approximately 50% higher than in females across the lifespan, and upper body strength 50% to 100% higher than in age-matched females [6]. Even when elite females, trained in sports where grip strength is an important component of performance Uudo and handball), do not outperform untrained males in a grip strength task, with the very best female performance corresponding to approximately the 58 $^{th}$ percentile for males, and a 26% advantage for untrained males compared to typical elite females. Punching performance, a composite movement reliant on strength, power, co-ordination and mass, has been found to be 162% higher in males than in females [18], and 17-year old boys are able to throw a ball further than 99% of adult females [19].

## Biological consideration for rugby union

The implications of biological and performance differences for rugby are two-fold. First, significant differences in strength, size, speed and power have potential consequences for the safety of participants in rugby, where much of the sport involves contacts in the form of tackles, rucks and mauls, as well as numerous periods of high force production during static contests for the ball, such as the scrum and ruck. Given the documented risk of injury in rugby from contact events in particular [20-24], the elevated possibility of all injuries, including serious injury, from large disparities in size, speed, power, and force, is of concern. Recent modelling of tackles using validated biomechanical models [25,26] suggests that the

greater risks for smaller and slower players, particularly when size and speed exist in combination.

Given that the typical male player mass is 20% to 40% greater than typical women mass, that males have strength 40% to 80% greater (unadjusted for mass), and that men are 10% to 15% faster than women despite being heavier, the risk of injury created by large imbalances in mass and speed may be considered significant. To explore this, we assessed the range of masses of players at the international level and applied the findings to a biomechanical model to explore possible implications for injury risk should cross-over scenarios occur.

With respect to mass, we documented the range of sizes of elite men's and women's players from the 2011 Rugby World Cup up to the 2019 Rugby World Cup, finding:

- Typical (median) men's players are 41.1% heavier than typical women's players (103 kg vs 73 kg)
- Among forwards, the heaviest 1% of women players are smaller than the typical men's forward (109kg for women vs 112kg for men)
- The heaviest 1% of women's backs are smaller than typical men's backs (89kg vs 92kg)
- The lightest 1% of men's forwards are approximately equal in mass to the heaviest 10% of women's forwards, while the lightest 2% of men's backs are approximately equal to the heaviest 10% of women's backs
- Figure 1 below shows the frequency histograms for men's and women's players in forward and back positions

***Figure 1: Frequency histograms of mass of forwards (left panel) and backs (right panel) in elite men's and women's rugby players. Dotted lines indicate the 50th percentile, while dashed lines indicate the 98th percentile for each group.***

## Implications for injury risk - head injury models

The differences observed between men and women with respects to mass may be combined with differences in speed to create a theoretical framework in which the inertial load and forces faced by smaller and slower player is significantly greater when in contact with a larger, faster player. this model is intended for illustrative purposes and demonstrates the impact of only one variable known to differ between biological males and females - namely mass - on head injury risk, in a basic parametric model, absent force application and complex movements, as a preliminary impact analysis. the principles illustrated by the model would apply to other injuries. The addition of speed, and strength or force exerted during contact would further increase the implications of the findings of this illustrative model, summarized below.

The representative figure below illustrates the concept of mass disparity as a risk of injury for ball carriers. It depicts the linear acceleration (A), angular acceleration (B), neck force (C) and neck moment (D) experienced by ball carriers of different masses when tackled by players with different masses. Using the known masses of international rugby player, the position of the average male (M50) and average female (F50) are plotted on each heat map. F90 shows the scenario where a tackler (T) corresponds to the 90th percentile for women's mass (see Figure 1) tackles a typical female mass ball carrier (BC). X indicates the 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.

***Figure 2. Graphical representations of linear acceleration (A), angular acceleration (B), neck force (C) and neck moment (D) for ball carriers of different masses during tackles by tacklers with different masses. Mso and Fso show the modelled situation when typical/median players tackle one another for men and women, respectively. F90 represents a female ball-carrier with typical mass against a tackler in the heaviest 10% of women's body mass. X denotes the cross-over situation that would hypothetically occur for a tackler at the men's median mass tackling a typical female ball carrier***

The modelling shows that a tackle involving players with typical or average mass produces slightly greater accelerations and forces in men (Mso) than in women (Fso). This is a function of the higher mass of men's players. Head and neck kinematic and kinetic variables increase significantly when the heaviest 10% of women's body mass is used for the tackler against a typical ball carrier (F90), but this extreme "within female-bodied" scenario produces smaller kinetic and kinematic outcomes than if the hypothetical cross-over scenario were to occur, where an average male-bodied player is the tackler and the average female-bodied player the ball carrier **(X).** The magnitude of the increase in neck forces, moments and accelerations for the ball carrier is between 20% and 30% for typical cross-over scenario compared to the typical female vs female scenario, and is 10% greater for the male-bodied vs female-bodied crossover scenario than a tackle where the heaviest 10% of women are matched against typical women's mass (F90).

compared to the typical tackle scenario in women's rugby. The magnitude of these extreme head accelerations and neck forces are not seen in women and are created by cross-over of male-bodied players to women's rugby. Similar differences are seen when examining the accelerations and forces for the tackler's head and neck.

The magnitude of the known risk factors for head injury are thus predicted by the size of the disparity in mass between players involved in the tackle. The addition of speed as a biomechanical variable further increases these disparities, which is relevant given that male players weighing 103kg (the median for men) would be expected to run between 10% and 15% faster than typical female players (mass 73kg), and thus considerably faster than female players who are heavier than the median (eg females at the 90 $^{th}$ percentile, Fig 1). This would further compound the disparity created.

Next, it is important to also consider that these models do not account for the ability of players to actively exert force at high rates during tackles. This would be a function of power and strength, which are similarly known to be 30% to 80% greater in biological males than females. When these active applications of force during contact are added to the mass and speed characteristics illustrated and described above, the resultant neck and head forces and accelerations will increase even further, such that the illustrative model shown above depicts the smallest possible risk increase for typical players involved in a tackle as a result of mass alone. The addition of speed and force disparities will increase the magnitude of these risk factors beyond the 20% to 30% we illustrate above.

The implication of these increases is complex to quantify but would result in increased injury risk for the player experiencing the elevated risk outcomes (force and acceleration). This is because head injuries occur when forces and accelerations on the head and neck reach a threshold necessary to cause injury, and which is unique to each tackle situation. A tackle situation that typically produces risk factors within 20% of this threshold would, in the circumstance of a typical male-bodied vs typical female-bodied player illustrated above, be sufficiently increased to cause an injury. The higher risk scenario involving heavier male-bodied players would further increase injury likelihood, since all tackle situations that normally produce kinetic and kinematic variables within 40% to 50% of an injury threshold would now exceed it, a

causing head injury.

Finally, it must also be considered that the ability to withstand or tolerate forces on the head and neck are required to avoid brain injury. This is the reason neck strength is critical in injury prevention. Since the strength disparities between males and females is so large, including a 50% lower neck isometric strength in females, the reduced ability of female players to tolerate or withstand the forces in tackles is a further risk factor for injury, including head injury as described above, but relevant to all injuries where the rapid application of force or load are responsible for injury.

## Implications for injury risk - scrum forces

The implication of greater mass and force-producing ability in males can be seen in forces measured during scrums in both elite and community level rugby. Research on the forces applied during scrums shows that at the elite level, males produce approximately twice the peak force of females in the scrum. Even at the community level, where peak force is 30% lower than in the elite game, males produce approximately 40% greater peak force during scrums than elite females. Given that force producing and receiving ability is likely to be significantly lower in female community players, the implication is that men's community level rugby scrums will be considerably more forceful than women's community level scrums.

The risk of particularly serious and catastrophic injuries during scrums has led to a number of law changes specifically designed to depower the scrum to reduce injury risk. This risk would be amplified by large mismatches in strength between opposing players, since the force applied must be withstood by a direct opponent. This is an illustration of how mismatches in strength and size are directly responsible for forces that result in injury.

It must be noted that the actual testosterone level, measurable in the body, is not a strong predictor of performance within men and within women [27-29]. This is because performance is multifactorial, and testosterone's androgenzing effects contribute to, but do not solely influence the biology and resultant performance outcomes within a group who are able to utilize it. The biological basis for male vs female differences is thus the result of testosterone, but it does not necessarily follow that within men and within women, the hormone is a predictor of performance.

Further, differences in the sensitivity to testosterone between individuals mean that a given level of testosterone is not a sensitive or specific predictor of performance within each group (males and females). This is in part because most males have elevated levels and some degree of sensitivity, while the level in females is significantly lower and rarely exceeds even the very low end of the male range [1]. Therefore, in two homogenous groups that are matched for either the presence or absence of a given variable (males and females for the presence or absence of testosterone, in this case), the predictive value of that variable within a group is greatly diminished, the same way that V02max is a significant predictor of running or cycling performance across the whole population, but not within a group of elite marathon runners or cyclists, who are already relatively homogenous for that characteristics [30]. Similarly, height is clearly advantageous for professional basketball, but within the National Basketball Association (NBA), where height has already been selected for and participants are in the extreme upper end of the overall population for that characteristic [31], it becomes a poorer predictor of performance.

However, when the same question -does testosterone predict performance across humans of both sexes - is asked of binary categories (males vs females in sport, rather than within males or females), then the predictive power of testosterone is strong, because "high testosterone" during adulthood is a very reliable indicator that the androgenizing effects of testosterone have occurred earlier during life. When understood and assessed this way, testosterone is necessary for peak performance (since the top performers within humans are all male), but it is not sufficient to attain it. It is here that the almost perfect sensitivity of biological sex emerges, since in a ranking list of the top thousand performances in sport, every year, every single one will be biologically male.

In summary, across all performance levels and ages after puberty, testosterone is primarily (though not exclusively) responsible for very large typical differences in the biology of males and females, and consequently, performances between the sexes. These are summarized in Figure 3 below, which combines the biological differences between males and females with their performance implications, and is reproduced from a recent article currently in review [10].

## *Figure 3: Summary comparison of biological (left table) and performance (right figure) differences between males and females for a range of biological variables and physical activities/events. Reproduced from Hilton & Lundberg [10]*

Given that the women's category exists to ensure protection, safety and equality for those who do not benefit from the biological advantage created by these biological performance attributes, the relevant and **crucial question is whether the suppression of testosterone for a period of 12 months, currently required for transgender women participation in women's sport, is sufficient to remove the biological differences summarized above?**

Current policies regulating the inclusion of transgender women in sport are based on the premise that reducing testosterone to levels found in biological females is sufficient to remove many of the biologically-based performance advantages described above. However, peer-reviewed evidence suggests that this is not the case, and particularly that the reduction in total mass, muscle mass, and strength variables of transgender women may not be sufficient in order to remove the differences between males and females, and thus assure other participants of safety or fairness in competition.

Based on the available evidence provided by studies where testosterone is reduced, the biological variables that confer sporting performance advantages and create risks as described previously appear to be only minimally affected. Indeed, most studies assessing mass, muscle mass and/or strength suggest that the reductions in these variables range between 5% and 10% (as described by Hilton & Lundberg [10]). Given that the typical male vs female advantage ranges from 30% to 100%, these reductions are small and the biological differences relevant to sport are largely retained.

For instance, bone mass is typically maintained in transgender women over the course of at least 24 months of testosterone suppression, with some evidence even indicating small but significant increases in bone mineral density at the lumbar spine [32-34]. Height and other skeletal measurements such as bone length and hip width have also not been shown to change with testosterone suppression, and nor is there any plausible biological mechanism by which this might occur, and so sporting advantages due to skeletal differences between males and females appear unlikely to change with testosterone reduction.

With respects to strength, **1** year of testosterone suppression and oestrogen supplementation has been found to reduce thigh muscle area by 9% compared to baseline measurement [35]. After 3 years, a further reduction of 3% from baseline measurement occurred [36]. The total loss of 12% over three years of treatment meant that transgender women retained significantly higher thigh muscle size ($p<0.05$) than the baseline measurement of thigh muscle area in transgender men (who are born female and experience female puberty), leading to a conclusion that testosterone suppression in transgender women does not [illegible]e muscle size to female levels [36].

find that **1** year of testosterone suppression to female typical reference levels results in a comparatively modest loss of lean body mass (LBM) or muscle size, with consistent changes between 3% and 5% reduction in LBM after 1 year of treatment (as summarized from source research studies by Hilton & Lundberg [10]).

Muscle force-producing capability is reduced after testosterone suppression, though as appears to be the case for muscle/lean mass, these reductions are considerably smaller in magnitude that the initial male-vs-female differences in these variables. For instance, hand-grip strength was reduced by 7% and 9% after **1** and 2 years, respectively, of cross hormone treatment in transgender women [39], and by 4% in 249 transwomen after 1 year of gender-affirming treatment, with no variation between different testosterone levels, age or BMI tertiles [45]. Transgender women retained a 17% grip-strength advantage over transgender men at baseline measurement, with a similarly large, retained advantage when compared to normative data from a reference or comparison group of biological females.

Most recently, Wiik et al found that isokinetic knee extension and flexion strength were not significantly reduced in 11 transgender women after 12 months of testosterone suppression, with a retained advantage of 50% compared to a reference group of biological females and the group of transgender men at baseline [41]. This absence of a reduction in strength occurred in conjunction with a 4% to 5% reduction in thigh volume, and no difference in the contractile density of the muscle, which suggests that the reduction of testosterone for a period of a year had no effect on the force-producing capacity per unit of cross sectional area [41], a variable that is known to be higher in males than females.

In conclusion, longitudinal research studies that have documented changes in lean mass, muscle mass/area and strength show consistently that small decreases occur as a result of testosterone suppression, with a resultant relatively large retained advantage in these variables compared to a group of biological females.

Testosterone exerts significant biological effects on biological males during puberty and adolescence. This creates large differences in strength, mass, speed, power, and endurance capacity. In turn, these create player welfare concerns and performance inequality in rugby, given the importance of physical contact and strength in the sport. Longitudinal research studies on the effect of reducing testosterone to female levels for periods of 12 months or more do not support the contention that variables such as mass, lean mass and strength are altered meaningfully in comparison to the original male-female differences in these variables. The lowering of testosterone removes only a small proportion of the documented biological differences, with large, retained advantages in these physiological attributes, with the safety and performance implications described previously. There is currently no basis with which safety and fairness can be assured to biologically female rugby players should they encounter contact situations with players whose biologically male advantages persist to a large degree.

While there is overlap in variables such as mass, strength, speed and the resultant kinetic and kinematic forces we have modelled to explore the risk factors, the situation where a typical player with male characteristics tackles a typical player with female characteristics increases the magnitude of known risk factors for head injuries by between 20% and 30%. In the event of smaller female players being exposed to that risk, or of larger male players acting as opponents, the risk factors increase significantly, and may reach levels twice as large, at the extremes. The basis for regulation is the typical scenario, though risk mitigation must be mindful of the potential for worst-case scenarios that may arise. Both are deemed unacceptably high, because they would result in a number of tackle situations that currently lie beneath a threshold required to cause injury increasing to exceed that threshold.

Thus, it is on the basis of male vs female biological differences, combined with no evidence for removal of their implications for safety and performance, that the guideline is that trans women should not compete in women's rugby.

It is acknowledged that the published studies currently available on testosterone suppression and physiological changes (compiled and described in Hilton and Lundberg, 2020 and reviewed individually in the proposed policy document) have been conducted in untrained transgender women. This invites questions over the validity and generalizability of the studies to a sports-playing population.

This is a valid question, and it is acknowledged that research is required to fully address questions arising out of this limitation. World Rugby is committed to supporting high quality research proposals in this area, should they be submitted as part of World Rugby's Research programme.

However, this limitation can also be assessed within an understanding of the physiological implications of trained compared to untrained individuals undergoing testosterone suppression. The application of insights from complementary studies leads to a conclusion that the available research is in fact sufficient to arrive at firm conclusions about safety, performance and retained advantages, and thus the recognized limitations are not sufficient to refrain from drawing a conclusion on the likely implications of the transgender research for athletes.

In assessing this issue, two primary questions may be asked:

1. How would training undertaken during the process of testosterone suppression affect the changes observed in muscle and lean body mass, and strength variables, compared to studies done in individuals who do not perform training?
2. How would training prior to a period of testosterone suppression influence:
1. The baseline or pre-suppression measures for muscle mass and strength in transgender women, and thus the differences in these variables compared to a reference or control group of biological women (cisgender women)?
2. The likely "end-point" for muscle and lean body mass as well as strength after the testosterone suppression for a period of at least twelve months, once again compared to a reference or comparison of cisgender women?

**Training during the intervention to lower testosterone levels** can reduce, eliminate, and even reverse any losses in muscle and lean body mass, muscle volume, and muscle strength. This is supported by evidence from various study models in which biological males reduce testosterone to within the female range, and are able to maintain or even increase these physiological variables through training [46-48].

> The implication is that any performance decline as a result of androgen deprivation is minimized or eliminated, and so the studies cited in support of the World Rugby Guideline, while conducted on non-training individuals, establish the minimum possible retained advantage for trans women. That is, they establish that in the absence of training during testosterone suppression, an advantage is retained compared to cisgender women. That advantage is either the same, or very plausibly increased as a result of training.

**Training prior to the intervention** will cause increased muscle mass and strength variables at baseline. This means that the initial or "pre-suppression" differences in these variables compared to biological females will be greater than in an untrained trans woman. This rebuts the assertion that trans women are weaker, less muscular and thus more similar to biological females at baseline, within a sporting context, since the transgender woman being considered by World Rugby is much more likely to be trained (or will train once transition begins, as described above).

> Further, once the period of testosterone suppression begins, then the degree to which muscle mass and strength decreases may be either the same or relatively greater in the trained trans women as a result of this higher baseline. Even if the relative loss of muscle mass and strength are higher than in untrained trans women, it is inconceivable, and even physiologically impossible, that a pre-trained athletic trans woman is going to lose so much muscle mass and strength that they end at a point where they are less muscular/lean and weaker than a theoretically untrained (and even 'self-starved') transgender woman.

advantage for a pre-trained trans woman. The effect of training can only be to increase this value or to achieve the same value of X percent retained advantage, but it cannot reduce it further, unless one argues that a trained trans woman will lose so much lean mass and strength that they end up weaker and less muscular than a completely non-athletic individual.

Finally, it is relevant that studies comparing untrained biological males and highly trained females, males retain an advantage despite the training status of biological females. For instance, in a study on grip strength, the strongest elite athletically trained females in sports where grip strength is a performance advantage Uudo and handball) are only as strong as untrained biological males at the 58th percentile, with a 26% difference in strength between typical elite females and typical untrained males [49]. Similarly, Morrow & Hosler (1981) found that untrained college-aged males were more than twice as strong as trained female basketball and volleyball players in a bench press task, with the top 5% strongest trained females equal in strength to the weakest 14% of untrained males This establishes that pre-trained biological females can increase strength beyond that of untrained females, but still do not compare to untrained biological males.

The implication is also that since even typical untrained biological males have a large strength advantage compared to elite and trained females, studies that have documented only small reductions in strength and thus persistence of strength advantages with androgen deprivation in untrained biological males (as in Kvorning et al [46], Chen et al [47] and in studies on transgender women cited herein) should be considered relevant for establishing the smallest possible retained advantage that would exist in the absence of training. As described above, and in studies where training is conducted while testosterone is suppressed [46-48], the advantage will only remain this size or increase.

Finally, it is also recognized that not all sports are affected similarly by the variables we have weighted as crucial for rugby (size, strength, speed, power). Indeed, in some sports, excess mass may be disadvantageous, and thus the model for retained advantage and persistent risk may present differently for different physical activities.

physiological and performance advantages and those to whom it does not are removed sufficiently to enable participation of transgender women in women's rugby. At the present time, however, based on the best published scientific evidence, that position is unsupported.

The referenced research used to support this position can be viewed **here**.

## Conclusion - Testosterone, Welfare and Performance

Having considered all of the currently available information, the working group determined that the best evidence **currently** available means that those who experienced the biological effects of testosterone during puberty and adolescence cannot safely or fairly compete in women's rugby. That means that currently, transgender women may not compete in women's rugby.

World Rugby is committed to encouraging transgender people to remain involved with rugby and is currently funding research to continue to review any evidence that may emerge to enable the participation of transgender women in women's rugby. Details of the research currently underway, along with details of how to apply for research funding for those who may be interested, is available **here**.

## How do I stay involved in rugby if I can no longer play in the category that I want to?

World Rugby acknowledges that the introduction of this Guideline will mean that some players can no longer play in the category that they want to. It is possible that will change in the future and World Rugby is funding research to try to find out if there are ways to allow that safely and fairly (see **here** for details). In the meantime, there are many other ways to stay involved with rugby:

- Other forms of the game: Many forms of non-contact Rugby exist such as: Tag; Touch; Flag etc all have

courses for coaches of children, adolescents, and adults. All courses are open to any participant.

- Refereeing: For many people who may not be able to play, refereeing is a viable alternative to stay close to the game. World Rugby and its member Unions offer several introductory courses and a pathway exists in all Unions for fast-tracking talented
- Administration: All clubs rely on volunteer administrators. As individuals enter the latter stages of the long-term participant model, then administration becomes a realistic outlet for many. A number of Unions have dedicated support resources for individuals who wish to pursue this path of staying involved.

World Rugby is currently exploring the possibility of an "open category" of rugby in which any player could play, regardless of gender identity. World Rugby has committed to exploring this option with its Unions, Associations, International Rugby Players, and trans advocate groups including Gendered Intelligence and International Gay Rugby.

## What if I have concerns about safety or fairness relating to someone I am playing with or against?

It is important to note that many people do not meet cultural or local norms or stereotypes related to the expression of gender identity. All players and Unions ought to take care to consider this when raising any concerns about another player. In the event that a player or Union has a genuine concern about safety or fairness in relation to another player, the concern should be dealt with as follows:

1. The concerned person should raise their concerns with their Union's Chief Medical Officer (CMO).
2. The Union's CMO should carefully consider the concerns raised, in the context of all of the known facts and if having done so, the CMO determines that the concerns are not frivolous or vexatious, the CMO should contact the World Rugby CMO setting out the basis for the
3. The World Rugby CMO will engage with the CMO of the Union of the player about whom the concerns h been raised, ensuring confidentiality for the player and involved team-mates and opponents ghout the

5. In some circumstances, such appropriate actions may include a recommendation that a standardised endocrinological assessment be performed [Appendix].

6. For the avoidance of doubt, no player should or would be forced to undergo any medical or other assessment. It is a player's responsibility to decide on whether he or she wishes to proceed with any assessment. However, it should be noted that deciding not to participate in an assessment, having been requested to do so, may have consequences in terms of the player's eligibility to participate in the category of competition that is consistent with his/her/their gender identity, since it may not be possible to determine whether issues of safety or fairness arise without such assessment.

# Exhibit 25

STATE OF ARIZONA
OFFICE OF THE GOVERNOR

DOUGLAS A. DUCEY
GOVERNOR

EXECUTIVE OFFICE

March 30, 2022

The Honorable Katie Hobbs
Secretary of State
1700 W. Washington, 7th Floor
Phoenix, AZ 85007

RE: Senate Bill 1138 irreversible gender reassignment surgery; minors & Senate Bill 1165 interscholastic; intramural athletics; biological sex

Dear Secretary Hobbs,

Today I signed S.B. 1138 and S.B. 1165, legislation to protect participation and fairness for female athletes, and to ensure that individuals undergoing irreversible gender reassignment surgery are of adult age. This legislation is common-sense and narrowly-targeted to address these two specific issues — while ensuring that transgender individuals continue to receive the same dignity, respect and kindness as every individual in our society.

S.B. 1138 delays any irreversible gender reassignment surgery until the age of 18. The reason is simple, and common sense – this is a decision that will dramatically affect the rest of an individual's life, including the ability of that individual to become a biological parent later in life.

Distinguishing between an adult and a child in law, as this bill does, is not unique. Throughout law, children are protected from making irreversible decisions, including buying certain products or participating in activities that can have lifelong health implications. These decisions should be made when an individual reaches adulthood. Further, many doctors who perform these procedures on adults agree it is not within the standards of care to perform these procedures on children.

The irreversible nature of these procedures underscores why such a decision should be made as an adult, not as a child, and further supports the importance of this legislation.

S.B. 1165 creates a statewide policy to ensure that biologically female athletes at Arizona public schools, colleges, and universities have a level playing field to compete. This bill does not deny student-athletes the eligibility to play on teams not designated as "female," and it doesn't impact club sports leagues offered outside of schools. Every young Arizona athlete should have the

opportunity to participate in extracurricular activities that give them a sense of belonging and allow them to grow and thrive.

This legislation simply ensures that the girls and young women who have dedicated themselves to their sport do not miss out on hard-earned opportunities including their titles, standings and scholarships due to unfair competition. This bill strikes the right balance of respecting all students while still acknowledging that there are inherent biological distinctions that merit separate categories to ensure fairness for all.

Sincerely,

Douglas A. Ducey
Governor
State of Arizona

cc: The Honorable Karen Fann
The Honorable Rusty Bowers
The Honorable Warren Petersen
The Honorable Nancy Barto