EXHIBIT 35

Cochrane Library

Cochrane Database of Systematic Reviews

Antiandrogen or estradiol treatment or both during hormone therapy in transitioning transgender women (Review)

Haupt C, Henke M, Kutschmar A, Hauser B, Baldinger S, Saenz SR, Schreiber G

Haupt C, Henke M, Kutschmar A, Hauser B, Baldinger S, Saenz SR, Schreiber G.
Antiandrogen or estradiol treatment or both during hormone therapy in transitioning transgender women.
Cochrane Database of Systematic Reviews 2020, Issue 11. Art. No.: CD013138.
DOI: 10.1002/14651858.CD013138.pub2.

www.cochranelibrary.com

TABLE OF CONTENTS

Cochrane Library
Trusted evidence.
Informed decisions.
Better health.

Cochrane Database of Systematic Reviews

[Intervention Review]

Antiandrogen or estradiol treatment or both during hormone therapy in transitioning transgender women

Claudia Haupt[1], Miriam Henke[2], Alexia Kutschmar[3], Birgit Hauser[4], Sandra Baldinger[5], Sarah Rafaela Saenz[6], Gerhard Schreiber[7]

[1]Medical Service for Transgenders Lucerne, Lucerne, Switzerland. [2]Independent Transgender Consultant, Rothenburg, Germany. [3]Independent Transgender Consultant, Gaimersheim, Germany. [4]Gynaecology and Obstetrics, HRT Transgender Center Medical Practice, Hohenstein, Germany. [5]Independent Transgender Consultant, Lucerne, Switzerland. [6]Independent Transgender Consultant, Ludwigshafen, Germany. [7]Institute for Theology and Social Ethics, Technical University Darmstadt, Darmstadt, Germany

Contact address: Claudia Haupt, ddr.haupt@hin.ch.

Editorial group: Cochrane Tobacco Addiction Group.
Publication status and date: New, published in Issue 11, 2020.

Citation: Haupt C, Henke M, Kutschmar A, Hauser B, Baldinger S, Saenz SR, Schreiber G. Antiandrogen or estradiol treatment or both during hormone therapy in transitioning transgender women. *Cochrane Database of Systematic Reviews* 2020, Issue 11. Art. No.: CD013138. DOI: 10.1002/14651858.CD013138.pub2.

ABSTRACT

Background

Gender dysphoria is described as a mismatch between an individual's experienced or expressed gender and their assigned gender, based on primary or secondary sexual characteristics. Gender dysphoria can be associated with clinically significant psychological distress and may result in a desire to change sexual characteristics. The process of adapting a person's sexual characteristics to their desired sex is called 'transition.'

Current guidelines suggest hormonal and, if needed, surgical intervention to aid transition in transgender women, i.e. persons who aim to transition from male to female. In adults, hormone therapy aims to reverse the body's male attributes and to support the development of female attributes. It usually includes estradiol, antiandrogens, or a combination of both. Many individuals first receive hormone therapy alone, without surgical interventions. However, this is not always sufficient to change such attributes as facial bone structure, breasts, and genitalia, as desired. For these transgender women, surgery may then be used to support transition.

Objectives

We aimed to assess the efficacy and safety of hormone therapy with antiandrogens, estradiol, or both, compared to each other or placebo, in transgender women in transition.

Search methods

We searched MEDLINE, the Cochrane Central Register of Controlled Trials (CENTRAL), Embase, Biosis Preview, PsycINFO, and PSYNDEX. We carried out our final searches on 19 December 2019.

Selection criteria

We aimed to include randomised controlled trials (RCTs), quasi-RCTs, and cohort studies that enrolled transgender women, age 16 years and over, in transition from male to female. Eligible studies investigated antiandrogen and estradiol hormone therapies alone or in combination, in comparison to another form of the active intervention, or placebo control.

Data collection and analysis

We used standard methodological procedures expected by Cochrane to establish study eligibility.

Main results

Our database searches identified 1057 references, and after removing duplicates we screened 787 of these. We checked 13 studies for eligibility at the full text screening stage. We excluded 12 studies and identified one as an ongoing study. We did not identify any completed studies that met our inclusion criteria. The single ongoing study is an RCT conducted in Thailand, comparing estradiol valerate plus cyproterone treatment with estradiol valerate plus spironolactone treatment. The primary outcome will be testosterone level at three month follow-up.

Authors' conclusions

We found insufficient evidence to determine the efficacy or safety of hormonal treatment approaches for transgender women in transition. This lack of studies shows a gap between current clinical practice and clinical research. Robust RCTs and controlled cohort studies are needed to assess the benefits and harms of hormone therapy (used alone or in combination) for transgender women in transition. Studies should specifically focus on short-, medium-, and long-term adverse effects, quality of life, and participant satisfaction with the change in male to female body characteristics of antiandrogen and estradiol therapy alone, and in combination. They should also focus on the relative effects of these hormones when administered orally, transdermally, and intramuscularly. We will include non-controlled cohort studies in the next iteration of this review, as our review has shown that such studies provide the highest quality evidence currently available in the field. We will take into account methodological limitations when doing so.

PLAIN LANGUAGE SUMMARY

Does hormone therapy help transgender women undergoing gender reassignment to transition?

Background

Transgender women may feel that they have been born in a body with the wrong sexual characteristics. This may result in significant psychological distress (gender dysphoria) and the desire to adapt their male physical and sexual characteristics to be more consistent with their experienced female gender. This is a process called transition. If measures to aid transition are not taken, this can result in greater psychological distress. One of the medical treatments given to help transgender women with male bodies to achieve transition is synthetic female hormones. These hormones can be taken by mouth, absorbed through the skin or injected into muscle.

Study characteristics

We looked for randomised controlled trials (RCTs) that included transgender women (age 16 and over) in transition from male to female. RCTs are a type of research study that can reduce the possibility of several types of bias. To be included in this review, studies needed to compare different hormone treatments used to support transgender women to transition (oestrogen alone, testosterone blockers alone, or oestrogen in combination with testosterone blockers), or compare these hormone treatments to placebos (fake or dummy treatments that appear to be the same as the actual treatment, but have no medical effects). We wanted to see whether hormone treatments help transgender women to make a transition that they are happy with. We also wanted to look at whether there were any health risks of the treatment.

Key results

We searched for studies up to 19 December 2019. We were unable to find any relevant completed studies that we could include. We did find one ongoing study that aimed to recruit all of the people taking part in the study by the end of 2020. This study is comparing the effects of estradiol valerate plus cyproterone treatment with estradiol valerate plus spironolactone treatment in transitioning transgender women in Thailand.

Quality of evidence

Our review found no RCTs that looked at whether hormone therapies are effective and safe when used to help transgender women to transition. Therefore, high-quality RCTs are needed to research these questions.

BACKGROUND

Description of the condition

There is a growing trend towards de-psychopathologisation of transgenderism (Drescher 2014; ATME 2015). There is an emerging consensus that transgenderism is not a psychiatric disorder (WPATH 2011). For instance, the 11th Revision of the International Classification of Diseases (ICD-11) (WHO 2018) no longer classifies transgenderism as a behavioural and personality disorder, but has instead drafted the term "gender incongruence" to describe gender dysphoria.

In contrast, the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM–5) (DSM-5 2013) describes gender dysphoria as a "marked incongruence between one's experienced/expressed gender and assigned gender, of at least six months duration, as manifested by at least two of the following" characteristics:

- A marked incongruence between one's experienced/expressed gender and primary and/or secondary sex characteristics (or, in young adolescents, the anticipated secondary sex characteristics);
- A strong desire to be rid of one's primary and/or secondary sex characteristics because of a marked incongruence with one's experienced/expressed gender (or, in young adolescents, a desire to prevent the development of the anticipated secondary sex characteristics);
- A strong desire for the primary and/or secondary sex characteristics of the other gender;
- A strong desire to be of the other gender (or some alternative gender different from one's assigned gender);
- A strong desire to be treated as the other gender (or some alternative gender different from one's assigned gender);
- A strong conviction that one has the typical feelings and reactions of the other gender (or some alternative gender different from one's assigned gender).

Gender dysphoria has been defined as associated with "clinically significant distress or impairment in social, occupational or other important areas of functioning" (Zucker 2016), which may lead to substantial suffering in affected people (Deutsch 2016a; Soll 2018). Gender dysphoria may result in the desire to modify one's physical and sexual characteristics to be consistent with those of the experienced gender. This process of adaptation is called transition.

The treatments applied in transition differ from those used for maintenance of the new sexual characteristics. Currently, there is uncertainty about the value of hormone therapy as a sole intervention, or when combined with surgery, for transition from male to female. This Cochrane Review specifically focuses on 'transgender women in transition from male to female,' a definition that includes biological males aiming to adapt their sexual characteristics to be consonant with those of females.

A meta-analysis that analyzed 21 studies on the prevalence of gender dysphoria (of which 12 studies contained evaluable data) estimated an overall prevalence of transgender women with gender dysphoria at 6.8 per 100,000 individuals (Arcelus 2015).

Description of the intervention

Current guidelines suggest hormonal and, if needed, surgical treatment of gender dysphoria in transgender women (WPATH 2011). Hormone therapy aims to suppress the development of, or to reverse, male attributes that have already developed. At the same time, hormones aim to develop female attributes. However, where male characteristics have already developed in adult males, such as in the bone structure of the face, hormones are not effective. Other treatments, such as surgery, would be required to change these (WPATH 2011).

The guidelines of the Endocrine Society working group suggest treatment with both oestrogens and antiandrogens (Hembree 2017). Oestrogens can be administered as either oral oestrogen, absorbed through transdermal estradiol patches, or by injection of estradiol valerate or estradiol cypionate. The application frequency differs depending on the patient's reaction to the agent and the administration regimen; it could be multiple times per day or once every two weeks. Meanwhile, antiandrogens such as spironolactone or cyproterone acetate (CPA) are commonly taken orally. Additionally, it is possible to block male puberty by treatment with gonadotropin-releasing hormone (GnRH) agonist injections (Hembree 2017).

While not every transgender woman undergoes hormone therapy in her transition, this intervention is still widely used (Hembree 2017). We know of no studies identifying the ratio of patients who undergo hormone therapy, nor do we know of studies investigating how much time passes between the start of transition (the decision to transition) and the start of hormone therapy. We are not aware of any studies on how often antiandrogens are being prescribed in addition to or instead of 17-beta-estradiol, how often they are being taken, or which kinds of androgens are in use besides CPA and spironolactone.

How the intervention might work

Several hormonal substances and combinations are used clinically for hormone therapy in transitioning women. CPA is a progestin, steroidal anti-androgen and anti-gonadotropin that blocks the receptors for testosterone (T) and dihydrotestosterone (DHT), and thereby prevents these steroidal hormones from exerting their androgenic effects. Hence, it stops processes like body hair growth, hair loss on the head, male body fat distribution and others (Figg 2010; WPATH 2011). According to the World Professional Association for Transgender Health (WPATH) guidelines, it is possible to suppress puberty with GnRH analogues or progestins such as medroxyprogesterone (WPATH 2011).

Spironolactone acts as a weak androgen receptor antagonist (Wenqing 2005). It also causes an increase in oestradiol levels (Thompson 1993), so that further virilisation is prevented and feminisation occurs (WPATH 2011).

17-beta-estradiol is used to feminise the external appearance (WPATH 2011). It binds to oestrogen receptors and thus ensures gene expression, which in turn feminises appearance (Hye-Rim 2012). In addition, estradiol suppresses gonadal testosterone production via the control systems of the hypothalamus (Hayes 2000).

Feminisation therapy aims to adapt the physical appearance and experience of the male body to that of a female body, by

inducing breast growth, softening facial features, and inducing other physical changes commonly considered to comprise a feminine appearance (WPATH 2011). For this purpose, oral or transdermal oestrogen is recommended, and therapy with oestrogen in combination with antiandrogens is most common. Co-treatment with antiandrogens minimises the required dose of oestrogen, and thereby reduces the potential risks of oestrogen identified in previous studies (Schürmeyer 1986; Prior 1989). Some antiandrogens are approved by WPATH, such as spironolactone, cyproterone acetate, GnRH analogues like goserelin, and 5-alpha-reductase inhibitors like finasteride (WPATH 2011).

Why it is important to do this review

Antiandrogens like CPA and spironolactone are prescribed to transgender women in transition by clinicians, including gynaecologists and endocrinologists (Schneider 2006; Flütsch 2015), and they are commonly considered to be valuable drugs to support transition (WPATH 2011; Hembree 2017). However, clinical evidence suggests that taking these drugs can result in adverse events; for example, CPA has significant potential for causing depression and for worsening depressive symptoms (Seal 2012). There is also some concern that CPA can lead to other psychiatric, neurological, and metabolic disorders (Griard 1978; Ramsay 1990; Oberhammer 1996; Giltay 2000; Calderón 2009; Bessone 2015). The most common adverse effects of spironolactone are hyperkalaemia, dehydration and hyponatraemia (Greenblatt 1973). Furthermore, spironolactone might have an influence on feelings of anxiety (Fox 2016).

Other studies from the 1980s and 90s reported that there were adverse effects from high-dose estradiol, but these studies used ethinyl estradiol or equine premarin (equine estradiol) instead of bioidentical 17-beta-estradiol; and used progestins, instead of bioidentical progesterone. This may have contributed to the adverse effect profile of these specific treatments (Prior 1989). Unlike the bioidentical alternatives used today (hormone preparations made from plant sources that are similar or identical to human hormones), substances administered in the past (e.g. equine oestrogens, ethinyl estradiol) were associated with more diverse adverse effects like thrombophilia, cardiovascular problems, breast and prostate cancer, as well as liver, adrenal gland and neural dysfunction (Griard 1978; Calderón 2009; Asscheman 2011). The health risks attributed to estradiol doses high enough to suppress androgens have not been found in the parenteral or transdermal application of bioidentical estradiol (Hembree 2017). Thus, it is unclear why those estradiol doses should be kept low in order to make the addition of androgen antagonists like CPA or spironolactone necessary.

In light of discussions among experts (Seal 2012; Wierckx 2014), and current recommendations for hormonal gender affirmation treatment (WPATH 2011) (which are strongly based on the values and preferences of health consumers), it is necessary to review the evidence from trials that show results for outcomes such as feminisation, satisfactory sexual function, reduced gender dysphoria, and improved quality of life (e.g. Murad 2010).

In 2017, the overall quality of evidence relating to these outcomes was classified as low (Hembree 2017). In 2011, WPATH summarised the situation as follows. "There is a need for further research on the effects of hormone therapy without surgery, and without the goal of maximum physical feminisation or masculinisation" (WPATH 2011). It is necessary to determine whether subsequent trials have provided additional evidence for efficacy, or whether there is still a lack of evidence for these desired outcomes.

OBJECTIVES

We aimed to assess the efficacy and safety of hormone therapy with antiandrogens, estradiol, or both, compared to each other or placebo, in transgender women in transition.

METHODS

Criteria for considering studies for this review

Types of studies

We aimed to include randomised controlled trials (RCTs), quasi-RCTs and controlled cohort studies.

We chose to include quasi-RCTs and cohort studies due to the low prevalence of the condition and the consequent current scarcity of RCTs (WPATH 2011).

Types of participants

We aimed to include studies that enrolled transgender women, age 16 years and over, in transition from male to female. Transitioning is defined as the process of changing one's gender profile or sexual characteristics (or both) to accord with one's sense of gender identity (WPATH 2011). Transition as a concept thus encompasses several aspects, e.g. social, psychological, or physical aspects, or a combination of these. There is consistency in the literature on when the transition begins: namely, with the decision to change a person's gender assignment (Brown 1996). However, we did not differentiate among any supposed phases of the respective types of transitions. Depending on the personal situation, the process of transition (which may include the decision to transition, gathering of information, gathering of experience, medical treatment and change of social role), can take very different periods of time, usually several months to years. Therefore, it is difficult to distinguish certain 'phases' of this process. When focusing on hormone therapy, the transition term can be more precisely defined. The transition process lasts as long as patients are in the process of changing their sexual characteristics (WPATH 2011).

We aimed to include studies with participants age 16 years and older because, according to currently applied guidelines, this is the age when patients start being treated with hormone therapy. Patients below this age are usually being treated with puberty blockers, which are outside the scope of this review (WPATH 2011).

Types of interventions

We considered studies evaluating hormone-based interventions only, excluding those that examined combined hormonal and either psychological or surgical treatments. We aimed to include studies reporting treatment with the following experimental interventions.

- Antiandrogens (cyproterone acetate or spironolactone) and estradiol
- Antiandrogens (cyproterone acetate or spironolactone) alone
- Estradiol alone

For the above interventions, we considered all types of administration: oral, sublingual, transdermal, subdermal and intramuscular. For estradiol, we also considered bioidentical 17-beta-estradiol, as well as synthetic derivatives.

We aimed to include the following comparator interventions.

- Any of the active interventions listed above
- Placebo

Although we consider placebo-controlled studies to be unethical (Bostick 2008), we made them eligible for inclusion in this review so that we could consider the evidence in its entirety. We did not consider interventions consisting purely of psychological treatment, spiritual support, or conversion therapy.

Types of outcome measures

For studies with repeated follow-up (i.e. reporting of outcomes at multiple time points), we regarded follow-up at three to six months as short term, six months to two years as medium term, and more than two years as long term (WPATH 2011).

We intended to include in the descriptive section of the review all studies that met the criteria for type of study, participants, intervention and comparator, regardless of outcomes reported or missing data.

Primary outcomes

- Quality of life (QoL) as measured by validated generic instruments, e.g. Quality of Life Inventory (QOLI) (Frisch 2005); or specific instruments, e.g. for body image, the Body Image Quality of Life Inventory (BIQLI) (Cash 2004); or for sexual life the Sexual Satisfaction Scale for Women (SSS-W) (Meston 2005).
- Satisfaction with change of male to female body characteristics, as measured with validated instruments
- Adverse events specific to hormone therapy, including serious adverse events

Secondary outcomes

- Severity of gender dysphoria/gender incongruence, e.g. as measured with the Utrecht Gender Dysphoria Scale (UGDS) (Schneider 2016)
- Measures of specific body changes, including:
  - breast size, e.g. by measurement of bust girth;
  - skin thickness, e.g. by echographic measurement (Laurent 2007);
  - skin sebum production, e.g. as measured by three-hour sebum collection with absorbent paper (Downing 1981; Giltay 2008; Ezerskaia 2016); and
  - hair growth, including hair density, diameter, growth rate and anagen/telogen ratio (Giltay 2000; Hoffmann 2013).
- Incidence or severity of depression.

We did not include surrogate outcomes, such as serum hormone levels (e.g. 17-beta-estradiol or testosterone). While these measures can help with monitoring the progress of hormone therapy, they are of little interest of themselves, especially since individuals require varying levels of these hormones to achieve a certain level of feminisation (Gooren 2017).

Search methods for identification of studies

Electronic searches

We searched the following electronic databases for relevant trials up to 19 December 2019 with no restrictions based on language of publication, date of publication, or publication status:

- MEDLINE via PubMed
- Cochrane Central Register of Controlled Trials (CENTRAL)
- Embase
- Biosis Preview
- PsycINFO
- PSYNDEX

Our search strategy is outlined in Appendix 1. We have successfully tested the screening methods for abstracts and titles.

Searching other resources

Had we identified any eligible studies through the electronic searches above we would have searched the reference lists of these in order to find additional relevant studies. We also searched the scientific abstracts of the last two meetings of each of the following organisations:

- American Association of Clinical Endocrinologists
- American Society of Andrology
- Berufsverband der deutschen Endokrinologen (Professional Association of the German Endocrinologists)
- Berufsverband der Frauenärzte e.V. (Professional Association of the Gynaecologists)
- Dachverband Reproduktionsbiologie und Medizin e.V. (Federal Association Reproductive Biology and Medicine)
- Deutsche Gesellschaft für Endokrinologie (German Society for Endocrinology)
- Deutsche Gesellschaft für Gynäkologie und Geburtshilfe (German Society for Gynaecology and Obstretics)
- Endocrine Society
- European Society of Gynaecological Oncology
- European Thyroid Association
- Nordrhein-Westfälische Gesellschaft für Endokrinologie und Diabetologie (North Rhine-Westphalian Society for Endocrinology and Diabetology)
- Royal College of Obstetricians and Gynaecologists
- Society for Endocrinology
- Society for Gynaecologic Investigation

We also searched the following grey literature databases:

- The New York Academy of Medicine Grey Literature Report (www.greylit.org/)
- OAIster (www.oclc.org/oaister.en.html)
- OpenGrey (www.opengrey.eu/)

Finally, in order to identify completed but unpublished or ongoing studies, we searched the following trial registries.

- ClinicalTrials.gov (www.clinicaltrials.gov/)
- metaRegister of Controlled Trials (mRCT; www.controlledtrials.com/mrct/)

- World Health Organization (WHO) International Clinical Trials Registry Platform (ICTRP) Search Portal (www.who.int/trialsearch/)
- Drugs@FDA (www.accessdata.fda.gov/scripts/cder/drugsatfda/)
- European Public Assessment Reports (EPAR; www.ema.europa.eu/ema/index.jsp?curl=pages/medicines/landing/epar_search.jsp)

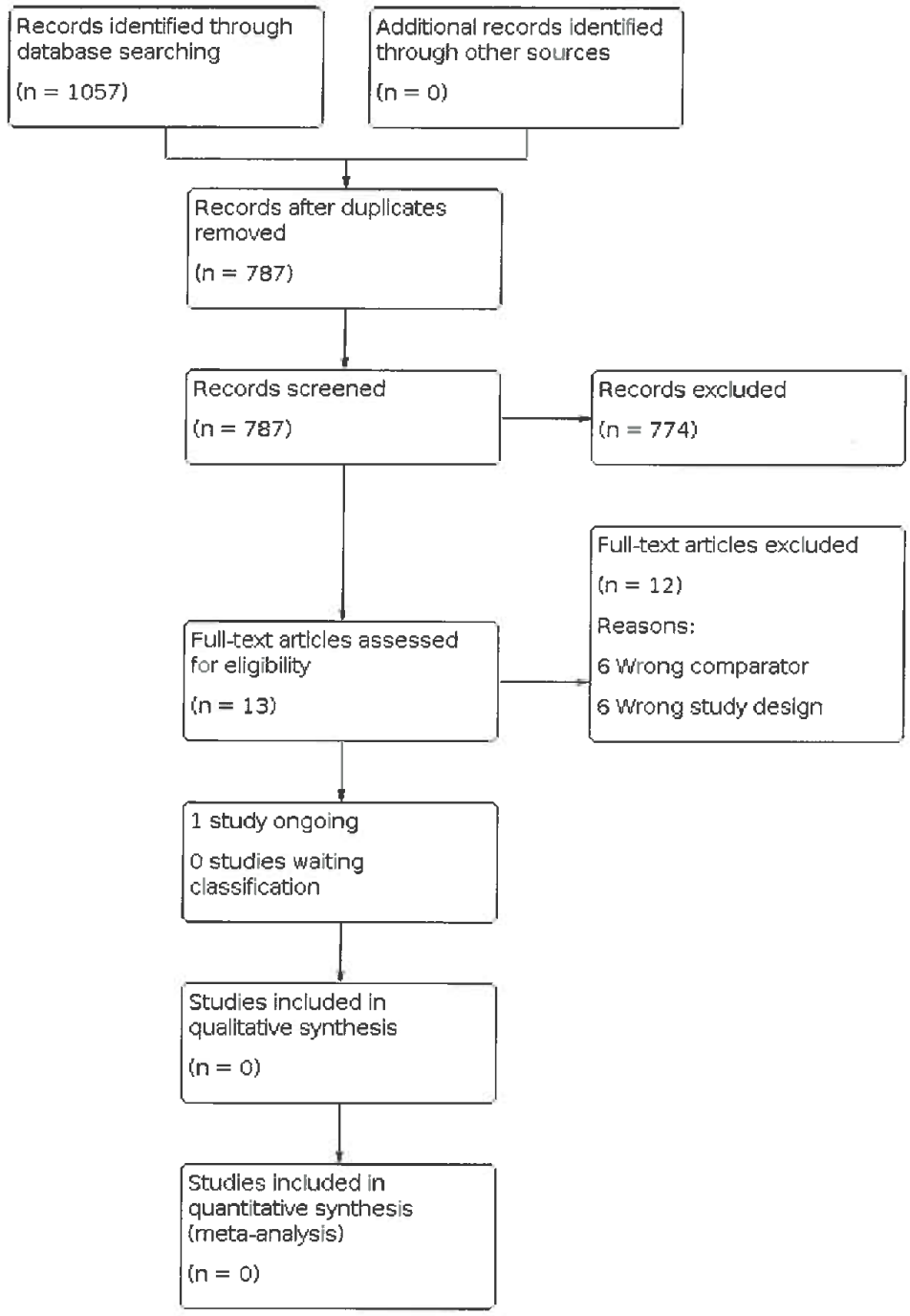
We contacted fifteen manufacturers of hormonal agents and experts in the field to identify unpublished or ongoing trials.

Data collection and analysis

Selection of studies

We used the reference management tool Covidence to identify and remove potential duplicate records of relevant studies (www.covidence.org). Two review authors (AKU and MHE) independently scanned titles and abstracts of the remaining records to compile a list of potential papers to potentially be included in the review. After this, the same review authors investigated the references in detail (as full text articles or matched records to studies), and categorised these as 'included studies,' 'excluded studies,' 'studies awaiting classification' and 'ongoing studies.' We executed this task in accordance with the criteria provided in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011a). If there had been discrepancies or if a consensus could not be reached, a third review author would have adjudicated (CHA). There were no disagreements that could not be thus resolved. Had this been the case, we would have designated the study as 'awaiting classification' and contacted the study authors for clarification. We listed studies excluded during the full text review stage, and documented the reasons for exclusion in Characteristics of excluded studies. We included an adapted PRISMA flow diagram outlining the study selection process (Moher 2009) (Figure 1).

Figure 1. Study Flow Diagram



Data extraction and management

If we had found relevant studies, two review authors (AKU and MHE) would have extracted data from all studies deemed eligible for inclusion independently, with the help of a standardized data extraction form that would have been pilot tested according to Chapter 7 of the *Cochrane Handbook* (Higgins 2011a). We have used Google Spreadsheets to manage all data gathered.

We would have collected data on the following items:

- General information on the study: first author, date of publication, study dates, publication type (full text article, abstract, unpublished), citation.
- Study methods: study design (e.g. parallel, factorial), number of study arms, study setting (single institution, multi-centre national, multi-centre international), study location, and length of follow-up.
- Participant characteristics: study inclusion/exclusion criteria, age (mean/median with range), ethnic distribution, number of participants randomised and included in analysis, participants lost to follow-up.
- Interventions: type of hormonal agents (for example CPA, estradiol, progesterone, spironolactone), dose, administration route, dosing schedule and any other associated therapies. We would have extracted data on the sample size for each intervention group.
- Outcomes: definition and method of assessment for each outcome (including the adverse event classification system used in individual studies), as well as any relevant subgroups. We would have extracted the number of events and participants per treatment group for dichotomous outcomes. We would also extract the mean, standard deviation or median and range, and number of participants per treatment group for continuous outcomes.
- Study funding sources.
- Declarations of potential conflicts of interest reported by study authors.

For each included study, we would have extracted the outcome data relevant for this review, and which would be required for the calculation of summary statistics and measures of variance. If there had been disagreements, we would have resolved them by discussion. If necessary, we would have consulted a third review author (CHA). We provided key information about potentially relevant ongoing studies, including trial identifiers, in the table of Characteristics of ongoing studies. We would have attempted to contact authors of included studies to obtain missing key data if needed.

Assessment of risk of bias in included studies

If relevant studies had been found, two review authors (AKU and MHE) would have examined all included studies to assess risk of bias (assessment of methodological quality) independently. We would have used the Cochrane 'Risk of bias' tool for assessing risk of bias in RCTs, as described in the *Cochrane Handbook* (Higgins 2011b). We would have resolved disagreements by consensus or by consulting a third review author (CHA). Our summary judgement would have included a rating (low, high or unclear risk of bias) for each domain (Higgins 2011b). We would have assessed the risk of bias for the following domains:

- Random sequence generation
- Allocation concealment
- Blinding of participants and personnel
- Blinding of outcome assessment
- Incomplete outcome data
- Selective reporting
- Other bias

We would have evaluated the risks of performance bias (blinding of participants and personnel) and detection bias (blinding of outcome assessment) separately for each outcome.

For any relevant cohort studies we would have used the ROBINS-I tool to assess risk of bias (Sterne 2016). We would have assessed each individual study in accordance with the guidance, documenting the results using a spreadsheet and providing details in 'Risk of bias' tables. We would have documented the reasons for our judgements, and would have included relevant quotations from the full-text articles or from information about the study provided by authors in the notes section of the 'Risk of bias' tables. We would have summarised the risk of bias across domains for each primary outcome in every included study, as well as across studies and domains for each primary outcome.

Measures of treatment effect

Dichotomous data

We planned to summarise dichotomous data using risk ratios (RRs), reported with 95% confidence intervals (CIs).

Continuous data

For continuous outcomes with a standard measure, we would have summarised the obtained data as mean differences (MDs) with 95% CIs. For continuous outcomes without a standard measure, we would have summarised data as standardized mean differences (SMDs) with 95% CIs. Alternatively, if the mean value and variance were missing, we would have estimated them using the methods described in Hozo 2005, which allows estimations for mean value and variance of a sample when only the median, range and size of the sample are known. We would also have considered the guidance in the *Cochrane Handbook* where appropriate (Higgins 2011c).

Unit of analysis issues

We planned to treat recurring events in individual participants as single events occurring in one participant (e.g. three episodes of major depressive disorder in one participant would have been recorded as one participant with major depressive disorder). We did not expect to include studies with interventions delivered at the cluster level.

Dealing with missing data

For studies with missing data, we would have followed the recommendations of the *Cochrane Handbook* (Higgins 2011d). We would have collected dropout rates for each study group and would have reported these in the 'Risk of bias' table. Our preferred option would have been to contact study authors in cases of missing data or statistics that were not due to participant dropout (e.g. missing statistics such as standard deviation (SD)). If missing outcome data were not provided, then we would have attempted to impute

data where possible and appropriate, and conduct sensitivity analyses to assess the effect of this on the analysis. However, where imputation is not appropriate, we would not have included the study in the respective meta-analysis, and would have discussed the potential impact of this in the text of the review. In the case of participants lost to follow-up, we would have performed meta-analyses on an intention-to-treat basis. We would have performed sensitivity analyses, excluding studies with missing outcome data, to evaluate the impact of missing data. We would have discussed the potential impact of missing data on review findings in the 'Discussion' section of the full review, using a summary table if appropriate.

Assessment of heterogeneity

We would have compared the characteristics of included studies to identify heterogeneity of content or methodology, and to determine the feasibility of performing a meta-analysis. We would have deemed meta-analyses unsuitable in cases where there was substantial content-related or methodological heterogeneity across studies. Instead, we would have used a narrative approach to data synthesis. Had meta-analyses been deemed appropriate, we would have assessed statistical heterogeneity by visually inspecting the scatter of individual study effect estimates on forest plots and by calculating the I^2 statistic (Higgins 2011c), which gives the percentage of variability in effect estimations that can be attributed to heterogeneity rather than to chance. We would have considered an I^2 of more than 50% to represent substantial heterogeneity. In the case of statistical heterogeneity, we would have conducted the prespecified subgroup and sensitivity analyses described below to investigate the source.

Assessment of reporting biases

If we had included 10 or more studies that investigated the same outcome, we would have used funnel plots to assess small-study effects and publication bias. Given that several explanations are possible for funnel plot asymmetry, we would have interpreted results carefully (Sterne 2011).

Data synthesis

Had we identified any eligible studies, we would have provided a narrative summary of the included studies. We would also have conducted meta-analyses of RCTs for all relevant outcomes, where possible, using data from studies that 1) compared the actual hormone therapy-relevant agents or combinations of agents to placebo, and 2) compared the actual hormone therapy-relevant agents or combinations of agents to other hormone therapy-agents or combinations of agents. Studies comparing two variations on the intervention would have been pooled separately to studies comparing the intervention to placebo. However, if there had been significant variability in the definition of outcomes across trials, we would have decided not to pool data.

Had we conducted meta-analyses, we would have used the Mantel-Haenszel approach to combine dichotomous data and calculate RRs with 95% CIs (Higgins 2011c). For continuous outcomes (e.g. quality of life) we would have calculated MDs or SMDs, with 95% CIs, using the inverse variance approach. Had studies reported the same outcome measure but some studies had reported data on the change from baseline (e.g. mean values and standard deviations) and others for final measurements of outcomes, they would have been placed in subgroups in the meta-analysis and pooled according to guidance in the *Cochrane Handbook* (Higgins 2011c).

For meta-analyses, we would have used a random-effects model, expecting the true effects to be related, but not the same, across all studies. We would have interpreted random-effects meta-analyses with due consideration of the whole distribution of effects, ideally by presenting a prediction interval (Higgins 2009). A prediction interval specifies a predicted range for the true treatment effect in an individual study (Riley 2011). In addition, we would have performed statistical analyses according to the statistical guidelines contained in the *Cochrane Handbook* (Higgins 2011c).

We would have summarised outcome data from cohort studies (e.g. change scores) narratively.

Subgroup analysis and investigation of heterogeneity

Wherever possible, we would have considered subgroup analyses that are structured by the following characteristics.

- Type of application of intervention (oral, transdermal, intramuscular, subcutaneous)
- Orchiectomy before or during hormone therapy

The justification for these analyses is as follows. Pharmacokinetic mechanisms lead to significant differences in the absorption and metabolism of an active substance depending on the type of application. Therefore, we would, if possible, have formed appropriate subgroups based on the application method of the intervention. Also, patients who have undergone an orchiectomy could have different outcomes than those patients without orchiectomy (Defreyne 2017).

Sensitivity analysis

We would have conducted sensitivity analyses to investigate any potential effect of removing studies judged to be at high risk of bias from meta-analyses. We would have classified studies as being at high risk of bias overall if one or more domains were judged to be at high risk. If appropriate, we would also have conducted sensitivity analyses excluding studies with missing outcome data, or where missing data have been imputed by the review author team. We would also have conducted a sensitivity analysis to compare a fixed-effect model to a random effects model where the studies in a meta-analysis appear more homogeneous than expected.

Summary of findings and assessment of the certainty of the evidence

Following standard Cochrane methodology, had we identified any included studies, we would have created a 'Summary of findings' table for all three primary outcomes. Also following standard Cochrane methodology, we would have used the five GRADE considerations (risk of bias, consistency of effect, imprecision, indirectness and publication bias) to assess the quality of the body of evidence for each outcome, and to draw conclusions about the quality of evidence within the text of the review.

RESULTS

Description of studies

Results of the search

We conducted our searches on 18 January 2019 and updated them on 19 December 2019. Through the database searches, we identified a total of 1057 references. After removing duplicates, we screened the titles and abstracts of 787 references. Through this screening, we identified 13 studies to assess as full text articles. We fully inspected these articles, and excluded 12 studies. The remaining study was still ongoing. Therefore, we did not include any studies in this review (Figure 1).

Of the manufacturers and experts in the field whom we contacted,15 responded but did not report any additional studies.

Included studies

None of the reports retrieved met the inclusion criteria for this review. Suggestions for future studies are given in Table 1

Excluded studies

We excluded all 12 of the full-text articles that we had assessed for eligibility, either because they used an ineligible comparator or because they used an ineligible study design. See Characteristics of excluded studies for further details.

Ongoing studies

We identified one ongoing RCT in Thailand, comparing spironolactone with CPA (Krasean 2019). This study started in April 2019. We describe this study in Characteristics of ongoing studies.

Risk of bias in included studies

As no studies met the inclusion criteria, it was not possible to assess risk of bias.

Effects of interventions

As no studies met the inclusion criteria, we were unable to calculate any effects of the interventions.

DISCUSSION

Summary of main results

No study met the inclusion criteria for this review. A total of 13 potentially eligible studies were identified, but ultimately all but one was excluded after we assessed the full text articles. The one remaining RCT is ongoing, and we are awaiting its publication (Krasean 2019). We conducted a comprehensive search to identify eligible studies for inclusion in this review. Despite more than four decades of ongoing efforts to improve the quality of hormone therapy for women in transition, we found that no RCTs or suitable cohort studies have yet been conducted to investigate the efficacy and safety of hormonal treatment approaches for transgender women in transition.

Overall completeness and applicability of evidence

The evidence is incomplete because no studies met the inclusion criteria for the review. This lack of studies shows a gap between current clinical practice and clinical research, which has been repeatedly emphasised (Hembree 2009; Hembree 2017). If hormone therapy is highly valued in the treatment of gender dysphoria (Hembree 2009; WPATH 2011; Hembree 2017), then this raises the question: why are there no RCTs or appropriate cohort studies for this clinical condition? There is also an ethical need for research into the efficacy and safety of hormone therapy, particularly comparing combination therapy with CPA/estradiol and spironolactone/estradiol to monotherapy with estradiol alone. In view of the reported but rather alarming side-effect profiles of CPA and spironolactone in other populations (De Bastos 2014; Khan 2016; PG12 2019), long-term clinical studies that aim to achieve adequate outcomes are urgently needed for the population of transgender women in transition.The lack of reliable data on hormone therapy for transitioning transgender women should encourage the development of well-planned RCTs and cohort studies to evaluate widespread empirical practice in the treatment of gender dysphoria.

The most common reason for the exclusion of studies from this review was the lack of a control group. We excluded some studies because they did not meet the eligibility requirements for study design (e.g. case series or case-control studies). Further, interventions were not clearly defined.

Among guideline developers in the field of transgender medicine, it has been discussed in recent years why the available evidence remains limited (Deutsch 2016a Reilly 2019). Deutsch 2016a has identified three main reasons, which they believe have hindered the development of evidence based healthcare guidelines. Firstly, a lack of research funding and institutional stigma means that the evidence currently centres around less robust study designs, such as retrospective studies, case series, and individual case reports (Bockting 2016 Reisner 2016a); secondly variation in the collection of gender identity data in observational data sets makes it difficult to identify relevant populations and monitor their health outcomes (Deutsch 2013 Bauer 2009); and finally, academic programmes focused on transgender medicine are in their infancy and few exist (Reisner 2016b), meaning there is a general lack of research and training on this topic.

Against this background, methodological problems such as inconsistent and missing comparison groups, uncontrolled confounding factors, small sample size, short follow-up time and difficulties in recording and evaluating a broad spectrum of health outcomes (physical and mental health, social functioning and QoL) have become apparent in hormone therapy (Deutsch 2016b). The performance of RCTs is controversial, especially with regard to placebo studies, and ethical and methodological objections have been raised (e.g. violation of the principle of equipoise, Miller 2003). However, the positive research potential of active-controlled RCTs is acknowledged, in order to compare different types, dosages and methods of administration of active treatments. Overall, there is a trend in the discussion to favour not only RCTs and quasi-RCTs, but also high-quality cohort studies conducted in a network of health centres, hospitals and practices (Deutsch 2016a; Deutsch 2016b).

Quality of the evidence

We could not appraise the quality of the evidence because no studies met our review's inclusion criteria.

Potential biases in the review process

We consider our search to have been consistent and comprehensive (including the fifteen contacts with manufacturers and experts in the field). At each stage, the review authors independently applied the inclusion criteria before comparing their judgements. Reliability testing was performed in the screening phase. Even though we were unable to test for publication bias, we think it is unlikely that there are studies that have been conducted but remained unpublished. The experts in the field we interviewed believed that there was a general lack of research activity by treatment manufacturers, and considered it very likely that no phase IV studies have ever been conducted in this population. For example, one expert stated that there was probably "nothing to be kept secret."

Agreements and disagreements with other studies or reviews

There are currently no systematic reviews in the Cochrane Library that evaluate the effectiveness of hormone therapy for transgender women in transition, nor are there systematic reviews that evaluate the clinical and economic impact of hormone therapy on transgender women in transition. The Endocrine Society's 2009 and 2017 guidelines addressed endocrine treatment of gender-dysphoric/gender-incongruent persons (Hembree 2009; Hembree 2017). The literature search included in these guidelines did not identify any RCTs of hormone therapy in transitioning transgender women. In the context of the preparation of UK National Health Service (NHS) guidelines (PG12 2019), the NHS Guideline Panel also found no RCTs. However, PG12 2019 includes a recommendation for the prescription of hormone therapy for transitioning transgender women.

Of the potentially relevant studies we excluded, some reported on relevant questions. Asscheman 2011 focused on the important outcome of mortality. Fisher 2016 investigated the important relationship between hormone therapy-related body changes and psychobiological well-being. Giltay 2000 focused on body related outcomes such as hormone therapy's effects on the skin (hair growth rate, density, and shaft diameter by image analysis; and sebum production). Toorians 2003 focused on the outcomes of different interventions (estradiol alone compared with combination therapy estradiol and antiandrogens). Miles 2006 was based on a cross-over design with the intention of comparing groups of individuals on and off oestrogen. Due to the reported deficits, we excluded these studies, although they addressed important questions.

AUTHORS' CONCLUSIONS

Implications for practice

We found insufficient evidence to determine the efficacy or safety of hormonal treatment approaches (estradiol alone or in combination with cyproterone acetate or spironolactone) for transgender women in transition. The evidence is very incomplete, demonstrating a gap between current clinical practice and clinical research.

Implications for research

This systematic review has shown that well-designed, sufficiently robust randomised controlled trials (RCTs) and controlled-cohort studies do not exist, and are needed, to assess the benefits and harms of hormone therapies (used alone or in combination) for transgender women in transition. The following questions should be addressed via RCTs and cohort studies:

1. What are the short-, medium-, and long-term effects (including adverse effects, benefits, and prognoses) of estradiol therapy alone, as opposed to combination therapy using estradiol together with cyproterone acetate or spironolactone?
2. What is the short-, medium-, and long-term clinical efficacy of hormone therapy when applied orally, transdermally, and intramuscularly?

Table 1 presents design components that we suggest could be used in future studies. Studies should be structured and reported according to the CONSORT Statement or the STROBE Statement in order to improve the quality of reporting on efficacy and to obtain better reports on harms in clinical research (von Elm 2007; Schulz 2010). There is an urgent need for research in this area, not least for ethical reasons.

We will include non-controlled cohort studies in the next iteration of this review, as this review has demonstrated that this is the highest quality evidence currently available in the field. We will take methodological limitations into account when doing so.

ACKNOWLEDGEMENTS

We would like to acknowledge the support of Dr. Jonathan Livingstone-Banks, Dr. Nicola Lindson, and Dr. Paul Aveyard from the Tobacco Addiction Group, as well as Dr. Erik von Elm from Cochrane Switzerland, whom we consulted in preparing this review.

We greatly appreciate Dr. Alissa Jones Nelson's support in reviewing the spelling and grammar of this review. We also gratefully acknowledge peer review comments from Igor Grabovac, Department of Social and Preventive Medicine, Centre for Public Health, Medical University of Vienna, Vienna, Austria and Dr. Barbara Nussbaumer-Streit, Cochrane Austria, Department for Evidence-based Medicine and Evaluation, Danube-University Krems, Krems, Austria, and consumer review comments from Sarah Stephenson-Hunter.

REFERENCES

References to studies excluded from this review

Asscheman 2011 *{published data only}*

Asscheman H, Giltay EJ, Megens JA, de Ronde WP, van Trotsenburg MA, Gooren LJ. A long-term follow-up study of mortality in transsexuals receiving treatment with cross-sex hormones. *European Journal of Endocrinology* 2011;**164**(4):635-42. [DOI: 1530/EJE-10-1038]

Colizzi 2015 *{published data only}*10.1016/j.jpsychores.2015.02.001

Colizzi M, Costa R, Scaramuzzi F. Concomitant psychiatric problems and hormonal treatment induced metabolic syndrome in gender dysphoria individuals: A 2 year follow-up study. *Journal of Psychosomatic Research* 2015;**78**:399–406. [DOI: 10.1016/j.jpsychores.2015.02.001]

Fighera 2018 *{published data only}*10.1111/cen.13607

Fighera TM, da Silva E, Lindenau JD, Spritzer PM. Impact of cross-sex hormone therapy on bone mineral density and body composition in transwomen. *Clinical Endocrinology* 2018;**88**(6):856-862. [DOI: 10.1111/cen.13607] [PMID: 29630732]

Fisher 2014 *{published data only}*10.1111/jsm.12413

Fisher AD, Castellini G, Bandini E, Casale H, Fanni E, Benni L, et al. Cross-sex hormonal treatment and body uneasiness in individuals with gender dysphoria. *International Society for Sexual Medicine* 2014;**11**:709–19.

Fisher 2016 *{published data only}*10.1210/jc.2016-1276

Fisher AD, Castellini G, Ristori J, Casale H, Cassioli E, Sensi C, et al. Cross-sex hormone treatment and psychobiological changes in transsexual persons: two-year follow-up data. *The Journal of Clinical Endocrinology and Metabolism* 2016;**101**:0000-0000.

Giltay 2000 *{published data only}*

Giltay EJ, Gooren L. Effects of sex steroid deprivation/administration on hair growth and skin sebum production in transsexual males and females. *The Journal of Clinical Endocrinology & Metabolism* 2000;**85**(8):2913-21. [DOI: 10.1210/jc.85.8.2913]

Haraldsen 2005 *{published data only}*

Haraldsen IR, Egeland T, Haug E, Finset A, Opjordsmoen S. Cross-sex hormone treatment does not change sex-sensitive cognitive performance in gender identity disorder patients. *Psychiatry Research* 2005;**137**:161-74. [10.1016/j.psychres.2005.05.014]

Haraldsen 2007 *{published data only}*10.1016/j.yhbeh.2007.05.012

Haraldsen IR, Haug E, Falch J, Egeland T, Opjordsmoen S. Cross-sex pattern of bone mineral density in early onset gender identity disorder. *Hormones and Behavior* 2007;**52**:334-43. [DOI: 10.1016/j.yhbeh.2007.05.012]

Miles 2006 *{published data only}*10.1016/j.yhbeh.2006.06.008

Miles C, Green R, Hines M. Estrogen treatment effects on cognition, memory and mood in male-to-female transsexuals. *Hormones and Behavior* 2006;**50**:708-17. [DOI: 10.1016/j.yhbeh.2006.06.008]

Schlatterer 1998 *{published data only}*

Schlatterer K, Auer DP, Yassouridis A, Von Werder K, Stalla GK. Transsexualism and osteoporosis. *Archives of Sexual Behavior* 1998;**27**(5):475-92. [0004-0002/98/1000-0475]

Toorians 2003 *{published data only}*

Toorians AW, Thomassen MCLGD, Zweegman S, Magdeleyns EJP, Tans G, Gooren L, et al. Venous thrombosis and changes of hemostatic variables during cross-sex hormone treatment in transsexual people. *The Journal of Clinical Endocrinology & Metabolism* 2003;**88**(12):5723-29. [DOI: 10.1210/jc.2003-030520]

Van Goozen 1995 *{published data only}*

Van Goozen SHM. Gender differences in behaviour: activating effects of cross-sex hormones. *Psychoneuroendocrinology* 1995;**20**(4):343-63. [DOI: 10.1016/0306-4530%2894%2900076-X]

References to ongoing studies

Krasean 2019 *{published data only}*

TCTR20190404001. Anti-androgenic effects comparison between Cyproterone acetate and Spironolactone in transgender women: a randomized controlled trial. Thai Clinical Trials Registry 2019. [THAI CLINICAL TRIALS REGISTRY: TCTR20190404001]

Additional references

Arcelus 2015

Arcelus J, Bouman WP, Van Den Noortgate W, Claes L, Witcomb G, Fernandez-Aranda F. Systematic review and meta-analysis of prevalence studies in transsexualism. *European Psychiatry* 2015;**30**(6):807-15.

ATME 2015

Aktion Transsexualität und Menschenrecht eV (ATME). Alternative recommendations for treatment in the presence of so-called "sex/gender variance". Medicine and psychotherapy without gender stereotyping. [STUTTGARTER ERKLÄRUNG - Alternative Handlungsempfehlungen bei geschlechtlichen Normvariationen]. In: In: v. Schreiber G editor(s). Transsexualität in Theologie und Neurowissenschaften - Ergebnisse, Kontroversen, Perspektiven. Vol. 1. Berlin: De Gruyter, 2015:77-8. [ISBN: 978-3110440805]

Bauer 2009

Bauer GR, Hammond R, TraversR, Kaay M, Hohenadel KM, Boyce M. "I don't think this is theoretical; this is our lives": how erasure impacts health care for transgender people. *Journal of the Association of Nurses in AIDS Care* 2009;**20**(5):348-61. [DOI: 10.1016/j.jana.2009.07.004]

Bessone 2015

Bessone F, Lucena MI, Roma MG, Stephens C, Medina-Cáliz I, Frider B, et al. Cyproterone acetate induces a wide spectrum of acute liver damage including corticosteroid-responsive hepatitis: report of 22 cases. *Liver International: Official Journal of the International Association for the Study of the Liver.* 2015;**36**(2):302-10. [DOI: 1111/liv.12899]

Bockting 2016

Bockting W, Coleman E, Deutsch MB, Guillamon A, Meyer I, Meyer III W, et al. Adult development and quality of life of transgender and gender nonconforming people. *Current Opinion in Endocrinology, Diabetes, and Obesity* 2016;**23**(2):188. [DOI: 10.1097/MED.0000000000000232]

Bostick 2008

Bostick NA, Sade R, Levine MA, Stewart Jr DM. Placebo use in clinical practice: report of the American Medical Association Council on Ethical and Judicial Affairs. *The Journal of Clinical Ethics* 2008;**19**(1):58-61. [PMID: 18552054]

Brown 1996

Brown ML, Rounsley CA. True Selves: Understanding Transsexualism - For Families, Friends, Coworkers, and Helping Professionals. 1 edition. Vol. **1**. San Francisco: Jossey-Bass, 1996. [ISBN: 0-7879-6702-5]

Calderón 2009

Calderón GD, Bratoeff E, Ramiréz LE, Osnaya BN, Garcia AR, Barragán MG, et al. Effects of two new steroids and cyproterone on some biomarkers of oxidative stress and serotonergic system on rat prostate and brain. Andrologie Feb 2009;**41**(1):29-34. [DOI: 1111/j.1439-0272.2008.00886.x]

Cash 2004

Cash TF, Jakatdar TA, Williams EF. The body image quality of life inventory: further validation with college men and women. *Body Image* 2004;**1**(3):279–87. [DOI: 1016/S1740-1445(03)00023-8]

De Bastos 2014

de Bastos M, Stegeman B, Rosendaal F. Combined oral contraceptives: venous thrombosis. *Cochrane Database of Systematic Reviews* 2014;**3**(Issue ID 2351):Art. No.: CD010813. [DOI: 10.1002/14651858.CD010813.pub2.] [PMID: 24590565]

Defreyne 2017

Defreyne J, Nota N, Pereira C, Schreiner T, Fisher AD, den Heijer M, et al. Transient elevated serum prolactin in trans women is caused by cyproterone acetate treatment. *LGBT Health* 2017;**4**(5):328-36. [DOI: 1089/lgbt.2016.0190]

Deutsch 2013

Deutsch MB, Green J, Keatley J, Mayer G, Hastings J, Hall AM, et al. Electronic medical records and the transgender patient: recommendations from the World Professional Association for Transgender Health. *Journal of the American Medical Informatics Association* 2013;**20**(4):700-3. [10.1136/amiajnl-2012-001472]

Deutsch 2016a

Deutsch M. Guidelines for the primary and gender-affirming care of transgender and gender nonbinary people. Center of Excellence for Transgender Health June 17th 2016;**2nd Edition**.

Deutsch 2016b

Deutsch MB, Radix A, Reisner S. What's in a guideline? Developing collaborative and sound research designs that substantiate best practice recommendations for transgender health care. *AMA Journal of Ethics* 2016;**18**(11):1098-106. [DOI: 10.1001/journalofethics.2016.18.11.stas1-1611]

Downing 1981

Downing DT, Stewart ME, Strauss JS. Estimation of sebum production rates in man by measurement of the squalene content of skin biopsies. *Journal of Investigative Dermatology* 1981;**77**(4):358-60. [PMID: 7276619]

Drescher 2014

Drescher J. Controversies in gender diagnoses. *LGBT Health* 2014;**1**(1):10-14. [DOI: 1089/lgbt.2013.1500]

DSM-5 2013

American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders (DSM-5), Fifth edition. 1 edition. Vol. **1**. Göttingen: Hogrefe, 2013. [ISBN: 9783801725990]

Ezerskaia 2016

Ezerskaia A, Pereira SF, Urbach HP, Verhagen R, Varghese B. Infrared spectroscopic measurement of skin hydration and sebum levels and comparison to corneometer and sebumeter. *Proceedings SPIE* 2016;**9887**(98872G):552-6. [DOI: 1117/12.2225434]

Figg 2010

Figg W, Chau CH, Cindy H, Small EJ. Drug Management of Prostate Cancer. 1 edition. New York: Springer, 2010. [ISBN: 978-1-60327-829-4]

Flütsch 2015

Flütsch N. Endocrinological treatment of the gender dysphoria in people with gender incongruence. [Endokrinologische Behandlung der Geschlechtsdysphorie bei Menschen mit Geschlechtsinkongruenz]. *Journal of Clinical Endocrinology and Metabolism* 2015;**8**(2):42-8.

Fox 2016

Fox LC, Davies DR, Scholl JL, Watt MJ, Forster GL. Differential effects of glucocorticoid and mineralocorticoid antagonism on anxiety behavior in mild traumatic brain injury. *Behavioural Brain Research* 2016;**312**:362-5. [DOI: 1016/j.bbr.2016.06.048]

Frisch 2005

Frisch MB, Clark MP, Rouse SV, Rudd MD, Paweleck JK, Greenstone A, et al. Predictive and treatment validity of life satisfaction and the quality of life inventory. *Assessment* 2005;**12**(1):66-78. [DOI: 1177/1073191104268006]

Giltay 2008

Giltay EJ, Bunck MC, Gooren L, Zitman FG, Diamant M, Teerlink T. Effects of sex steroids on the neurotransmitter-

specific aromatic amino acids phenylalanine, tyrosine, and tryptophan in transsexual subjects. *Neuroendocrinology* 2008;**88**(2):103-10. [DOI: 1159/000135710]

Gooren 2017

den Heijer M, Bakker A, Gooren L. Long term hormonal treatment for transgender people. *The BMJ* 2017;**359**(j5027):n/a. [DOI: 1136/bmj.j5027]

Greenblatt 1973

Greenblatt DJ, Koch-Weser J. Adverse reactions to spironolactone. *JAMA* 1973;**225**(1):40-3. [DOI: 1001/jama.1973.03220280028007]

Griard 1978

Griard J, Bühler U, Zuppinger K, Haas HG, Staub JJ, Wyss HI. Cyproterone acetate and ACTH adrenal function. *The Journal of Clinical Endocrinology and Metabolism* 1978;**47**(3):581-6. [DOI: 1210/jcem-47-3-581]

Hayes 2000

Hayes FJ, Seminara SB, Decruz S, Boepple PA, Crowley WF Jr. Aromatase inhibition in the human male reveals a hypothalamic site of estrogen feedback. *The Journal of Clinical Endocrinology & Metabolism* 2000;**85**(9):3027-35. [DOI: 0021-972X/00/S03.00/0]

Hembree **2009**

Hembree WC, Cohen-Kettenis P, Delemarre-van de Waal HA, Gooren LJ, Meyer III WJ, Spack NP, et al. Endocrine treatment of transsexual persons: an endocrine society clinical practice guideline. *Journal of Clinical Endocrinology and Metabolism* 2009;**94**(9):3132–54. [DOI: 10 1210/jc 2009-0345]

Hembree 2017

Hembree WC, Cohen-Kettenis P, Delemarre-van de Waal HA, Gooren L, Hannema SE, Meyer III WJ, et al. Endocrine treatment of transsexual persons: an Endocrine Society clinical practice guideline. *The Journal of Clinical Endocrinology & Metabolism* 2017;**102**(11):3869-903. [DOI: 1210/jc 2017-01658]

Higgins 2009

Higgins JPT, Thompson SG, Spiegelhalter DJ. A re-evaluation of random-effects meta-analysis. Journal of the Royal Statistical Society: Series A (Statistics in Society) 2009;**172**(1):137-59. [DOI: 172(1):137–159]

Higgins 2011a

Higgins JPT, Deeks JJ. Chapter 7: Selecting studies and collecting data. In: Higgins JPT, Deeks JJ, editor(s), Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from handbook.cochrane.org.

Higgins 2011b

Higgins JPT, Deeks JJ. Chapter 8: Assessing risk of bias in included studies. In: Higgins JPT, Altman DG, Sterne JAC, editor(s), Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from handbook.cochrane.org.

Higgins 2011c

Higgins JPT, Deeks JJ. Chapter 9: Analysing data and undertaking meta-analyses. In: Deeks JJ, Higgins JPT, Altman DG, editor(s), Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from handbook.cochrane.org.

Higgins 2011d

Higgins JPT, Deeks JJ. Chapter 16: General principles for dealing with missing data. In: Higgins JPT, Deeks JJ, Altman DG, editor(s), Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from handbook.cochrane.org.

Hoffmann 2013

Hoffman R. TrichoScan: a novel tool for the analysis of hair growth in vivo. *Journal of Investigative Dermatology Symposium Proceedings* 2003;**8**(1):109-15. [DOI: 1046/j.1523-1747.2003.12183.x]

Hozo 2005

Hozo SP, Djulbegovic B, Iztok I. Estimating the mean and variance from the median, range, and the size of a sample. BMC Medical Research Methodology 2005;**13**(5). [DOI: 1186/1471-2288-5-13]

Hye-Rim 2012

Hye-Rim L, Tae-Hee K, Kyung-Chul C. Functions and physiological roles of two types of estrogen receptors, ERα and ERβ, identified by estrogen receptor knockout mouse. *Laboratory Animal Research* 2012;**28**(2):71-6. [DOI: 5625/lar.2012.28.2.71]

Khan 2016

Khan O, Mashru A. The efficacy, safety and ethics of the use of testosterone-suppressing agents in the management of sex offending. *Current Opinion in Endocrinology, Diabetes and Obesity* 2016;**23**(3):271-8. [DOI: 10.1097/MED.0000000000000257]

Laurent 2007

Laurent A, Mistretta F, Bottigioli D, Dahel K, Goujon C, Nicolas JF, et al. Echographic measurement of skin thickness in adults by high frequency ultrasound to assess the appropriate microneedle length for intradermal delivery of vaccines. Vaccine 2007;**25**(34):6423-30. [DOI: 25(34):6423-6430]

Meston 2005

Meston C, Trapnell P. Development and validation of a five-factor sexual satisfaction and distress scale for women: the Sexual Satisfaction Scale for Women (SSS-W). *The Journal of Sexual Medicine* 2005;**2**(1):66-81.

Miller 2003

Miller FG, Brody H. A critique of clinical equipoise: therapeutic misconception in the ethics of clinical trials. *Hastings Center Report* 2003;**33**(3):19-28. [PMID: 12854452]

Moher 2009

Moher D, Liberati A, Tetzlaff J, Altman DG. The PRISMA Group (2009). Preferred reporting items for systematic

reviews and meta analyses: The PRISMA statement. Annals of Internal Medicine 2009;**151**(4):264-9. [DOI: 7326/0003-4819-151-4-200908180-00135]

Murad 2010

Murad MH, Elamin MB, Garcia MZ, Mullan RJ, Murad A, Erwin PJ, et al. Hormonal therapy and sex reassignment:a systematic review and meta-analysis of quality of life and psychosocial outcomes. Clinical Endocrinology 2010;**72**(2):214-31. [DOI: 1111/j.1365-2265.2009.03625.x]

Oberhammer 1996

Oberhammer F, Nagy P, Tiefenbacher R, Fröschl G, Bouzahzah B, Thorgeirsson SS, et al. The antiandrogen cyproterone acetate induces synthesis of transforming growth factor beta 1 in the parenchymal cells of the liver accompanied by an enhanced sensitivity to undergo apoptosis and necrosis without inflammation. *Hepatology* 1996;**23**(2):329-37. [DOI: 1002/hep.510230220]

PG12 2019

Sullivan C, Dean J. Prescribing Guideline PG12 Pharmacological Treatment of Gender Dysphoria. Devon Partnership NHS Trust 2019.

Prior 1989

Prior CJ, Vigna YM, Watson D. Spironolactone with physiological female steroids for presurgical therapy of male-to-female transsexualism. *Archives of Sexual Behavior* 1989;**18**(1):49-57. [DOI: 18(1):49-57]

Ramsay 1990

Ramsay ID, Rushton DH. Reduced serum vitamin B12 levels during oral cyproterone-acetate and ethinyl-oestradiol therapy in women with diffuse androgen-dependent alopecia. *Clinical and Experimental Dermatology* 1990;**15**(4):277-81. [DOI: 1111/j.1365-2230.1990.tb02089.x]

Reilly 2019

Reilly ZP, Fruhauf TF, Martin SJ. Barriers to evidence-based transgender care: knowledge gaps in gender-affirming hysterectomy and oophorectomy. *Obstetrics & Gynecology* 2019;**134**(4):714-17. [DOI: 10.1097/AOG.0000000000003472]

Reisner 2016a

Reisner SL, Deutsch MB, Bhasin S, Bockting W, Brown GR, Feldman J, et al. Advancing methods for US transgender health research. *Current opinion in endocrinology, diabetes, and obesity* 2016;**23**(2):198. [DOI: 10.1097/MED.0000000000000229]

Reisner 2016b

Reisner SL, Radix A, Deutsch MB. Integrated and gender-affirming transgender clinical care and research. *Journal of acquired immune deficiency syndromes* 1999;**72**(3):235. [DOI: 10.1097/QAI.0000000000001088]

Riley 2011

Riley RD, Higgins JP, Deeks JJ. Interpretation of random effects meta-analyses. BMJ 2011;**342**. [DOI: 1136/bmj.d549.]

Schneider 2006

Schneider H, Stalla G. Hormonal Therapy [Hormonelle Therapie]. In: Therapieleitfaden Transsexualität. 1 edition. Bremen: Uni-Med Science, 2006:85-9. [ISBN: 3-89599-888-5]

Schneider 2016

Schneider C, Cerwenka S, Nieder TO, Briken P, Cohen-Kettenis PT, de Cuypere G, et al. Measuring gender dysphoria: a multicenter examination and comparison of the Utrecht gender dysphoria scale and the gender identity/gender dysphoria questionnaire for adolescents and adults. *Archives of Sexual Behavior* 2013;**45**(3):551-8. [DOI: 1007/s10508-016-0702-x]

Schulz 2010

Schulz KF, Altman DG, Moher D. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. *The BMJ* 2010;**340**:698-702. [DOI: 10.1136/bmj.c332]

Schürmeyer 1986

Schürmeyer T, Graff J, Senge T, Nieschlag E. Effect of oestrogen or cyproterone acetate treatment on adrenocortical function in prostate carcinoma patients. Acta Endocrinologica 1986;**111**(3):360-7. [PMID: 2421511]

Seal 2012

Seal LJ, Franklin S, Richards C, Shishkareva A, Sinclaire C, Barret J. Predictive markers for mammoplasty and a comparison of side effect profiles in transwomen taking various hormonal regimens. *The Journal of Clinical Endocrinology & Metabolism* 2012;**97**(12):4422-8. [DOI: 1210/jc.2012-2030]

Soll 2018

Soll BM, Robles-García R, Brandelli-Costa A, Mori D, Mueller A, Vaitses-Fontanari AM, et al. Gender incongruence: a comparative study using ICD-10 and DSM-5 diagnostic criteria. *Brazilian Journal of Psychiatry* 2018;**40**(2):174-80.

Sterne 2011

Sterne JA, Sutton AJ, Ioannidis JP, Terrin N, Jones DR, Lau J, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. BMJ 2011;**343**. [343:d4002]

Sterne 2016

Sterne JAC, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ* 2016;**355**:i4919. [DOI: 10.1136/bmj.i4919]

Thompson 1993

Thompson DF, Carter JR. Drug-induced gynecomastia. *Pharmacotherapy* Jan-Feb 1993;**13**(1):37-45. [PMID: 8094898]

von Elm 2007

von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies.. *Annals of Internal Medicine* 2007;**147**(8):573-7. [DOI: 10.7326/0003-4819-147-8-200710160-00010] [PMID: 17938396]

Wenqing 2005

Wenqing G, Bohl CE, Dalton JT. Chemistry and structural biology of androgen receptor. *Chemical Reviews* 2005;**105**(9):3352-70. [DOI: 10.1021/cr020456u]

WHO 2018

World Health Organisation. International classification of diseases for mortality and morbidity statistics (11th Revision). icd.who.int/en/ (accessed 30 October 2020).

Wierckx 2014

Wierckx K, van Caenegem E, Schreiner T, Haraldsen I, Fisher AD, Toye K, et al. Cross-sex hormone therapy in trans persons is safe and effective at short-time follow-up: results from the European network for the investigation of gender incongruence. *International Society for Sexual Medicine* 2014;**11**(8):1999-2011. [DOI: 10.1111/jsm.12571]

WPATH 2011

Coleman E, Bockting W, Botzer M, Cohen-Kettenis P, DeCuypere G, Feldman J, et al. Standards of care for the health of transsexual, transgender, and gender-nonconforming people version 7. *International Journal of Transgenderism* 2011;**13**(4):165–232. [DOI: 10.1080/15532739.2011.700873]

Zucker 2016

Zucker KJ. The DSM-5 diagnostic criteria for gender dysphoria. In: Trombetta C, Liguori G, Bertolotto M, editors(s). Management of Gender Dysphoria - A Multidisciplinary Approach. First edition. Vol. 1. Mailand: Springer, 2016:33-7. [ISBN: 978-88-470-5695-4]

References to other published versions of this review

Haupt 2018

Haupt C, Henke M, Kutschmar A, Hauser B, Baldinger S, Schreiber G. Antiandrogens or estradiol treatments or both during hormone replacement therapy in transitioning transgender women. *Cochrane Database of Systematic Reviews* 2018, Issue 10. Art. No: CD013138. [DOI: 10.1002/14651858.CD013138]

CHARACTERISTICS OF STUDIES

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
Asscheman 2011	Mortality rates in transgender people receiving long-term cross-sex hormones. A cohort study. Adequate controls are missing. Interventions are not clearly defined
Colizzi 2015	Increased prevalence of metabolic syndrome among individuals with gender dysphoria treated by cross-sex hormonal treatment. Study without adequate comparator group.
Fighera 2018	Hormone therapy has been associated with changes in bone and lean/fat mass. This study assessed bone mineral density, appendicular lean mass, and total fat mass in transwomen undergoing cross-sex hormone therapy. Study without adequate comparator group.
Fisher 2014	This study aimed to assess differences in body uneasiness and psychiatric symptoms between gender dysphoria clients taking hormone therapy and those not taking hormones (no hormone therapy). A second aim was to assess whether length of hormone treatment and daily dose provided an explanation for levels of body uneasiness and psychiatric symptoms. Cross-sectional design.
Fisher 2016	The objective of the study was to assess whether hormone therapy-related body changes affect psychobiological well-being in gender dysphoria. Study without adequate comparator group.
Gillay 2000	Hormone therapy effects on the skin (hair growth rate, density, and shaft diameter by image analysis; and sebum production) of transsexual patients receiving cross-sex hormones. It is a case series, adequate controls are missing.
Haraldsen 2005	Hormone therapy effects on cognitive performance. Study without adequate comparator group.
Haraldsen 2007	The effects of cross-sex hormones on bone metabolism (bone mineral density, total body fat, total lean body mass) in patients with early onset gender identity disorder. Study without adequate comparator group.
Miles 2006	The study was designed to examine associations between oestrogen and cognition (memory, including visual, spatial, object and location memory, other cognitive abilities that show reliable sex

Study	Reason for exclusion
	differences, including verbal and visual-spatial abilities, and mood variables). The cross-over design used was comparative, but did not used randomization or quasi-randomisation.
Schlatterer 1998	This follow-up study was carried out to validate the effectiveness of cross-gender hormone therapy embedded in a multistep treatment concept for transgender patients. Study without adequate comparator group. This study lacks adequate controls.
Toorians 2003	To find an explanation for the different thrombotic risks of oral ethinyl estradiol and transdermal 17-beta-estradiol use, the researchers compared the effects of treatment of male-to-female transgender people with cyproterone acetate only, and with cyproterone acetate in combination with transdermal 17-beta-estradiol, oral ethinyl estradiol, or oral 17-beta-estradiol on a number of haemostatic variables. There is no adequate control group.
Van Goozen 1995	Effects of sex hormones to the establishment of gender differences in behaviour, a large battery of tests on aggression, sexual motivation and cognitive functioning was administered twice: shortly before and three months after the start of cross-sex hormone treatment. The study does not have an adequate comparator group.

Characteristics of ongoing studies *[ordered by study ID]*

Krasean 2019

Study name	Anti-androgenic effects comparison between cyproterone acetate and spironolactone in transgender women: a randomised controlled trial (Trial ID: TCTR20190404001)
Methods	Allocation: randomised Study design: randomised controlled trial Control: active Study endpoint classification: efficacy study Intervention model: Parallel Number of arms: 2 Masking: double blind (Masked roles: participant caregiver, investigator) Primary purpose: treatment Study phase: phase 4
Participants	Gender: male Age limit: minimum 18 years: maximum 40 years Condition: Gender dysphoria patients diagnosed from DSM V Male to female transgender Not undergone orchidectomy No psychological disease or mental disability
Interventions	Arm 1: Intervention name: cyproterone acetate Type: active comparator

Krasean 2019 *(Continued)*

	Classification: drug Descriptions: participants (gender dysphoria patients) will receive estradiol valerate (4 mg daily) combined with cyproterone acetate (25 mg daily) for cross-sex hormone treatment. Arm: 2 Intervention name: spironolactone Type: experimental Classification: drug Descriptions: participants (gender dysphoria patients) will be received estradiol valerate (4 mg daily) combined with spironolactone (100 mg daily) for cross-sex hormone treatment.
Outcomes	Primary outcome(s): Outcome name: testosterone level Measurement: Electrochemiluminescent Immunoassay (ECLIA) of total testosterone level Time point: three months after intervention Safety issue: no Key secondary outcomes: Outcome name: physical and metabolic changes Measurement: physical examination, metabolic profile parameters Time point: three months after intervention Safety Issue: no
Starting date	April 3, 2019 (estimated end date: June 16, 2020)
Contact information	Contact: Krasean Panyakhamlerd Degree: Assoc. Prof. Phone: 0926536415 Email: krasean@hotmail.com Postal Address: 1873 Rama 4 Road, Patumwan State/Province: Bangkok Postal Code: 10400 Country: Thailand
Notes	Source(s) of monetary or material supports: Ratchadapisek Sompoch Fund, Faculty of Medicine, Chulalongkorn University Declarations of interest not reported

ADDITIONAL TABLES

Table 1. Suggested design of future studies

Methods	RCT or controlled cohort study
Participants	Transgender women experiencing gender dysphoria, in transition N* Age: from the age of 16 years
Intervention	• Antiandrogens (cyproterone acetate or spironolactone) and estradiol • Antiandrogens (cyproterone acetate or spironolactone) alone • Estradiol alone All types of administration: oral, sublingual, transdermal, subdermal and intramuscular. For estradiol and bioidentical 17-beta-estradiol, as well as synthetic derivatives.
Comparator	Any of the active interventions listed above
Outcomes	Primary outcomes • Quality of life (QoL) • Satisfaction with change of male to female body characteristics, • Adverse events specific to hormone therapy, including serious adverse events
Notes	* Size of study with sufficient power to detect a ~ 10% difference between the two groups for primary outcome

APPENDICES

Appendix 1. OvidSP search strategy

Search	Query
#1	(transsexual* OR transgender OR "gender dysphoria" OR transident* OR "trans women" OR "trans woman").mp.
#2	("cyproterone acetate" OR CPA OR androcur).mp. or cyproterone Acetate/
#3	(spironolactone OR Aldactone OR Jenaspiron OR Osyrol OR Spirobene OR Verospiron OR Xenalon).mp. or spironolactone/
#4	(estradiol* OR oestradiol* OR estrifam OR gynocadin OR neofollin OR lenzetto).mp. or Estradiol/
#5	2 OR 3 OR 4
#6	1 AND 5

HISTORY

Protocol first published: Issue 10, 2018
Review first published: Issue 11, 2020

CONTRIBUTIONS OF AUTHORS

All authors contributed to the Abstract, Background, Methods, Results, Discussion, and Authors' conclusions.
Claudia Haupt, Alexia Kutschmar and Miriam Henke conducted the study selection.

DECLARATIONS OF INTEREST

Claudia Haupt declares no competing interest.

Miriam Henke declares no competing interest.

Alexia Kutschmar declares no competing interest.

Birgit Hauser (BH) declares no competing interest. BH is a clinical practitioner in private practice, who also prescribes hormone therapy.

Sandra Baldinger declares no competing interest.

Sarah Rafaela Saenz declares no competing interest.

Gerhard Schreiber declares no competing interest.

None of the review authors' incomes depends on the prescription of drugs. The review authors did not receive any financial support for this project, but paid for all related expenses themselves. They worked voluntarily and free of charge.

INDEX TERMS

Medical Subject Headings (MeSH)

Androgen Antagonists [*therapeutic use]; Drug Therapy, Combination [methods]; Estradiol [*therapeutic use]; Estrogens [*therapeutic use]; Placebos [therapeutic use]; Sex Reassignment Procedures [*methods]; *Transgender Persons

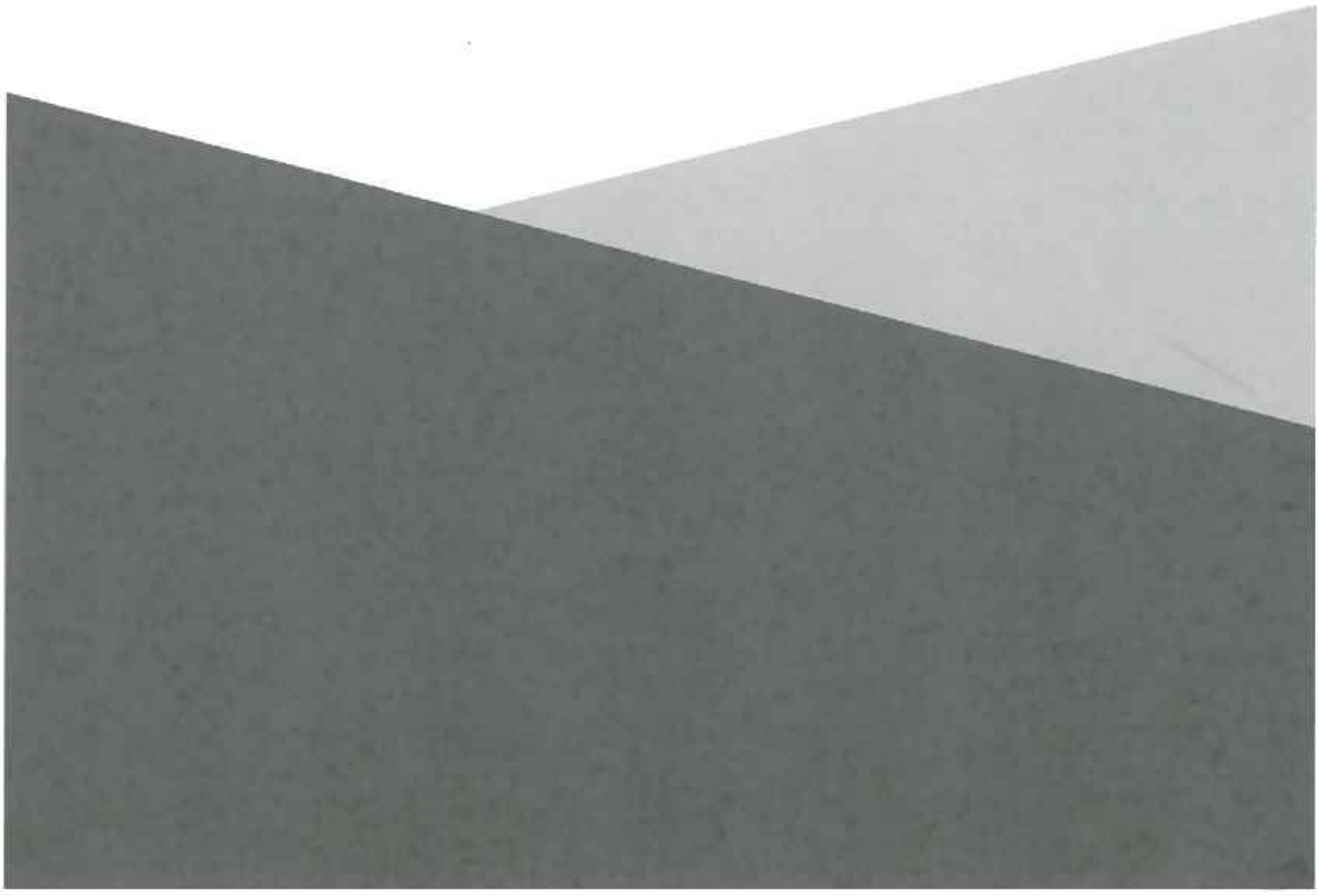
MeSH check words

Female; Humans; Male

Socialstyrelsen
THE NATIONAL BOARD OF HEALTH AND WELFARE

Care of children and adolescents with gender dysphoria

Summary



Summary

The National Board of Health and Welfare (NBHW) has been commissioned by the Swedish government to update the national guidelines on care of children and adolescents with gender dysphoria, first published in 2015 [1]. Guidelines chapters are updated stepwise and this report contains revised guidance on psychosocial support and diagnostic assessment, and on puberty suppressing treatment with GnRH-analogues and gender-affirming hormonal treatment. This report thus replaces the corresponding chapters in the publication from 2015. Remaining chapters and the updated guidelines as a whole will be published later in 2022. In response to comments received during external review, two new chapters have been added, named *New recommendations on hormonal treatment – their reasons and consequences and Non-binary gender identity – current knowledge and a need for clarification*. Another difference compared to the guidelines from 2015 [1] is that the term "gender incongruence" is used alongside the term "gender dysphoria". For explanations of terms and abbreviations, see Appendix 2. For a description of the scientific evidence and clinical experience underlying the recommendations and the work process, see Appendices 3 and 4.

The guidelines apply to children and adolescents, i.e. people under 18 years of age. In the medical text sections, the term children (barn) refers to persons who have not yet entered puberty, while the term adolescents (ungdomar) refers to people whose puberty has started. In the text sections relating to juridical regulations, only the term children (barn) is used and denotes people younger than 18 years of age. Finally, the term "young people" (unga) is sometimes used in text sections addressing both children and adolescents.

Introductory comment

The summary that follows and the introductory chapter describe that the updated recommendations for puberty suppression with GnRH-analogues and gender-affirming hormonal treatment have become more restrictive compared to 2015, and the reasons that they have changed. The new recommendations entail that a larger

proportion than before, among adolescents with gender incongruence referred for diagnostic assessment of gender dysphoria, will need to be offered other care than hormonal treatments. Questions on how to ensure that all young people suffering from gender dysphoria be taken seriously and confirmed in their gender identity, well received and offered adequate care are becoming increasingly relevant, and will need to be answered during the ongoing restructuring of certain care for gender dysphoria into three national specialised medical care services (NBHW decision in December 2020). The care for children, adolescents and adults with gender dysphoria in these three national specialised units aims to improve equality in care, coordination and dialogue, and may enhance the implementation of national guidelines.

Recommendations and criteria for hormonal treatment

For adolescents with gender incongruence, the NBHW deems that the risks of puberty suppressing treatment with GnRH-analogues and gender-affirming hormonal treatment currently outweigh the possible benefits, and that the treatments should be offered only in exceptional cases. This judgement is based mainly on three factors: the continued lack of reliable scientific evidence concerning the efficacy and the safety of both treatments [2], the new knowledge that detransition occurs among young adults [3], and the uncertainty that follows from the yet unexplained increase in the number of care seekers, an increase particularly large among adolescents registered as females at birth [4].

A systematic review published in 2022 by the Swedish Agency for Health Technology Assessment and Assessment of Social Services [2] shows that the state of knowledge largely remains unchanged compared to 2015. High quality trials such as RCTs are still lacking and the evidence on treatment efficacy and safety is still insufficient and inconclusive for all reported outcomes. Further, it is not possible to determine how common it is for adolescents who undergo gender-affirming treatment to later change their perception of their gender identity or interrupt an ongoing treatment. An important difference compared to 2015 however, is that the occurrence of

detransition among young adults is now documented [3], meaning that the uncertain evidence that indicates a low prevalence of treatment interruptions or any aspects of regret is no longer unchallenged. Although the prevalence of detransition is still unknown, the knowledge that it occurs and that genderconfirming treatment thus may lead to a deteriorating of health and quality of life (i.e. harm), is important for the overall judgement and recommendation.

To minimize the risk that a young person with gender incongruence later will regret a gender-affirming treatment, the NBHW deems that the criteria for offering GnRH-analogue and gender-affirming hormones should link more closely to those used in the Dutch protocol, where the duration of gender incongruence over time is emphasized [5-7]. Accordingly, an early (childhood) onset of gender incongruence, persistence of gender incongruence until puberty and a marked psychological strain in response to pubertal development is among the recommended criteria. The publications that describe these criteria and the treatment outcomes when given in accordance [5, 6, 8] consitute the best available knowledge and should be used as guidance.

To ensure that new knowledge is gathered, the NBHW further deems that treatment with GnRH-analogues and sex hormones for young people should be provided within a research context, which does not necessarily imply the use of randomized controlled trials (RCTs). As in other healthcare areas where it is difficult to conduct RCTs while retaining sufficient internal validity, it is also important that other prospective study designs are considered for ethical review and that register studies are made possible. Until a research study is in place, the NBHW deems that treatment with GnRH-analogues and sex hormones may be given in exceptional cases, in accordance with the updated recommendations and criteria described in the guidelines. The complex multidisciplinary assessments will eventually be carried out in the three national units that are granted permission to provide highly specialized care services.

In accordance with the DSM-5, the recommendations in the guidelines from 2015 applied to young people with gender dysphoria in general, i.e. also young people with a non-binary gender identity. Another criterion within the Dutch protocol is that the child has had a binary ("cross-gender") gender identity since childhood [5, 6].

It has emerged during the review process, that the clinical experience and documentation of puberty-suppressing and hormonal treatment for young people with non-binary gender identity is lacking, and also that it is limited for adults. The NBHW still considers that gender dysphoria rather than gender identity should determine access to care and treatment. An urgent work thus remains, to clarify criteria under which adolescents with non-binary gender identity may be offered puberty-suppressing and gender-affirming hormonal treatment within a research framework.

References

1. Socialstyrelsen. God vård av barn och ungdomar med könsdysfori. Nationellt kunskapsstöd. https://www.socialstyrelsen.se/globalassets/sharepoint-dokument/artikelkatalog/kunskapsstod/2015-4-6.pdf; 2015.
2. Statens beredning för medicinsk och social utvärdering. Hormonbehandling vid könsdysfori - barn och unga. En systematisk översikt och utvärdering av medicinska aspekter: SBU; 2022.
3. Littman L. Individuals Treated for Gender Dysphoria with Medical and/or Surgical Transition Who Subsequently Detransitioned: A Survey of 100 Detransitioners. Arch Sex Behav. 2021; 50(8):3353-69.
4. Socialstyrelsen. Utvecklingen av diagnosen könsdysfori - förekomst, samtidiga psykiatriska diagnoser och dödlighet i suicid. https://www.socialstyrelsen.se/globalassets/sharepoint-dokument/artikelkatalog/ovrigt/2020-2-6600.pdf.; 2020.
5. Cohen-Kettenis PT, van Goozen SH. Sex reassignment of adolescent transsexuals: a follow-up study. J Am Acad Child Adolesc Psychiatry. 1997; 36(2):263-71.
6. Smith YL, van Goozen SH, Cohen-Kettenis PT. Adolescents with gender identity disorder who were accepted or rejected for sex reassignment surgery: a prospective follow-up study. J Am Acad Child Adolesc Psychiatry. 2001; 40(4):472-81.
7. Delemarre-van de Waal H, Cohen-Kettenis P. Clinical management of gender identity disorder in adolescents: a protocol on

psychological and paediatric endocrinology aspects. European Journal of Endocrinology. 2006; 155:S131-7.

8. Schagen SE, Cohen-Kettenis PT, Delemarre-van de Waal HA, Hannema SE. Efficacy and Safety of Gonadotropin-Releasing Hormone Agonist Treatment to Suppress Puberty in Gender Dysphoric Adolescents. J Sex Med. 2016; 13(7):1125-32.

Bilaga till rapport

Hormonbehandling vid könsdysfori - barn och unga/ Hormone treatment of children and adolescents with gender dysphoria, rapport 342 (2022)

Bilaga 3. Inkluderade studier
Appendix 3. Characteristics of included studies: Extracted data

Innehåll

SBU Statens beredning för medicinsk och social utvärdering • www.sbu.se
Telefon 08-412 32 00 • **Fax** 08-411 32 60 • **Organisationsnummer** 202100-4417
Besöksadress S:t Eriksgatan 117, Stockholm • **Postadress** Box 6183, 102 33 Stockholm

Table 1. Effects on mental health by puberty suppression in adolescents

Author, Year (ref) Title Country Study design	**De Vries et al 2014 (1)** *Young Adult Psychological Outcome After Puberty Suppression and Gender Reassignment* The Netherlands Longitudinal cohort study, before-after 2008-2012
POPULATION (ages) Age at start Age in cohort Tanner stage	Age at assessment pre-treatment: Range 11.1–17.0 years 13.6 years (SD 1.9) At start of puberty suppression: Range 11.5–18.5 14.8 years (SD 1.8) At start of cross-sex hormones: Range 13.9–19.0 years 16.7 years (SD 1.1)
POPULATION (n) n patients natal male (M-t-F) natal female (F-t-M)	196 referred 111 prescribed puberty suppression 15 non-participating 1 death after vaginoplasty 55 individuals evaluated: 22 transwomen 33 transmen 40 complete data 15 missing data
INTERVENTION (type) Puberty suppression (GnRH) Cross-sex hormone treatment (CSHT)	Puberty suppression (GnRH) Cross-sex hormone treatment (CSHT) Gender reassignment surgery: vaginoplasty, mastectomy, hysterectomy, ovariectomy, (phalloplasty)
INTERVENTION (time) Treatment duration Follow-up time, Follow-up age	GnRH duration: Not specified CSHT duration: Not specified Age at Follow-up: at assessment Post-Treatment Mean 20.7 years (SD 1.0) Range 19.5–22.8
OUTCOMES – Reported outcomes	Gender Dysphoria Utrecht Gender Dysphoria Scale (UGDS) Global functioning Children's Global Assessment Scale (CGAS) Depressive symptoms: The Beck Depression Inventory (BDI) Anger Spielberger's Trait Anger (TPI) Anxiety: Spielberger's Trait Anxiety (STAI) Body Image Scale (BIS) Child Behavior Checklist (CBCL)
RESULTS Extracted outcomes	**Before start / During puberty suppression / After gender reassignment** (mean (SD)) <u>Gender dysphoria (UGDS)</u> Total 53.51 (8.29) / 54.39 (7.70) / 15.81 (2.78) MtF 47.07 (11.05) / 48.95 (10.80) / 17.27 (2.57) FtM 56.74 (3.74) / 57.11 (3.40) / 15.08 (2.64) <u>Global functioning (CGAS)</u> Total 71.13 (10.46) / 74.81 (9.86) / 79.94 (11.56) MtF 74.33 (7.53) / 78.20 (9.56) / 82.40 (8.28) FtM 67.65 (11.87) / 70.65 (9.89) / 76.29 (14.48) <u>Depression (BDI)</u> Total 7.89 (7.52) / 4.10 (6.17) / 5.44 (8.40) MtF 4.73 (4.20) / 2.25 (3.54) / 3.38 (4.40) FtM 10.09 (8.34) / 5.05 (7.08) / 6.95 (9.83) <u>Anxiety (STAI)</u> Total 39.57 (10.53)/ 37.52 (9.87)/ 37.61 (10.39) MtF 31.87 (7.42)/ 31.71 (8.36)/ 35.83 (10.22) FtM 44.41 (9.06)/ 41.59 (9.03)/ 39.20 (10.53) <u>Anger (TPI)</u> Total 17.55 (5.72)/ 17.22 (5.61)/ 16.01 (5.28) MtF 14.17 (3.01)/ 14.00 (3.36)/ 5.58 (3.92) FtM 19.55 (5.96)/ 19.25 (5.69)/ 16.56 (6.06)

Author, Year (ref) Title Country Study design	**Costa et al 2015 (2)** *Psychological Support, Puberty Suppression, and Psychosocial Functioning in Adolescents with Gender Dysphoria.* The UK Longitudinal cohort study, before-after, 2010-2014
POPULATION (ages) Age at start Age in cohort Tanner stage	Age at baseline: Range 12-17 years 15.6 years (SD 1.7) natal male 15.4 years (SD 1.2) natal female Age at start of GnRH: Range 13-17 years 16.6 years (SD 1.22) natal male 16.4 years (SD 1.3) natal female
POPULATION (n) n patients natal male (M-t-F) natal female (F-t-M)	436 referred [1: 1.7 natal male/natal female ratio] 235 did not complete diagnostic procedure 201 completed diagnostic procedure [1: 1.6 natal male/natal female ratio] 121 eligeable for puberty suppression 80 not eligeable for puberty suppression after 6 months psychological support* 101 GnRH treated "Immediate eligible": 35 GnRH treated evaluated at end of study 100 GnRH untreated "Delayed eligible": 36 GnRH untreated evaluated at end of study
INTERVENTION (type) Puberty suppression (GnRH) Cross-sex hormone treatment (CSHT)	GnRH: Drug, dose and treatment frequency not indicated. Start after 6 months of psychological assessment and support (mean 0.75 + 0.6 years), referred as "diagnostic procedure". Psychotherapeutic interventions: "Individual or family or group therapy, carried out on a regular basis (at least one a month)"
INTERVENTION (time) Treatment duration Follow-up time, Follow-up age	GnRH duration: 12 months Psychological support: 18 months total Follow-up times: 6 months, 12 months, 18 months
OUTCOMES – Reported outcomes	UGDS Children's Global Assessment Scale (CGAS) [high score=better psychosocial functioning]
RESULTS Extracted outcomes	**Psychosocial functioning:** <u>Children's Global Assessment Scale score:</u> All GD adolescents, during diagnostic procedure (n=201): 57.7 (SD 12.3) at enrolment 60.7 (SD 12.5) 6 months after psychological support only GnRH treated group: (n= 101 at baseline) 60.9 (SD 12.2) after 6 months psychological support only (n= 61) 67.4 (SD 13.9) at 18 months psychological support + GnRHa (7-18 months) (n= 35) Delayed group: (n= 100 at baseline) 60.3 after 6 months psychological support only 62.5 after 18 months (n= 36)

Author, Year (ref) Title Country Study design	**Becker-Hebly et al 2020** (3) *Psychosocial health in adolescents and young adults with gender dysphoria before and after gender-affirming medical interventions* Germany Retrospective cohort study, before-after 2013-2018
POPULATION (ages) Age at start Age in cohort Tanner stage	Age at baseline (intake): Minimum 11 years Mean 15.5 years (SD 1.2) Range 11.2 - 18.0 years Age at Follow-up: Mean 17.4 years (SD 1.7) Range 11.95 - 21.0 years
POPULATION (n) n patients natal male (M-t-F) natal female (F-t-M)	434 adolescents 164 dropouts at baseline 129 dropouts during follow-up 75 evaluated: 64 birth assigned female 11 birth assigned male 21 no hormone 11 GnRH 32 GnRH + CSHT 11 CSHT + surgery (type not specified) Excluded severe psychiatric problems (psychosis, suicidality)
INTERVENTION (type) Puberty suppression (GnRH) Cross-sex hormone treatment (CSHT)	GnRH: Drug, dose and treatment frequency not indicated. CSHT: Drug, dose and treatment frequency not indicated. Groups: No hormone treatment (no GnRH, no CSHT) GnRH GnRH + CSHT CSHT + surgery (surgery type not specified, "mainly mastectomy") Psychotherapy (79%)
INTERVENTION (time) Treatment duration Follow-up time, Follow-up age	Duration of GnRH or CSHT: not specified. Possible range 7-49 months, "time since first referral" GnRH: minimum 7 months CSHT: up to 40 or 47 months Follow-up time: Mean 21.4 (SD 12.2) months Range 6 months - 4 years
OUTCOMES - Reported outcomes	**Psychological functioning:** Children's Global Assessment Scale (CGAS, clinician-rated) HR QoL (mental and physical dimensions): assessed by Kidscreen-27 (>18 years) SF-8 (<18 years) Youth Self Report (YSR, ages 11-18y) Adult version (ASR, >18y)

RESULTS Extracted outcomes	**Psychosocial functioning:** CGAS Global functioning Baseline/ Follow-up (mean (SD)) No medical treatment (diagnostics or psychosocial interventions) 68.10 (11.23) / 70.00 (12.25) Puberty suppression (GnRH) 67.27 (11.91) / 81.82 (7.51) GA hormones (GnRH and GAH) 73.13 (10.91) / 85.63 (9.14) GA surgery (at least one operation and GAH) 66.36 (14.33) / 83.64 (8.09) Health-related quality of life (mean ± SD) Baseline T Mental dimension/T Physical dimension No medical treatment (diagnostics of psychosocial interventions) 34.86 (6.27) / 37.51 (8.27) Puberty suppression (GnRH) 39.04 (9.25) / 43.43 (8.61) GA hormones (GAH and GnRH) 36.16 (6.78) / 39.12 (7.10) GA surgery (at least one operation and GAH) 37.88 (6.53) / 39.88 (8.49) Follow-up T Mental dimension/T Physical dimension No medical treatment (diagnostics or psychosocial interventions) 36.37 (7.71) / 42.51 (10.40) Puberty suppression (GnRH) 43.17 (10.20) / 49.57 (11.64) GA hormones (GAH and GnRH) 42.07 (10.74) / 49.36 (9.81) GA surgery (at least one operation and GAH) 43.44 (9.57) / 53.87 (6.15)

Author, Year (ref) Title Country Study design	**Cantu et al 2020 (4)** *Changes in Anxiety and Depression from Intake to First Follow-Up Among Transgender Youth in a Pediatric Endocrinology Clinic* USA Retrospective cohort study chart review, before-after, 2017 - 2019
POPULATION (ages) Age at start Age in cohort Tanner stage	Age at start: Min 11 years Max 18 years Age in cohort: Mean 15.1 years (SD 1.8)
POPULATION (n) n patients natal male (M-t-F) natal female (F-t-M)	80 15 female affirmed 58 male affirmed 7 nonbinary In Follow-up cohort: 13 hormone blockers 25 hormone treatment (HT) 4 hormone blockers + HT 38 no treatment
INTERVENTION (type) Puberty suppression (GnRH) Cross-sex hormone treatment (CSHT)	Previous intervention: Drug, dose and treatment frequency not indicated. Hormone blockers only Hormone treatment (HT) only (feminizing; masculinizing) Both hormone blockers and HT Neither hormone blockers nor HT Of 28 youth: 6 feminizing hormones 22 masculinizing hormones
INTERVENTION (time) Treatment duration Follow-up time, Follow-up age	Duration of GnRH or CSHT: Not specified. Time between initial visit and follow-up appointment: Mean 4.7 months Range < 1 - 11 months
OUTCOMES – Reported outcomes	Depression: assessed with PHQ-9 (Patient Health Questionnaire-9) Anxiety: assessed with GAD-7 (Generalized Anxiety Disorder-7)
RESULTS Extracted outcomes	**Psychosocial functioning:** Acute distress (not defined) Baseline/follow-up Mean (SD) PHQ-9 HT initiated (n=28) 9.8 (7.1)/ 10.3 (7.3) No HT (n=51) 11.1 (6.3)/ 10.1 (5.9) GAD-7 HT initiated (n=27) 8.4 (6.4)/ 8.5 (5.5) No HT (n=50) 9.6 (5.9)/ 9.1 (5.8) Suicidality "Of the 27 (34%) youth who endorsed suicidality at intake, 22 (81%) continued to endorse suicidality at their follow-up visit, and only 4 (4%) no longer endorsed suicidality at follow-up".

Author, Year (ref) Title Country Study design	**Carmichael et al 2021 (5)** *Short-term outcomes of pubertal suppression in a selected cohort of 12 to 15 year old young people with persistent gender dysphoria in the UK* The UK Prospective cohort, 2011 -2015
POPULATION (ages) Age at start Age in cohort Tanner stage	Age at consent (median, IQR): 13.6 years (12.8 - 14.6) Range 12.0 - 15.3 years At end of pathway (median, IQR): 16.1 years (16.0 - 16.4)
POPULATION (n) n patients natal male (M-t-F) natal female (F-t-M)	44 recruited: 25 birth registered males 19 birth-registered females Tanner stage: (n (%), birth registered males, birth registered females): Stage 2: 0, 0 Stage 3: 17 (68%), 2 (10%) Stage 4: 5 (20%), 11 (58%) Stage 5: 3 (12%), 6 (32%) 1 discontinued GnRH
INTERVENTION (type) Puberty suppression (GnRH) Cross-sex hormone treatment (CSHT)	GnRHa: triptorelin Psychosocial assessment and support: Before entering the study for a median of 2.0 years (IQR 1.4 to 3.2; range 0.7 to 6.6 years). Continued regular attendance for psychological support and therapy throughout the study was a precondition of GnRHa prescription. Local psychological services provided support for co-occurring difficulties as required. No interview conducted before young people started GnRHa
INTERVENTION (time) Treatment duration Follow-up time, Follow-up age	Follow-up time: 12 months follow-up (n=44), 24 months (n=24), 36 months (n=14) Median time in study: 31 months (IQR 20 to 42, range 12 to 59 months). Age at end of pathway (IQR): 16.1 years (16.0, 16.4)
OUTCOMES - Reported outcomes	Child Behaviour Checklist (CBCL) (parent report) Youth Self Report (YSR) Kidscreen-52 questionnaire Body Image Scale (BIS) is Utrecht Gender Dysphoria Scale (UGDS) Children's Global Assessment Scale (CGAS) Semi-structured qualitative interviews. Participant experience and satisfaction with GnRHa No interview conducted before young people started GnRHa

RESULTS – Extracted outcomes	**CBCL** Parent report, Total problems t-score: mean (95% CI): Baseline; 12 months, change; 24 months, change; 36 months, change 61.6 (58.4, 64.7); 61.8 (58.4, 65.1), 0.3 (-2.0, 2.6); 60.2 (54.6, 65.8), -1.0 (-4.0, 2.1); 61.1 (52.3, 69.9), -1.3 (-6.6, 4.0) **CBCL** Parent report, Self-harm: median (IQR): Baseline; 12 months; 24 months; 36 months 0 (0,1) ; 0 (0,1) ; 0 (0,1) ; 0 (0,1) ; 0 (0,1) **YSR** Self-report, Total problems t-score: mean (95% CI): Baseline; 12 months, change; 24 months, change 57.9 (55.0, 60.8); 58.4 (54.6, 62.2), 0.8 (-3.1, 4.8); 56.5 (50.6, 62.5), 1.5 (-3.4, 6.3) **YSR** Self-report, Self-harm: median (IQR): Baseline; 12 months; 24 months 0 (0,1) ; 0 (0,2) ; 0 (0,0) **Kidscreen-52, HRQOL,** Parent report, Psychological wellbeing, t-score, mean (95% CI) Baseline; 12 months; 24 months 43.0 (39.6, 46.4); 41.1 (37.0, 45.2); 51 (45.8, 56.2) **Kidscreen-52, HRQOL,** Self-report, Psychological wellbeing, t-score, mean (95% CI) Baseline; 12 months; 24 months 39.8 (36.7, 42.8) ; 39.0 (35.4, 42.6) ; 42.4 (36.9, 48) **Body image scale,** Overall score: mean (95% CI) Baseline; 12 months; 24 months; 36 months 3.1 (2.8, 3.3) ; 3.2 (3.0, 3.4) ; 3.0 (2.7, 3.2) ; 3.1 (2.4, 3.7) **Utrecht Gender dysphoria score:** median (IQR) Baseline; 12 months; 24 months 4.8 (4.6, 5.0) ; 4.7 (4.6, 5.0) ; 4.7 (4.3, 5.0) **CGAS global score,** mean (95% CI) Baseline; 12 months; 24 months; 36 months 62.9 (59.6, 66.2) ; 64.1 (59.9, 68.3) ; 65.7 (59.6, 71.8) ; 66.0 (58.1, 73.9) No changes from baseline to 12 or 24 months in CBCL or YSR total t-scores or for CBCL or YSR self-harm indices, nor for CBCL total t-score or self-harm index at 36 months. Most participants reported positive or a mixture of positive and negative life changes on GnRHa.

Author, Year (ref) Title Country Study design	**Hisle-Gorman et al 2021** (6) *Mental Healthcare Utilization of Transgender Youth Before and After Affirming Treatment* USA Retrospective cohort study (military healthcare data), 2010–2018
POPULATION (ages) Age at start Age in cohort Tanner stage	Age at Study Initiation: years (median (IQR)) 10 years (8–13) transgender 9 years (4–14) siblings Age of First Affirming Medication (CSHT), years (median (IQR)) 18.2 years (16.6–19.8) Age at Study Completion, years (median (IQR)) 18 years (16–21) transgender 17 years (11–21) siblings
POPULATION (n) n patients natal male (M-t-F) natal female (F-t-M)	3754 transgender 1193 (31.8%) male at birth 2561 (68.2%) female at birth 963 transgender adolescents receiving hormone treatment (before-after data) 6603 cisgender siblings
INTERVENTION (type) Puberty suppression (GnRH) Cross-sex hormone treatment (CSHT)	<u>Hormone treatment (n=963)</u> Puberty Suppressant n=96 (7.2%) Masculinizing Hormone n=591 (61.4%) Feminizing Hormone n=276 (28.7%) Psychotropic medication n=857 (89%)
INTERVENTION (time) Treatment duration Follow-up time, Follow-up age	Full study period: 8.5 years in total follow-up time Hormone treatment: Years followed (median (IQR)) 7.1 years (5.6–7.9) before HT 1.5 years (0.7-2.7) after HT
RESULTS Reported outcomes	
RESULTS Extracted outcomes	<u>Mental health over full 8-year study period*:</u> *TGD adolescents compared to siblings were more likely to have a mental health diagnosis, be prescribed more psychotropic medications and use more mental healthcare services:* Mental health diagnosis (n (%)): 3352 (89.3%) transgender vs 3308 (50.1%) siblings; adjusted OR 5.45 (4.77–6.24) On psychotropics (n (%)): 2820 (75.1%) transgender vs 2425 (37.7%) siblings Psychotropic medication days: All mental health meds (medications days per year): 111.4 transgender vs 42.5 siblings; adjusted IRR 2.57 (2.36-2.80) <u>Mental health diagnoses at some point during the 8-year study period:</u> Transgender vs Siblings (n (%); adjusted odds of mental health diagnosis* aOR (95% CI)) **after adjustment for age at study initiation, assigned sex at birth, parent rank, and number of outpatient visits per year, odds of having any mental health diagnosis:* All Mental Health 3352 (89.3%) vs 3308 (50.1%); aOR 5.45 (4.77–6.24) Mood 2413 (64.3%) vs 1182 (18.9%); aOR 6.12 (5.51-6.8) Anxiety 1908 (50.8%) vs 1216 (18.4%); aOR 3.30 (2.98–3.65) ADHD 1119 (29.8%) vs 1229 (18.6%); aOR 1.77 (1.59–1.97) Adjustment 1687 (44.9%) vs 1191 (18.0%); aOR 1.09 (1.80–3.41) Psychotic 363 (9.7%) vs 104 (1.6%); aOR 5.38 (4.20-6.88) Personality disorders 86 (2.3%) vs 43 (0.7%); aOR 2.54 (1.71–3.78) Suicide 683 (18.2%) vs 162 (2.5%); aOR 7.45 (6.11–9.08) (suicidal ideation/attempted suicide/self-harm) <u>Psychotropic medication*:</u>

	**including antidepressants (wellbutrin, SSRI, SNRI, other antidepressant) benzodiazepines, sleep medications, anti-psychotics, litium* Transgender vs Siblings (medication days per year): All mental health medications: 1114 days vs 425 days; adjusted IRR 2.57 (2.36-2.80) After hormone treatment: (n=963 individuals-initiated puberty suppression or CSHT, median age 18.2 years): Crude rate of medication days (number of days, Before - After hormone treatment)) All Mental Health Medications: (days) 119.7 before vs 211.5 after; aIRR 1.67 (1.46–1.91) Psychotropic medication use: increased from mean 120 days per year to mean 212 days per year following gender affirming pharmaceutical care. <u>Medication days by type of medication:</u> <u>(number of medication days: Before vs After hormone treatment):</u> Wellbutrin 6.3 before vs 16.2 after; aIRR 2.51 (2.71–3.69) SSRI 44.8 before vs 73.9 after; aIRR 1.72 (1.47–2.00) SNRI 4.7 before vs 14.0 after; aIRR 2.59 (1.52–4.38) other antidepressant 9.2 before vs 18.9 after; aIRR 1.61 (1.18–2.21) sleep medications 6.4 before vs 16.2 after; aIRR 2.23 (1.61–3.10) benzodiazepines 3.0 before vs 12.7 after; aIRR 3.01 (1.95–4.65) anti-psychotics 15.9 before vs 30.1 after; aIRR 1.77 (1.34–2.35) lithium 1.3 before vs 2.3 after; aIRR 1.11 (0.48–2.59) stimulants 26.4 before vs 25.1 after; aIRR 0.96 (0.72–1.26) migraine medications 1.5 before vs 2.2 after; aIRR 0.76 (0.37–1.53)

Author, Year (ref) Title Country Study design	**Staphorsius et al 2015 (7)** *Puberty suppression and executive functioning: An fMRI-study in adolescents with gender dysphoria* The Netherlands Functional MRI study, Cross-sectional, up to 2014
POPULATION (ages) Age at start Age in cohort Tanner stage	Age at start: Minimum 12 years, Tanner B2, Tanner G2-G3 Age at GnRH start: Not indicated Age in cohort: (mean ± SD) Age at scan: 15.1 years ± 2.4 M-t-F 15.8 years ± 1.9 F-t-M Control group age: 14.9 years ± 1.5 (boys) 14.4 years ± 1.8 (girls)
POPULATION (n) n patients natal male (M-t-F) natal female (F-t-M)	41 adolescents 22 F-t-M (natal females): (12 using GnRH, "suppressed FM") (10 untreated, "untreated FM") 18 M-to-F (natal males): (8 using GnRH, "suppressed FM") (10 untreated, "untreated FM") Control group* (siblings, friends): 24 girls (F) 21 boys (M) 10 not investigated due to brain scan problems
INTERVENTION (type) Puberty suppression (GnRH) Cross-sex hormone treatment (CSHT)	GnRH: triptorelin (Decapeptyl-CR®) 3,75 mg/4w, s.c. or i.m Study intervention: MRI scan (3.0 T) axial T2*-weighted whole-brain volumes sensitive to BOLD contrast, sagittal T1-weighted Tasks in MRI: 1 executive function task: event-related parametric version of the Tower-of-London (ToL) task 3 cognitive tasks: verbal fluency task, mental rotation task, face recognition task
INTERVENTION (time) Treatment duration Follow-up time, Follow-up age	Puberty suppression duration (mean ± SD): 1.6±1.0 years: 1.8 years ± 0.8 MtF 1.4 years ± 1.1 FtM
OUTCOMES - Reported outcomes	Executive function: Tower-of-London (ToL) performance scores: reaction times, accuracy Region-of-interest (ROI) analyses: left DLPFC (dorsolateral prefrontal cortex), bilateral RLPFC (rostrolateral prefrontal cortex), precuneus Psychological functioning: Child Behaviour Checklist (CBCL) IQ: Wechsler Intelligence Scales (WISC-III®, Wechsler, 1991; WAIS-III®,Wechsler, 1997)
RESULTS Extracted outcomes	Executive function: Functional task (ToL): Accuracy (%) mean ± SD) 88.5 ± 6.8 boys (M) ; 87.2 ± 11.9 girls (F) 79.1 ± 10.3 M-t-F (total) 73.9 ± 9.1 suppressed ; 83.4 ± 9.5 untreated 87.1 ± 10.0 F-t-M (total) 85.7 ± 10.5 suppressed ; 88.8 ± 9.7 untreated Reaction time (sec) mean ± SD 9.6 ± 2.5 boys (M) ; 9.0 ± 1.8 girls (F) 10.4 ± 3.5 M-t-F (total) 10.9 ± 4.1 suppressed ; 9.9 ± 3.1 untreated 10.0 ± 2.6 F-t-M (total) 9.9 ± 3.1 suppressed ; 10.0 ± 2.0 untreated Psychological functioning: CBCL scores, mean ± SD 48.4 ± 10.5 boys (M) ; 48.4 ± 10.3 girls (F) 57.8 ± 9.2 M-t-F (total) 57.4 ± 9.8 suppressed ; 58.2 ± 9.3 untreated 60.4 ± 10.2 F-t-M (total) 57.5 ± 9.4 suppressed ; 63.9 ± 10.5 untreated

Table 2. Effects on bone health by puberty suppression in adolescents

Author, Year (ref) Title Country Study design	**Joseph et al 2019 (8)** *The effect of GnRH analogue treatment on bone mineral density in young adolescents with gender dysphoria: findings from a large national cohort* UK Retrospective review of national cohort, before-after, 2011–2016
POPULATION (ages) Age at start Age in cohort Tanner stage	Age at GnRH start: Range 12–14 years Age in First year cohort: Age at treatment start: (mean (SD) 13.2 (1.4) trans girls 12.6 (1.0) trans boys Age at 1 year scan: 14.4 (1.5) trans girls 13.8 (1.1) trans boys Age in Longitudinal cohort Age at treatment start: 13.0 (1.1) trans girls 12.9 (3.0) trans boys Age at 2 years scan: 15.8 (1.3) trans girls 15.6 (3.5) trans boys
POPULATION (n) n patients natal male (M-t-F) natal female (F-t-M)	First year cohort: 70 31 trans girls 39 trans boys Longitudinal cohort: 31 10 trans girls 21 trans boys
INTERVENTION (type) Puberty suppression (GnRH) Cross-sex hormone treatment (CSHT)	GnRH Study intervention: DXA - dual energy X-ray absorptiometry Z-scores [calculated from Crabtree et al. from ALPHABET study using UK norms for Caucasian subjects]. Hip BMAD Z-scores not calculated (no reference ranges available)
INTERVENTION (time) Treatment duration Follow-up time, Follow-up age	GnRH duration: 1 year (1st year cohort) 2.8 years (longitudinal cohort) Follow-up time: 1–2.8 years
OUTCOMES - Reported outcomes	Bone health: Hip (femoral neck) and lumbar spine (L1-L4): BMD - bone mineral density BMAD - bone mineral apparent density Z-score compared to natal sex (birth sex, age) Hip BMD g/cm^2 Hip BMD Z score Spine BMD g/cm^2 Spine BMD Z score Spine BMAD g/cm^3 Spine BMAD Z score

RESULTS – Extracted outcomes	**Characteristics,** mean (SD) **Baseline / 1 year** Trans girls (n=31/31) Age, year 13.2 (1.4) / 14.4 (1.5) Height, cm 161.0 (8.0) / 163.7 (8.1) Weight, kg 64.7 (17.1) / 70.3 (21.2) BMI, kg/m² 24.8 (5.3) / 26.1 (6.9) Hip BMD, kg/m² 0.894 (0.118) / 0.905 (0.104) Hip Z-score 0.157 (0.905) / –0.340 (0.816) Spine BMD, kg/m² 0.860 (0.154) / 0.859 (0.129) Spine BMD Z-score –0.016 (1.106) / –0.461 (1.121) Spine BMAD, g/cm³ 0.235 (0.030) / 0.233 (0.029) Spine BMAD Z-score 0.859 (0.154) / –0.228 (1.027) Trans boys (n=39/39) Age, years 12.6 (1.0) / 13.8 (1.1) Height, cm 158.4 (9.5) / 163.3 (8.7) Weight, kg 51.0 (13.7) / 56.2 (13.4) BMI, kg/m² 20.1 (4.1) / 21.4 (5.4) Hip BMD, kg/m² 0.772 (0.137) / 0.785 (0.120 Hip Z-score –0.863 (1.215) / –1.440 (1.075) Spine BMD, kg/m² 0.694 (0.149) / 0.718 (0.124) Spine Z-score –0.395 (1.428) / –1.276 (1.410) Spine BMAD, g/cm³ 0.196 (0.035) / 0.201 (0.033) Spine BMAD Z-score –0.186 (1.230) / –0.541 (1.396) **Baseline / 2.8 years** Trans girls (n=10/10) Age, years 13.0 (1.1) / 15.8 (1.3) Height, cm 160.3 (5.4) / 165.1 (5.7) Weight, kg 66.4 (14.6) / 82.9 (30.5) BMI, kg/m² 25.8 (5.3) / 30.5 (8.6) Hip BMD, kg/m² 0.920 (0.116) / 0.910 (0.125) Hip Z-score 0.45 (0.781) / –0.600 (1.059) Spine BMD, kg/m² 0.867 (0.141) / 0.878 (0.130) Spine BMD Z-score 0.130 (0.972) / 0.890 (1.075) Spine BMAD, g/cm³ 0.240 (0.027) / 0.240 (0.030) Spine BMAD Z-score 0.486 (0.809) / –0.279 (0.93) Trans boys (n=21/21) Age, years 12.9 (3.0) / 15.6 (3.5) Height, cm 159.0 (35.8) / 168.7 (37.5) Weight, kg 49.8 (17.1) / 59.5 (19.6) BMI, kg/m² 19.4 (5.9) / 20.9 (6.6) Hip BMD, kg/m² 0.766 (0.215) / 0.773 (0.197) Hip Z-score –1.075 (1.145) / –1.779 (0.816) Spine BMD, kg/m² 0.695 (0.220) / 0.731 (0.209) Spine BMD Z-score –0.715 (1.406) / –2.000 (1.384) Spine BMAD, g/cm3 0.195 (0.058) / 0.198 (0.05) Spine BMAD Z-score –0.361 (1.439) / –0.913 (1.318)

Author, Year (ref) Title Country Study design	Klink et al (9) 2015 *Bone mass in young adulthood following gonadotropin-releasing hormone analog treatment and cross-sex hormone treatment in adolescents with gender dysphoria* The Netherlands Retrospective longitudinal cohort study , before-after, 1998–2012
POPULATION (ages) Age at start Age in cohort Tanner stage	**Age at start of GnRH:** Range 11.4–18.3 years Transwomen: Tanner G5 Mean: 14.9 years ± 1.9 SD Transmen: Tanner B4 Mean: 15.0 years ± 2.0 SD **At start of CSHT:** Range 15.6–19 years Transwomen: Mean: 16.6 years ± 1.4 SD Transmen: Median: 16.4 years (2.3 IQR)
POPULATION (n) n patients natal male (M-t-F) natal female (F-t-M)	34 15 MtF 19 FtM
INTERVENTION (type) Puberty suppression (GnRH) Cross-sex hormone treatment (CSHT)	**GnRH:** Triptorelin (Decapeptyl-CR): 3.75 mg/4 weeks s.c. **CSHT:** 17-estradiol p.o. (incremental dosing), dose not indicated. Mixed testosterone esters i.m. 250 mg/ml/ 2–4 weeks (incremental dosages), dose not indicated. Surgery: gonadectomy (min age 18 years) **Study intervention:** DXA (dual energy x-ray absorptiometry Lumbar spine (LS), Femoral region (FN) aBMD *Z*-scores according to natal sex, age, and ethnicity based on the *National Health and Nutrition Examination Survey* reference in Manitoba, Canada. LS *Z* scores available from start of the study. FN *Z* scores available in 2003, 5 years after the start of the study. Volumetric BMD (bone mineral apparent density (BMAD)) of the LS and FN calculated as previously described, *Z* scores determined using UK reference population. Reference values of BMAD in young adulthood are not available. In females lumbar peak bone mass (PBM) expressed as BMAD is attained at age 18–20 years and in males between 18 and 23 years (8). To calculate the *Z* score of the LS BMAD at age 22 years, the reference of LS BMAD of 17 years was used.
INTERVENTION (time) Treatment duration Follow-up time, Follow-up age	**GnRH duration** Median: 1.3 years natal boys, Range: 0.5–3.8 years Median: 1.5 y natal girls, Range: 0.25–5.2 years **CSHT duration** Median: 5.8 years natal boys, Range: 3.0–8.0 years Median: 5.4 years natal girls, Range: 2.8–7.8 years **GnRH + CSHT duration:** Median: 3.1 years natal boys, Range: 2.1–4.5 years Median: 2.2 years natal girls, Range: 1.4–3.1 years After gonadectomy: GnRH terminated and CSHT continued. **FU until age 22 years**
OUTCOMES - Reported outcomes	**Bone health** Bone mineral density (BMD): Bone mineral apparent density (BMAD) Areal BMD (aBMD, g/cm^2) lumbar spine and femoral region: BMAD (g/cm^3) BMAD Z-score aBMD (g/cm^2) aBMD Z-score T-score Z-score relative natal sex

RESULTS – Extracted outcomes	**Start GnRH / Start CSH / Age 22 years** (mean ± SD) **Transwomen** Height cm 174.6 8.9 / 179.9/ 181 ± 9.3 *Lumbar spine* BMAD, g/cm³ 0.22 ± 0.03 / 0.22 ± 0.02 / 0.23 ± 0.03 BMAD *Z* score –0.44 ± 1.10 / –0.90 ± 0.80 / –0.78 ± 1.03 aBMD, g/cm² 0.84 ± 0.13 / 0.84 ± 0.11 / 0.93 ± 0.10 aBMD *Z* score –0.77 ± 0.89 / –1.01 ± 0.98 / –1.36 ± 0.83 T–score at 22 years: –1.5 ± 1.10 *Femoral neck* BMAD, g/cm³ 0.28 ± 0.04 / 0.26 ± 0.04 / 0.28 ± 0.05 BMAD *Z* score –0.93 ± 1.22 / –1.57 ± 1.74 / -- aBMD, g/cm² 0.88 ± 0.1 / 0.87 ± 0.08 / 0.94 ± 0.11 aBMD *Z* score –0.66 ± 0.77 / –0.95 ± 0.63 / –0.69 ± 0.74 T-score at 22 years: –0.75 ± 0.78 **Transmen** *Height cm* 165.2 ± 9.1 / 168.4 ± 8.3 / 170.6 ± 7.9 *Lumbar spine* BMAD, g/cm³ 0.25 ± 0.03 / 0.24 ± 0.02 / 0.25 ± 0.28 BMAD *Z* score 0.28 ± 0.90 / –0.50 ± 0.81 / –0.033 ± 0.95 aBMD, g/cm2 0.95 ± 0.12 / 0.91 ± 0.10 / 0.99 ± 0.13 aBMD *Z* score 0.17 ± 1.18 / –0.72 ± 0.99 / –0.33 ± 1.12 T-score at 22 years: –0.43 ± 1.2 *Femoral neck* BMAD, g/cm³ 0.32 ± 0.04 / 0.31 ± 0.04 / 0.33 ± 0.05 BMAD Z score 0.01 ± 0.70 / –0.28 ± 0.74 / -- aBMD, g/cm² 0.92 ± 0.10 / 0.88 ± 0.09 / 0.95 ± 0.10 aBMD *Z* score 0.36 ± 0.88/ –0.35 ± 0.79/ –0.35 ± 0.74 T-score at 22 years: 0.005 ± 0.87

Author, Year (ref) Title Country Study design	**Vlot, et al 2017 (10)** *Effect of pubertal suppression and cross-sex hormone therapy on bone turnover markers and bone mineral apparent density (BMAD) in transgender adolescents* The Netherlands Retrospective, cohort study, before after 2001-2011
POPULATION (ages) Age at start Age in cohort Tanner stage	**Age at start of GnRH:** Min Tanner B2 or G2 **Age in cohort:** *Transmen:* Median: 15.1 years Range: 11.7–18.6 years Tanner B2-B5 *Transwomen:* Median: 13.5 years Range: 11.5–18.3 years Tanner G2-G5 **Age at start of CSHT** (min age 16 years): *Transmen:* Median: 16.3 years Range: 15.9–19.5 years *Transwomen:* Median: 16.0 years Range: 14.0–18.9 years
POPULATION (n) n patients natal male (M-t-F) natal female (F-t-M)	*In Table 1:* 70 42 female-to-male (transmen) 28 male-to-female (transwomen) *In abstract:* 56 34 female-to-male (transmen) 22 male-to-female (transwomen)
INTERVENTION (type) Puberty suppression (GnRH) Cross-sex hormone treatment (CSHT)	**GnRH:** Triptorelin (Decapeptyl–CR ®) 3.75 mg s.c. /4 weeks **CSHT:** Testosterone esters (Sustanon) i.m.: 25 mg/m² /2 weeks, 6-month increment until 250 mg/4 w 17-β estradiol: 5 µg/kg/day, 6-months increments until 2 mg/day **Study intervention:** DXA- dual energy X-ray absorptiometry BMAD Z-scores calculated for sex assigned at birth using UK reference population, due to the lack of consensus with regard to the use of either sex assigned at birth or desired sex reference values in transgender adolescents. The lack of validated reference values of bone age needed to calculate the BMAD, and Z-scores limits the use of bone age and therefore the chronological calendar age of the transgender adolescents was used. Reference values of L- M- and S-values of 17-year-old biological males and females were used to calculate the BMAD for patients older than 17 years, due to the lack of reference values of adolescents exceeding the age of 17 years. Two groups: Young group: bone age <15 years in transwomen or <14 years in transmen Old group: bone age ≥15 years in transwomen or ≥14 years in transmen
INTERVENTION (time) Treatment duration Follow-up time, Follow-up age	**GnRH** Approximately 1 year in transmen Approximately 2–3 years in transwomen **CSHT** Up to 24 months.
OUTCOMES - Reported outcomes	**Bone mineral turnover markers:** N-terminal propertied of type I collagen (PINP) Osteocalcin (OC) Carboxy terminal cross linked telopeptide of type I collagen (ICTP) Bone mineral apparent density (BMAD) of lumbar spine (LS) and femoral neck (FM) Z-scores

RESULTS – Extracted outcomes	**At start GnRH / at start CHST / at 24 months** Height, cm, median (range) *Transmen:* 164.2 (149.6–180.1) / 165.8 (152.6–181.2) / 168.6 (155.6–183) *Transwomen:* 166.9 (153.9–185.7) / 176.3 (165.1–186.4) / 180.7 (167.4-195.0) **Transmen, "young"** P1NP median/range: 783 (516–1090) / 324 (194–402) / 186 (163–334) OC median/range: 5 (2.2–11.7) / 6.8 (1.8–7.7) / 4.9 (4.2–7.8) ICTP median/range: 24 (17–29.9) / 11 (7.8–12) / 12 (11–14) BMAD HIP: 0.31 (0.26–0.36) / 0.30 (0.22–0.35) / 0.33 (0.23–0.37) BMAD HIP Z-score: −0.01 (−1.30–0.91) / −0.37 (−2.28–0.47) / −0.37 (−2.03–0.85) BMAD LS: 0.23 (0.20–0.29) / 0.23 (0.19–0.28) / 0.25 (0.22–0.28) BMAD LS Z-score: −0.05 (−0.78–2.94) / −0.84 (−2.2–0.87) / −0.15 (−1.38–0.94) **Transmen, "old"** P1NP median/range: 110 (38–471) / 127 (61–321) / 101 (44–181) OC median/range: 2.4 (0.4–4.6) / 3.9 (0.4–8.6) / 2.9 (0.8–5) ICTP median/range: 7 (5.2–15) / 6.9 (4.6–14) / 8.2 (4.1–16) BMAD HIP: 0.33 (0.25–0.39) / 0.30 (0.23–0.41) / 0.32 (0.23–0.41) BMAD HIP Z-score: 0.27 (−1.39–1.32) / −0.27 (−1.91–1.29) / 0.02 (−2.1–1.35) BMAD LS: 0.26 (0.21–0.29) / 0.24 (0.20–0.28) / 0.25 (0.21–0.30) BMAD LS Z-score: 0.27 (−1.6–1.8) / −0.29 (−2.28–0.90) / −0.06 (−1.76–1.61) **Transwomen, "young"** P1NP median/range: 935 (617–1348) / 363 (185–643) / 204 (137–314) OC median/range: 4.8 (2.6–21.9) / 6.4 (0.7–12.8) / 5.4 (3.9–12.5) ICTP median/range: 23 (15–34) / 13 (8.7–21) / 10 (8.5–13) BMAD HIP: 0.29 (0.20–0.33) / 0.27 (0.20–0.33) / 0.27 (0.20–0.36) BMAD HIP Z-score: −0.71 (−3.35–0.37) / −1.32 (−3.39–0.21) / −1.3 (−3.51–0.92) BMAD LS: 0.21 (0.17–0.25) / 0.20 (0.18–0.24) / 0.22 (0.19–0.27) BMAD LS Z-score: −0.2 (−1.82–1.18) / −1.52 (−2.36–0.42) / −1.10 (−2.44–0.69) **Transwomen," old"** P1NP median/range: 191 (96–792) / 140 (111–467) / 119 (55–296) OC median/range: 2.29 (0.8–11) / 2.2 (0.5–6.1) / 3.3 (1.8–6.8) ICTP median/range: 12 (6.9–21) / 7.4 (6.9–13) / 6.8 (4.8–15) BMAD HIP: 0.30 (0.26–0.36) / 0.30 (0.26–0.34) / 0.29 (0.24–0.38) BMAD HIP Z-score: −0.44 (1.37–0.93) / −0.36 (−1.5–0.46) / −0.56 (−2.17–1.29) BMAD LS: 0.22 (0.18–0.25) / 0.22 (0.19–0.24) / 0.23 (0.21–0.26) BMAD LS Z-score: −1.18 (−1.78–1.09) / −1.15 (−2.21–0.08) / −0.66 (−1.66–0.54)

Author, Year (ref) Title Country Study design	**Schagen et al 2020 (11)** *Bone Development in Transgender Adolescents Treated With GnRH Analogues and Subsequent Gender-Affirming Hormones* The Netherlands Prospective observational study, 1998 - 2009
POPULATION (ages) Age at start Age in cohort Tanner stage	At the start of GnRHa: Early pubertal group: Tanner stage 2 or 3 Late pubertal group: Tanner stage 4 or 5 At start of GnRH: (mean ± SD) 14.1 ± 1.7 trans girls 14.5 ± 2.0 trans boys At start of CSHT: 16.2 ± 1.2 trans girls 16.9 ± 1.1 trans boys
POPULATION (n) n patients natal male (M-t-F) natal female (F-t-M)	**GnRHa group:** 121 51 trans girls 70 trans boys Pubertal group: Early (Tanner 2-3) / Late (Tanner 4-5) 15 / 36 trans girls 14 / 56 trans boys **GnRHa + CSHT group:** 78 36 trans girls 42 trans boys Pubertal group: Early (Tanner 2-3) / Late (Tanner 4-5) 10 / 26 trans girls 5 / 37 trans boys
INTERVENTION (type) Puberty suppression (GnRH) Cross-sex hormone treatment (CSHT)	GnRHa i.m. 3.75 mg/ 4 weeks (Triptorelin) CSHT: Oestrogens oral Testosterone i.m. (Sustanon) In subjects > 16 years at the start of pubertal suppression: CSHT started at half the adult dose and increased to the adult dose after 6 months. (2 mg 17beta-estradiol/day, 125 mg testosterone-esters/ 2 weeks considered an adult dose). Study intervention: Dual-energy x-ray absorptiometry (DXA) Calculate z-scores based on age and sex using *National Health and Nutrition Examination Surveys (NHANES)* references values; reference population of the birth-assigned sex was used. BMAD (g/cm3) calculated as described by Ward [Ward et al. 2007 *UK reference data for the Hologic QDR Discovery dual-energy x ray absorptiometry scanner in healthy children and young adults aged 6-17 years*. Arch Dis Child. 92(1): 53-59). BMAD Z-scores calculated using LMS data from an English reference population [Ward et al. 2007).
INTERVENTION (time) Treatment duration Follow-up time, Follow-up age	Duration of GnRH: (years) 1.9 ± 1.03 mean 2.0 ± 0.94 transgirls 1.8 ± 1.11 transboys Early pubertal groups were on GnRHa for a significantly longer time (2.5 years in transgirls (n = 7) and 4.0 years in transboys (n = 3)) when compared with both late-pubertal groups (1.5 years in transgirls and 1.7 years in transboys) Duration of CSHT: 3 years (not further detailed)
OUTCOMES - Reported outcomes	Bone mineral apparent density (BMAD) BMAD Z-scores (age- and sex-specific) Serum bone markers: P1NP, P3NP, osteocalcin, 1CTP Areal BMD (aBMD, g/cm2) lumbar spine, nondominant hip, whole body; Bone mineral content of the whole body (BMC-WB, g)

RESULTS – Extracted outcomes	**aBMD 2 Years of GnRHa Treatment, Baseline / 24 months** **Transgirls** Early Pubertal (n=15) aBMD_hip g/cm² 0.81 (0.03) / 0.86 (0.03) Z-score −0.49 (0.24) / −0.93 (0.21) Late-Pubertal (n=36) aBMD_hip g/cm² 0.87 (0.02) / 0.89 (0.02) Z-score −0.43 (0.16) / −1.01 (0.15) **Transboys** Early-pubertal (n=14) aBMD_hip g/cm² 0.79 (0.03) / 0.83 (0.03) Z-score 0.09 (0.26) /−0.50 (0.24) **Transboys** Late-pubertal (n=56) aBMD_hip g/cm² 0.93 (0.01) / 0.89 (0.02) Z-score 0.46 (0.13) / −0.56 (0.13) **aBMD GnRHa + 3 Years of Gender-Affirming Hormone Treatment, Baseline / 36 months** **Transgirls** Early-Pubertal: (n=10) aBMD_hip g/cm² 0.87 (0.03) / 1.02 (0.04) Z-score −0.99 (0.23) / −0.09 (0.28) **Transgirls** Late-Pubertal: (n=26) aBMD_hip g/cm² 0.88 (0.02) / 0.96 (0.02) Z-score −0.86 (0.14) / −0.70 (0.18) **Transboys** Early-pubertal: (n=5) aBMD_hip g/cm² 0.83 (0.04) / 1.02 (0.06) Z-score −0.82 (0.33) / 0.59 (0.43) **Transboys** Late-pubertal: (n=37) aBMD_hip g/cm² 0.88 (0.02) / 0.96 (0.02) Z-score −0.50 (0.12) / 0.12 (0.16)

Author, Year (ref) Title Country Study design	**Stoffers et al 2019 (12)** *Physical changes, laboratory parameters, and bone mineral density during testosterone treatment in adolescents with gender dysphoria* The Netherlands Retrospective, cohort study before-after, 2010-2018
POPULATION (ages) Age at start Age in cohort Tanner stage	**At start of GnRH:** Median: 16.5 years Range: 11.8–18.0 years **At start of testosterone:** Median: 17.2 years Range: 14.9–18.4 years
POPULATION (n) n patients natal male (M-t-F) natal female (F-t-M)	62 trans males (FtM) 17 evaluated 0 discontinued testosterone "Excluded psychological, medical, or social problems that might interfere with treatment"
INTERVENTION (type) Puberty suppression (GnRH) Cross-sex hormone treatment (CSHT)	GnRH (Decapeptyl-CR®): 3.75 mg /4 weeks s.c. for at least 6 months Testosterone (Sustanon®); start at 250 mg i.m. Age 15–16 years: increased every 6 months using 25 mg/m²/2 weeks, 50 mg/m²/2 weeks, and 75 mg/m²/2 weeks, leading up to a standard adult dose of 125 mg every 2 weeks. ≥16 years: start 75 mg/m²/2 weeks for 6 months, thereafter 125mg/m²/2 weeks Study intervention: Dual energy x-ray absorptiometry. Lumbar spine (LS) and hip (n=17) BMD Z-scores calculated using female reference data from *Bone Mineral Density in Childhood Study* (USA) for those >16 years of age, reference data from the *Third National Health and Nutrition Examination Survey* for the neck area of the hip and Hologic adult reference data for the LS were used. Bone mineral apparent density (BMAD) calculated and Z- scores determined for lumbar spine and left femoral neck as described by Ward et al. (UK). Reference values provided for up to 17 years, reference values for 17-year-olds used for those >17 y.
INTERVENTION (time) Treatment duration Follow-up time, Follow-up age	**GnRH duration** Median: 8 months Range: 3–39 months (3.25 years) **Testosterone duration** Min: 6 months Mean: 12 months Range: 5–33 months (2.75 years)
OUTCOMES - Reported outcomes	Virilization (acne, hair growth, voice deepening, absence of menses) height, weight, BMI, BP, hematcrit, cholesterol, ALP, triglycerides, Hb Hormone levels: FSH, LH, DHAES, FT4, testosterone, estradiol, TSH, prolactin, androstenedione, sex-hormone binding globulin (SHBP) Bone mineral density (BMD) lumbar spine, femoral neck BMD Z-scores
RESULTS – Extracted outcomes	**Bone health:** **At start GnRH** (n=62) **/ at start testosterone** (n=62) **/ at 24 months** (n=15) Blood pressure, mm Hg (median (IQR) Systolic 124 (115-129) / 118 (114-126) / 126 (117-129) Diastolic 68 (65-73) / 72 (66-77) / 74 (63-76) Height (cm (mean ± SD)) 167.1 ± 6.9 / 168.2 ± 6.2 / 167.8 ± 5.3 BMD, g/cm² (mean ± SD) Lumbar spine 0.96 ± 0.11 / 0.90 ± 0.11/ 0.95 ± 0.11 Left hip 0.84 ± 0.11 / 0.76±0.09/ 0.86 ± 0.09 Right hip 0.84 ± 0.11 / 0.77 ±0.08/ 0.85 ± 0.11 BMD Z-score (mean ± SD) Lumbar spine: 0.02 ± 1.00 / –0.81 ± 1.02/ –0.74 ± 1.1 Left hip –0.19 ± 1.04 / –1.07 ± 0.85/ –0.20 ± 0.70 Right hip –0.16 ± 1.00/ –0.97 ± 0.79/ –0.31 ± 0.84

Author, Year (ref) Title Country Study design	**Navabi et al 2021 (13)** *Pubertal Suppression, Bone Mass, and Body Composition in Youth With Gender Dysphoria* Canada Retrospective review of medical records 2006 - 2017
POPULATION (ages) Age at start Age in cohort Tanner stage	Age in cohort: (years ± SD) 15.2 (± 1.8) transgender males 15.4 (± 2.0) transgender females 90.7 % Tanner 4–5 transgender males 80.3 % Tanner 4–5 transgender females
POPULATION (n) n patients natal male (M-t-F) natal female (F-t-M)	198 youth 172 included 119 transgender males (female at birth) 51 transgender females (male at birth) 2 nonbinary Pre-Post GnRH analysis: 116 individuals: 80 transgender males 36 transgender females
INTERVENTION (type) Puberty suppression (GnRH) Cross-sex hormone treatment (CSHT)	GnRHa: leuprolide acetate i.m. start at 7.5 mg/4 weeks (3 doses), followed by 11.25 mg/ 12 w. calcium carbonate 500 mg twice daily (advised for youth with poor calcium intake) vitamin D 1000 to 2000 IU daily (advised for all youth) Dual-energy radiograph absorptiometry
INTERVENTION (time) Treatment duration Follow-up time, Follow-up age	FU times: 6, 12 and 18 months Pre-GnRHa DXA: at -51.4 ± 41.3 days (range -158 to +28 days) relative to GnRHa initiation. Post-GnRHa DXA: at 355.2 ± 96.7 days (range 188–676 days) after GnRHa initiation (median 352.5 (294.5, 385.8) Mean time interval between pre- and post-DXA scans: 406.7 ± 98.3 days (range 210–720 days).
OUTCOMES - Reported outcomes	aBMD areal bone mineral density aBMD z scores Lumbar spine (LS) (L2–L4) left total hip (LTH) aBMD z scores Vitamin D status
RESULTS – Extracted outcomes	At baseline: Transgender females had lower z scores at lumbar spine aBMD, LS BMAD, left total hip aBMD, and bone mineral content (BMC) than transgender males. 55.2 % of transgender youth had vitamin D deficiency or insufficiency. **Post-pre-GnRH** mean difference (95% CI) Transgender males: Lumbal spine aBMD z score -0.74 (-0.85 to - 0.63) BMAD z score -0.59 (-0.74 to - 0.45) Left total hip aBMD z score -0.33 (-0.40 to -0.26) Total body less head aBMD z score -0.34 (-0.43 to -0.25) Transgender females: Lumbal spine aBMD z score -0.33 (-0.46 to -0.19) BMAD z score -0.37 (-0.61 to -0.14) Left total hip aBMD z score -0.46 (-0.60 to -0.31) Total body less head aBMD z score -0.34 (-0.48 to -0.21)
Author, Year (ref) Title Country Study design	**van der Loos et al 2021** (14) *Development of Hip Bone Geometry During Gender-Affirming Hormone Therapy in Transgender Adolescents Resembles That of the Experienced Gender When Pubertal Suspension Is Started in Early Puberty* The Netherlands Retrospective cohort, 2011-2018

POPULATION (ages) Age at start Age in cohort Tanner stage	Age at start of GnRH: (min Tanner B2, Tanner G2–G3): 11-17 years Age at start of CSHT: 15 – 17 years At start of GnRH: early, mid or late puberty groups: Tanner stage: early: B2; mid: B3; late: B4 and B5 Testicular volume: early: ≤9 mL; mid: 10–19 mL; late: ≥20 mL
POPULATION (n) n patients natal male (M-t-F) natal female (F-t-M)	322 included 106 transwomen (early: n=32; mid: n=30; late: n=44) 216 transmen (early: n=8; mid: n= 22; late: n=186) 115 gonadectomy
INTERVENTION (type) Puberty suppression (GnRH) Cross-sex hormone treatment (CSHT)	GnRHa: triptorelin s.c. 3.75 mg / 4 weeks, or 11.25 mg /12 weeks CSHT (GAH- gender affirming hormone treatment): 17-beta-estradiol oral, start at 5 µg/kg, increased up to 2 to 4 mg/day. Testosterone ester mixture i.m. 25 mg/m2, increased up to 250 mg / 3 to 4 weeks. Surgery: Gonadectomy at earliest age 18 years (if performed, GnRH was stopped afterwards) Study intervention: DXA: narrow neck hip structure analysis (HSA)
INTERVENTION (time) Treatment duration Follow-up time, Follow-up age	GnRH duration (min 6 months): range 1-4 years CSHT duration: range 2-6 years DXA after ≥2years of CSHT
OUTCOMES - Reported outcomes	Subperiostal width Endocortical diameter BMI, Height, Hormone levels
RESULTS – Extracted outcomes	**Subperiosteal Width and Endocortical Diameter,** Change in Centimeters, mean (95% CI) Δ between start of GnRHa and start of GAH / Δ between the start of GnRHa and after ≥2 years of GAH / Δ between the start of GAH and after ≥2 years of GAH / <u>Trans women</u> Early puberty Subperiosteal width 0.38 (0.16; 0.60) / 0.44 (0.23; 0.65) / 0.06 (–0.15; 0.27) Endocortical diameter 0.39 (0.16; 0.61) / 0.38 (0.17; 0.60) /–0.00 (–0.21; 0.21) Mid puberty Subperiosteal width 0.33 (0.15; 0.50) / 0.57 (0.39; 0.75) / 0.25 (0.11; 0.38) Endocortical diameter 0.34 (0.17; 0.51) / 0.55 (0.37; 0.72) / 0.21 (0.08; 0.34) Late puberty Subperiosteal width 0.06 (–0.08; 0.20) / 0.27 (0.16; 0.39) / 0.21 (0.09; 0.34) Endocortical diameter 0.08 (–0.06; 0.22) / 0.27 (0.15; 0.40) / 0.19 (0.06; 0.33) <u>Trans men</u> Early puberty Subperiosteal width 0.63 (0.58; 0.68) / 0.79 (0.72; 0.85) / 0.15 (0.12; 0.19) Endocortical diameter 0.62 (0.57; 0.67) / 0.73 (0.67; 0.79) /0.11 (0.08; 0.14) Mid puberty Subperiosteal width 0.10 (–0.09; 0.29) / 0.31 (0.11; 0.50) / 0.21 (0.03; 0.38) Endocortical diameter 0.09 (–0; 11; 0.30) / 0.27 (0.06; 0.48) / 0.18 (–0.01; 0.36) Late puberty Subperiosteal width 0.07 (–0.03; 0.18) / 0.15 (0.04; 0.26) / 0.07 (–0.04; 0.18) Endocortical diameter 0.10 (–0.01; 0.21) / 0.17 (0.05; 0.28) / 0.07 (–0.04; 0.17) "development of hip bone geometry in transgender adolescents resembled that of the experienced gender if the GnRHa treatment was initiated during early puberty and was followed by a start of GAH. Only participants starting during early puberty showed more resemblance to the reference curves of their experienced gender. Participants starting GnRHa and GAH treatments during mid or late puberty continued within the curve of their gender assigned at birth"
Author, Year (ref) Title Country Study design	**Lee et al 2020 (15)** *Low Bone Mineral Density in Early Pubertal Transgender/Gender Diverse Youth: Findings From the Trans Youth Care Study* USA Cross-sectional analysis of prospective, observational, longitudinal cohort, multicenter

POPULATION (ages) Age at start Age in cohort Tanner stage	Age at start of GnRH: 11.0 ± 1.4 years designated females at birth (DFAB) 12.1 ± 1.3 years designated males at birth (DMAB)
POPULATION (n) n patients natal male (M-t-F) natal female (F-t-M)	63 transgender youth 30 designated females at birth (DFAB) 33 designated males at birth (DMAB) Tanner stages 2-3: 40 (63.5%) Tanner 2 23 (36.5%) Tanner 3
INTERVENTION (type) Puberty suppression (GnRH) Cross-sex hormone treatment (CSHT)	GnRH (not further specified) Study intervention: DXA (before or 2 months after start of GnRH): DXA scans: total body less head (TBLH) lumbar spine total hip femoral neck Quantitative computed tomography (QCT): cortical and trabecular vBMD: midshaft femur L1-L3 vertebral bodies.
INTERVENTION (time) Treatment duration Follow-up time, Follow-up age	GnRH duration before DXA: 0-2 months
OUTCOMES - Reported outcomes	Areal and volumetric BMD Z-scores dietary calcium serum 25-hydroxy-vitamin D physical activity (assessed with Physical Activity Questionnaire for Older Children (PAQ-C))
RESULTS – Extracted outcomes	Bone health: Areal and volumetric BMD Z-scores. BMD assessed before initiation of GnRHa: 90% (57/63) of participants Low aBMD or vBMD Z-score, defined as < -2: in 30% (95% CI 15.6-48.7) of DMAB (10/33) in 13% (95% CI 3.8-30.7)) of DFAB (4/30) At least 1 BMD Z-score was < -2 in: 30% of DMAB 13% of DFAB Designated males at birth (DMAB): BMD Z-scores below-average compared with male reference standards. Designated females at birth (DFAB): BMD Z-scores below-average when compared with female reference standards except at hip sites. Physical Activity Questionnaire for Older Children: low score in youth with low BMD than youth with normal BMD. Dietary calcium intake: suboptimal in all youth. Vitamin D: no significant deficiencies.

Table 3 Effects on anthropometric measures and metabolism by puberty suppression in adolescents

Author, Year (ref) Title Country Study design	**Schagen et al 2016** (16) *Efficacy and Safety of Gonadotropin-Releasing Hormone Agonist Treatment to Suppress Puberty in Gender Dysphoric Adolescents* The Netherlands Prospective cohort study, before-after, 1998 – 2009
POPULATION (ages) Age at Tx start Age in cohort Tanner stage	Age at start: M-t-F: Range 11.6–17.9 years Median 13.6 years Tanner G2–G5 F-t-M: Range 11.1–18.6 years Median 14.2 years Tanner B2-B5
POPULATION (n) n patients natal male (M-t-F) natal female (F-t-M)	116 49 M-t-F 67 F-t-M 77 analyzed: 36 M-t-F 41 F-t-M 0 discontinued GnRH treatment
INTERVENTION (type) Puberty suppression (GnRH) Cross-sex hormone treatment (CSHT)	GnRH: Triptorelin (Decapeptyl-CR) 3.75 mg i.m. at 0, 2, and 4 weeks, followed by every 4 weeks. Study intervention: Dual energy x-ray absorptiometry (DEXA)
INTERVENTION (time) Treatment duration Follow-up time, Follow-up age	GnRH duration: 3 to 12 months (depended on when the individual reached the age at which CSHT could be added)
OUTCOMES - All reported outcomes	Physical examination Tanner stage (breast development, testicular volume) Height and weight, height SD score Body mass index (BMI), BMI SD score Body composition: (fat mass, fat %, lean body mass %) Hormone levels: LH, FSH, testosterone, estradiol Liver enzymes: alkaline phosphatase (AP), aspartate aminotransferase (AST), alanine aminotransferase (ALT), gamma-glutamyl transferase Creatinine
RESULTS Extracted outcomes	**At start GnRH / at 1 y GnRH** (mean (SD)) <u>M-t-F (n=36):</u> Height (cm) 167.8 (7.5) / 172.3 (6.5) Weight (kg) 57.4 (11.1) / 63.3 (11.9) BMI (kg/m2) 20.3 (3.0) / 21.2 (3.2) Lean body mass (%) 74.6 (6.4) / 70.9 (7.3) Alkaline phosphatase (U/L) 303 (109) / 216 (79) Creatinine (mmol/L) 70 (12) / 66 (13) <u>F-t-M (n=41) :</u> Height (cm) 161.4 (8.4) / 163.5 (7.9) Weight (kg) 55.1 (14.7) / 59.5 (14.4) BMI (kg/m2) 21.0 (4.5) / 22.1 (4.6) Lean body mass (%) 71.5 (6.7) / 67.7 (6.7) Alkaline phosphatase (U/L) 215 (101) / 168 (58) Creatinine (mmol/L) 73 (8) / 68 (13)

Author, Year (ref) Title Country Study design	**Klaver et al. 2018** (17) *Early Hormonal Treatment Affects Body Composition and Body Shape in Young Transgender Adolescents* The Netherlands Retrospective cohort study of medical records, before-after, 1998–2014
POPULATION (ages) Age at Tx start Age in cohort Tanner stage	Age at start of GnRH: Min age: 12 years Min Tanner B2 (girls) Min Tanner G3 (boys) 14.5 ± 1.8 years transwomen 15.3 ± 2.0 years transmen Age at start of CSHT: Min age 16 years 16.4 ± 1.1 years transwomen 16.9 ± 0.9 years transmen
POPULATION (n) n patients natal male (M-t-F) natal female (F-t-M)	192 71 transwomen (MtF) (birth-assigned boys) 121 transmen (FtM) (birth-assigned girls)
INTERVENTION (type) Puberty suppression (GnRH) Cross-sex hormone treatment (CSHT)	GnRH: 3.75 mg for 4 weeks until gonadectomy Cross-sex hormonal treatment (CSHT): 17b-estradiol oral (5 mg/kg/day, increased by 5 mg/kg/day every 6 months until 2 mg/day) mixed testosterone esters i.m. (25 mg/m2/ 2 weeks, increased by 25 mg/m2 every 6 months until 250 mg/m2/3-to 4 weeks) Surgery: Gonadectomy Study intervention: Whole-body dual-energy x-ray absorptiometry
INTERVENTION (time) Treatment duration Follow-up time, Follow-up age	GnRH duration: until gonadectomy, at earliest age 18 Follow-up time: *GnRH monotherapy:* 2.1 years (1.0–2.8) transwomen (M-t-F) 1.0 years (0.5–2.9) transmen (F-t-M) *GnRH + CSHT:* 3.1 years (2.5–3.6) transwomen (M-t-F) 2.4 years (2.0–3.1) transmen (F-t-M) *CSHT monotherapy:* 2.8 years (1.6–3.4) transwomen (M-t-F) 3.0 years (1.9–3.4) transmen (F-t-M) Follow-up age: 22 years
OUTCOMES - All reported outcomes	Body weight, BMI Waist circumference (cm), Hip circumference Change in waist-hip ratio (WHR) total body fat (TBF), android (%), gynoid (%) total lean body mass (LBM)
RESULTS Extracted outcomes	**At start of GnRH** (±4months) **/ at start of CSHT** (±4months) **/ at age 22** (±1.5 years) <u>Transwomen (MtF):</u> Body weight (kg) 58 (56–61) / 66 (63–69) / 76 (71–82) BMI (kg/m²) 20.2 (19.4–20.9) / 21.3 (20.5–22.0) / 23.2 (21.6–24.8) WHR 0.81 (0.79–0.82) / 0.79 (0.78–0.80) / 0.77 (0.75–0.79) LBM (%) 75 (74–77) / 69 (68–71) / 66 (64–68) <u>Transmen (FtM):</u> Body weight (kg) 58 (56–61) / 63 (60–65) / 69 (66–71) BMI (kg/m2) 21.6 (20.9–22.3) / 22.5 (21.7–23.2) / 23.9 (23.0–24.7) WHR 0.77 (0.76–0.78) / 0.76 (0.75–0.77) / 0.80 (0.78-0.82) LBM (%) 70 (69–71) / 67 (66–68) / 73 (72–74)

Author, Year (ref) Title Country Study design	**Klaver et al. 2020** (18) *Hormonal Treatment and Cardiovascular Risk Profile in Transgender Adolescents* The Netherlands Retrospective cohort study, before after, 1998–2015
POPULATION (ages) Age at Tx start Age in cohort Tanner stage	At min age 12 years Tanner B2 (girls) Tanner G3 (boys) Age at start of GnRHa (mean (SD)): 14.6 years (1.8) transwomen 15.2 years (2.0) transmen Age at start of CSHT: (mean (SD)): 16.4 years (1.1) transwomen 16.9 years (0.9) transmen
POPULATION (n) n patients natal male (M-t-F) natal female (F-t-M)	192 71 transwomen (M-t-F) 121 transmen (F-t-M)
INTERVENTION (type) Puberty suppression (GnRH) Cross-sex hormone treatment (CSHT)	GnRH: 3.75 mg/4 weeks s.c. Cross sex hormonal treatment (CSHT): (from age 16 years): 17-b estradiol (E2) oral (5 µg/kg/day, increased every 6 months until 2 mg/day) mixed testosterone esters i.m. (25 mg/m²/2 weeks, increased every 6 months until 250 mg/3–4 weeks. When GnRHs were started after age 16: Cross-sex hormones added: after 3 to 6 months: start dose 1 mg E2 daily or 75 mg of testosterone esters i.m weekly after 6 months: 2 mg E2 daily or 250 mg of testosterone esters /3–4 weeks
INTERVENTION (time) Treatment duration Follow-up time, Follow-up age	GnRHa monotherapy duration (median (IQR)): 2.1 (1.0–2.7) transwomen 1.0 (0.5–2.9) transmen GnRHa + CSHT duration (median (IQR)): 3.1 (2.5–3.6) transwomen 2.3 (1.8–2.8) transmen CSHT monotherapy duration (median (IQR): 2.2 (1.1–3.1) transwomen 2.9 (1.7–3.4) transmen Follow-up age: 22 years: Range 20.5–23.5 years
OUTCOMES - All reported outcomes	Changes in body mass index (BMI) systolic blood pressure (SBP) diastolic blood pressure (DBP) glucose homeostatic model assessment for insulin resistance (HOMA-IR) lipid values prevalence of obesity dyslipidaemia

RESULTS Extracted outcomes	**At start of GnRH / at 22 years/ change during GnRH / change between start of CSHT and age 22** (mean (95% CI) <u>Transwomen (n=71):</u> BMI 20.2 (19.4 to 20.9) / 23.2 (21.6 to 24.8) / +1.1 (0.7 to 1.5) / +1.9 (0.6 to 3.2) SBP (mmHg) 120 (116 to 123) / 117 (113 to 122) / +1 (–3 to 5) / –3 (–8 to 2) DBP (mmHg) 65 (63 to 67) / 75 (72 to 78) / +4 (1 to 7) / +6 (3 to 10) Glucose (mmol/L) 5.0 (4.8 to 5.2) / 5.0 (4.8 to 5.1) / –0.1 (–0.3 to 0.1) / +0.1 (–0.1 to 0.2) Insulin (mU/L) 9.5 (6.7 to 12.2) / 13.0 (8.4 to 17.6) / +0.8 (–2.5 to 4.1) / +2.7 (–1.7 to 7.1) HOMA-IR 2.3 (1.2 to 3.4) / 2.9 (1.9 to 3.9) / 0.0 (–1.2 to 1.2) /+0.7 (–0.2 to 1.5) Total cholesterol (mmol/L) 3.7 (3.5 to 3.9) / 4.1 (3.8 to 4.4) / 0.3 (0.2 to 0.5) / 0.1 (20.2 to 0.4) HDL cholesterol (mmol/L) 1.4 (1.3 to 1.5) / 1.6 (1.4 to 1.7) / +0.2 (0.1 to 0.3) / 0.0 (–0.1 to 0.2) LDL cholesterol mmol/L 1.9 (1.7 to 2.1) / 2.0 (1.8 to 2.3) / +0.2 (0.0 to 0.3) / 0.0 (–0.3 to 0.2) Triglycerides (mmol/L) 0.8 (0.7 to 0.9) / 1.1 (0.9 to 1.4) / +0.1 (–0.1 to 0.2) / +0.2 (0.0 to 0.5) <u>Transmen (n=121):</u> BMI 21.6 (20.9 to 22.3) / 23.9 (23.0 to 24.7) / +0.9 (0.5 to 1.3) / +1.4 (0.8 to 2.0) SBP (mmHg) 120 (118 to 122) / 126 (122 to 130) / +2 (–1 to 4) / +5 (1 to 9) DBP (mmHg) 67 (66 to 69) / 74 (72 to 77) / +1 (–1 to 3) / +6 (4 to 9) Glucose (mmol/L) 4.8 (4.7 to 4.9) / 4.8 (4.7 to 5.0) / +0.1 (–0.1 to 0.2) / 0.0 (–0.2 to 0.2) Insulin (mU/L) 9.5 (8.0 to 11.0) / 8.6 (6.9 to 10.2) / +1.2 (–0.6 to 3.0) / –2.1 (–3.9 to –0.3) HOMA-IR 2.1 (1.6 to 2.5) / 1.8 (1.4 to 2.2) / +0.3 (–0.2 to 0.8) / –0.5 (–1.0 to –0.1) Total cholesterol (mmol/L) 3.9 (3.7 to 4.0) / 4.6 (4.3 to 4.8) / +0.3 (0.2 to 0.4) / +0.4 (0.2 to 0.6) HDL cholesterol (mmol/L) 1.5 (1.4 to 1.5) / 1.3 (1.2 to 1.3) / +0.1 (0.1 to 0.2) / –0.3 (–0.4 to –0.2) LDL cholesterol (mmol/L) 2.1 (1.9 to 2.2) / 2.6 (2.4 to 2.8) / +0.2 (0.1 to 0.3) / +0.4 (0.2 to 0.6) Triglycerides (mmol/L) 0.8 (0.7 to 0.8) / 1.3 (1.1 to 1.5) / 0.0 (0.0 to 0.1) / +0.5 (0.3 to 0.7) **Obesity prevalence** (at age 22): BMI ≥30 in both sexes 9.9% in transwomen (M-t-F) 6.6% in transmen (F-f-M) 2.2% in ciswomen (females) 3.0% in cismen (males)

Author, Year (ref) Title Country Study design	**Perl et** al 2020 **(19)** *Blood Pressure Dynamics After Pubertal Suppression with Gonadotropin-Releasing Hormone Analogs Followed by Testosterone Treatment in Transgender Male Adolescents: A Pilot Study* Israel Retrospective pilot study, 2013 - 2018
POPULATION (ages) Age at Tx start Age in cohort Tanner stage	Age at start of GnRH: 14.4 ±1.0 years Tanner stage 4/5 Age at start of testosterone: 15.1 ± 0.9
POPULATION (n) n patients natal male (M-t-F) natal female (F-t-M)	48 transgender male adolescents 15 included 15 GnRH subsequently were 9 treated with testosterone
INTERVENTION (type) Puberty suppression (GnRH) Cross-sex hormone treatment (CSHT)	Previous intervention: GnRHa D-Trp-6-LHRH depot (3.75mg/4 weeks intramuscular injection) CSHT: (patients who reached ≥14 years of age) testosterone enanthate intramuscular injection (250 mg/mL), starting dose of 50–100 mg /4 weeks. Medical nutrition counseling, not further specified. Psychosocial support , not further specified
INTERVENTION (time) Treatment duration Follow-up time, Follow-up age	GnRHa duration: 3 ± 1 months. Testosterone duration: 4 ± 2 months
OUTCOMES - All reported outcomes	BMI BP (procedure for measurement not given) luteinizing hormone (LH) follicle-stimulating hormone (FSH) estradiol testosterone
RESULTS Extracted outcomes	Anthropometric (before GnRH; after GnRH; before testosterone; after testosterone) mean – SD BMI (kg/m2), mean ± SD 21.3 ± 4.7 ; 22.0 ± 4.8 ; 23.3 ± 5.6 ; 24.2 ± 4.6 BMI-SDS did not increase significantly during GnRHa therapy. Diastolic BP percentiles: mean ± SD 56% ± 26 ; 74% ± 9.0 ; 74% ± 9.0 ; 56% ± 17 DBP percentiles increased significantly after GnRHa treatment and remained significant after adjusting for the change in BMI-SDS. DBP percentile decreased after adding testosterone. BP levels did not meet criteria for hypertension. Systolic BP percentiles: mean ± SD 71% – 19 ; 76% – 14 ; 76% – 14 ; 72% – 21 BP levels within the normal range and did not meet criteria for pediatric hypertension.

Author, Year (ref) Title Country Study design	**Schulmeister et al. 2021** (20) *Growth in Transgender / Gender-Diverse Youth in the First Year of Treatment with Gonadotropin-Releasing Hormone Agonists* USA Multisite prospective observational study, 2016 - 2018
POPULATION (ages) Age at Tx start Age in cohort Tanner stage	Age at GnRHa start (mean (range)): 11.5 years (9.0-14.5) total 11.9 years (10.2-14.5) male at birth 11.1 years (9.0-13.9) female at birth Comparison group: 11.0 ± 2.8 years, Tanner I Tanner stage at GnRHa start (n (%)): Tanner II 34 (62%) total; 21 (81%) male at birth; 13 (45%), female at birth Tanner III 16 (29%) total; 3 (12%) male at birth; 13 (45%) female at birth Tanner IV 5 (9%) total; 2 (8%) male at birth; 3 (10%) female at birth
POPULATION (n) n patients natal male (M-t-F) natal female (F-t-M)	92 enrolled 55 in cohort 26 male at birth 29 female at birth Comparison group: 226 participants: 118 males 108 female Prepubertal, presumed cisgender youth not receiving hormonal intervention from the *Bone Mineral Density in Childhood Study* (BMDCS) (Age-based reference ranges for annual height velocity in US children. Kelly, Winer, Kalkwarf, Oberfield, Lappe, Gilsanz, Zemel; J Clin Endocrinol Metab 2014 Jun; 99(6): 2104-12). Exclusions: Serious psychiatric symptoms.
INTERVENTION (type) Puberty suppression (GnRH) Cross-sex hormone treatment (CSHT)	GnRH: Drug, dose and frequency not reported. Full description of study protocol published in [Olson-Kennedy J, Chan YM, Garofalo R, et al. Impact of early medical treatment for transgender youth: Protocol for the longitudinal, observational trans youth care study. J Med Internet Res 2019; 21: e14434]
INTERVENTION (time) Treatment duration Follow-up time, Follow-up age	Duration: GnRHa: min 10 months max 14 months. FU time: Prior to beginning GnRHa (baseline), 6- and 12-month follow-up visits.
OUTCOMES - All reported outcomes	HV (height velocity) BMI FSH (follicle-stimulating hormone) LH (luteinizing hormone) estradiol testosterone

RESULTS Extracted outcomes	Height velocity (HV) in the first year of GnRHa use: 5.1 (3.7-5.6) cm/year (median (IQR)). Later Tanner stage at GnRHa initiation was associated with lower HV: 5.3 (4.4-5.6) cm/year for Tanner stage II 4.4 (3.3-6.0) cm/year for Tanner stage III 1.6 (1.5-2.9) cm/year for Tanner stage IV Height velocity by Tanner stage at baseline ((cm/year) median (IQR)) (total; designated male at birth; designated female at birth) Tanner stage II 5.3 (4.4-5.6) total; 5.6 (4.7-5.7) male at birth; 5.0 (4.2-5.4) female at birth Tanner stage III 4.4 (3.3-6.0) total; 4.2 (2.3-6.4) male at birth; 4.4 (4.0-5.5) female at birth Tanner stage IV 1.6 (1.5-2.9) total; 1.5 (1.4-1.6) male at birth: 2.9 (1.5-3.5) female at birth BMI z-score (mean (SD)) (total; designated male at birth; designated female at birth) Baseline visit 0.46 (0.89) total; 0.56 (0.84) male at birth; 0.38 (0.94) female at birth 12-month visit 0.66 (0.97) total; 0.68 (1.00) male at birth; 0.63 (0.95) female at birth When controlled for age, there was not a significant difference in mean height velocity between transgender youth and prepubertal youth (comparison group);

Author, Year (ref) Title Country Study design	**Nokoff et al 2020 (21)** *Body Composition and Markers of Cardiometabolic Health in Transgender Youth Compared With Cisgender Youth* USA Cross-sectional study, controlled, 2016-2019
POPULATION (ages) Age at Tx start Age in cohort Tanner stage	Age at start of GnRH (mean ± SD): 12.1 ± 1.9 years transgender males 12.8 ± 1.3 years transgender females Age in cohort (mean ± SD): 13.8 ± 1.7 years (range 10.1–16.0) transgender males 13.7 ± 1.2 years (range 12.6–16.1) transgender females Comparator groups: 10.6–16.2 years cisgender females 12.5–15.5 years cisgender males
POPULATION (n) n patients natal male (M-t-F) natal female (F-t-M)	17 youth 9 transgender males on GnRHa 8 transgender females on GnRHa Comparator groups: 31 youth 14 cisgender females 17 cisgender males Exclusions: Significant medical or psychiatric comorbidities (incl. diabetes or antipsychotic treatment)
INTERVENTION (type) Puberty suppression (GnRH) Cross-sex hormone treatment (CSHT)	GnRH: Drug, dose and frequency not reported.
INTERVENTION (time) Treatment duration Follow-up time, Follow-up age	GnRHa duration (mean ± SD): 20.9 ± 19.8 months transgender males (range 17.5-70.4 months) 11.3 ± 7 months transgender females (range 4.7-24.2 months)
OUTCOMES - All reported outcomes	insulin sensitivity and body composition insulin sensitivity (1/ (fasting insulin), homeostatic model of insulin resistance [HOMA-IR]), glycemia (hemoglobin A1C (HbA1c), fasting glucose), BMI, body mass index BP, blood pressure AST, aspartate aminotransferase ALT, alanine aminotransferase HDL, high-density lipoprotein LDL, low-density lipoprotein SHBG, sex hormone-binding globulin LH, luteinizing hormone FSH, follicle stimulating hormone estradiol testosterone
RESULTS Extracted outcomes	Transgender males vs cisgender females: 1/fasting insulin (0,067 ± 0,02 vs 0,103 ± 0,049 mL/µU) HOMA-IR (3,7 ± 1,7 vs 2,3 ± 1,1) fasting glucose (89 ± 4 vs 79 ± 13 mg/dL) HbA1c (5.4 ± 0.2 vs. 5.2 ± 0.2%) percent body fat (36 ± 7 vs 32 ± 5%) Transgender females vs cisgender males: 1/fasting insulin (0,076 ± 0,029 vs 0,135 ± 0,049 mL/µU) HOMA-IR (3,5 ± 1,4 vs 2,2 ± 1,3) HbA1c (5.4 ± 0.1% vs 5.1 ± 0.2%) percent body fat (31 ± 9 vs 24 ± 10%) lower percent lean mass (66 ± 8 vs 74 ± 10%)

Table 4 Effects of cross-sex hormonal treatment started before age of 18 years without previous puberty suppression

Author, Year (ref) Title Country Study design	**Tack et al 2016 (22)** *Consecutive lynestrenol and cross-sex hormone treatment in biological female adolescents with gender dysphoria: a retrospective analysis.* Belgium Retrospective cohort study, 2010–2015
POPULATION (ages) Age at Tx start Age in cohort Tanner stage	Age at start of lynestrenol: Min Tanner B4 (post menarche) 15 years and 10 months (mean) Age at start of testosterone: 17 years and 5 months (mean)
POPULATION (n) n patients natal male (M-t-F) natal female (F-t-M)	45 initials 43 in cohort (F-t-M) Of 45 subjects: 25 testosterones added later 11 psychiatric comorbidities (unspecified) 1 suicide during follow-up 1 did not consent use of data
INTERVENTION (type) Puberty suppression (GnRH) Cross-sex hormone treatment (CSHT)	Hormone treatment: Androgenic progestin: lynestrenol (L) (Orgametril®) monotherapy: dose not reported Testosterone esters (Sustanon®): added from age 16: start at 50 mg (16 years) or 100 mg (17–19 years)/ 2 weeks (injection); incremental increases (+25 mg) up to 125 mg/2 weeks, up to 18 months. Vitamin D and calcium supplements Psychiatric intervention: During treatment, patients seen every 3 months by the team child psychologist. In the absence of psychiatric comorbidity, evaluated twice by the team child psychiatrist during this phase; once before initiation of lynestrenol and once more at start of lynesterol + testosterone.
INTERVENTION (time) Treatment duration Follow-up time, Follow-up age	Treatment duration: (min 6 months, up to 18 months) Mean 12.6 months Lynestrenol (L) Mean 11.4 months Lynestrenol (L) + testosterone esters (T):
OUTCOMES - All reported outcomes	Anthropometry Safety parameters, side effects Biochemical analysis: complete blood count, electrolytes, liver, and renal function, fasting glucose, insulin, lipid metabolism Hormone levels: Thyroid stimulating hormone (TSH), free thyroxin (fT4), luteinizing hormone (LH), follicular stimulating hormone (FSH), estradiol (E2), total and free testosterone (T and free T), sex hormone-binding globulin (SHBG), anti-Müllerian hormone (AMH)
RESULTS Extracted outcomes	**At start of lynestrenol / at 12 months of L / at start of testosterone / at 12 months of T** Mean height 164.6 / --/ -- / 167.6 / -- Weight 61.48 / 61.03 / 58.65 / 65.10 BMI 22.58 / 22.39 / 20.69 / 23.26 Triglycerides (mmol/L) 0.838 / 0.661 / 0.651 / 1.394 Total cholesterol (mmol/l) 4.153 / 4.237 / 4.212 / 4.450 HDL (mmol/l) 1.481 / 1.017 / 1.098 / 1.085 Side effects: Metrorrhagia: in L+T long term Acne: In L no increase, in L+T significant increase Headaches: in L Hot flushes: in L Fatigue: in L+T

Author, Year (ref) Title Country Study design	Jarin et al 2017 (23) *Cross-Sex Hormones and Metabolic Parameters in Adolescents With Gender Dysphoria* USA Retrospective, cohort study, 2008-2014
POPULATION (ages) Age at Tx start Age in cohort Tanner stage	Age in cohort: Range: 13 – 25 years Affirmed male: mean 16 years range 13 - 22 Affirmed female: mean 18 years range 14 - 25
POPULATION (n) n patients natal male (M-t-F) natal female (F-t-M)	161 adolescents: 72 affirmed males (FtM) 44 affirmed females (MtF) 7 affirmed males on GnRHa before treatment 2 affirmed females on GnRHa before treatment 2 affirmed males reported hormone use outside medical practice (street hormones) 5 affirmed females reported exogenous street hormone use. Comorbidities: 35 depression 11 anxiety 8 ADHD 10 HIV
INTERVENTION (type) Puberty suppression (GnRH) Cross-sex hormone treatment (CSHT)	CSHT: Testosterone (s.c.): 25 mg/ week, weekly doses of 25, 50, or 100 mg at subsequent visits. Oestrogen (± testosterone blocker spironolactone): orally at 1, 2, 3, 4, 6, and 8 mg daily; or intramuscularly at 20, 40, or 80 mg monthly; or trans dermally at 0.025, 0.05, 0.100, or 0.200 mg weekly
INTERVENTION (time) Treatment duration Follow-up time, Follow-up age	Follow-up time: Up to 35 months. Follow-up groups: 1 to 3 months after initiation 4 to 6 months after initiation 6 months and beyond
OUTCOMES - All reported outcomes	Body mass index (BMI) Systolic blood pressure (SBP), Diastolic blood pressure (DBP) Hematokrit, Haemoglobin Total testosterone Estradiol Total cholesterol, Low density lipoprotein (LDL), High density lipoprotein (HDL), Triglycerides (TG) TG : HDL ratio Creatinine Prolactin Aspartate aminotransferase, (AST), Alanine aminotransferase (ALT) HbA1c
RESULTS Extracted outcomes	<u>Affirmed male (FtM):</u> BMI: increased at 6 months (from 26.0 to 27.3) DBP: reduced at 6 months (from 71 to 67 mm Hg) Hematokrit: increased at 6 months (from 39.4% to 44.5%) 2 subjects had supraphysiologic hematokrit levels (>50%) after 3 months of treatment, 1 subject maintained elevated hematokrit levels after 6 and 9 months (51.0% and 52.7%) Haemoglobin: increased at 6 months. Cholesterol: nonsignificant increase at 6 months (nonsignificant), plateau after 3 months. (6 subjects had cholesterol levels >200 mg/dL). LDL: nonsignificant increase at 6 months, plateau after 3 months. HDL: level decreased at 6 months (from of 50.2 to 45.0 mg/dL). <u>Affirmed female (MtF):</u> No significant changes in any other parameter tested were found. No statistically significant difference in measured metabolic parameters among the various methods of oestrogen administration (patch, oral, or intramuscular).

Author, Year (ref) Title Country Study design	**Mullins et al 2021 (24)** *Thrombosis Risk in Transgender Adolescents Receiving Gender-Affirming Hormone Therapy.* USA Retrospective chart review, 2013 - 2019
POPULATION (ages) Age at Tx start Age in cohort Tanner stage	Age at start of CSHT: range 13 - 24 years 17 years (IQR 15–19) total cohort 18 years (IQR 15.5–20) estrogen 17 years (IQR 15–19) testosterone
POPULATION (n) n patients natal male (M-t-F) natal female (F-t-M)	611 participants 428 female at birth 183 male at birth
INTERVENTION (type) Puberty suppression (GnRH) Cross-sex hormone treatment (CSHT)	Estrogen: 4.0 mg (2.0–6.0mg): oral (90.7%), transdermal (5.5%), intramuscular (3.8%) Testosterone: 70.0 mg (60.0–80.0) s.c (72.7%), i.m. (24.4%), gel (2.8%), transdermal (0.7%) Previous hormones used (%): Norethindrone contraceptive pill (24.2%) Depo-medroxyprogesterone acetate (18.5%) Combined oral contraceptive pill (5.7%) Norethindrone acetate (2.5%) LNG-IUS (2.5%) Etonogestrel implant (0.3%)
INTERVENTION (time) Treatment duration Follow-up time, Follow-up age	Treatment duration, days (median, IQR): 554 days (283.0–1037.5) estrogen 577 days (283.0–923.0) testosterone
OUTCOMES - All reported outcomes	Incidence of arterial or venous thrombosis during GAHT Prevalence of thrombosis risk factors, risk factors for thrombosis (migraine with aura, elevated BMI, tobacco use, medical diagnoses associated with increased risk of thrombosis, family history of thrombosis (arterial or venous) and laboratory measures of risk factors for thrombosis) testosterone and estradiol levels complete blood counts coagulation testing result thrombophilia evaluation arterial or venous thrombosis therapeutic anticoagulation treatment prophylactic anticoagulation treatment concurrent with CSHT duration of anticoagulation treatment
RESULTS Extracted outcomes	<u>Hematologic Evaluation and Incidence of Thrombosis</u> 17 (2.8%) referred to haematology Thrombophilia evaluation: 4 (23.5%) elevated factor VIII (>150%) 10 (2.0%) erythrocytosis (>17.7 g/dL) 1 (6.3%) activated protein C resistance ratio (<0.78) 5 (31.3%) PAI-1 (<16.3 IU/mL) 2 (11.8%) Factor V Leiden heterozygous 2 (12.5%) prothrombin G20210A heterozygous 3 (21.4%) MTHFR 677 homozygous 5 (35.7%) PAI-1 4G homozygous 2 (20.0%) elevated homocysteine (>10.7 µmol/L) Thromboprophylaxis before GAHT: 5 (0.8%) Overall cohort 2 (0.3%) History of thrombosis before GAHT 3 (0.5%) No history of thrombosis before GAHT 0 Thrombosis on GAHT Multiple thrombotic risk factors were noted among the cohort, including obesity, tobacco use, and personal and family history of thrombosis. BMI median IQR: 26.0 (22.1–32.0) 40 (6.5%) BMI <18.5 212 (34.7%) BMI 18.5–25 148 (24.2%) BMI 25–30 211 (34.5%) BMI >30

Table 5. Studies investigating discontinuation of treatment and regret in adolescents with gender dysphoria

Author, Year Country	Inclusion period	Population	Treatment	Follow-up method	Follow-up time	Regret
Pullen Sansfaçon et al 2019 (25) Canada	November 2017 – August 2018	35 trans and gender diverse young people aged 9 to17 years	Puberty blockers, hormone therapy, surgery	Semi-structured interviews	Follow-up-time not reported	0/35
Segev-Becker et al 2020 (26) Israel	March 2013 – January 2019	106 (10 prepubertal) consecutive children and adolescents with gender dysphoria, aged <18 years	77 (80%) pubertal patients began GnRH. 61 of these (83%) started gender affirming treatment	Chart review	Median 1.2 years (range, 0 to 5.1 years)	2/96 (pubertal at start) 16/77 (21%) on GnRH did not start gender affirming treatment
Cohen-Kettenis et al 1997 (27) The Netherlands	Time period not given	22 patients (15 FtM, 7 MtF) Mean age at pretest: 17.5 years (range 15-20) Mean age at follow-up: 22.0 years (range 19-27) Post-treatment sample: 14 FtM, 5 MtF	Surgically reassigned (various procedures)	Questionnaires and interview	1 year or more	0/19
Olson-Kennedy et al 2018 (28) USA	June – December 2016	68 FtM undergoing chest surgery Mean age 18.9 (SD 2.5) (range 14–25)	Chest surgery	Chest dysphoria score,	1–5 years after surgery	1/68
Smith et al 2001 (29) The Netherlands	Not given Follow-up interviews from March 1995 until July 1999	Prospective 20 treated adolescent transsexuals Mean age at pretest 16.6 years (range15–19) Mean age at follow-up 21.0 years (range 19–23)	Surgical reassignment Not specified	Semi-structured interview	1–4 years post-surgery	0/20
Mehringer et al 2021 (30) USA	Not given	30 transmasculine 13 to 21 years mean age 17.5 (14-21) 14 had undergone chest surgery. Mean age 16.4 years	Chest surgery/ dysphoria	Interview transcripts coded employing modified grounded theory	19 (6–48) months after surgery	0/14 All post-surgery youth reported near or total resolution of chest dysphoria, lack of regret, improved quality of life and functioning
Nieder et al 2021 (31) Germany	Sept 2013 – June 2017	75 11-21 years	Varying, hormones, various surgery	Clinical follow-up	2 years	0/75
Carmichael et al 2021 (5) The UK	April 2011 – April 2014	44 25 trans women 19 trans male 11-15 years	GnRH	Clinical follow-up	Median 31 months	No data on regret 1/44 did not start gender affirming treatment
Littman 2021 (32) USA	Dec 2016 – April 2017	100 detransitioners, mean age at detransition 26 years Mean age at transition 22 years	Varying gender affirming treatments	Open survey over Internet		

References

1. de Vries AL, McGuire JK, Steensma TD, Wagenaar EC, Doreleijers TA, Cohen-Kettenis PT. Young adult psychological outcome after puberty suppression and gender reassignment. Pediatrics. 2014;134:696-704.
2. Costa R, Dunsford M, Skagerberg E, Holt V, Carmichael P, Colizzi M. Psychological Support, Puberty Suppression, and Psychosocial Functioning in Adolescents with Gender Dysphoria. J Sex Med. 2015;12:2206-14.
3. Becker-Hebly I, Fahrenkrug S, Campion F, Richter-Appelt H, Schulte-Markwort M, Barkmann C. Psychosocial health in adolescents and young adults with gender dysphoria before and after gender-affirming medical interventions: a descriptive study from the Hamburg Gender Identity Service. Eur Child Adolesc Psychiatry. 2021;30:1755–1767
4. Cantu AL, Moyer, D.N., Connely, K.J., Holley, A.L. Changes in Anxiety and Depression from Intake to First Follow-Up Among Transgender Youth in a Pediatric Endocrinology Clinic. Trangender Health. 2020;5: 196–200
5. Carmichael P, Butler G, Masic U, Cole TJ, De Stavola BL, Davidson S, et al. Short-term outcomes of pubertal suppression in a selected cohort of 12 to 15 year old young people with persistent gender dysphoria in the UK. PLoS One. 2021;16:e0243894.
6. Hisle-Gorman E, Schvey NA, Adirim TA, Rayne AK, Susi A, Roberts TA, et al. Mental Healthcare Utilization of Transgender Youth Before and After Affirming Treatment. J Sex Med. 2021;18:1444-54.
7. Staphorsius AS, Kreukels BP, Cohen-Kettenis PT, Veltman DJ, Burke SM, Schagen SE, et al. Puberty suppression and executive functioning: An fMRI-study in adolescents with gender dysphoria. Psychoneuroendocrinology. 2015;56:190-9.
8. Joseph T, Ting J, Butler G. The effect of GnRH analogue treatment on bone mineral density in young adolescents with gender dysphoria: findings from a large national cohort. J Pediatr Endocrinol Metab. 2019;32:1077-81.
9. Klink D, Caris M, Heijboer A, van Trotsenburg M, Rotteveel J. Bone mass in young adulthood following gonadotropin-releasing hormone analog treatment and cross-sex hormone treatment in adolescents with gender dysphoria. J Clin Endocrinol Metab. 2015;100:E270-5.
10. Vlot MC, Klink DT, den Heijer M, Blankenstein MA, Rotteveel J, Heijboer AC. Effect of pubertal suppression and cross-sex hormone therapy on bone turnover markers and bone mineral apparent density (BMAD) in transgender adolescents. Bone. 2017;95:11-9.
11. Schagen SEE, Wouters FM, Cohen-Kettenis PT, Gooren LJ, Hannema SE. Bone Development in Transgender Adolescents Treated With GnRH Analogues and Subsequent Gender-Affirming Hormones. J Clin Endocrinol Metab. 2020;105(12).
12. Stoffers IE, de Vries MC, Hannema SE. Physical changes, laboratory parameters, and bone mineral density during testosterone treatment in adolescents with gender dysphoria. J Sex Med. 2019;16(9):1459-68.
13. Navabi B, Tang K, Khatchadourian K, Lawson ML. Pubertal Suppression, Bone Mass, and Body Composition in Youth With Gender Dysphoria. Pediatrics. 2021;148(4).
14. van der Loos MA, Hellinga I, Vlot MC, Klink DT, den Heijer M, Wiepjes CM. Development of Hip Bone Geometry During Gender-Affirming Hormone Therapy in Transgender Adolescents Resembles That of the Experienced Gender When Pubertal Suspension Is Started in Early Puberty. J Bone Miner Res. 2021;36(5):931-41.
15. Lee JY, Finlayson C, Olson-Kennedy J, Garofalo R, Chan YM, Glidden DV, et al. Low Bone Mineral Density in Early Pubertal Transgender/Gender Diverse Youth: Findings From the Trans Youth Care Study. J Endocr Soc. 2020;4(9):bvaa065.
16. Schagen SE, Cohen-Kettenis PT, Delemarre-van de Waal HA, Hannema SE. Efficacy and Safety of Gonadotropin-Releasing Hormone Agonist Treatment to Suppress Puberty in Gender Dysphoric Adolescents. J Sex Med. 2016;13(7):1125-32.

17. Klaver M, de Mutsert R, Wiepjes CM, Twisk JWR, den Heijer M, Rotteveel J, et al. Early Hormonal Treatment Affects Body Composition and Body Shape in Young Transgender Adolescents. J Sex Med. 2018;15(2):251-60.

18. Klaver M, de Mutsert R, van der Loos M, Wiepjes CM, Twisk JWR, den Heijer M, et al. Hormonal Treatment and Cardiovascular Risk Profile in Transgender Adolescents. Pediatrics. 2020;145(3).

19. Perl L, Perl, Anat Segev-Becker, Galit Israeli, Erella Elkon-Tamir, Asaf Oren. Blood Pressure Dynamics After Pubertal Suppression with Gonadotropin-Releasing Hormone Analogs Followed by Testosterone Treatment in Transgender Male Adolescents: A Pilot Study. LGBT Health Aug/Sep 2020;7(6):340-344.

20. Schulmeister C, Millington K, Kaufman M, Finlayson C, Kennedy JO, Garofalo R, et al. Growth in Transgender/Gender-Diverse Youth in the First Year of Treatment With Gonadotropin-Releasing Hormone Agonists. J Adolesc Health. 2021.

21. Nokoff NJ, Scarbro SL, Moreau KL, Zeitler P, Nadeau KJ, Juarez-Colunga E, et al. Body Composition and Markers of Cardiometabolic Health in Transgender Youth Compared With Cisgender Youth. J Clin Endocrinol Metab. 2020;105(3):e704-14.

22. Tack LJ, Craen M, Dhondt K, Vanden Bossche H, Laridaen J, Cools M. Consecutive lynestrenol and cross-sex hormone treatment in biological female adolescents with gender dysphoria: a retrospective analysis. Biol Sex Differ. 2016;7:14.

23. Jarin J, Pine-Twaddell E, Trotman G, Stevens J, Conard LA, Tefera E, et al. Cross-Sex Hormones and Metabolic Parameters in Adolescents With Gender Dysphoria. Pediatrics. 2017;139(5).

24. Mullins ES, Geer R, Metcalf M, Piccola J, Lane A, Conard LAE, et al. Thrombosis Risk in Transgender Adolescents Receiving Gender-Affirming Hormone Therapy. Pediatrics. 2021;147(4).

25. Pullen Sansfaçon A, Temple-Newhook J, Suerich-Gulick F, Feder S, Lawson ML, Ducharme J, et al. The experiences of gender diverse and trans children and youth considering and initiating medical interventions in Canadian gender-affirming speciality clinics. International Journal of Transgenderism. 2019;20:371-87.

26. Segev-Becker A, Israeli G, Elkon-Tamir E, Perl L, Sekler O, Amir H, et al. Children and Adolescents with Gender Dysphoria in Israel: Increasing Referral and Fertility Preservation Rates. Endocr Pract. 2020;26:423-8.

27. Cohen-Kettenis PT, van Goozen SH. Sex reassignment of adolescent transsexuals: a follow-up study. J Am Acad Child Adolesc Psychiatry. 1997;36:263-71.

28. Olson-Kennedy J, Warus J, Okonta V, Belzer M, Clark LF. Chest Reconstruction and Chest Dysphoria in Transmasculine Minors and Young Adults: Comparisons of Nonsurgical and Postsurgical Cohorts. JAMA Pediatr. 2018;172:431-6.

29. Smith YL, van Goozen SH, Cohen-Kettenis PT. Adolescents with gender identity disorder who were accepted or rejected for sex reassignment surgery: a prospective follow-up study. J Am Acad Child Adolesc Psychiatry. 2001;40:472-81.

30. Mehringer JE, Harrison JB, Quain KM, Shea JA, Hawkins LA, Dowshen NL. Experience of Chest Dysphoria and Masculinizing Chest Surgery in Transmasculine Youth. Pediatrics. 2021;147. e2020013300

31. Nieder TO, Mayer TK, Hinz S, Fahrenkrug S, Herrmann L, Becker-Hebly I. Individual Treatment Progress Predicts Satisfaction With Transition-Related Care for Youth With Gender Dysphoria: A Prospective Clinical Cohort Study. J Sex Med. 2021;18:632-45.

32. Littman L. Individuals Treated for Gender Dysphoria with Medical and/or Surgical Transition Who Subsequently Detransitioned: A Survey of 100 Detransitioners. Arch Sex Behav. 2021;50:3353-69.

Bilaga 2 Studier exkluderade på grund av hög risk för snedvridning (bias) /Appendix 2 Studies excluded due to high risk of bias.

Studier med hög risk för bias/ Studies with high risk of bias

Achille, C., T. Taggart, N.R. Eaton, J. Osipoff, K. Tafuri, A. Lane and T. A. Wilson (2020). Longitudinal impact of gender-affirming endocrine intervention on the mental health and well-being of transgender youths: preliminary results." Int J Pediatr Endocrinol **2020**: 8.

Allen, L. R., L. B. Watson, A. M. Egan and C. N. Moser (2019). "Well-Being and Suicidality Among Transgender Youth After Gender-Affirming Hormones." Clinical Practice in Pediatric Psychology **7**(3): 302-311.

de Vries, A. L., T. D. Steensma, T. A. Doreleijers and P. T. Cohen-Kettenis (2011). "Puberty suppression in adolescents with gender identity disorder: a prospective follow-up study." J Sex Med **8**(8): 2276-2283.

Ghelani, R., C. Lim, C. Brain, M. Fewtrell and G. Butler (2020). "Sudden sex hormone withdrawal and the effects on body composition in late pubertal adolescents with gender dysphoria." J Pediatr Endocrinol Metab **33**(1): 107-112.

Hannema, S. E., S. E. E. Schagen, P. T. Cohen-Kettenis and H. A. Delemarre-Van De Waal (2017). "Efficacy and safety of pubertal induction using 17β-estradiol in transgirls." Journal of Clinical Endocrinology and Metabolism **102**(7): 2356-2363.

Jensen, R. K., J. K. Jensen, L. K. Simons, D. Chen, I. Rosoklija and C. A. Finlayson (2019). "Effect of Concurrent Gonadotropin-Releasing Hormone Agonist Treatment on Dose and Side Effects of Gender-Affirming Hormone Therapy in Adolescent Transgender Patients." Transgend Health **4**(1): 300-303.

Karalexi, M. A., M. K. Georgakis, N. G. Dimitriou, T. Vichos, A. Katsimpris, E. T. Petridou and F. C. Papadopoulos (2020). "Gender-affirming hormone treatment and cognitive function in transgender young adults: a systematic review and **meta-analysis**." Psychoneuroendocrinology **119**: 104721.

López de Lara, D., O. Pérez Rodríguez, I. Cuellar Flores, J. L. Pedreira Masa, L. Campos-Muñoz, M. Cuesta Hernández and J. T. Ramos Amador (2020). "Psychosocial assessment in transgender adolescents." Anales de Pediatria.

Millington, K., C. Schulmeister, C. Finlayson, R. Grabert, J. Olson-Kennedy, R. Garofalo, S. M. Rosenthal and Y. M. Chan (2020). "Physiological and Metabolic Characteristics of a Cohort of Transgender and Gender-Diverse Youth in the United States." J Adolesc Health.

Neyman, A., J. S. Fuqua and E. A. Eugster (2019). "Bicalutamide as an Androgen Blocker With Secondary Effect of Promoting Feminization in Male-to-Female Transgender Adolescents." J Adolesc Health **64**(4): 544-546.

Zucker, K. J., S. J. Bradley, A. Owen-Anderson, D. Singh, R. Blanchard and J. Bain (2011). "Puberty-blocking hormonal therapy for adolescents with gender identity disorder: A descriptive clinical study." Journal of Gay and Lesbian Mental Health **15**(1): 58-82.

Evidence review: Gonadotrophin releasing hormone analogues for children and adolescents with gender dysphoria

This document will help inform Dr Hilary Cass' independent review into gender identity services for children and young people. It was commissioned by NHS England and Improvement who commissioned the Cass review. It aims to assess the evidence for the clinical effectiveness, safety and cost-effectiveness of gonadotrophin releasing hormone (GnRH) analogues for children and adolescents aged 18 years or under with gender dysphoria.

The document was prepared by NICE in October 2020.

The content of this evidence review was up to date on 14 October 2020. See summaries of product characteristics (SPCs), British National Formulary (BNF) or the Medicines and Healthcare products Regulatory Agency (MHRA) or NICE websites for up-to-date information.

Contents

1. Introduction

This review aims to assess the evidence for the clinical effectiveness, safety and cost-effectiveness of gonadotrophin releasing hormone (GnRH) analogues for children and adolescents aged 18 years or under with gender dysphoria. The review follows the NHS England Specialised Commissioning process and template and is based on the criteria outlined in the PICO framework (see appendix A). This document will help inform Dr Hilary Cass' independent review into gender identity services for children and young people.

Gender dysphoria in children, also known as gender identity disorder or gender incongruence of childhood (World Health Organisation 2020), refers to discomfort or distress that is caused by a discrepancy between a person's gender identity (how they see themselves[1] regarding their gender) and that person's sex assigned at birth and the associated gender role, and/or primary and secondary sex characteristics (Diagnostic and Statistical Manual of Mental Disorders 2013).

GnRH analogues suppress puberty by delaying the development of secondary sexual characteristics. The intention is to alleviate the distress associated with the development of secondary sex characteristics, thereby providing a time for on-going discussion and exploration of gender identity before deciding whether to take less reversible steps. In England, the GnRH analogue triptorelin (a synthetic decapeptide analogue of natural GnRH, which has marketing authorisations for the treatment of prostate cancer, endometriosis and precocious puberty [onset before 8 years in girls and 10 years in boys]) is used for this purpose. The use of triptorelin for children and adolescents with gender dysphoria is off-label.

For children and adolescents with gender dysphoria it is recommended that management plans are tailored to the needs of the individual, and aim to ameliorate the potentially negative impact of gender dysphoria on general developmental processes, support young people and their families in managing the uncertainties inherent in gender identity development and provide on-going opportunities for exploration of gender identity. The plans may also include psychological support and exploration and, for some individuals, the use of GnRH analogues in adolescence to suppress puberty; this may be followed later with gender-affirming hormones of the desired sex (NHS England 2013).

2. Executive summary of the review

Nine observational studies were included in the evidence review. Five studies were retrospective observational studies (Brik et al. 2020, Joseph et al. 2019, Khatchadourian et al. 2014, Klink et al. 2015, Vlot et al. 2017), 3 studies were prospective longitudinal observational studies (Costa et al. 2015, de Vries et al. 2011, Schagen et al. 2016) and 1 study was a cross-sectional study (Staphorsius et al. 2015). Two studies (Costa et al. 2015 and Staphorsius et al. 2015) provided comparative evidence and the remaining 7 studies used within-person, before and after comparisons.

[1] Gender refers to the roles, behaviours, activities, attributes and opportunities that any society considers appropriate for girls and boys, and women and men (World Health Organisation, Health Topics, Gender)

The terminology used in this topic area is continually evolving and is different depending on stakeholder perspectives. In this evidence review we have used the phrase 'people's assigned sex at birth' rather than natal or biological sex, gonadotrophin releasing hormone (GnRH) analogues rather than 'puberty blockers' and gender-affirming hormones rather than 'cross sex hormones'. The research studies included in this evidence review may use historical terms which are no longer considered appropriate.

In children and adolescents with gender dysphoria, what is the clinical effectiveness of treatment with GnRH analogues compared with one or a combination of psychological support, social transitioning to the desired gender or no intervention?

Critical outcomes

The critical outcomes for decision making are the impact on gender dysphoria, mental health and quality of life. The quality of evidence for these outcomes was assessed as very low certainty using modified GRADE.

Impact on gender dysphoria

The study by de Vries et al. 2011 in 70 adolescents with gender dysphoria found that treatment with GnRH analogues before starting gender-affirming hormones does not affect gender dysphoria (measured using the Utrecht Gender Dysphoria Scale [UGDS]). The mean (±SD) gender dysphoria (UGDS) score was not statistically significantly different at baseline compared with follow-up (n=41, 53.20 [±7.91] versus 53.9 [±17.42], p=0.333).

Impact on mental health

The study by de Vries et al. 2011 in 70 adolescents with gender dysphoria found that treatment with GnRH analogues before starting gender-affirming hormones may reduce depression (measured using the Beck Depression Inventory-II [BDI-II]). The mean [±SD] BDI score was statistically significantly lower (improved) from baseline compared with follow-up (n=41, 8.31 [±7.12] versus 4.95 [±6.72], p=0.004).

The study by de Vries et al. 2011 in 70 adolescents with gender dysphoria found that treatment with GnRH analogues before starting gender-affirming hormones does not affect anger (measured using the Trait Anger Scale [TPI]). The mean [±SD] anger (TPI) score was not statistically significantly different at baseline compared with follow-up (n=41, 18.29 [±5.54] versus 17.88 [±5.24], p=0.503).

The study by de Vries et al. 2011 in 70 adolescents with gender dysphoria found that treatment with GnRH analogues before starting gender-affirming hormones does not affect anxiety (measured using the Trait Anxiety Scale [STAI]). The mean [±SD] anxiety (STAI) score was not statistically significantly different at baseline compared with follow-up (n=41, 39.43 [±10.07] versus 37.95 [±9.38], p=0.276).

Impact on quality of life

No evidence was identified.

Important outcomes

The important outcomes for decision making are impact on body image, psychosocial impact, engagement with health care services, impact on extent of and satisfaction with surgery and stopping treatment. The quality of evidence for all these outcomes was assessed as very low certainty using modified GRADE.

Impact on body image

The study by de Vries et al. 2011 in 70 adolescents with gender dysphoria found that treatment with GnRH analogues before starting gender-affirming hormones does not affect body image (measured using the Body Image Scale [BIS]). The mean [±SD] body image (BIS) scores were not statistically significantly different from baseline compared with follow-up for primary sexual characteristics (n=57, 4.10 [±0.56] versus 3.98 [±0.71], p=0.145), secondary sexual characteristics (n=57, 2.74 [±0.65] versus 2.82 [±0.68], p=0.569) or neutral body characteristics (n=57, 2.41 [±0.63] versus 2.47 [±0.56], p=0.620).

Psychosocial impact

The study by de Vries et al. 2011 in 70 adolescents with gender dysphoria found that treatment with GnRH analogues before starting gender-affirming hormones may improve psychosocial impact over time (measured using the Children's Global Assessment Scale [CGAS]). The mean [±SD] CGAS score was statistically significantly higher (improved) from baseline compared with follow-up (n=41, 70.24 [±10.12] versus 73.90 [±9.63], p=0.005).

This study also found that psychosocial functioning may improve over time (measured using the Child Behaviour Checklist [CBCL] and the self-administered Youth Self-Report [YSR]). The mean [±SD] CBCL scores were statistically significantly lower (improved) from baseline compared with follow-up for Total T score (n=54, 60.70 [±12.76] versus 54.46 [±11.23], p<0.001), internalising T score (n=54, 61.00 [±12.21] versus 52.17 [±9.81], p<0.001) and externalising T score (n=54, 58.04 [±12.99] versus 53.81 [±11.86], p=0.001). The mean [±SD] YSR scores were statistically significantly lower (improved) from baseline compared with follow-up for Total T score (n=54, 55.46 [±11.56] versus 50.00 [±10.56], p<0.001), internalising T score (n=54, 56.04 [±12.49] versus 49.78 [±11.63], p<0.001) and externalising T score (n=54, 53.30 [±11.87] versus 49.98 [±9.35], p=0.009). The proportion of adolescents scoring in the clinical range decreased from baseline to follow up on the CBCL total problem scale (44.4% versus 22.2%, p=0.001) and the internalising scale of the YSR (29.6% versus 11.1%, p=0.017).

The study by Costa et al. 2015 in 201 adolescents with gender dysphoria who had 6 months of psychological support followed by either GnRH analogues and continued psychological support or continued psychological support only, found that during treatment with GnRH analogues psychosocial impact in terms of global functioning may improve over time (measured using the CGAS). In the group receiving GnRH analogues, the mean [±SD] CGAS score was statistically significantly higher (improved) after 6 months (n=60, 64.70 [±13.34]) and 12 months (n=35, 67.40 [±13.39]) compared with baseline (n=101, 58.72 [±11.38], p=0.003 and p<0.001, respectively). However, there was no statistically significant difference in global functioning (CGAS scores) between the group receiving GnRH analogues plus psychological support and the group receiving psychological support only at any time point.

The study by Staphorsius et al. 2015 in 40 adolescents with gender dysphoria (20 of whom were receiving GnRH analogues) gave mean [±SD] CBCL scores for each group, but statistical analysis is unclear (transfemales receiving GnRH analogues 57.4 [±9.8], transfemales not receiving GnRH analogues 58.2 [±9.3], transmales receiving GnRH analogues 57.5 [±9.4], transmales not receiving GnRH analogues 63.9 [±10.5]).

Engagement with health care services

The study by Brik et al. 2018 in 143 children and adolescents with gender dysphoria receiving GnRH analogues found that 9 adolescents in the original sampling frame (9/214, 4.2%) were excluded from the study because they stopped attending appointments.

The study by Costa et al. 2015 in 201 adolescents with gender dysphoria who had 6 months of psychological support followed by either GnRH analogues and continued psychological support or continued psychological support only had a large loss to follow-up over time. The sample size at baseline and 6 months was 201, which dropped by 39.8% to 121 after 12 months and by 64.7% to 71 at 18 months follow-up. No explanation of the reasons for loss to follow-up are reported.

Impact on extent of and satisfaction with surgery

No evidence was identified.

Stopping treatment

The study by Brik et al. 2018 in 143 children and adolescents with gender dysphoria receiving GnRH analogues reported the reasons for stopping GnRH analogues. During the follow-up period 6.2% (9/143) of adolescents had stopped GnRH analogues after a median duration of 0.8 years (range 0.1 to 3.0). Five adolescents stopped treatment because they no longer wished to receive gender-affirming treatment for various reasons. In 4 adolescents (all transmales), GnRH analogues were stopped mainly because of adverse effects (such as mood and emotional lability), although they wanted to continue treatments for gender dysphoria.

The study by Khatchadourian et al. 2014 in 27 adolescents with gender dysphoria who started GnRH analogues reported the reasons for stopping them. Eleven out of 26 where data was available (42%) stopped GnRH analogues during follow up.

In children and adolescents with gender dysphoria, what is the short-term and long-term safety of GnRH analogues compared with one or a combination of psychological support, social transitioning to the desired gender or no intervention?

Evidence was available for bone density, cognitive development or functioning, and other safety outcomes. The quality of evidence for all these outcomes was assessed as very low certainty using modified GRADE.

Bone density

The study by Joseph et al. 2019 in 70 adolescents with gender dysphoria found that GnRH analogues may reduce the expected increase in lumbar or femoral bone density (measured with the z-score). However, the z-scores were largely within 1 standard deviation of normal,

and actual lumbar or femoral bone density values were not statistically significantly different between baseline and follow-up.

- The mean z-score [±SD] for lumbar bone mineral apparent density (BMAD) was statistically significantly lower at 1 year compared with baseline in transfemales (baseline 0.859 [±0.154], 1 year −0.228 [±1.027], p=0.000) and transmales (baseline −0.186 [±1.230], 1 year −0.541 [±1.396], p=0.006).
- The mean z-score [±SD] for lumbar BMAD was statistically significantly lower after receiving GnRH analogues for 2 years compared with baseline in transfemales (baseline 0.486 [±0.809], 2 years −0.279 [±0.930], p=0.000) and transmales (baseline −0.361 [±1.439], 2 years −0.913 [±1.318], p=0.001).
- The mean z-score [±SD] for femoral neck bone mineral density (BMD) was statistically significantly lower after receiving GnRH analogues for 2 years compared with baseline in transfemales (baseline 0.0450 [±0.781], 2 years −0.600 [±1.059], p=0.002) and transmales (baseline −1.075 [±1.145], 2 years −1.779 [±0.816], p=0.001).

The study by Klink et al. 2015 in 34 adolescents with gender dysphoria found that GnRH analogues may reduce the expected increase in lumbar (transmales only), but not femoral bone density. However, the z-scores are largely within 1 standard deviation of normal. Actual lumbar or femoral bone density values were not statistically significantly different between baseline and follow-up (apart from BMD measurements in transmales):

- The mean z-score [±SD] for lumbar BMAD was not statistically significantly different between starting GnRH analogues and starting gender-affirming hormones in transfemales, but was statistically significantly lower when starting gender-affirming hormones in transmales (GnRH analogues 0.28 [±0.90], gender-affirming hormones −0.50 [±0.81], p=0.004).

The study by Vlot et al. 2017 in 70 adolescents with gender dysphoria found that GnRH analogues may reduce the expected increase in lumbar or femoral bone density. However, the z-scores were largely within 1 standard deviation of normal. Actual lumbar or femoral bone density values were not statistically significantly different between baseline and follow-up (apart from in transmales with a bone age ≥14 years). This study reported change in bone density from starting GnRH analogues to starting gender-affirming hormones by bone age:

- The median z-score [range] for lumbar BMAD in transfemales with a bone age of <15 years was statistically significantly lower at starting gender-affirming hormones than at starting GnRH analogues (GnRH analogues −0.20 [−1.82 to 1.18], gender-affirming hormones −1.52 [−2.36 to 0.42], p=0.001) but was not statistically significantly different in transfemales with a bone age ≥15 years.
- The median z-score [range] for lumbar BMAD in transmales with a bone age of <14 years was statistically significantly lower at starting gender-affirming hormones than at starting GnRH analogues (GnRH analogues −0.05 [−0.78 to 2.94], gender-affirming hormones −0.84 [−2.20 to 0.87], p=0.003) and in transmales with a bone age ≥14 years (GnRH analogues 0.27 [−1.60 to 1.80], gender-affirming hormones −0.29 [−2.28 to 0.90], p≤0.0001).

- The median z-score [range] for femoral neck BMAD in transfemales with a bone age of <15 years was not statistically significantly lower at starting gender-affirming hormones than at starting GnRH analogues (GnRH analogues −0.71 [−3.35 to 0.37], gender-affirming hormones −1.32 [−3.39 to 0.21], p≤0.1) or in transfemales with a bone age ≥15 years (GnRH analogues −0.44 [−1.37 to 0.93], gender-affirming hormones −0.36 [−1.50 to 0.46]).
- The z-score for femoral neck BMAD in transmales with a bone age of <14 years was not statistically significantly lower at starting gender-affirming hormones than at starting GnRH analogues (GnRH analogues −0.01 [−1.30 to 0.91], gender-affirming hormone −0.37 [−2.28 to 0.47]) but was statistically significantly lower in transmales with a bone age ≥14 years (GnRH analogues 0.27 [−1.39 to 1.32], gender-affirming hormones −0.27 [−1.91 to 1.29], p=0.002).

Cognitive development or functioning

The study by Staphorsius et al. 2015 in 40 adolescents with gender dysphoria (20 of whom were receiving GnRH analogues) measured cognitive development or functioning (using an IQ test, and reaction time and accuracy measured using the Tower of London task):

- The mean (±SD) IQ in transfemales receiving GnRH analogues was 94.0 (±10.3) and 109.4 (±21.2) in the control group. In transmales receiving GnRH analogues the mean (±SD) IQ was 95.8 (±15.6) and 98.5 (±15.9) in the control group.
- The mean (±SD) reaction time in transfemales receiving GnRH analogues was 10.9 (±4.1) and 9.9 (±3.1) in the control group. In transmales receiving GnRH analogue it was 9.9 (±3.1) and 10.0 (±2.0) in the control group.
- The mean (±SD) accuracy score in transfemales receiving GnRH analogues was 73.9 (±9.1) and 83.4 (±9.5) in the control group. In transmales receiving GnRH analogues it was 85.7 (±10.5) and 88.8 (±9.7) in the control group.

No statistical analyses or interpretation of the results was reported.

Other safety outcomes

The study by Schagen et al. 2016 in 116 adolescents with gender dysphoria found that GnRH analogues do not affect renal or liver function:

- There was no statistically significant difference between baseline and 1 year results for serum creatinine in transfemales, but there was a statistically significant decrease between baseline and 1 year in transmales (p=0.01).
- Glutamyl transferase, alanine aminotransferase (ALT), and aspartate aminotransferase (AST) levels did not significantly change from baseline to 12 months of treatment.

The study by Khatchadourian et al. 2014 in 27 adolescents with gender dysphoria who started GnRH analogues narratively reported adverse effects from GnRH analogues in 26 adolescents:

- 1 transmale developed sterile abscesses; they were switched from leuprolide acetate to triptorelin, and this was well tolerated
- 1 transmale developed leg pains and headaches, which eventually resolved
- 1 participant gained 19 kg within 9 months of starting GnRH analogues.

In children and adolescents with gender dysphoria, what is the cost-effectiveness of GnRH analogues compared to one or a combination of psychological support, social transitioning to the desired gender or no intervention?

No cost-effectiveness evidence was found for GnRH analogues in children and adolescents with gender dysphoria.

From the evidence selected, are there any subgroups of children and adolescents with gender dysphoria that may benefit from GnRH analogues more than the wider population of interest?

Some studies reported data separately for the following subgroups of children and adolescents with gender dysphoria: sex assigned at birth males (transfemales) and sex assigned at birth females (transmales). This included some direct comparisons of these subgroups, and differences were largely seen at baseline as well as follow up. No evidence was found for other specified subgroups.

Sex assigned at birth males (transfemales)

Impact on gender dysphoria

The study by Costa et al. 2015 in 201 adolescents with gender dysphoria who had 6 months of psychological support followed by either GnRH analogues and continued psychological support or continued psychological support only, found that gender dysphoria (measured using the UGDS) in sex assigned at birth males is lower than in sex assigned at birth females. Sex assigned at birth males had a statistically significantly lower (improved) mean [±SD] UGDS score of 51.6 [±9.7] compared with sex assigned at birth females (56.1 [±4.3], p<0.001), but it was not reported if this was at baseline or follow-up.

The study by de Vries et al. 2011 in 70 adolescents with gender dysphoria found that gender dysphoria (measured using the UGDS) in sex assigned at birth males is lower than in sex assigned at birth females at baseline and follow up. The mean [±SD] UGDS score was statistically significantly lower (improved) in sex assigned at birth males compared with sex assigned at birth females at baseline (n=not reported, mean UGDS score: 47.95 [±9.70] versus 56.57 [±3.89]) and follow up (n=not reported, 49.67 [±9.47] versus 56.62 [±4.00]); between sex difference p<0.001).

Impact on mental health

The study by de Vries et al. 2011 in 70 adolescents with gender dysphoria found that the impact on mental health (depression, anger and anxiety) may be different in sex assigned at birth males compared with sex assigned at birth females. Over time there was no statistically significant difference between sex assigned at birth males and sex assigned at birth females for depression, but sex assigned at birth males had statistically significantly lower levels of anger and anxiety than sex assigned at birth females at baseline and follow up.

- The mean [±SD] depression (BDI-II) score was not statistically significantly different in sex assigned at birth males compared with sex assigned at birth females at baseline (n=not reported, mean BDI score [±SD]: 5.71 [±4.31] versus 10.34 [±8.24]) and follow-up (n=not reported, 3.50 [±4.58] versus 6.09 [±7.93]), between sex difference p=0.057

- The mean [±SD] anger (TPI) score was statistically significantly lower (improved) in sex assigned at birth males compared with sex assigned at birth females at baseline (n=not reported, mean TPI score [±SD]: 5.22 [±2.76] versus 6.43 [±2.78]) and follow-up (n=not reported, 5.00 [±3.07] versus 6.39 [±2.59]), between sex difference p=0.022
- The mean [±SD] anxiety (STAI) score was statistically significantly lower (improved) in sex assigned at birth males compared with sex assigned at birth females at baseline (n=not reported, mean STAI score [±SD]: 4.33 [±2.68] versus 7.00 [±2.36]) and follow-up (n=not reported, 4.39 [±2.64] versus 6.17 [±2.69]), between sex difference p<0.001.

Impact on body image

The study by de Vries et al. 2011 in 70 adolescents with gender dysphoria found that the impact on body image may be different in sex assigned at birth males compared with sex assigned at birth females. Sex assigned at birth males are less dissatisfied with their primary and secondary sex characteristics than sex assigned at birth females at both baseline and follow up, but the satisfaction with neutral body characteristics is not different.

- The mean [±SD] BIS score for primary sex characteristics was statistically significantly lower (improved) in sex assigned at birth males compared with sex assigned at birth females at baseline (n=not reported, mean BIS score [±SD]: 4.02 [±0.61] versus 4.16 [±0.52]) and follow up (n=not reported, 3.74 [±0.78] versus 4.17 [±0.58]) between sex difference p=0.047.
- The mean [±SD] BIS score for secondary sex was statistically significantly lower (improved) in sex assigned at birth males compared with sex assigned at birth females at baseline (n=not reported, mean BIS score [±SD]: 2.66 [±0.50] versus 2.81 [±0.76]) and follow up (n=not reported, 2.39 [±0.69] versus 3.18 [±0.42]), between sex difference p=0.001.
- The mean [±SD] BIS score for neutral body characteristics was not statistically significantly different in sex assigned at birth males compared with sex assigned at birth females at baseline (n=not reported, 2.60 [±0.58] versus 2.24 [±0.62], between sex difference p=0.777).

Psychosocial impact

The study by Costa et al. 2015 in 201 adolescents with gender dysphoria who had 6 months of psychological support followed by either GnRH analogues and continued psychological support or continued psychological support only, found that sex assigned at birth males had statistically significant lower mean [±SD] CGAS scores at baseline compared with sex assigned at birth females (n=201, 55.4 [±12.7] versus 59.2 [±11.8], p=0.03), but no conclusions could be drawn.

The study by de Vries et al. 2011 in 70 adolescents with gender dysphoria found that psychosocial impact in terms of global functioning (CGAS) and psychosocial functioning (CBCL and YSR) may be different in sex assigned at birth males compared with sex assigned at birth females, but no conclusions could be drawn.

- There was no statistically significant difference between sex assigned at birth males and sex assigned at birth females (at baseline or follow up) for the CBCL Total T

score, the CBCL internalising T score, the YSR Total T score or the YSR internalising T score.

- Sex assigned at birth males had statistically higher mean [±SD] CGAS scores compared with sex assigned at birth females at baseline (n=54, 73.10 [±8.44] versus 67.25 [±11.06]) and follow up (n=54, 77.33 [±8.69] versus 70.30 [±9.44]), between sex difference p=0.021.
- Sex assigned at birth males had statistically lower mean [±SD] CBCL externalising T scores compared with sex assigned at birth females at baseline (n=54, 54.71 [±12.91] versus 60.70 [±12.64]) and follow up (n=54, 48.75 [±10.22] versus 57.87 [±11.66]), between sex difference p=0.015.
- Sex assigned at birth males had statistically lower mean [±SD] YSR externalising T scores compared with sex assigned at birth females at both baseline (n=54, 48.72 [±11.38] versus 57.24 [±10.59]) and follow up (n=54, 46.52 [±9.23] versus 52.97 [±8.51]), between sex difference p=0.004.

Bone density
The studies by Joseph et al. 2019, Klink et al. 2015 and Vlot et al. 2017 provided evidence on bone density in sex assigned at birth males (see above for details).

Cognitive development or functioning
The study by Staphorsius et al. 2015 provided evidence on cognitive development or functioning in sex assigned at birth males (see above for details).

Other safety outcomes
The study by Schagen et al. 2016 provided evidence on renal function in sex assigned at birth males (see above).

Sex assigned at birth females (transmales)
Impact on gender dysphoria
The studies by de Vries et al. 2011 and Costa et al. 2015 found that gender dysphoria (measured using the UGDS) in sex assigned at birth females is higher than in sex assigned at birth males at baseline and follow up (see above for details).

Impact on mental health
The study by de Vries et al. 2011 found that the impact on mental health (depression, anger and anxiety) may be different in sex assigned at birth females compared with sex assigned at birth males. Over time there was no statistically significant difference between sex assigned at birth females and sex assigned at birth males for depression, but sex assigned at birth females had statistically significantly greater levels of anger and anxiety than sex assigned at birth males at both baseline and follow up (see above for details).

Impact on body image
The study by de Vries et al. 2011 found that the impact on body image may be different in sex assigned at birth females compared with sex assigned at birth males. Sex assigned at birth females are more dissatisfied with their primary and secondary sex characteristics than sex assigned at birth males at both baseline and follow up, but the satisfaction with neutral body characteristics is not different (see above for details).

Psychosocial impact
The studies by de Vries et al. 2011 and Costa et al. 2015 found that psychosocial impact in terms of global functioning (CGAS) and psychosocial functioning (CBCL and YSR) may be different in sex assigned at birth females compared with sex assigned at birth males, but no conclusions could be drawn (see above for details).

Bone density
The studies by Joseph et al. 2019, Klink et al. 2015 and Vlot et al. 2017 provided evidence on bone density in sex assigned at birth females (see above for details).

Cognitive development or functioning
The study by Staphorsius et al. 2015 provided evidence on cognitive development or functioning in sex assigned at birth females (see above for details).

Other safety outcomes
The study by Schagen et al. 2016 provided evidence on renal function in sex assigned at birth females (see above for details).

From the evidence selected:
(a) what are the criteria used by the research studies to define gender dysphoria, gender identity disorder and gender incongruence of childhood?
(b) what were the ages at which participants commenced treatment with GnRH analogues?
(c) what was the duration of treatment with GnRH analogues?

All studies that reported diagnostic criteria for gender dysphoria (6/9 studies) used the version of the Diagnostic and Statistical Manual of Mental Disorders (DSM) criteria that was in use at the time. In 5 studies (Costa et al. 2015, Klink et al. 2015, Schagen et al. 2016, Staphorsius et al. 2015 and Vlot et al. 2017) the DSM-fourth edition, text revision (IV-TR) criteria were used. The study by Brik et al. 2020 used DSM-V criteria. It was not reported how gender dysphoria was defined in the remaining 3 studies.

The studies show variation in the age (11 to 18 years old) at which children and adolescents with gender dysphoria started GnRH analogues.

Most studies did not report the duration of treatment with GnRH analogues (Joseph et al. 2019, Khatchadourian et al. 2014, Vlot et al. 2017, Costa et al. 2015, de Vries et al. 2011, Schagen et al. 2016), but where this was reported (Brik et al. 2020, Klink et al. 2015, Staphorsius et al. 2015) there was a wide variation ranging from a few months to about 5 years.

Discussion

A key limitation to identifying the effectiveness and safety of GnRH analogues for children and adolescents with gender dysphoria is the lack of reliable comparative studies. The lack of clear, expected outcomes from treatment with a GnRH analogue (the purpose of which is to suppress secondary sexual characteristics which may cause distress from unwanted pubertal changes) also makes interpreting the evidence difficult.

The studies included in this evidence review are all small, uncontrolled observational studies, which are subject to bias and confounding, and all the results are of very low certainty using modified GRADE. They all reported physical and mental health comorbidities and concomitant treatments very poorly. All the studies are from a limited number of, mainly European, care facilities. They are described as either tertiary referral or expert services but the low number of services providing such care and publishing evidence may bias the results towards the outcomes in these services only and limit extrapolation.

Many of the studies did not report statistical significance or confidence intervals. Changes in outcome scores for clinical effectiveness and bone density were assessed with regards to statistical significance. However, there is relatively little interpretation of whether the changes in outcomes are clinically meaningful.

In the observational, retrospective studies providing evidence on bone density, participants acted as their own controls and change in bone density was determined between starting GnRH analogues and follow up. Observational studies such as these can only show an association with GnRH analogues and bone density; they cannot show that GnRH analogues caused any differences in bone density seen. Because there was no comparator group and participants acted as their own controls, it is not known whether the findings are associated with GnRH analogues or due to changes over time.

Conclusion

The results of the studies that reported impact on the critical outcomes of gender dysphoria and mental health (depression, anger and anxiety), and the important outcomes of body image and psychosocial impact (global and psychosocial functioning), in children and adolescents with gender dysphoria are of very low certainty using modified GRADE. They suggest little change with GnRH analogues from baseline to follow-up.

Studies that found differences in outcomes could represent changes that are either of questionable clinical value, or the studies themselves are not reliable and changes could be due to confounding, bias or chance. It is plausible, however, that a lack of difference in scores from baseline to follow-up is the effect of GnRH analogues in children and adolescents with gender dysphoria, in whom the development of secondary sexual characteristics might be expected to be associated with an increased impact on gender dysphoria, depression, anxiety, anger and distress over time without treatment. The study by de Vries et al. 2011 reported statistically significant reductions in the Child Behaviour Checklist (CBCL) and Youth Self-Report (YSR) scores from baseline to follow up, which include measures of distress. As the aim of GnRH analogues is to reduce distress caused by the development of secondary sexual characteristics, this may be an important finding. However, as the studies all lack appropriate controls who were not receiving GnRH analogues, any positive changes could be a regression to mean.

The results of the studies that reported bone density outcomes suggest that GnRH analogues may reduce the expected increase in bone density (which is expected during puberty). However, as the studies themselves are not reliable, the results could be due to confounding, bias or chance. While controlled trials may not be possible, comparative studies are needed to understand this association and whether the effects of GnRH analogues on bone density are seen after they are stopped. All the studies that reported safety outcomes provided very low certainty evidence.

No cost-effectiveness evidence was found to determine whether or not GnRH analogues are cost-effective for children and adolescents with gender dysphoria.

The results of the studies that reported outcomes for subgroups of children and adolescents with gender dysphoria, suggest there may be differences between sex assigned at birth males (transfemales) and sex assigned at birth females (transmales).

3. Methodology

Review questions

The review question(s) for this evidence review are:

1. For children and adolescents with gender dysphoria, what is the clinical effectiveness of treatment with GnRH analogues compared with one or a combination of psychological support, social transitioning to the desired gender or no intervention?
2. For children and adolescents with gender dysphoria, what is the short-term and long-term safety of GnRH analogues compared with one or a combination of psychological support, social transitioning to the desired gender or no intervention?
3. For children and adolescents with gender dysphoria, what is the cost-effectiveness of GnRH analogues compared to one or a combination of psychological support, social transitioning to the desired gender or no intervention?
4. From the evidence selected, are there any subgroups of children and adolescents with gender dysphoria that may derive more (or less) advantage from treatment with GnRH analogues than the wider population of children and adolescents with gender dysphoria?
5. From the evidence selected,
   a) what are the criteria used by the research studies to define gender dysphoria, gender identity disorder and gender incongruence of childhood?
   b) what were the ages at which participants commenced treatment with GnRH analogues?
   c) what was the duration of treatment with GnRH analogues?

See appendix A for the full review protocol.

Review process

The methodology to undertake this review is specified by NHS England in their 'Guidance on conducting evidence reviews for Specialised Services Commissioning Products' (2020).

The searches for evidence were informed by the PICO document and were conducted on 23 July 2020

See appendix B for details of the search strategy.

Results from the literature searches were screened using their titles and abstracts for relevance against the criteria in the PICO framework. Full text references of potentially

relevant evidence were obtained and reviewed to determine whether they met the inclusion criteria for this evidence review.

See appendix C for evidence selection details and appendix D for the list of studies excluded from the review and the reasons for their exclusion.

Relevant details and outcomes were extracted from the included studies and were critically appraised using a checklist appropriate to the study design. See appendices E and F for individual study and checklist details.

The available evidence was assessed by outcome for certainty using modified GRADE. See appendix G for GRADE Profiles.

4. Summary of included studies

Nine observational studies were identified for inclusion. Five studies were retrospective observational studies (Brik et al. 2020, Joseph et al. 2019, Khatchadourian et al. 2014, Klink et al. 2015, Vlot et al. 2017), 3 studies were prospective longitudinal observational studies (Costa et al. 2015, de Vries et al. 2011, Schagen et al. 2016) and 1 study was a cross-sectional study (Staphorsius et al. 2015).

The terminology used in this topic area is continually evolving and is different depending on stakeholder perspectives. In this evidence review we have used the phrase 'people's assigned sex at birth' rather than natal or biological sex, gonadotrophin releasing hormone (GnRH) analogues rather than 'puberty blockers' and gender-affirming hormones rather than 'cross sex hormones'. The research studies included in this evidence review may use historical terms which are no longer considered appropriate.

Table 1 provides a summary of these included studies and full details are given in appendix E.

Table 1 Summary of included studies

Study	Population	Intervention and comparison	Outcomes reported
Brik et al. 2020 Retrospective observational single-centre study Netherlands	The study was conducted at the Curium-Leiden University Medical Centre gender clinic in Leiden, the Netherlands and involved adolescents with gender dysphoria. The sample size was 143 adolescents (median age at start of treatment was 15.0 years, range 11.1 to 18.6 years in transfemales; 16.1 years, range 10.1 to 17.9 years in transmales) from a sampling frame of 269 children and adolescents registered at the clinic between November 2010 and January 2018.	**Intervention** 143 children and adolescents receiving GnRH analogues (no specific treatment, dose, route or frequency of administration reported). The median duration was 2.1 years (range 1.6–2.8 years). **Comparison** No comparator.	**Critical Outcomes** • No critical outcomes reported **Important outcomes** • Stopping treatment
Costa et al 2015 Prospective longitudinal observational single centre cohort study United Kingdom	The study was conducted at the Gender Identity Development Service in London and involved adolescents with gender dysphoria. The sample size was 201 adolescents (mean [±SD] age 15.52±1.41 years, range 12 to 17 years) from a sampling frame of 436 consecutive adolescents referred to the service between 2010 and 2014. The mean [±SD] age at the start of GnRH analogues was 16.48 [±1.26] years, range 13 to 17 years. Participants were invited to participate following a 6-month diagnostic process using DSM-IV-TR criteria. No concomitant treatments were reported.	**Intervention** 101 adolescents assessed as being immediately eligible for GnRH analogues (no specific treatment, dose or route of administration reported) plus psychological support. The average duration of treatment was approximately 12 months (no exact figure given). **Comparison** 100 adolescents assessed as not immediately eligible for GnRH analogues (more time needed to make the decision to start GnRH analogues) who had psychological support only. None received GnRH analogues throughout the study.	**Critical Outcomes** • No critical outcomes reported **Important outcomes** • Psychosocial impact
de Vries et al. 2011 Prospective longitudinal observational single centre before and after study Netherlands	The study was conducted at the Amsterdam gender identity clinic of the VU University Medical Centre and involved adolescents who were defined as "transsexual". The sample size was 70 adolescents receiving GnRH analogues (mean age [±SD] at assessment 13.6±1.8 years) from a sampling frame of 196 consecutive adolescents referred to the service between 2000 and 2008. Participants were invited to participate if they subsequently started gender-affirming hormones between 2003 and 2009. No diagnostic criteria or concomitant treatments were reported.	**Intervention** 70 individuals assessed at baseline (T0) before the start of GnRH analogues (no specific treatment, dose or route of administration reported). **Comparison** No comparator.	**Critical Outcomes** • Gender dysphoria • Mental health (depression, anger and anxiety) **Important outcomes** • Body image • Psychosocial impact

Study	Population	Intervention and comparison	Outcomes reported
Joseon et al 2019 Retrospective longitudinal observational single centre study United Kingdom	This study was conducted at the Early intervention clinic at University College London Hospital (all participants had been seen at the Gender Identity Development Service in London) and involved adolescents with gender dysphoria. The sample size was 70 adolescents with gender dysphoria (no diagnostic criteria described) all offered GnRH analogues. The mean age at the start of treatment was 13.2 years (SD ±1.4) for transfemales and 12.6 years (SD ±1.0) for transmales. Details of the sampling frame were not reported. Further details of how the sample was drawn are not reported. No concomitant treatments were reported.	**Intervention** GnRH analogues. No specific treatment, duration, dose or route of administration reported. **Comparison** No comparator.	**Critical Outcomes** • No critical outcomes reported **Important outcomes** • Safety: bone density
Khatchadourian et al 2014 Retrospective observational chart review single centre study Canada	This study was conducted at the Endocrinology and Diabetes Unit at British Columbia Children's Hospital, Canada and involved youths with gender dysphoria. The sample size was 27 young people with gender dysphoria who started GnRH analogues (at mean age 14.7 [SD ±1.9] years) out of 84 young people seen at the unit between 1998 and 2011. Diagnostic criteria and concomitant treatments were not reported.	**Intervention** 84 young people with gender dysphoria. For GnRH analogues no specific treatment, duration, dose or route of administration reported. **Comparison** No comparator	**Critical Outcomes** • No critical outcomes reported **Important outcomes** • Stopping treatment • Safety: adverse effects
Klink et al 2015 Retrospective longitudinal observational single centre study Netherlands	This study was conducted in the Netherlands at a tertiary referral centre. It is unclear which centre this was The sample size was 34 adolescents (mean age 14.9 [SD ±1.9] years for transfemales and 15.0 [SD ±2.0] years for transmales at start of GnRH analogues). Details of the sampling frame are not reported. Participants were included if they met DSM-IV-TR criteria for gender identity disorder of adolescence and had been treated with GnRH analogues and gender-affirming hormones during their pubertal years. No concomitant treatments were reported.	**Intervention** The intervention was GnRH analogue monotherapy (triptorelin 3.75 mg subcutaneously every 4 weeks) followed by gender-affirming hormones with discontinuation of GnRH analogues after gonadectomy. Duration of GnRH analogues was 1.3 years (range 0.5 to 3.8 years) in transfemales and 1.5 years (0.25 to 5.2 years in transmales **Comparison** No comparator.	**Critical Outcomes** • No critical outcomes reported **Important outcomes** • Safety: bone density

Study	Population	Intervention and comparison	Outcomes reported
Schagen et al 2016 Prospective longitudinal study Netherlands	This study was conducted at the Centre of Expertise on Gender Dysphoria at the VU University Medical Centre (Amsterdam, Netherlands) and involved adolescents with gender dysphoria. The sample size was 116 adolescents (median age [range] 13.6 years [11.6 to 17.9] in transfemales and 14.2 years [11.1 to 18.6] in transmales during first year of GnRH analogues) out of 128 adolescents who started GnRH analogues. Participants were included if they met DSM-IV-TR criteria for gender dysphoria, had lifelong extreme gender dysphoria, were psychologically stable and were living in a supportive environment. No concomitant treatments were reported.	**Intervention** The intervention was GnRH analogue monotherapy (triptorelin 3.75 mg at 0, 2 and 4 weeks followed by intramuscular injections every 4 weeks, for at least 3 months). **Comparison** No comparator	**Critical Outcomes** • No critical outcomes reported **Important outcomes** • Safety: liver and renal function.
Staphorsius et al 2015 Cross-sectional (single time point) assessment single centre study Netherlands	This study was conducted at the VU University Medical Centre (Amsterdam, Netherlands) and involved adolescents with gender dysphoria. The sample size was 85, of whom 40 were adolescents with gender dysphoria (20 of whom were being treated with GnRH analogues) and 45 were controls without gender dysphoria (not further reported here). Mean (±SD) age 15.1 (±2.4) years in transfemales and 15.8 (±1.9) years in transmales. Details of the sampling frame are not reported. Participants were included if they were diagnosed with Gender Identity Disorder according to the DSM-IV-TR and at least 12 years old and Tanner stage of at least B2 or G2 to G3 with measurable oestradiol and testosterone levels in girls and boys, respectively. No concomitant treatments were reported.	**Intervention** The intervention was a GnRH analogue (triptorelin 3.75 mg every 4 weeks subcutaneously or intramuscularly). The mean duration of treatment was 1.6 years (SD ±1.0) **Comparison** Adolescents with gender dysphoria not treated with GnRH analogues.	**Critical Outcomes** • No critical outcomes reported **Important outcomes** • Psychosocial impact • Safety: cognitive functioning
Vlot et al 2017 Retrospective observational data analysis study	This study was conducted at the VU University Medical Centre (Amsterdam, Netherlands) and involved adolescents with gender dysphoria. The sample size was 70 adolescents (median age [range] 15.1 years [11.7 to 18.6] for	**Intervention** The intervention was a GnRH analogue (triptorelin 3.75 mg every 4 weeks subcutaneously). **Comparison** No comparator.	**Critical Outcomes** • No critical outcomes reported **Important outcomes**

Study	Population	Intervention and comparison	Outcomes reported
Netherlands	transmales and 13.5 years [11.5 to 18.3] for transfemales at start of GnRH analogues) Details of the sampling frame are not reported. Participants were included if they had a diagnosis of gender dysphoria according to DSM-IV-TR criteria who were receiving GnRH analogues and then gender-affirming hormones. No concomitant treatments were reported.		• Safety, bone density

Abbreviations: DSM-IV-TR, Diagnostic and Statistical Manual of Mental Disorders, fourth edition, text revision; GnRH, Gonadotrophin releasing hormone; SD, Standard deviation.

5 Results

In children and adolescents with gender dysphoria, what is the clinical effectiveness of treatment with GnRH analogues compared with one or a combination of psychological support, social transitioning to the desired gender or no intervention?

Outcome	Evidence statement
Clinical Effectiveness	
Critical outcomes	
Impact on gender dysphoria **Certainty of evidence: very low**	This is a critical outcome because gender dysphoria in children and adolescents is associated with significant distress and problems with functioning. One uncontrolled, prospective observational longitudinal study (de Vries et al. 2011) provided evidence relating to the impact on gender dysphoria in adolescents, measured using the Utrecht Gender Dysphoria Scale (UGDS). The UGDS is a validated screening tool for both adolescents and adults to assess gender dysphoria. It consists of 12 items, to be answered on a 1- to 5-point scale, resulting in a sum score between 12 and 60. The higher the UGDS score the greater the gender dysphoria. The study measured the impact on gender dysphoria at 2 time points: • before starting a GnRH analogue (mean [±SD] age: 14.75 [±1.92] years), and • shortly before starting gender-affirming hormones (mean [±SD] age: 16.64 [±1.90] years). The mean (±SD) UGDS score was not statistically significantly different at baseline compared with follow-up (n=41, 53.20 [±7.91] versus 53.9 [±17.42], p=0.333) **(VERY LOW)**.

	This study provides very low certainty evidence that treatment with GnRH analogues, before starting gender-affirming hormones, does not affect gender dysphoria.
Impact on mental health: depression **Certainty of evidence: very low**	This is a critical outcome because self-harm and thoughts of suicide have the potential to result in significant physical harm and, for completed suicides, the death of the young person. One uncontrolled, prospective observational longitudinal study (de Vries et al. 2011) provided evidence relating to the impact on depression in children and adolescents with gender dysphoria. Depression was measured using the Beck Depression Inventory-II (BDI-II). The BDI-II is a valid, reliable, and widely used tool for assessing depressive symptoms. There are no specific scores to categorise depression severity, but it is suggested that 0 to 13 is minimal symptoms, 14 to 19 is mild depression, 20 to 28 is moderate depression, and severe depression is 29 to 63. The study provided evidence for depression measured at 2 time points: • before starting a GnRH analogue (mean [±SD] age: 14.75 [±1.92] years), and • shortly before starting gender-affirming hormones (mean [±SD] age: 16.64 [±1.90] years). The mean (±SD) depression (BDI) score was statistically significantly lower (improved) from baseline compared with follow-up (n=41, 8.31 [±7.12] versus 4.95 [±6.72], p=0.004) **(VERY LOW)**. **This study provides very low certainty evidence that treatment with GnRH analogues, before starting gender-affirming hormones, may reduce depression.**
Impact on mental health: anger **Certainty of evidence: very low**	This is a critical outcome because self-harm and thoughts of suicide have the potential to result in significant physical harm and, for completed suicides, the death of the young person. One uncontrolled, prospective observational longitudinal study (de Vries et al. 2011) provided evidence relating to the impact on anger in children and adolescents with gender dysphoria. Anger was measured using the Trait Anger Scale of the State-Trait Personality Inventory (TPI). This is a validated 20-item inventory tool which measures the intensity of anger as the disposition to experience angry feelings as a personality trait. Higher scores indicate greater anger. The study provided evidence for anger measured at 2 time points: • before starting a GnRH analogue (mean [±SD] age: 14.75 [±1.92] years), and • shortly before starting gender-affirming hormones (mean [±SD] age: 16.64 [±1.90] years). The mean (±SD) anger (TPI) score was not statistically significantly different at baseline compared with follow-up (n=41, 18.29 [±5.54] versus 17.88 [±5.24], p=0.503) **(VERY LOW)**. **This study provides very low certainty evidence that treatment with GnRH analogues, before starting gender-affirming hormones, does not affect anger.**

Impact on mental health: anxiety **Certainty of evidence: very low**	This is a critical outcome because self-harm and thoughts of suicide have the potential to result in significant physical harm and, for completed suicides, the death of the young person. One uncontrolled, prospective observational longitudinal study (de Vries et al 2011) provided evidence relating to the impact on anxiety in children and adolescents with gender dysphoria. Anxiety was measured using the Trait Anxiety Scale of the State-Trait Personality Inventory (STAI). This is a validated and commonly used measure of trait and state anxiety. It has 20 items and can be used in clinical settings to diagnose anxiety and to distinguish it from depressive illness. Higher scores indicate greater anxiety. The study provided evidence for anxiety at 2 time points: • before starting a GnRH analogue (mean [±SD] age: 14.75 [±1.92] years), and • shortly before starting gender-affirming hormones (mean [±SD] age: 16.64 [±1.90] years). The mean (±SD) anxiety (STAI) score was not statistically significantly different at baseline compared with follow-up (n=41, 39.43 [±10.07] versus 37.95 [±9.38], p=0.276) **(VERY LOW)**. **This study provides very low certainty evidence that treatment with GnRH analogues, before starting gender-affirming hormones, does not affect levels of anxiety.**
Quality of life	This is a critical outcome because gender dysphoria in children and adolescents may be associated with a significant reduction in health-related quality of life. No evidence was identified.
Important outcomes	
Impact on body image **Certainty of evidence: very low**	This is an important outcome because some children and adolescents with gender dysphoria may want to take steps to suppress features of their physical appearance associated with their sex assigned at birth or accentuate physical features of their desired gender. One uncontrolled, prospective observational longitudinal study provided evidence relating to the impact on body image (de Vries et al. 2011). Body image was measured using the Body Image Scale (BIS) which is a validated 30-item scale covering 3 aspects: primary, secondary and neutral body characteristics. Higher scores represent a higher degree of body dissatisfaction. The study (de Vries et al. 2011) provided evidence for body image measured at 2 time points: • before starting a GnRH analogue (mean [±SD] age: 14.75 [±1.92] years), and • shortly before starting gender-affirming hormones (mean [±SD] age: 16.64 [±1.90] years). The mean (±SD) body image (BIS) scores for were not statistically significantly different from baseline compared with follow-up for:

	• primary sexual characteristics (n=57, 4.10 [±0.56] versus 3.98 [±0.71], p=0.145) • secondary sexual characteristics (n=57, 2.74 [±0.65] versus 2.82 [±0.68], p=0.569) • neutral body characteristics (n=57, 2.41 [±0.63] versus 2.47 [±0.56], p=0.620) **(VERY LOW)**. **This study provides very low certainty evidence that treatment with GnRH analogues, before starting gender affirming hormones, does not affect body image.**
Psychosocial impact: global functioning **Certainty of evidence: very low**	This is an important outcome because gender dysphoria in children and adolescents is associated with internalising and externalising behaviours, and emotional and behavioural problems which may impact on social and occupational functioning. One uncontrolled, observational, prospective cohort study (de Vries et al 2011) and one prospective cross-sectional cohort study (Costa et al 2015) provided evidence relating to psychosocial impact in terms of global functioning. Global functioning was measured using the Children's Global Assessment Scale (CGAS). The CGAS tool is a validated measure of global functioning on a single rating scale from 1 to 100. Lower scores indicate poorer functioning. One study (de Vries et al. 2011) provided evidence for global functioning (CGAS) at 2 time points: • before starting a GnRH analogue (mean [±SD] age: 14.75 [±1.92] years), and • shortly before starting gender-affirming hormones (mean [±SD] age: 16.64 [±1.90] years). The mean (±SD) CGAS score was statistically significantly higher (improved) from baseline compared with follow-up (n=41, 70.24 [±10.12] versus 73.90 [±9.63], p=0.005) **(VERY LOW)**. One study (Costa et al 2015) in adolescents with gender dysphoria who had 6 months of psychological support followed by either GnRH analogues and continued psychological support (the immediately eligible group) or continued psychological support only (the delayed eligible group who did not receive GnRH analogues) provided evidence for global functioning (CGAS) measured at 4 time points: • at baseline (T0) in both groups, • after 6 months of psychological support in both groups (T1), • after 6 months of GnRH analogues and 12 months of psychological support in the immediately eligible group and 12 months of psychological support only in the delayed eligible group (T2), and • after 18 months of psychological support and 12 months of GnRH analogues in the immediately eligible group and after 18 months of psychological support only in the delayed eligible group (T3). The mean [±SD] CGAS score was statistically significantly higher (improved) for all adolescents (including those not receiving GnRH analogues) at T1, T2 or T3 compared with baseline (T0).

	For the immediately eligible group (who received GnRH analogues) versus the delayed eligible group (who did not receive GnRH analogues) there were no statistically significant differences in CGAS scores between the 2 groups at baseline T0 (n=201, p=0.23), T1 (n=201, p=0.73), T2 (n=121, p=0.49) or T3 (n=71, p=0.14) time points. For the immediately eligible group (who received GnRH analogues), the mean (±SD) CGAS score was not statistically significantly different at: • T1 compared with T0 • T2 compared with T1 • T3 compared with T2. The mean (±SD) CGAS score was statistically significantly higher (improved) at: • T2 compared with T0 (n=60, 64.70 [±13.34] versus n=101, 58.72 [±11.38], p=0.003) • T3 compared with T0 (n=35, 67.40 [±13.39] versus n=101, 58.72 [±11.38], p<0.001) • T3 compared with T1 (n=35, 67.40 [±13.93] versus n=101, 60.89 [±12.17], p<0.001) **(VERY LOW)**. **These studies provide very low certainty evidence that during treatment with GnRH analogues, global functioning may improve over time. However, there was no statistically significant difference in global functioning between GnRH analogues plus psychological support compared with psychological support only at any time point.**
Psychosocial impact: psychosocial functioning **Certainty of evidence: very low**	This is an important outcome because gender dysphoria in children and adolescents is associated with internalising and externalising behaviours, and emotional and behavioural problems which may impact on social and occupational functioning. Two studies provided evidence for this outcome. One uncontrolled, observational, prospective cohort study (de Vries et al. 2011) and 1 cross-sectional observational study (Staphorsius et al. 2015) assessed psychosocial functioning using the Child Behaviour Checklist (CBCL) and the self-administered Youth Self-Report (YSR). The CBCL is a checklist parents complete to detect emotional and behavioural problems in children and adolescents. YSR is similar but is self-completed by the child or adolescent. The scales consist of a Total problems score, which is the sum of the scores of all the problem items. An internalising problem scale sums the anxious/depressed, withdrawn-depressed, and somatic complaints scores while the externalising problem scale combines rule-breaking and aggressive behaviour. The standard scores are scaled so that 50 is average for the child or adolescent's age and gender, with a SD of 10 points. Higher scores indicate greater problems, with a T-score above 63 considered to be in the clinical range. One study (de Vries et al. 2011) provided evidence for psychosocial functioning (CBCL and YSR scores) at 2 time points: • before starting a GnRH analogue (mean [±SD] age 14.75 [±1.92] years), and

	• shortly before starting gender-affirming hormones (mean [±SD] age: 16.64 [±1.90] years). At follow up, the mean (±SD) CBCL scores were statistically significantly lower (improved) compared with baseline for: • Total T score (n=54, 60.70 [±12.76] versus 54.46 [±11.23], p<0.001 • Internalising T score (n=54, 61.00 [±12.21] versus 52.17 [±9.81], p<0.001) • Externalising T score (n=54, 58.04 [±12.99] versus 53.81 [±11.86], p=0.001) At follow up, the mean (±SD) YSR scores were statistically significantly lower (improved) compared with baseline for: • Total T score (n=54, 55.46 [±11.56] versus 50.00 [±10.56], p<0.001) • Internalising T score (n=54, 56.04 [±12.49] versus 49.78 [±11.63], p<0.001) • Externalising T score (n=54, 53.30 [±11.87] versus 49.98 [±9.35], p=0.009). The proportion of adolescents scoring in the clinical range decreased from baseline to follow up on the CBCL total problem scale (44.4% versus 22.2%, p=0.001) and the internalising scale of the YSR (29.6% versus 11.1%, p=0.017) **(VERY LOW)**. One study (Staphorsius et al. 2015) assessed CBCL in a cohort of adolescents with gender dysphoria (transfemale: n=18, mean [±SD] age 15.1 [±2.4] years and transmale: n=22, mean [±SD] age 15.8 [±1.9] years) either receiving GnRH analogues (transfemale, n=8 and transmale, n=12), or not receiving GnRH analogues (transfemale, n=10 and transmale, n=10). The mean (±SD) CBCL scores for each group were (statistical analysis unclear): • transfemales (total) 57.8 [±9.2] • transfemales receiving GnRH analogues 57.4 [±9.8] • transfemales not receiving GnRH analogues 58.2 [±9.3] • transmales (total) 60.4 [±10.2] • transmales receiving GnRH analogues 57.5 [±9.4] • transmales not receiving GnRH analogues 63.9 [±10.5] **(VERY LOW)**. **These studies provide very low certainty evidence that during treatment with GnRH analogues psychosocial functioning may improve, with the proportion of adolescents in the clinical range for some CBCL and YSR scores decreasing over time.**
Engagement with health care services **Certainty of evidence: very low**	This is an important outcome because patient engagement with health care services will impact on their clinical outcomes. Two uncontrolled observational cohort studies provided evidence relating to loss to follow up, which could be a marker of engagement with health care services (Brik et al. 2018 and Costa et al. 2015).

	In one retrospective study (Brik et al. 2018), 9 adolescents (9/214, 4.2%) who had stopped attending appointments were excluded from the study between November 2010 and July 2019 **(VERY LOW)**. One prospective study (Costa et al. 2015) had evidence for a large loss to follow-up over time. The sample size at baseline (T0) and 6 months (T1) was 201, which dropped by 39.8% to 121 after 12 months (T2) and by 64.7% to 71 at 18 months follow-up (T3). No explanation of the reasons for loss to follow-up are reported **(VERY LOW)**. Due to their design there was no reported loss to follow-up in the other 3 effectiveness studies (de Vries et al 2011; Khatchadourian et al. 2014; Staphorsius et al. 2015). **These studies provide very low certainty evidence about loss to follow up, which could be a marker of engagement with health care services, during treatment with GnRH analogues. Due to the large variation in rates between studies no conclusions could be drawn.**
Impact on extent of and satisfaction with surgery	This is an important outcome because some children and adolescents with gender dysphoria may proceed to transitioning surgery. No evidence was identified.
Stopping treatment **Certainty of evidence: very low**	This is an important outcome because there is uncertainty about the short- and long-term safety and adverse effects of GnRH analogues in children and adolescents with gender dysphoria. Two uncontrolled, retrospective, observational cohort studies provided evidence relating to stopping GnRH analogues. One study had complete reporting of the cohort (Brik et al. 2018), the other (Khatchadourian et al. 2014) had incomplete reporting of its cohort, particularly for transfemales where outcomes for only 4/11 were reported. Brik et al. 2018 narratively reported the reasons for stopping GnRH analogues in a cohort of 143 adolescents (38 transfemales and 105 transmales). Median age at the start of GnRH analogues was 15.0 years (range, 11.1–18.6 years) in transfemales and 16.1 years (range, 10.1–17.9 years) in transmales. Of these adolescents, 125 (87%, 36 transfemales, 89 transmales) subsequently started gender-affirming hormones after 1.0 (0.5–3.8) and 0.8 (0.3–3.7) years of GnRH analogues. At the time of data collection, the median duration of GnRH analogue use was 2.1 years (1.6–2.8). During the follow-up period 6.3% (9/143) of adolescents had discontinued GnRH analogues after a median duration of 0.8 years (range 0.1 to 3.0). The percentages and reasons for stopping were: • 2.8% (4/143) stopped GnRH analogues although they wanted to continue endocrine treatments for gender dysphoria: ○ 1 transmale stopped due to increase in mood problems, suicidal thoughts and confusion attributed to GnRH analogues ○ 1 transmale had hot flushes, increased migraines, fear of injections, stress at school and unrelated medical issues, and temporarily stopped treatment (after 4 months) and restarted 5 months later.

	○ 1 transmale had mood swings 4 months after starting GnRH analogues. After 2.2 years had unexplained severe nausea and rapid weight loss and discontinued GnRH analogues after 2.4 years ○ 1 transmale stopped GnRH analogues because of inability to regularly collect medication and attend appointments for injections. • 3.5% (5/143) stopped treatment because they no longer wished to receive gender-affirming treatment for various reasons **(VERY LOW)**. Khatchadourian et al. 2014 narratively reported the reasons for stopping GnRH analogues in a cohort of 26 adolescents (15 transmales and 11 transfemales), 42% (11/26) discontinued GnRH analogues during follow-up between 1998 and 2011. Of 15 transmales receiving GnRH analogues, 14 received testosterone during the observation period, of which: • 7 continued GnRH analogues after starting testosterone • 7 stopped GnRH analogues after a median of 3.0 years (range 0.2 to 9.2 years), of which: ○ 5 stopped after hysterectomy and salpingo-oophorectomy ○ 1 stopped after 2.2 years (transitioned to gender-affirming hormones) ○ 1 stopped after <2 months due to mood and emotional lability **(VERY LOW)**. Of 11 transfemales receiving GnRH analogues, 5 received oestrogen during the observation period, of which: • 4 continued GnRH analogues after starting oestrogen • 1 stopped GnRH analogues when taking oestrogen (no reason reported) **(VERY LOW)**. Of the remaining 6 transfemales taking GnRH analogues: • 1 stopped GnRH analogues after a few months due to emotional lability • 1 stopped GnRH analogues before taking oestrogen (the following year delayed due to heavy smoking) • 1 stopped GnRH analogues after 13 months due not to pursuing transition **(VERY LOW)**. **These studies provide very low certainty evidence for the number of adolescents who stop GnRH analogues and the reasons for this.**

Abbreviations: GnRH, gonadotrophin releasing hormone; SD, standard deviation.

In children and adolescents with gender dysphoria, what is the short-term and long-term safety of GnRH analogues compared with one or a combination of psychological support, social transitioning to the desired gender or no intervention?

Outcome	Evidence statement

Safety	
Change in bone density: lumbar **Certainty of evidence: very low**	This is an important outcome because puberty is an important time for bone development and puberty suppression may affect bone development, as shown by changes in lumbar bone density. Three uncontrolled, observational, retrospective studies provided evidence relating to the effect of GnRH analogues on bone density (based on lumbar BMAD) between starting with a GnRH analogue and at 1 and 2 year intervals (Joseph et al. 2019), and between starting GnRH analogues and starting gender-affirming hormones (Klink et al. 2015 and Vlot et al. 2017). All outcomes were reported separately for transfemales and transmales; also see subgroups table below. BMAD is a size adjusted value of BMD incorporating body size measurements using UK norms in growing adolescents. It was reported as g/cm³ and as z-scores. Z-scores report how many standard deviations from the mean a measurement sits. A z-score of 0 is equal to the mean, a z-score of −1 is equal to 1 standard deviation below the mean, and a z-score of +1 is equal to 1 standard deviation above the mean. One retrospective observational study (Joseph et al. 2019, n=70) provided non-comparative evidence on change in lumbar BMAD increase using z-scores. • The z-score for lumbar BMAD was statistically significantly lower at 2 years compared with baseline in transfemales (z-score [±SD]: baseline 0.486 [0.809], 2 years −0.279 [0.930], p=0.000) and transmales (baseline −0.361 [1.439], 2 years −0.913 [1.318], p=0.001) **(VERY LOW)**. • The z-score for lumbar BMAD was statistically significantly lower at 1 year compared with baseline in transfemales (baseline 0.859 [0.154], 1 year −0.228 [1.027], p=0.000) and transmales (baseline −0.186 [1.230], 1 year −0.541 [1.396], p=0.006) **(VERY LOW)**. • Actual lumbar BMAD values in g/cm³ were not statistically significantly different between baseline and 1 or 2 years in transfemales or transmales **(VERY LOW)**. Two retrospective observational studies (Klink et al. 2015 and Vlot et al. 2017, n=104 in total) provided non-comparative evidence on change in lumbar BMAD between starting GnRH analogues and starting gender-affirming hormones. All outcomes were reported separately for transfemales and transmales; also see subgroups table below. In Klink et al. 2015 the z-score for lumbar BMAD was not statistically significantly different between starting GnRH analogues and starting gender-affirming hormones in transfemales but was statistically significantly lower when starting gender-affirming hormones in transmales (z-score mean [±SD]: GnRH analogue 0.28 [±0.90], gender-affirming hormone −0.50 [±0.81], p=0.004). Actual lumbar BMAD values in g/cm³ were not statistically significantly different between starting GnRH analogues and starting gender-affirming hormones in transfemales or transmales **(VERY LOW)**.

	Vlot et al. 2017 reported change from starting GnRH analogues to starting gender-affirming hormones in lumbar BMAD by bone age. • The z-score for lumbar BMAD in transfemales with a bone age of <15 years was statistically significantly lower at starting gender-affirming hormone treatment than at starting GnRH analogues (z-score median [range]: GnRH analogue −0.20 [−1.82 to 1.18], gender-affirming hormone −1.52 [−2.36 to 0.42], p=0.001) but was not statistically significantly different in transfemales with a bone age ≥15 years **(VERY LOW)**. • The z-score for lumbar BMAD in transmales with a bone age of <14 years was statistically significantly lower at starting gender-affirming hormone treatment than at starting GnRH analogues (z-score median [range]: GnRH analogue −0.05 [−0.78 to 2.94], gender-affirming hormone −0.84 [−2.20 to 0.87], p=0.003) and in transmales with a bone age ≥14 years (GnRH analogue 0.27 [−1.60 to 1.80], gender-affirming hormone −0.29 [−2.28 to 0.90], p≤0.0001) **(VERY LOW)**. • Actual lumbar BMAD values in g/cm³ were not statistically significantly different between starting GnRH analogues and starting gender-affirming hormones in transfemales or transmales with young or old bone age **(VERY LOW)**. Two uncontrolled, observational, retrospective studies provided evidence for the effect of GnRH analogues on bone density (based on lumbar BMD) between starting GnRH analogues and either at 1 or 2 year intervals (Joseph et al. 2019), or starting gender-affirming hormones (Klink et al. 2015). All outcomes were reported separately for transfemales and transmales; also see subgroups table below. One retrospective observational study (Joseph et al. 2019, n=70) provided non-comparative evidence on change in lumbar BMD increase using z-scores. • The z-score for lumbar BMD was statistically significantly lower at 2 years compared with baseline in transfemales (z-score mean [±SD]: baseline 0.130 [0.972], 2 years −0.890 [±1.075], p=0.000) and transmales (baseline −0.715 [±1.406], 2 years −2.000 [1.384], p=0.000) **(VERY LOW)**. • The z-score for lumbar BMD was statistically significantly lower at 1 year compared with baseline in transfemales (z-score mean [±SD]: baseline −0.016 [±1.106], 1 year −0.461 [±1.121], p=0.003) and transmales (baseline −0.395 [±1.428], 1 year −1.276 [±1.410], p=0.000) **(VERY LOW)**. • With the exception of transmales, where lumbar BMD in kg/m² increased between baseline and 1 year (mean [±SD]: baseline 0.694 [±0.149], 1 year 0.718 [±0.124], p=0.006), actual lumbar BMD values were not statistically significantly different between baseline and 1 or 2 years in transfemales or between 0 and 2 years in transmales **(VERY LOW)**. One retrospective observational study (Klink et al. 2015, n=34) provided non-comparative evidence on change in lumbar BMD between starting GnRH analogues and starting gender-affirming hormones. • The z-score for lumbar BMD was not statistically significantly different between starting GnRH analogue and starting gender-affirming hormone treatment in transfemales, but was

	statistically significantly lower when starting gender-affirming hormones in transmales (z-score mean [±SD]: GnRH analogue 0.17 [±1.18], gender-affirming hormone −0.72 [±0.99], p<0.001) **(VERY LOW)**. • Actual lumbar BMD in g/cm² was not statistically significantly different between starting GnRH analogues and starting gender-affirming hormones in transfemales but was statistically significantly lower when starting gender-affirming hormones in transmales (mean [±SD]: GnRH analogues 0.95 [±0.12], gender-affirming hormones 0.91 [±0.10], p=0.006) **(VERY LOW)** **These studies provide very low certainty evidence that GnRH analogues reduce the expected increase in lumbar bone density (BMAD or BMD) compared with baseline (although some findings were not statistically significant). These studies also show that GnRH analogues do not statistically significantly decrease actual lumbar bone density (BMAD or BMD).**
Change in bone density: femoral **Certainty of evidence: very low**	This is an important outcome because puberty is an important time for bone development and puberty suppression may affect bone development, as shown by changes in femoral bone density. Two uncontrolled, observational, retrospective studies provided evidence relating to the effect of GnRH analogues on bone density (based on femoral BMAD) between starting treatment with a GnRH analogue and starting gender-affirming hormones (Klink et al. 2015 and Vlot et al. 2017). All outcomes were reported separately for transfemales and transmales; also see subgroups table below. One retrospective observational study (Klink et al. 2015, n=34) provided non-comparative evidence on change in femoral area BMAD between starting GnRH analogues and starting gender-affirming hormones. All outcomes were reported separately for transfemales and transmales. • The z-score for femoral area BMAD was not statistically significantly different between starting GnRH analogues and starting gender-affirming hormones in transfemales or transmales **(VERY LOW)**. • Actual femoral area BMAD values were not statistically significantly different between starting GnRH analogues and starting gender-affirming hormones in transmales or transfemales **(VERY LOW)**. One retrospective observational study (Vlot et al. 2017, n=70) provided non-comparative evidence on change in femoral neck (hip) BMAD between starting GnRH analogues and starting gender-affirming hormones. All outcomes were reported separately for transfemales and transmales; also see subgroups table below. • The z-score for femoral neck BMAD in transfemales with a bone age of <15 years was not statistically significantly lower at starting gender-affirming hormones than at starting GnRH analogues (z-score median [range]: GnRH analogue −0.71 [−3.35 to 0.37], gender-affirming hormone −1.32 [−3.39 to 0.21], p≤0.1) or in transfemales with a bone age ≥15 years (GnRH analogue −0.44 [−1.37 to 0.93], gender-affirming hormone −0.36 [−1.50 to 0.46]) **(VERY LOW)**.

	• The z-score for femoral neck BMAD in transmales with a bone age of <14 years was not statistically significantly lower at starting gender-affirming hormones than at starting GnRH analogues (z-score median [range]: GnRH analogue −0.01 [−1.30 to 0.91], gender-affirming hormone −0.37 [−2.28 to 0.47]) but was statistically significantly lower in transmales with a bone age ≥14 years (GnRH analogue 0.27 [−1.39 to 1.32], gender-affirming hormone −0.27 [−1.91 to 1.29], p=0.002) **(VERY LOW)**. • Actual femoral neck BMAD values were not statistically significantly different between starting GnRH analogues and starting gender-affirming hormones in transfemales or in transmales with a young bone age, but were statistically significantly lower in transmales with a bone age ≥14 years (GnRH analogue 0.33 [0.25 to 0.39), gender-affirming hormone 0.30 [0.23 to 0.41], p≤0.01) **(VERY LOW)** Two uncontrolled, observational, retrospective studies provided evidence for the effect of GnRH analogues on bone density (based on femoral BMD) between starting GnRH analogues and either at 1 or 2 year intervals (Joseph et al. 2019), or starting gender-affirming hormones (Klink et al. 2015). All outcomes were reported separately for transfemales and transmales; also see subgroups table below. One retrospective observational study (Joseph et al. 2019, n=70) provided non-comparative evidence on change in femoral neck BMD increase using z-scores. All outcomes were reported separately for transfemales and transmales. • The z-score for femoral neck BMD was statistically significantly lower at 2 years compared with baseline in transfemales (z-score mean [±SD]: baseline 0.0450 [±0.781], 2 years −0.600 [±1.059], p=0.002) and transmales (baseline −1.075 [±1.145], 2 years −1.779 [±0.816], p=0.001) **(VERY LOW)**. • The z-score for femoral neck BMD was statistically significantly lower at 1 year compared with baseline in transfemales (z-score mean [±SD]: baseline 0.157 [±0.905], 1 year −0.340 [±0.816], p=0.002) and transmales (baseline −0.863 [±1.215], 1 year −1.440 [±1.075], p=0.000) **(VERY LOW)**. • Actual femoral neck BMD values in kg/m² were not statistically significantly different between baseline and 1 or 2 years in transmales or transfemales **(VERY LOW)**. One retrospective observational study (Klink et al. 2015, n=34) provided non-comparative evidence on change in femoral area BMD between starting GnRH analogues and starting gender-affirming hormones. All outcomes were reported separately for transfemales and transmales. • The z-score for femoral area BMD was not statistically significantly different between starting GnRH analogues and starting gender-affirming hormones in transfemales, but was statistically significantly lower in transmales (z-score mean [±SD]: GnRH analogue 0.36 [±0.88], gender-affirming hormone −0.35 [±0.79], p=0.001) **(VERY LOW)**. • Actual femoral area BMD values were not statistically significantly different between starting GnRH analogues and starting gender-affirming hormones in transfemales, but were

	statistically significantly lower in transmales (mean [±SD] GnRH analogue 0.92 [±0.10], gender-affirming hormone 0.88 [±0.09], p=0.005) **(VERY LOW)** **These studies provide very low certainty evidence that GnRH analogues may reduce the expected increase in femoral bone density (femoral neck or area BMAD or BMD) compared with baseline (although some findings were not statistically significant). These studies also show that GnRH analogues do not statistically significantly decrease actual femoral bone density (femoral area BMAD or femoral neck BMD), apart from actual femoral area BMD in transmales.**
Cognitive development or functioning **Certainty of evidence: very low**	This is an important outcome because puberty is an important time for cognitive development and puberty suppression may affect cognitive development or functioning. One cross-sectional observational study (Staphorsius et al. 2015, n=70) provided comparative evidence on cognitive development or functioning in adolescents with gender dysphoria on GnRH analogues compared with adolescents with gender dysphoria not on GnRH analogues. Cognitive functioning was measured using an IQ test. Reaction time (in seconds) and accuracy (percentage of correct trials) were measured using the Tower of London (ToL) task. All outcomes were reported separately for transfemales and transmales; also see subgroups table below. No statistical analyses or interpretation of the results in these groups were reported. • IQ in transfemales (mean [±SD] GnRH analogue 94.0 [±10.3], control 109.4 [±21.2]). IQ transmales (GnRH analogue 95.8 [±15.6], control 98.5 [±15.9]. • Reaction time in transfemales (mean [±SD] GnRH analogue 10.9 [±4.1], control: 9.9 [±3.1]). Reaction time transmales (GnRH analogue 9.9 [±3.1], control 10.0 [±2.0]). • Accuracy score in transfemales (GnRH analogue 73.9 [±9.1], control 83.4 [±9.5]. Accuracy score in transmales (GnRH analogue 85.7 [±10.5], control 88.8 [±9.7]. **This study provides very low certainty evidence (with no statistical analysis) on the effects of GnRH analogues on cognitive development or functioning. No conclusions could be drawn.**
Other safety outcomes: kidney function **Certainty of evidence: very low**	This is an important outcome because if renal damage (raised serum creatinine is a marker of this) is suspected, GnRH analogues may need to be stopped. One prospective observational study (Schagen et al. 2016, n=116) provided non-comparative evidence on change in serum creatinine between starting GnRH analogues and at 1 year. All outcomes were reported separately for transfemales and transmales; also see subgroups table below. • There was no statistically significant difference between baseline and 1 year for serum creatinine in transfemales (mean [±SD] baseline 70 [±12], 1 year 66 [±13], p=0.20). • There was a statistically significant decrease between baseline and 1 year for serum creatinine in transmales (baseline 73 [±8], 1 year 68 [±13], p=0.01).

	This study provides very low certainty evidence that GnRH analogues do not affect renal function.
Other safety outcomes: liver function **Certainty of evidence: very low**	This is an important outcome because if treatment-induced liver injury (raised liver enzymes are a marker of this) is suspected, GnRH analogues may need to be stopped. One prospective observational study (Schagen et al. 2016, n=116) provided non-comparative evidence on elevated liver enzymes between starting GnRH analogues and during use. No comparative values or statistical analyses were reported. • Glutamyl transferase was not elevated at baseline or during use in any person. • Mild elevations of AST and ALT above the reference range were present at baseline but were not more prevalent during use than at baseline. • Glutamyl transferase, AST, and ALT levels did not significantly change from baseline to 12 months of use. **This study provides very low certainty evidence (with no statistical analysis) that GnRH analogues do not affect liver function.**
Other safety outcomes: adverse effects **Certainty of evidence: very low**	This is an important outcome because if there are adverse effects, GnRH analogues may need to be stopped. One uncontrolled, retrospective, observational cohort study (Khatchadourian et al. 2014) provided evidence relating to adverse effects from GnRH analogues. It had incomplete reporting of its cohort, particularly for transfemales where outcomes for only 4/11 were reported. Khatchadourian et al. 2014 reported adverse effects in a cohort of 26 adolescents (15 transmales and 11 transfemales) receiving GnRH analogues. Of these: • 1 transmale developed sterile abscesses; they were switched from leuprolide acetate to triptorelin, and this was well tolerated. • 1 transmale developed leg pains and headaches, which eventually resolved. • 1 participant gained 19 kg within 9 months of starting GnRH analogues. **This study provides very low certainty evidence about potential adverse effects of GnRH analogues. No conclusions could be drawn.**

Abbreviations: ALT, alanine aminotransferase; AST, aspartate aminotransferase; BMAD, bone mineral apparent density; BMD, bone mineral density; GnRH, gonadotrophin releasing hormone; IQ, intelligence quotient; NS, not significant; SD, standard deviation.

In children and adolescents with gender dysphoria, what is the cost-effectiveness of GnRH analogues compared to one or a combination of psychological support, social transitioning to the desired gender or no intervention?

Outcome	Evidence statement

Cost-effectiveness	No studies were identified to assess the cost-effectiveness of GnRH analogues for children and adolescents with gender dysphoria.

From the evidence selected, are there any subgroups of children and adolescents with gender dysphoria that may benefit from GnRH analogues more than the wider population of interest?

Subgroup	Evidence statement
Sex assigned at birth males (transfemales) **Certainty of evidence: Very low**	Some studies reported data separately for sex assigned at birth males (transfemales). This included some direct comparisons with sex assigned at birth females (transmales). **Impact on gender dysphoria** One uncontrolled prospective observational longitudinal study (de Vries et al. 2011) provided evidence for gender dysphoria in sex assigned at birth males. See the clinical effectiveness results table above for a full description of the study The mean (±SD) UGDS score was statistically significantly lower (improved) in sex assigned at birth males compared with sex assigned at birth females at both baseline (T0) (n=not reported, mean UGDS score [±SD]: 47.95 [±9.70] versus 56.57 [±3.89]) and T1 (n=not reported, 49.67 [±9.47] versus 56.62 [±4.00]); between sex difference p<0.001 **(VERY LOW)**. One further prospective observational longitudinal study (Costa et al. 2015) provided evidence for the impact on gender dysphoria in sex assigned at birth males. See the clinical effectiveness results table above for a full description of the study. Sex assigned at birth males had a statistically significantly lower (improved) mean (±SD) UGDS score of 51.6 [±9.7] compared with sex assigned at birth females (56.1 [±4.3], p<0.001). However, it was not reported if this was baseline or follow-up **(VERY LOW)**. **These studies provide very low certainty evidence that in sex assigned at birth males (transfemales), gender dysphoria is lower than in sex assigned at birth females (transmales).** **Impact on mental health** One uncontrolled prospective observational longitudinal study (de Vries et al. 2011) provided evidence for the impact on mental health (depression, anger and anxiety) in sex assigned at birth males. See the clinical effectiveness results table above for a full description of the study. • The mean (±SD) depression (BDI-II) score was not statistically significantly different in sex assigned at birth males compared with sex assigned at birth females at both baseline (T0) (n=not reported, mean BDI score [±SD]: 5.71 [±4.31] versus 10.34 [±8.24]) and T1 (n=not reported, 3.50 [±4.58] versus 6.09 [±7.93]), between sex difference p=0.057 • The mean (±SD) anger (TPI) score was statistically significantly lower (improved) in sex assigned at birth males compared with sex assigned at birth females at both baseline (T0) (n=not reported, mean TPI score [±SD]: 5.22 [±2.76] versus 6.43 [±2.78]) and T1 (n=not reported, 5.00 [±3.07] versus 6.39 [±2.59]), between sex difference p=0.022 • The mean (±SD) anxiety (STAI) score was statistically significantly lower (improved) in sex assigned at birth males compared with sex assigned at birth females at both baseline (T0) (n=not reported, mean STAI score [±SD]: 4.33 [±2.68] versus 7.00 [±2.36]) and T1 (n=not reported, 4.39 [±2.64] versus 6.17 [±2.69]), between sex difference p<0.001 **(VERY LOW)**. **This study provides very low certainty evidence that the impact on mental health (depression, anger and anxiety) may be different in sex assigned at birth males (transfemales) compared with sex assigned at birth females (transmales). Over time there was no statistically significant difference between sex assigned at birth males and sex assigned at birth females for depression. However, sex assigned at birth males had statistically significantly lower levels of anger and anxiety than sex assigned at birth females at both baseline and follow up.** **Impact on body image** One uncontrolled prospective observational longitudinal study (de Vries et al. 2011) provided evidence relating to the impact on body image in sex assigned at birth males. • The mean (±SD) BIS score for primary sex characteristics was statistically significantly lower (improved) in sex assigned at birth males compared with sex assigned at birth females at both baseline (T0) (n=not reported, mean BIS score [±SD] 4.02 [±0.61] versus 4.16 [±0.52]) and T1 (n=not reported, 3.74 [±0.78] versus 4.17 [±0.58]), between sex difference p=0.047 • The mean (±SD) BIS score for secondary sex was statistically significantly lower (improved) in sex assigned at birth males compared with sex assigned at birth females at both baseline (T0) (n=not reported, mean BIS score [±SD]: 2.66 [±0.50] versus 2.81 [±0.76]) and T1 (n=not reported, 2.39 [±0.69] versus 3.18 [±0.42]), between sex difference p=0.001 • The mean (±SD) BIS score for neutral body characteristics was not statistically significantly different in sex assigned at birth males compared with sex assigned at birth females at both baseline (T0) (n=not reported, mean BIS score [±SD]: 2.60 [±0.58] versus 2.24 [±0.62]) and T1 (n=not reported, 2.32 [±0.59] versus 2.61 [±0.50]), between sex difference p=0.777 **(VERY LOW)**. **This study provides very low certainty evidence that the impact on body image may be different in sex assigned at birth males (transfemales) compared with sex assigned at birth females (transmales). Sex assigned at birth males are less dissatisfied with their primary and secondary sex characteristics than sex assigned at birth females at both baseline and follow up, but the satisfaction with neutral body characteristics is not different.** **Psychosocial impact** One uncontrolled prospective observational longitudinal study (de Vries et al. 2011) provided evidence for psychosocial impact in terms

	of global functioning (CGAS) and psychosocial functioning (CBCL and YSR) in sex assigned at birth males. • Sex assigned at birth males had statistically higher mean (±SD) CGAS scores compared with sex assigned at birth females at both baseline (T0) (n=54, 73.10 [±8.44] versus 67.25 [±11.06]) and T1 (n=54, 77.33 [±8.69] versus 70.30 [±9.44]), between sex difference p=0.021 • There was no statistically significant difference between sex assigned at birth males and sex assigned at birth females for the CBCL Total T score at T0 or T1 (n=54, p=0.110) • There was no statistically significant difference between sex assigned at birth males and sex assigned at birth females for the CBCL internalising T score at T0 or T1 (n=54, p=0.286) • Sex assigned at birth males had statistically lower mean (±SD) CBCL externalising T scores compared with sex assigned at birth females at both T0 (n=54, 54.71 [±12.91] versus 60.70 [±12.64]) and T1 (n=54, 48.75 [±10.22] versus 57.87 [±11.66]), between sex difference p=0.015 • There was no statistically significant difference between sex assigned at birth males and sex assigned at birth females for the YSR Total T score at T0 or T1 (n=54, p=0.164) • There was no statistically significant difference between sex assigned at birth males and sex assigned at birth females for the YSR internalising T score at T0 or T1 (n=54, p=0.825) • Sex assigned at birth males had statistically lower mean (±SD) YSR externalising T scores compared with sex assigned at birth females at both T0 (n=54, 48.72 [±11.38] versus 57.24 [±10.59]) and T1 (n=54, 46.52 [±9.23] versus 52.97 [±8.51]), between sex difference p=0.004 **(VERY LOW)**. One uncontrolled, observational, prospective cohort study (Costa et al. 2015) provided evidence for psychosocial impact in terms of global functioning (CGAS) in sex assigned at birth males. • Sex assigned at birth males had statistically significant lower mean (±SD CGAS scores at baseline) compared with sex assigned at birth females (n=201, 55.4 [±12.7] versus 59.2 [±11.8], p=0.03) **(VERY LOW)** **These studies provide very low certainty evidence that psychosocial impact may be different in sex assigned at birth males (transfemales) compared with sex assigned at birth females (transmales). However, no conclusions could be drawn.** **Change in bone density: lumbar** Three uncontrolled, observational, retrospective studies provided evidence relating to the effect of GnRH analogues on lumbar bone density in sex assigned at birth males (Joseph et al. 2019, Klink et al. 2015 and Vlot et al. 2017). See the safety results table above for a full description of the results. **These studies provide very low certainty evidence that GnRH analogues reduce the expected increase in lumbar bone density (BMAD or BMD) in sex assigned at birth males (transfemales; although some findings were not statistically significant). These studies also show that GnRH analogues do not statistically**

	significantly decrease actual lumbar bone density (BMAD or BMD) in sex assigned at birth males (transfemales). **Change in bone density: femoral** Three uncontrolled, observational, retrospective studies provided evidence for the effect of GnRH analogues on femoral bone density in sex assigned at birth males (Joseph et al. 2019, Klink et al. 2015 and Vlot et al. 2017). See the safety results table above for a full description of the results. **These studies provide very low certainty evidence that GnRH analogues may reduce the expected increase in femoral bone density (femoral neck or area BMAD or BMD) in sex assigned at birth males (transfemales; although some findings were not statistically significant). These studies also show that GnRH analogues do not statistically significantly decrease actual femoral bone density (femoral area BMAD or femoral neck BMD) in sex assigned at birth males (transfemales).** **Cognitive development or functioning** One cross-sectional observational study (Staphorsius et al. 2015) provided comparative evidence on cognitive development or functioning in sex assigned at birth males. See the safety results table above for a full description of the results. **This study provides very low certainty evidence (with no statistical analysis) on the effects of GnRH analogues on cognitive development or functioning in sex assigned at birth males (transfemales). No conclusions could be drawn.** **Other safety outcomes: kidney function** One prospective observational study (Schagen et al. 2016) provided non-comparative evidence on change in serum creatinine in sex assigned at birth males. See the safety results table above for a full description of the results. **This study provides very low certainty evidence that GnRH analogues do not affect renal function in sex assigned at birth males (transfemales).**
Sex assigned at birth females (transmales) **Certainty of evidence: Very low**	Some studies reported data separately for sex assigned at birth females (transmales). This included some direct comparisons with sex assigned at birth males (transfemales). **Impact on gender dysphoria** One uncontrolled prospective observational longitudinal study (de Vries et al. 2011) and one prospective observational longitudinal study (Costa et al. 2015) provided evidence for gender dysphoria in sex assigned at birth females. See the sex assigned at birth males (transfemales) row above for a full description of the results. **These studies provide very low certainty evidence that in sex assigned at birth females (transmales), gender dysphoria is higher than in sex assigned at birth males (transfemales) at both baseline and follow up.**

	Impact on mental health One uncontrolled prospective observational longitudinal study (de Vries et al. 2011) provided evidence relating to the impact on mental health (depression, anger and anxiety) in sex assigned at birth females. See the sex assigned at birth males (transfemales) row above for a full description of the results. **This study provides very low certainty evidence that the impact on mental health (depression, anger and anxiety) may be different in sex assigned at birth females (transmales) compared with sex assigned at birth males (transfemales). Over time there was no statistically significant difference between sex assigned at birth females and sex assigned at birth males for depression. However, sex assigned at birth females had statistically significantly greater levels of anger and anxiety than sex assigned at birth males at baseline and follow up.** **Impact on body image** One uncontrolled prospective observational longitudinal study (de Vries et al. 2011) provided evidence relating to the impact on body image in sex assigned at birth females. See the sex assigned at birth males (transfemales) row above for a full description of the results. **This study provides very low certainty evidence that the impact on body image may be different in sex assigned at birth females (transmales) compared with sex assigned at birth males (transfemales). Sex assigned at birth females are more dissatisfied with their primary and secondary sex characteristics than sex assigned at birth males at both baseline and follow up, but the satisfaction with neutral body characteristics is not different.** **Psychosocial impact** One uncontrolled prospective observational longitudinal study (de Vries et al. 2011) provided evidence for psychosocial impact in terms of global functioning (CGAS) and psychosocial functioning (CBCL and YSR) in sex assigned at birth females. One uncontrolled, observational, prospective cohort study (Costa et al. 2015) provided evidence for psychosocial impact in terms of global functioning (CGAS) in sex assigned at birth females. See the sex assigned at birth males (transfemales) row above for a full description of the results. **These studies provide very low certainty evidence that psychosocial impact may be different in sex assigned at birth females (transmales) compared with sex assigned at birth males (transfemales). However, no conclusions could be drawn.** **Change in bone density: lumbar** Three uncontrolled, observational, retrospective studies provided evidence relating to the effect of GnRH analogues on lumbar bone density in sex assigned at birth females (Joseph et al. 2019, Klink et al. 2015 and Vlot et al. 2017). See the safety results table above for a full description of the results.

	These studies provide very low certainty evidence that GnRH analogues reduce the expected increase in lumbar bone density (BMAD or BMD) in sex assigned at birth females (transmales; although some findings were not statistically significant). These studies also show that GnRH analogues do not statistically significantly decrease actual lumbar bone density (BMAD or BMD) in sex assigned at birth females (transmales). **Change in bone density: femoral** Three uncontrolled, observational, retrospective studies provided evidence relating to the effect of GnRH analogues on femoral bone density in sex assigned at birth females (Joseph et al. 2019, Klink et al. 2015 and Vlot et al. 2017). See the safety results table above for a full description of the results. **These studies provide very low certainty evidence that GnRH analogues may reduce the expected increase in femoral bone density (femoral neck or area BMAD or BMD) in sex assigned at birth females (transmales; although some findings were not statistically significant). These studies also show that GnRH analogues do not statistically significantly decrease actual femoral bone density (femoral area BMAD or femoral neck BMD) in sex assigned at birth females (transmales), apart from actual femoral area.** **Cognitive development or functioning** One cross-sectional observational study (Staphorsius et al. 2015) provided comparative evidence on cognitive development or functioning in sex assigned at birth females. See the safety results table above for a full description of the results. **This study provides very low certainty evidence (with no statistical analysis) on the effects of GnRH analogues on cognitive development or functioning in sex assigned at birth females (transmales). No conclusions could be drawn.** **Other safety outcomes: kidney function** One prospective observational study (Schagen et al. 2016) provided non-comparative evidence on change in serum creatinine in sex assigned at birth females (transmales). See the safety results table above for a full description of the results. **This study provides very low certainty evidence that GnRH analogues do not affect renal function in sex assigned at birth females (transmales).**
Duration of gender dysphoria	No evidence was identified.
Age at onset of gender dysphoria	No evidence was identified.
Age at which GnRH analogue started	No evidence was identified.
Age at onset of puberty	No evidence was identified.

Tanner stage at which GnRH analogue started	No evidence was identified
Diagnosis of autistic spectrum disorder	No evidence was identified
Diagnosis of mental health condition	No evidence was identified

Abbreviations: BDI-II, Beck Depression Inventory-II; BIS, Body Image Scale; CBCL, Child Behaviour Checklist; CGAS, Children's Global Assessment Scale; SD, standard deviation; STAI, Trait Anxiety Scale of the State-Trait Personality Inventory; TPI, Trait Anger Scale of the State-Trait Personality Inventory; UGDS, Utrecht Gender Dysphoria Scale; YSR, Youth Self-Report

From the evidence selected,

(a) what are the criteria used by the research studies to define gender dysphoria, gender identity disorder and gender incongruence of childhood?

(b) what were the ages at which participants commenced treatment with GnRH analogues?

(c) what was the duration of treatment with GnRH analogues?

Outcome	Evidence statement
Diagnostic criteria	In **5 studies** (Costa et al. 2015, Klink et al. 2015, Schagen et al. 2016, Staphorsius et al. 2015 and Vlot et al. 2017) the DSM-IV-TR criteria of gender identity disorder was used. The study by Brik et al. 2020 used DSM-V criteria. The DSM-V has one overarching definition of gender dysphoria with separate specific criteria for children and for adolescents and adults. The general definition describes a conflict associated with significant distress and/or problems functioning associated with this conflict between the way they feel and the way they think of themselves which must have lasted at least 6 months. It was not reported how gender dysphoria was defined in the remaining 3 studies (**VERY LOW**). **From the evidence selected, all studies that reported diagnostic criteria for gender dysphoria (6/9 studies) used the DSM criteria in use at the time the study was conducted.**
Age when GnRH analogues started	8/9 studies reported the age at which participants started GnRH analogues, either as the mean age (with SD) or median age (with the range): **Study** — **Mean age (±SD)** Costa et al. 2015 — **16.5 years (±1.3)** de Vries et al. 2011 — **13.6 years (±1.8)** Joseph et al. 2019 — **13.2 years (±1.4) in transfemales; 12.6 years (±1.0) in transmales** Khatchadourian et al. 2014 — **14.7 years (±1.9)**

	Klink et al. 2015 — **14.9 years (±1.9) in transfemales; 15.0 years (±2.0) in transmales** **Study** — **Median age (range)** Brik et al. 2020 — **15.5 years (11.1–18.6) in transfemales; 16.1 years (10.1–17.9) in transmales** Schagen et al. 2016 — **13.6 years (11.6–17.9) in transfemales; 14.2 years (11.1–18.6) in transmales** Vlot et al. 2017 — **13.5 years (11.5–18.3) in transfemales; 15.1 years (11.7–18.6) in transmales** Age at the start of GnRH analogues was not reported in Staphorsius et al. 2015, but participants were required to be at least 12 years (**VERY LOW**). **The evidence included showed wide variation in the age (11 to 18 years old) at which children and adolescents with gender dysphoria started GnRH analogues.**
Duration of treatment	The duration of treatment with GnRH analogues was reported in 3/9 studies. The median duration was: • 2.1 years (range 1.6–2.8) in Brik et al. 2020. • 1.3 years (range 0.5–3.8) in transfemales and 1.5 years (range 0.25–5.2) in transmales in Klink et al. 2015. In Staphorsius et al. 2015, the mean duration was 1.6 years (SD ±1.0). In de Vries et al. 2011, the mean duration of time between starting GnRH analogues and gender-affirming hormones was 1.88 years (SD ±1.05). **The evidence included showed wide variation in the duration of treatment with GnRH analogues, but most studies did not report this information. Treatment duration ranged from a few months up to about 5 years.**

Abbreviations: DSM, Diagnostic and Statistical Manual of Mental Disorders criteria; SD, standard deviation.

6. Discussion

A key limitation to identifying the effectiveness and safety of GnRH analogues for children and adolescents with gender dysphoria is the lack of reliable comparative studies. The lack of clear, expected outcomes from treatment with a GnRH analogue (the purpose of which is to suppress secondary sexual characteristics which may cause distress from unwanted pubertal changes) also makes interpreting the evidence difficult. The size of the population with gender dysphoria means conducting a prospective trial may be unrealistic, at least on a single centre basis. There may also be ethical issues with a 'no treatment arm' in comparative trials of GnRH analogues, where there may be poor mental health outcomes if treatment is withheld. However, the use of an active comparator such as close psychological support may reduce ethical concerns in future trials.

The studies included in this evidence review are all small, uncontrolled observational studies, which are subject to bias and confounding, and are of very low certainty as

assessed using modified GRADE. All the included studies reported physical and mental health comorbidities and concomitant treatments very poorly. For example, very little data are reported on how many children and adolescents needed additional mental health support, and for what reasons, or whether additional interventions, and what form and duration (for example drug treatment or counselling) that took. This is a possible confounder for the treatment outcomes in the studies because changes in critical and important outcomes may be attributable to external care rather than the psychological support or GnRH analogues used in the studies.

The studies that reported diagnostic criteria for gender dysphoria (6/9 studies) used the Diagnostic and Statistical Manual of Mental Disorders (DSM) criteria in use at the time the study was conducted (either DSM-IV-TR or DSM-V). The definition was unclear in the remaining studies. There was wide variation in the ages at which participants started a GnRH analogue, typically ranging from about 11 to 18 years. Similarly, there was a wide variation in the duration of use, but few studies reported this.

Changes in outcome scores for clinical effectiveness were assessed for statistical significance in the 3 studies reporting these outcomes (Costa et al. 2015; de Vries et al. 2011; Staphorsius et al. 2015). However, there is relatively little interpretation of whether the changes in outcome scores seen in these studies are clinically meaningful.

For some outcomes there was no statistically significant difference from before starting GnRH analogues until just before starting gender-affirming hormones. These were the Utrecht Gender Dysphoria Scale (UGDS) (which was assessed in 1 study de Vries et al. 2011), the Trait Anger (TPI) and Trait Anxiety (STAI) Scales (which were assessed in 1 study de Vries et al. 2011), and Body Image Scale (BIS) which was assessed in 1 study (de Vries et al. 2011).

The Beck Depression Inventory (BDI-II) was used in 1 study (de Vries et al. 2011) to assess change in depression from before starting GnRH analogues to just before starting gender-affirming hormones. The result is statistically significant, with the mean (±SD) BDI-II score decreasing from 8.31 (±7.12) at baseline to 4.95 (±6.27) at follow up (p=0.004). However, both scores fall into the minimal range using the general guidelines for interpretation of BDI-II (0 to 13 minimal, 14 to 19 mild depression, 20 to 28 moderate depression and 29 to 63 severe depression), suggesting that while statistically significant, it is unclear if this is a clinically meaningful change.

Psychosocial outcomes were assessed in 3 studies (Costa et al. 2015; de Vries et al. 2011; Staphorsius et al. 2015) using the Children's Global Assessment Scale (CGAS) and Child Behavior Checklist/Youth Self-Report (CBCL/YSR). The CGAS score was assessed in 2 studies (Costa et al. 2015; de Vries et al. 2011). In de Vries et al. 2011 the mean (±SD) CGAS score statistically significantly increased over time from 70.24 [±10.12] at baseline to 73.90 [±9.63] at follow up. CGAS scores are clinically categorised into 10 categories (10 to 1, 20 to 11 and so on until 100 to 91) and both scores reported were in a single category (71 to 80, no more than slight impairment) suggesting that while statistically significant, it is unclear if this is a clinically meaningful change. The Costa et al. 2015 study does highlight a larger change in CGAS scores from baseline to follow-up (mean [±SD] 58.72 [±11.38] compared with 67.40 [±13.39]), but whether this is clinically meaningful is unclear. The average score moved from the clinical category of 60 to 51 (variable functioning with sporadic difficulties) at baseline to 70 to 61 (some difficulty in a single area, but generally

functioning pretty well) at follow up, but the large standard deviations suggest clinically significant overlaps between the scores from baseline to follow-up.

Psychosocial functioning using the CBCL/YSR was assessed in 2 studies (de Vries et al. 2011; Staphorsius et al. 2015). In de Vries et al. 2011 there was a statistically significant reduction in both CBCL and YSR scores from before starting GnRH analogues to just before starting gender-affirming hormones. The study interpreted the CBCL/YSR with a proportion of adolescents who scored in the clinical range (a T-score above 63), which allows changes in clinically meaningful scores to be assessed, and proportions of adolescents in the clinical range for some CBCL and YSR scores decreased over time. One cross-sectional study (Staphorsius et al. 2015) assessed CBCL scores only, but it was unclear if this was the Total T score, or whether subscales of internalising or externalising scores were also assessed, and whether the results were statistically significant.

The 2 prospective observational studies (Costa et al. 2015; de Vries et al. 2011) are confounded by a number of common factors. Firstly, the single assessment of scores at baseline means it is unclear if scores were stable, already improving or declining before starting treatment. Secondly, in an uncontrolled study any changes in scores from baseline to follow-up could be attributed to a regression-to-mean, for example getting older has been positively associated with maturity and wellbeing. The studies use mean and standard deviations in the descriptive statistics and analyses; however, they do not report testing the normality of data which would support the use of parametric measures. The study by de Vries et al. 2011 used general linear models (regression) to examine between and within group variances (changes in outcomes). In using such models, the data is assumed to be balanced (measured at regular intervals and without missing data), but the large ranges in ages at which participants were assessed and started on various interventions suggests that ascertainment of outcome was unlikely to be regular and missing data was likely. Missing data was handled through listwise deletion (omits those cases with the missing data and analyses the remaining data) which is acceptable if data loss is completely random but for some outcomes where there was incomplete data for individual items this was not random (items were introduced by the authors after the first eligible adolescents had started GnRH analogues). The study provided no detail on whether these assumptions for the modeling were met, they also provided no adequate assessment of whether any regression diagnostics (analysis that seek to assess the validity of a model) or model fit (how much of the variance in outcome is explained by the between and within group variance) were undertaken.

The 2 retrospective observational studies (Brik et al. 2020; Khatchadourian et al. 2014) both only report absolute numbers for each trajectory along with reasons for stopping GnRH analogues. It is difficult to assess outcomes from such single centre studies because there is little comparative data for outcomes from other such services. A lack of any critical or other important outcomes also means the success of the treatment across all the participants is difficult to judge.

Three uncontrolled, observational, retrospective studies provided evidence relating to the effect of GnRH analogues on bone density (Joseph et al. 2019; Klink et al. 2015; Vlot et al. 2017). In all 3 studies, the participants acted as their own controls and change in bone density was determined between starting GnRH analogues and either after 1 and 2 year follow-up timepoints (Joseph et al. 2019) or when gender-affirming hormones were started

(Klink et al. 2015 and Vlot et al. 2017). Observational studies such as these can only show an association with GnRH analogues and bone density; they cannot show that GnRH analogues caused any differences in bone density seen. Because there was no comparator group and participants acted as their own controls, it is unclear whether the findings are associated with GnRH analogues or due to changes over time. The authors reported z-scores which allows for comparison with the expected increase in bone density in the general population. However, because no concomitant treatments or comorbidities were reported it is possible that the findings may not be because of GnRH analogues and there is another way in which the study population differs from the general population.

All the studies are from a limited number of, mainly European, care facilities. They are described as either tertiary referral or expert services but the low number of services providing such care and publishing evidence may bias the results towards the outcomes in these services only and limit extrapolation.

The first study (Brik et al. 2020) was an uncontrolled, retrospective, observational study that assessed the outcome trajectories of adolescents receiving GnRH analogues for gender dysphoria. This study followed-up 143 individuals who had received GnRH analogues (38 transfemales and 105 transmales) using clinical records to show outcomes for up to 9 years (continuing use of GnRH analogues, reasons for stopping GnRH analogues and onward care such as gender-affirming hormone use). The methods and results are well reported, but no analysis of data was undertaken. The views of adolescents and their parents are particularly difficult to interpret because no data on how many responded to each question and in what ways are reported.

The second study (Costa et al. 2015) was an uncontrolled, prospective observational study which assessed global functioning in adolescents with gender dysphoria using CGAS every 6 months, including during the first 6 months where statistically significant improvements were seen without GnRH analogues. The study is confounded by significant unexplained loss to follow-up (64.7%: from n=201 adolescents to n=71 after 18 months). Missing data for those lost to follow-up maybe more than sufficient to change the direction of effects seen in the study if the reasons for loss to follow-up are systematic (such as deriving little or no benefit from treatment). The study uses clustered data in its analysis, a single outcome (CGAS) measured in clusters (at different visits), and the analysis does not take account of the correlation of scores (data at different time points are not independent) as a significant change in scores early in the study means the successive changes measured against baseline were also significant. The study relies on multiple (>20) pairwise independent *t*-tests to examine change in CGAS between the 4 time points, increasing the possibility of type-I error (a false positive which occurs when a researcher incorrectly rejects a true null hypothesis) because the more tests performed the more likely a statistically significant result will be observed by chance alone.

The Costa et al. 2015 study compares immediately eligible and delayed eligible cohorts, however, it is highly likely that they are non-comparable groups because the immediately eligible group were those able to start GnRH analogues straight away whilst those in the delayed eligible group were either not ready to make a decision about starting treatment (no age comparison was made between the 2 groups so it is unclear if they were a younger cohort than the immediately eligible group) or had comorbid mental health or psychological difficulties. The authors report that those with concomitant problems (such as mental health

problems, substantial problems with peers, or conflicts with parents or siblings) were referred to local mental health services but no details are provided.

The third study (de Vries et al. 2011) was an uncontrolled, prospective observational study which assessed gender dysphoria and psychological functioning before and after puberty suppression in adolescents with gender dysphoria. Although the study mentions the DSM-IV-TR there is no explicit discussion of this, or any other criteria, being used as the diagnostic criteria for study entry. There are no details reported for how the outcomes in the study were assessed, and by whom. The length of follow-up for the outcomes in the model are questionable in relation to whether there was sufficient time for GnRH analogues to have a measurable effect. The time points used are start of GnRH analogues and start of gender-affirming hormones. Overall, the mean time between starting GnRH analogues and gender-affirming hormones was 1.88 (±1.05) years, but the range is as low as just 5 months between the 2 time points, which may be insufficient for any difference in outcome to have occurred in some individuals.

The fourth study (Joseph et al. 2019) was a retrospective, longitudinal observational single centre study which assessed bone mineral density in adolescents with gender dysphoria in the UK. For inclusion in the study, participants had to have been assessed by the Gender Identity Development Service multi-disciplinary psychosocial health team for at least 4 assessments over a minimum of 6 months. No other diagnostic criteria, such as the DSM-IV-TR, are discussed. Bone density was assessed using dual energy X-ray absorptiometry (DAXA) scan of the lumbar spine (L1-L4) and the femoral neck at baseline (n=70), 1 year (n=70) and 2 years after starting GnRH analogues (n=39). The results suggest a possible association between GnRH analogues and bone mineral apparent density. However, the evidence is of poor quality, and the results could be due to bias or chance. No concomitant treatments or comorbidities were reported.

The fifth study (Khatchadourian et al. 2014) was an uncontrolled retrospective observational study which describes patient characteristics at presentation, treatment, and response to treatment in 84 adolescents with gender dysphoria, of whom 27 received GnRH analogues. The study used clinical records to show outcomes for up to 13 years (continuing use of GnRH analogues, reasons for stopping GnRH analogues and onward care such as gender-affirming hormone use). The methods are well reported but the results for those taking GnRH analogues are poorly and incompletely reported, particularly for transfemales, and no analysis of data was undertaken. It is difficult to assess the results for stopping GnRH analogues due to incomplete reporting of this outcome.

The sixth study (Klink et al. 2015) was a retrospective longitudinal observational single centre study which assessed bone mineral density in adolescents with gender dysphoria, diagnosed with the DSM-IV-TR criteria. Bone density was assessed when starting GnRH analogues and then when starting gender-affirming hormones. Results are reported for transmales and transfemales separately and no results for the whole cohort are given. Statistical analyses were reported for all outcomes of interest but, because there was no comparator group and participants acted as their own controls, it is not known whether the findings are associated with GnRH analogues or due to changes over time. The authors reported z-scores which allows for comparison with the expected increase in bone density in the general population. However, because no concomitant treatments or comorbidities were

reported it is possible that the findings may not be because of GnRH analogues and there is another way in which the study population differs from the general population.

The seventh study (Schagen et al. 2016) was a prospective observational study of 116 adolescents which provided very low certainty non-comparative evidence on change in serum creatinine between starting GnRH analogues and 1 year, and liver function during treatment. Statistical analyses were reported for changes in serum creatinine but not for liver function. Because there was no comparator group and participants acted as their own controls, it is not known whether the findings are associated with GnRH analogues or due to changes over time, or concomitant treatments.

The eighth study (Staphorsius et al. 2015) was a cross-sectional study of 85 adolescents, 40 with gender dysphoria (of whom 20 were receiving GnRH analogues) and 45 matched controls (not further reported in this evidence review). The study includes 1 outcome of interest for clinical effectiveness (CBCL) and 1 outcome of interest for safety (cognitive development or functioning). The mean (±SD) CBCL, IQ test, reaction time and accuracy scores were given for each group, but the statistical analysis is unclear. It is not reported what analysis was used or which of the groups were compared, therefore it is difficult to interpret the results.

The ninth study (Vlot et al. 2017) was a retrospective observational study which assessed bone mineral apparent density in adolescents with DSM-IV-TR gender dysphoria. Measurements were taken at the start of GnRH analogues and at the start of gender-affirming hormones. Results are reported for young bone age and old bone age in transmales and transfemales separately, and no results for the whole cohort are given. Statistical analyses were reported for all outcomes of interest but, because there was no comparator group and participants acted as their own controls, it is not known whether the findings are associated with GnRH analogues or due to changes over time. The authors reported z-scores which allows for comparison with the expected increase in bone density in the general population. However, because no concomitant treatments or comorbidities were reported it is possible that the findings may not be because of GnRH analogues and there is another way in which the study population differs from the general population.

7 Conclusion

The results of the studies that reported impact on the critical outcomes of gender dysphoria and mental health (depression, anger and anxiety), and the important outcomes of body image and psychosocial impact (global and psychosocial functioning) in children and adolescents with gender dysphoria are of very low certainty using modified GRADE. They suggest little change with GnRH analogues from baseline to follow-up.

Studies that found differences in outcomes could represent changes that are either of questionable clinical value, or the studies themselves are not reliable and changes could be due to confounding, bias or chance. It is plausible, however, that a lack of difference in scores from baseline to follow-up is the effect of GnRH analogues in children and adolescents with gender dysphoria, in whom the development of secondary sexual characteristics might be expected to be associated with an increased impact on gender dysphoria, depression, anxiety, anger and distress over time without treatment. One study reported statistically significant reductions in the Child Behaviour Checklist/Youth Self-Report (CBCL/YSR) scores from

baseline to follow up, and given that the purpose of GnRH analogues is to reduce distress caused by the development of secondary sexual characteristics and the CBCL/YSR in part measures distress, this could be an important finding. However, as the studies all lack reasonable controls not receiving GnRH analogues, the natural history of the outcomes measured in the studies is not known and any positive changes could be a regression to mean.

The results of the studies that reported bone density outcomes suggest that GnRH analogues may reduce the increase in bone density which is expected during puberty. However, as the studies themselves are not reliable, the results could be due to confounding, bias or chance. While controlled trials may not be possible, comparative studies are needed to understand this association and whether the effects of GnRH analogues on bone density are seen after treatment is stopped. All the studies that reported safety outcomes provided very low certainty evidence.

No cost-effectiveness evidence was found to determine whether or not GnRH analogues are cost-effective for children and adolescents with gender dysphoria.

The results of the studies that reported outcomes for subgroups of children and adolescents with gender dysphoria, suggest there may be differences between sex assigned at birth males (transfemales) and sex assigned at birth females (transmales).

Appendix A PICO document

The review questions for this evidence review are:

1. For children and adolescents with gender dysphoria, what is the clinical effectiveness of treatment with GnRH analogues compared with one or a combination of psychological support, social transitioning to the desired gender or no intervention?
2. For children and adolescents with gender dysphoria, what is the short-term and long-term safety of GnRH analogues compared with one or a combination of psychological support, social transitioning to the desired gender or no intervention?
3. For children and adolescents with gender dysphoria, what is the cost-effectiveness of GnRH analogues compared to one or a combination of psychological support, social transitioning to the desired gender or no intervention?
4. From the evidence selected, are there any subgroups of children and adolescents with gender dysphoria that may derive more (or less) advantage from treatment with GnRH analogues than the wider population of children and adolescents with gender dysphoria?
5. From the evidence selected,
   a) what are the criteria used by the research studies to define gender dysphoria, gender identity disorder and gender incongruence of childhood?
   b) what were the ages at which participants commenced treatment with GnRH analogues?
   c) what was the duration of treatment with GnRH analogues?

PICO table

P – Population and Indication	Children and adolescents aged 18 years or less who have gender dysphoria, gender identity disorder or gender incongruence of childhood as defined by study The following subgroups of children and adolescents with gender dysphoria, gender identity disorder or gender incongruence of childhood need to be considered: • Sex assigned at birth males. • Sex assigned at birth females. • The duration of gender dysphoria: less than 6 months, 6-24 months, and more than 24 months. • The age of onset of gender dysphoria. • The age at which treatment was initiated. • The age of onset of puberty. • Tanner stage at which treatment was initiated. • Children and adolescents with gender dysphoria who have a pre-existing diagnosis of autistic spectrum disorder. • Children and adolescents with gender dysphoria who had a significant mental health symptom load at diagnosis including anxiety, depression (with or without a history of self-harm and suicidality), suicide attempts, psychosis, personality disorder, Attention Deficit Hyperactivity Disorder and eating disorders.
I – Intervention	Any GnRH analogue including: triptorelin*; buserelin, histrelin; goserelin (Zoladex), leuprorelin/leuprolide (Prostap); nafarelin

	* Triptorelin (brand names Gonapeptyl and Decapeptyl) are used in Leeds Hospital, England. The search should include brand names as well as generic names.
C – Comparator(s)	One or a combination of: • Psychological support. • Social transitioning to the gender with which the individual identifies. • No intervention.
O – Outcomes	There are no known minimal clinically important differences and there are no preferred timepoints for the outcome measures selected. **All outcomes should be stratified by:** • The age at which treatment with GnRH analogues was initiated • The length of treatment with GnRH analogues where possible. **A: Clinical Effectiveness** *Critical to decision making* • **Impact on Gender Dysphoria** This outcome is critical because gender dysphoria in adolescents and children is associated with significant distress and problems functioning. Impact on gender dysphoria may be measured by the Utrecht Gender Dysphoria Scale. Other measures as reported in studies may be used as an alternative to the stated measure. • **Impact on mental health** Examples of mental health problems include self-harm, thoughts of suicide, suicide attempts, eating disorders, depression/low mood and anxiety. These outcomes are critical because self-harm and thoughts of suicide have the potential to result in significant physical harm and for completed suicides the death of the young person. Disordered eating habits may cause significant morbidity in young people. Depression and anxiety are also critical outcomes because they may impact on social, occupational, or other areas of functioning of children and adolescents. The Child and Adolescent Psychiatric Assessment (CAPA) may be used to measure depression and anxiety. The impact on self-harm and suicidality (ideation and behaviour) may be measured using the Suicide Ideation Questionnaire Junior. Other measures may be used as an alternative to the stated measures. • **Impact on Quality of Life** This outcome is critical because gender dysphoria in children and adolescents may be associated with a significant reduction in health-related quality of life. Quality of Life may be measured by the KINDL questionnaire, Kidscreen 52. Other measures as reported in studies may be used as an alternative to the stated measure. *Important to decision making* • **Impact on body image** This outcome is important because some transgender young people may desire to take steps to suppress features of their physical appearance associated with their sex assigned at birth or accentuate physical features of their desired gender. The Body Image Scale could be used as a measure. Other measures

	as reported in studies may also be used as an alternative to the stated measure. • **Psychosocial Impact** Examples of psychosocial impact are: coping mechanisms which may impact on substance misuse; family relationships; peer relationships. This outcome is important because gender dysphoria in adolescents and children is associated with internalising and externalising behaviours and emotional and behavioural problems which may impact on social and occupational functioning. The child behavioural check list (CBCL) may be used to measure the impact on psychosocial functioning. Other measures as reported in studies may be used as an alternative to the stated measure. • **Engagement with health care services** This outcome is important because patient engagement with healthcare services will impact on their clinical outcomes. Engagement with health care services may be measured using the Youth Health Care measure-satisfaction, utilization, and needs (YHC-SUN) questionnaire. Loss to follow up should also be ascertained as part of this outcome. Alternative measures to the YHC-SUN questionnaire may be used as reported in studies. • **Transitioning surgery – Impact on extent of and satisfaction with surgery** This outcome is important because some children and adolescents with gender dysphoria may proceed to transitioning surgery. Stated measures of the extent of transitioning surgery and satisfaction with surgery in studies may be reported. • **Stopping treatment** The proportion of patients who stop treatment with GnRH analogues and the reasons why. This outcome is important to patients because there is uncertainty about the short- and long-term safety and adverse effects of GnRH analogues in children and adolescents being treated for gender dysphoria. **B: Safety** • Short and long-term safety and adverse effects of taking GnRH analogues are important because GnRH analogues are not licensed for the treatment of adolescents and children with gender dysphoria. Aspects to be reported on should include: ○ Impact of the drug use such as its impact on bone density, arterial hypertension, cognitive development/functioning ○ Impact of withdrawing the drug such as, slipped upper femoral epiphysis, reversibility on the reproductive system, and any others as reported. **C: Cost effectiveness** Cost effectiveness studies should be reported
Inclusion criteria	
Study design	Systematic reviews, randomised controlled trials, controlled clinical trials, cohort studies If no higher level quality evidence is found, case series can be considered.

Language	English only
Patients	Human studies only
Age	18 years or less
Date limits	2000-2020
Exclusion criteria	
Publication type	Conference abstracts, non-systematic reviews, narrative reviews, commentaries, letters, editorials, guidelines and pre-publication prints
Study design	Case reports, resource utilisation studies

Appendix B Search strategy

Medline, Embase, the Cochrane Library, HTA and APA PsycInfo were searched on 23 July 2020, limiting the search to papers published in English language in the last 20 years. Conference abstracts and letters were excluded.

Database: Medline
Platform: Ovid
Version: Ovid MEDLINE(R) <1946 to July 21, 2020>
Search date: 23/7/2020
Number of results retrieved: 144
Search strategy:

1 Gender Dysphoria/ (485)
2 Gender Identity/ (18452)
3 "Sexual and Gender Disorders"/ (75)
4 Transsexualism/ (3758)
5 Transgender Persons/ (3143)
6 Health Services for Transgender Persons/ (136)
7 exp Sex Reassignment Procedures/ (838)
8 (gender* adj3 (dysphori* or affirm* or incongruen* or identi* or disorder* or confus* or minorit* or queer*)).tw. (7435)
9 (transgend* or transex* or transsex* or transfem* or transwom* or transma* or transmen* or transperson* or transpeopl*).tw. (12678)
10 (trans or crossgender* or cross-gender* or crossex* or cross-sex* or genderqueer*).tw. (102343)
11 ((sex or gender*) adj3 (reassign* or chang* or transform* or transition*)).tw. (6974)
12 (male-to-female or m2f or female-to-male or f2m).tw. (114841)
13 or/1-12 (252702)
14 exp Infant/ or Infant Health/ or Infant Welfare/ (1137479)
15 (prematur* or pre-matur* or preterm* or pre-term* or infan* or newborn* or new-born* or perinat* or peri-nat* or neonat* or neo-nat* or baby* or babies or toddler*).ti,ab,in,jn. (852400)
16 exp Child/ or exp Child Behavior/ or Child Health/ or Child Welfare/ (1913257)

17 Minors/ (2574)
18 (child* or minor or minors or boy* or girl* or kid or kids or young*).ti,ab,in,jn. (2361686)
19 exp pediatrics/ (58118)
20 (pediatric* or paediatric* or peadiatric*).ti,ab,in,jn. (836269)
21 Adolescent/ or Adolescent Behavior/ or Adolescent Health/ (2024207)
22 Puberty/ (13278)
23 (adolescen* or pubescen* or prepubescen* or pre-pubescen* or pubert* or prepubert* or pre-pubert* or teen* or preteen* or pre-teen* or juvenil* or youth* or under*age*).ti,ab,in,jn. (424246)
24 Schools/ (38104)
25 Child Day Care Centers/ or exp Nurseries/ or Schools, Nursery/ (7199)
26 (pre-school* or preschool* or kindergar* or daycare or day-care or nurser* or school* or pupil* or student*).ti,ab,jn. (468992)
27 ({"eight" or "nine" or "ten" or "eleven" or "twelve" or "thirteen" or "fourteen" or "fifteen" or "sixteen" or "seventeen" or "eighteen" or "nineteen") adj2 (year or years or age or ages or aged)).ti,ab. (89353)
28 ({"8" or "9" or "10" or "11" or "12" or "13" or "14" or "15" or "16" or "17" or "18" or "19") adj2 (year or years or age or ages or aged)).ti,ab. (887838)
29 or/14-28 (5534171)
30 13 and 29 (79263)
31 (transchild* or transyouth* or transteen* or transadoles* or transgirl* or transboy*).tw. (7)
32 30 or 31 (79263)
33 Gonadotropin-Releasing Hormone/ (27588)
34 (pubert* adj3 block*).ti,ab. (78)
35 ((gonadotrophin or gonadotropin) and releasing).ti,ab. (17299)
36 (GnRH adj2 analog*).ti,ab. (2541)
37 GnRH*.ti,ab. (20991)
38 "GnRH agonist*".ti,ab. (4040)
39 Triptorelin Pamoate/ (1906)
40 triptorelin.ti,ab. (677)
41 arvekap.ti,ab. (1)
42 ("AY 25650" or AY25650).ti,ab. (1)
43 ("BIM 21003" or BIM21003).ti,ab. (0)
44 ("BN 52014" or BN52014).ti,ab. (0)
45 ("CL 118532" or CL118532).ti,ab. (0)
46 Debio.ti,ab. (83)
47 diphereline.ti,ab. (17)
48 moapar.ti,ab. (0)
49 pamorelin.ti,ab. (0)
50 trelstar.ti,ab. (3)
51 triptodur.ti,ab. (1)
52 ("WY 42422" or WY42422).ti,ab. (0)
53 ("WY 42462" or WY42462).ti,ab. (0)
54 gonapeptyl.ti,ab. (0)
55 decapeptyl.ti,ab. (210)
56 salvacyl.ti,ab. (0)
57 Buserelin/ (2119)
58 buserelin.ti,ab. (1304)

59 bigonist.ti,ab. (0)
60 ("hoe 766" or hoe-766 or hoe766).ti,ab. (69)
61 profact.ti,ab. (2)
62 receptal.ti,ab. (30)
63 suprecur.ti,ab. (4)
64 suprefact.ti,ab. (22)
65 tiloryth.ti,ab. (0)
66 histrelin.ti,ab. (55)
67 "LHRH-hydrogel implant".ti,ab. (1)
68 ("RL 0903" or RL0903).ti,ab. (1)
69 ("SPD 424" or SPD424).ti,ab. (1)
70 goserelin.ti,ab. (875)
71 Goserelin/ (1612)
72 ("ici 118630" or ici118630).ti,ab. (51)
73 ("ZD-9393" or ZD9393).ti,ab. (0)
74 zoladex.ti,ab. (379)
75 leuprorelin.ti,ab. (413)
76 carcinil.ti,ab. (0)
77 enanton*.ti,ab. (23)
78 ginecrin.ti,ab. (0)
79 leuplin.ti,ab. (13)
80 Leuprolide/ (2900)
81 leuprolide.ti,ab. (1743)
82 lucrin.ti,ab. (11)
83 lupron.ti,ab. (162)
84 provren.ti,ab. (0)
85 procrin.ti,ab. (3)
86 ("tap 144" or tap144).ti,ab. (40)
87 (a-43818 or a43818).ti,ab. (3)
88 Trenantone.ti,ab. (1)
89 staladex.ti,ab. (0)
90 prostap.ti,ab. (6)
91 Nafarelin/ (327)
92 nafarelin.ti,ab. (251)
93 ("76932-56-4" or "76932564").ti,ab. (0)
94 ("76932-60-0" or "76932600").ti,ab. (0)
95 ("86220-42-0" or "86220420").ti,ab. (0)
96 ("rs 94991 298" or rs94991298).ti,ab. (0)
97 synarel.ti,ab. (12)
98 deslorelin.ti,ab. (263)
99 gonadorelin.ti,ab. (201)
100 ("33515-09-2" or "33515092").ti,ab. (0)
101 ("51952-41-1" or "51952411").ti,ab. (0)
102 ("52699-48-6" or "52699486").ti,ab. (0)
103 cetrorelix.ti,ab. (463)
104 cetrotide.ti,ab. (41)
105 ("NS 75A" or NS75A).ti,ab. (0)
106 ("NS 75B" or NS75B).ti,ab. (0)

107 ("SB 075" or SB075).ti,ab. (0)
108 ("SB 75" or SB75).ti,ab. (63)
109 gonadoliberin.ti,ab. (143)
110 kryptocur.ti,ab. (6)
111 cetrorelix.ti,ab. (463)
112 cetrotide.ti,ab. (41)
113 antagon.ti,ab. (17)
114 ganirelix.ti,ab. (138)
115 ("ORG 37462" or ORG37462).ti,ab. (3)
116 orgalutran.ti,ab. (20)
117 ("RS 26306" or RS26306).ti,ab. (5)
118 ("AY 24031" or AY24031).ti,ab. (0)
119 factrel.ti,ab. (11)
120 fertagyl.ti,ab. (11)
121 lutrelef.ti,ab. (5)
122 lutrepulse.ti,ab. (3)
123 relefact.ti,ab. (10)
124 fertiral.ti,ab. (0)
125 (hoe471 or "hoe 471").ti,ab. (6)
126 relisorm.ti,ab. (4)
127 cystorelin.ti,ab. (18)
128 dirigestran.ti,ab. (5)
129 or/33-128 (42216)
130 32 and 129 (416)
131 limit 130 to english language (393)
132 limit 131 to (letter or historical article or comment or editorial or news or case reports) (36)
133 131 not 132 (357)
134 animals/ not humans/ (4686361)
135 133 not 134 (181)
136 limit 135 to yr="2000 -Current" (144)

Database: Medline in-process
Platform: Ovid
Version: Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations <1946 to July 21, 2020>
Search date 23/7/2020
Number of results retrieved
Search strategy: 42

1 Gender Dysphoria/ (0)
2 Gender Identity/ (0)
3 "Sexual and Gender Disorders"/ (0)
4 Transsexualism/ (0)
5 Transgender Persons/ (0)
6 Health Services for Transgender Persons/ (0)
7 exp Sex Reassignment Procedures/ (0)

8 (gender* adj3 (dysphori* or affirm* or incongruen* or identi* or disorder* or confus* or minorit* or queer*)).tw. (1645)
9 (transgend* or transex* or transsex* or transfem* or transwom* or transma* or transmen* or transperson* or transpeopl*).tw. (2333)
10 (trans or crossgender* or cross-gender* or crossex* or cross-sex* or genderqueer*).tw. (20884)
11 ((sex or gender*) adj3 (reassign* or chang* or transform* or transition*)).tw. (968)
12 (male-to-female or m2f or female-to-male or f2m).tw. (15513)
13 or/1-12 (39905)
14 exp Infant/ or Infant Health/ or Infant Welfare/ (0)
15 (prematur* or pre-matur* or preterm* or pre-term* or infan* or newborn* or new-born* or perinat* or peri-nat* or neonat* or neo-nat* or baby* or babies or toddler*).ti,ab,in,jn. (80723)
16 exp Child/ or exp Child Behavior/ or Child Health/ or Child Welfare/ (0)
17 Minors/ (0)
18 (child* or minor or minors or boy* or girl* or kid or kids or young*).ti,ab,in,jn. (321871)
19 exp pediatrics/ (0)
20 (pediatric* or paediatric* or peadiatric*).ti,ab,in,jn. (119783)
21 Adolescent/ or Adolescent Behavior/ or Adolescent Health/ (0)
22 Puberty/ (0)
23 (adolescen* or pubescen* or prepubescen* or pre-pubescen* or pubert* or prepubert* or pre-pubert* or teen* or preteen* or pre-teen* or juvenil* or youth* or under*age*).ti,ab,in,jn. (60264)
24 Schools/ (0)
25 Child Day Care Centers/ or exp Nurseries/ or Schools, Nursery/ (0)
26 (pre-school* or preschool* or kindergar* or daycare or day-care or nurser* or school* or pupil* or student*).ti,ab,jn. (69233)
27 (("eight" or "nine" or "ten" or "eleven" or "twelve" or "thirteen" or "fourteen" or "fifteen" or "sixteen" or "seventeen" or "eighteen" or "nineteen") adj2 (year or years or age or ages or aged)).ti,ab. (10319)
28 (("8" or "9" or "10" or "11" or "12" or "13" or "14" or "15" or "16" or "17" or "18" or "19") adj2 (year or years or age or ages or aged)).ti,ab. (112800)
29 or/14-28 (525529)
30 13 and 29 (9196)
31 (transchild* or transyouth* or transteen* or transadoles* or transgirl* or transboy*).tw. (3)
32 30 or 31 (9197)
33 Gonadotropin-Releasing Hormone/ (0)
34 (pubert* adj3 block*).ti,ab. (19)
35 ((gonadotrophin or gonadotropin) and releasing).ti,ab. (1425)
36 (GnRH adj2 analog*).ti,ab. (183)
37 GnRH*.ti,ab. (1695)
38 "GnRH agonist*".ti,ab. (379)
39 Triptorelin Pamoate/ (0)
40 triptorelin.ti,ab. (72)
41 arvekap.ti,ab. (0)
42 ("AY 25650" or AY25650).ti,ab. (0)
43 ("BIM 21003" or BIM21003).ti,ab. (0)
44 ("BN 52014" or BN52014).ti,ab. (0)
45 ("CL 118532" or CL118532).ti,ab. (0)

46 Debio.ti,ab. (11)
47 diphereline.ti,ab. (6)
48 moapar.ti,ab. (0)
49 pamorelin.ti,ab. (0)
50 trelstar.ti,ab. (0)
51 triptodur.ti,ab. (0)
52 ("WY 42422" or WY42422).ti,ab. (0)
53 ("WY 42462" or WY42462).ti,ab. (0)
54 gonapeptyl.ti,ab. (0)
55 decapeptyl.ti,ab. (8)
56 salvacyl.ti,ab. (0)
57 Buserelin/ (0)
58 buserelin.ti,ab. (59)
59 bigonist.ti,ab. (0)
60 ("hoe 766" or hoe-766 or hoe766).ti,ab. (3)
61 profact.ti,ab. (0)
62 receptal.ti,ab. (0)
63 suprecur.ti,ab. (1)
64 suprefact.ti,ab. (2)
65 tiloryth.ti,ab. (0)
66 histrelin.ti,ab. (9)
67 "LHRH-hydrogel implant".ti,ab. (0)
68 ("RL 0903" or RL0903).ti,ab. (0)
69 ("SPD 424" or SPD424).ti,ab. (0)
70 goserelin.ti,ab. (68)
71 Goserelin/ (0)
72 ("ici 118630" or ici118630).ti,ab. (0)
73 ("ZD-9393" or ZD9393).ti,ab. (0)
74 zoladex.ti,ab. (6)
75 leuprorelin.ti,ab. (47)
76 carcinil.ti,ab. (0)
77 enanton*.ti,ab. (1)
78 ginecrin.ti,ab. (0)
79 leuplin.ti,ab. (1)
80 Leuprolide/ (0)
81 leuprolide.ti,ab. (121)
82 lucrin.ti,ab. (4)
83 lupron.ti,ab. (10)
84 provren.ti,ab. (0)
85 procrin.ti,ab. (0)
86 ("tap 144" or tap144).ti,ab. (0)
87 (a-43818 or a43818).ti,ab. (0)
88 Trenantone.ti,ab. (1)
89 staladex.ti,ab. (0)
90 prostap.ti,ab. (0)
91 Nafarelin/ (0)
92 nafarelin.ti,ab. (5)
93 ("76932-56-4" or "76932564").ti,ab. (0)

94 ("76932-60-0" or "76932600").ti,ab. (0)
95 ("86220-42-0" or "86220420").ti,ab. (0)
96 ("rs 94991 298" or rs94991298).ti,ab. (0)
97 synarel.ti,ab. (0)
98 deslorelin.ti,ab. (14)
99 gonadorelin.ti,ab. (13)
100 ("33515-09-2" or "33515092").ti,ab. (0)
101 ("51952-41-1" or "51952411").ti,ab. (0)
102 ("52699-48-6" or "52699486").ti,ab. (0)
103 cetrorelix.ti,ab. (31)
104 cetrotide.ti,ab. (5)
105 ("NS 75A" or NS75A).ti,ab. (0)
106 ("NS 75B" or NS75B).ti,ab. (0)
107 ("SB 075" or SB075).ti,ab. (0)
108 ("SB 75" or SB75).ti,ab. (2)
109 gonadoliberin.ti,ab. (4)
110 kryptocur.ti,ab. (1)
111 cetrorelix.ti,ab. (31)
112 cetrotide.ti,ab. (5)
113 antagon.ti,ab. (0)
114 ganirelix.ti,ab. (8)
115 ("ORG 37462" or ORG37462).ti,ab. (0)
116 orgalutran.ti,ab. (3)
117 ("RS 26306" or RS26306).ti,ab. (0)
118 ("AY 24031" or AY24031).ti,ab. (0)
119 factrel.ti,ab. (2)
120 fertagyl.ti,ab. (1)
121 lutrelef.ti,ab. (0)
122 lutrepulse.ti,ab. (0)
123 relefact.ti,ab. (0)
124 fertiral.ti,ab. (0)
125 (hoe471 or "hoe 471").ti,ab. (0)
126 relisorm.ti,ab. (0)
127 cystorelin.ti,ab. (1)
128 dirigestran.ti,ab. (0)
129 or/33-128 (2332)
130 32 and 129 (45)
131 limit 130 to english language (45)
132 limit 131 to yr="2000 -Current" (42)

Database: Medline epubs ahead of print
Platform: Ovid
Version: Ovid MEDLINE(R) Epub Ahead of Print <July 21, 2020>
Search date: 23/7/2020
Number of results retrieved: 8
Search strategy:

1 Gender Dysphoria/ (0)

2 Gender Identity/ (0)
3 "Sexual and Gender Disorders"/ (0)
4 Transsexualism/ (0)
5 Transgender Persons/ (0)
6 Health Services for Transgender Persons/ (0)
7 exp Sex Reassignment Procedures/ (0)
8 (gender* adj3 (dysphori* or affirm* or incongruen* or identi* or disorder* or confus* or minorit* or queer*)).tw. (486)
9 (transgend* or transex* or transsex* or transfem* or transwom* or transma* or transmen* or transperson* or transpeopl*).tw. (640)
10 (trans or crossgender* or cross-gender* or crossex* or cross-sex* or genderqueer*).tw. (1505)
11 ((sex or gender*) adj3 (reassign* or chang* or transform* or transition*)).tw. (178)
12 (male-to-female or m2f or female-to-male or f2m).tw. (2480)
13 or/1-12 (4929)
14 exp Infant/ or Infant Health/ or Infant Welfare/ (0)
15 (prematur* or pre-matur* or preterm* or pre-term* or infan* or newborn* or new-born* or perinat* or peri-nat* or neonat* or neo-nat* or baby* or babies or toddler*).ti,ab,in,jn. (15496)
16 exp Child/ or exp Child Behavior/ or Child Health/ or Child Welfare/ (0)
17 Minors/ (0)
18 (child* or minor or minors or boy* or girl* or kid or kids or young*).ti,ab,in,jn. (53563)
19 exp pediatrics/ (0)
20 (pediatric* or paediatric* or peadiatric*).ti,ab,in,jn. (22796)
21 Adolescent/ or Adolescent Behavior/ or Adolescent Health/ (0)
22 Puberty/ (0)
23 (adolescen* or pubescen* or prepubescen* or pre-pubescen* or pubert* or prepubert* or pre-pubert* or teen* or preteen* or pre-teen* or juvenil* or youth* or under*age*).ti,ab,in,jn. (13087)
24 Schools/ (0)
25 Child Day Care Centers/ or exp Nurseries/ or Schools, Nursery/ (0)
26 (pre-school* or preschool* or kindergar* or daycare or day-care or nurser* or school* or pupil* or student*).ti,ab,jn. (12443)
27 (("eight" or "nine" or "ten" or "eleven" or "twelve" or "thirteen" or "fourteen" or "fifteen" or "sixteen" or "seventeen" or "eighteen" or "nineteen") adj2 (year or years or age or ages or aged)).ti,ab. (1416)
28 (("8" or "9" or "10" or "11" or "12" or "13" or "14" or "15" or "16" or "17" or "18" or "19") adj2 (year or years or age or ages or aged)).ti,ab. (20166)
29 or/14-28 (88366)
30 13 and 29 (1638)
31 (transchild* or transyouth* or transteen* or transadoles* or transgirl* or transboy*).tw. (1)
32 30 or 31 (1638)
33 Gonadotropin-Releasing Hormone/ (0)
34 (pubert* adj3 block*).ti,ab. (2)
35 ((gonadotrophin or gonadotropin) and releasing).ti,ab. (176)
36 (GnRH adj2 analog*).ti,ab. (30)
37 GnRH*.ti,ab. (223)
38 "GnRH agonist*".ti,ab. (49)
39 Triptorelin Pamoate/ (0)

57

40 triptorelin.ti,ab. (12)
41 arvekap.ti,ab. (0)
42 ("AY 25650" or AY25650).ti,ab. (0)
43 ("BIM 21003" or BIM21003).ti,ab. (0)
44 ("BN 52014" or BN52014).ti,ab. (0)
45 ("CL 118532" or CL118532).ti,ab. (0)
46 Debio.ti,ab. (2)
47 diphereline ti,ab. (1)
48 moapar.ti,ab. (0)
49 pamorelin.ti,ab. (0)
50 trelstar.ti,ab. (0)
51 triptodur.ti,ab. (0)
52 ("WY 42422" or WY42422).ti,ab. (0)
53 ("WY 42462" or WY42462).ti,ab. (0)
54 gonapeptyl.ti,ab. (0)
55 decapeptyl.ti,ab. (0)
56 salvacyl.ti,ab. (0)
57 Buserelin/ (0)
58 buserelin.ti,ab. (7)
59 bigonist.ti,ab. (0)
60 ("hoe 766" or hoe-766 or hoe766).ti,ab. (0)
61 profact.ti,ab. (0)
62 receptal.ti,ab. (0)
63 suprecur.ti,ab. (0)
64 suprefact.ti,ab. (1)
65 tiloryth.ti,ab. (0)
66 histrelin.ti,ab. (2)
67 "LHRH-hydrogel implant".ti,ab. (0)
68 ("RL 0903" or RL0903).ti,ab. (0)
69 ("SPD 424" or SPD424).ti,ab. (0)
70 goserelin.ti,ab. (11)
71 Goserelin/ (0)
72 ("ici 118630" or ici118630).ti,ab. (0)
73 ("ZD-9393" or ZD9393).ti,ab. (0)
74 zoladex.ti,ab. (1)
75 leuprorelin.ti,ab. (13)
76 carcinil.ti,ab. (0)
77 enanton*.ti,ab. (1)
78 ginecrin.ti,ab. (0)
79 leuplin.ti,ab. (0)
80 Leuprolide/ (0)
81 leuprolide.ti,ab. (22)
82 lucrin.ti,ab. (0)
83 lupron.ti,ab. (2)
84 provren.ti,ab. (0)
85 procrin.ti,ab. (0)
86 ("tap 144" or tap144).ti,ab. (1)
87 (a-43818 or a43818).ti,ab. (0)

58

88 Trenantone.ti,ab. (0)
89 staladex.ti,ab. (0)
90 prostap.ti,ab. (0)
91 Nafarelin/ (0)
92 nafarelin.ti,ab. (4)
93 ("76932-56-4" or "76932564").ti,ab. (0)
94 ("76932-60-0" or "76932600").ti,ab. (0)
95 ("86220-42-0" or "86220420").ti,ab. (0)
96 ("rs 94991 298" or rs94991298).ti,ab. (0)
97 synarel.ti,ab. (0)
98 deslorelin.ti,ab. (3)
99 gonadorelin.ti,ab. (3)
100 ("33515-09-2" or "33515092").ti,ab. (0)
101 ("51952-41-1" or "51952411").ti,ab. (0)
102 ("52699-48-6" or "52699486") ti,ab. (0)
103 cetrorelix.ti,ab. (6)
104 cetrotide.ti,ab. (2)
105 ("NS 75A" or NS75A).ti,ab. (0)
106 ("NS 75B" or NS75B).ti,ab. (0)
107 ("SB 075" or SB075) ti,ab. (0)
108 ("SB 75" or SB75).ti,ab. (0)
109 gonadoliberin.ti,ab. (0)
110 kryptocur.ti,ab. (0)
111 cetrorelix.ti,ab. (6)
112 cetrotide.ti,ab. (2)
113 antagon.ti,ab. (1)
114 ganirelix.ti,ab. (1)
115 ("ORG 37462" or ORG37462) ti,ab (0)
116 orgalutran.ti,ab. (0)
117 ("RS 26306" or RS26306).ti,ab. (0)
118 ("AY 24031" or AY24031).ti,ab. (0)
119 factrel.ti,ab. (0)
120 fertagyl.ti,ab. (0)
121 lutrelef.ti,ab (0)
122 lutrepulse.ti,ab. (0)
123 relefact.ti,ab. (0)
124 fertiral.ti,ab. (0)
125 (hoe471 or "hoe 471") ti,ab (0)
126 relisorm.ti,ab. (0)
127 cystorelin.ti,ab. (0)
128 dirigestran.ti,ab. (0)
129 or/33-128 (310)
130 32 and 129 (8)
131 limit 130 to english language (8)
132 limit 131 to yr="2000 -Current" (8)

Database: Medline daily update
Platform: Ovid

Version: Ovid MEDLINE(R) Daily Update <July 21, 2020>
Search date: 23/7/2020
Number of results retrieved: 1
Search strategy

1 Gender Dysphoria/ (4)
2 Gender Identity/ (38)
3 "Sexual and Gender Disorders"/ (0)
4 Transsexualism/ (2)
5 Transgender Persons/ (26)
6 Health Services for Transgender Persons/ (1)
7 exp Sex Reassignment Procedures/ (3)
8 (gender* adj3 (dysphori* or affirm* or incongruen* or identi* or disorder* or confus* or minorit* or queer*)).tw. (24)
9 (transgend* or transex* or transsex* or transfem* or transwom* or transma* or transmen* or transperson* or transpeopl*).tw. (39)
10 (trans or crossgender* or cross-gender* or crossex* or cross-sex* or genderqueer*).tw. (87)
11 ((sex or gender*) adj3 (reassign* or chang* or transform* or transition*)).tw. (15)
12 (male-to-female or m2f or female-to-male or f2m).tw. (181)
13 or/1-12 (358)
14 exp Infant/ or Infant Health/ or Infant Welfare/ (932)
15 (prematur* or pre-matur* or preterm* or pre-term* or infan* or newborn* or new-born* or perinat* or peri-nat* or neonat* or neo-nat* or baby* or babies or toddler*).ti,ab,in,jn. (981)
16 exp Child/ or exp Child Behavior/ or Child Health/ or Child Welfare/ (1756)
17 Minors/ (3)
18 (child* or minor or minors or boy* or girl* or kid or kids or young*).ti,ab,in,jn (3672)
19 exp pediatrics/ (75)
20 (pediatric* or paediatric* or peadiatric*).ti,ab,in,jn. (1658)
21 Adolescent/ or Adolescent Behavior/ or Adolescent Health/ (2006)
22 Puberty/ (8)
23 (adolescen* or pubescen* or prepubescen* or pre-pubescen* or pubert* or prepubert* or pre-pubert* or teen* or preteen* or pre-teen* or juvenil* or youth* or under*age*).ti,ab,in,jn (732)
24 Schools/ (56)
25 Child Day Care Centers/ or exp Nurseries/ or Schools, Nursery/ (5)
26 (pre-school* or preschool* or kindergar* or daycare or day-care or nurser* or school* or pupil* or student*).ti,ab,jn. (622)
27 (("eight" or "nine" or "ten" or "eleven" or "twelve" or "thirteen" or "fourteen" or "fifteen" or "sixteen" or "seventeen" or "eighteen" or "nineteen") adj2 (year or years or age or ages or aged)).ti,ab. (98)
28 (("8" or "9" or "10" or "11" or "12" or "13" or "14" or "15" or "16" or "17" or "18" or "19") adj2 (year or years or age or ages or aged)).ti,ab. (1301)
29 or/14-28 (6705)
30 13 and 29 (130)
31 (transchild* or transyouth* or transteen* or transadoles* or transgirl* or transboy*) tw. (0)
32 30 or 31 (130)
33 Gonadotropin-Releasing Hormone/ (11)

34 (pubert* adj3 block*) ti,ab. (0)
35 ((gonadotrophin or gonadotropin) and releasing) ti,ab. (10)
36 (GnRH adj2 analog*).ti,ab. (2)
37 GnRH*.ti,ab. (14)
38 "GnRH agonist*" ti,ab. (4)
39 Triptorelin Pamoate/ (1)
40 triptorelin.ti,ab. (1)
41 arvekap.ti,ab. (0)
42 ("AY 25650" or AY25650).ti,ab. (0)
43 ("BIM 21003" or BIM21003).ti,ab. (0)
44 ("BN 52014" or BN52014).ti,ab. (0)
45 ("CL 118532" or CL118532).ti,ab. (0)
46 Debio.ti,ab. (1)
47 diphereline.ti,ab. (0)
48 moapar.ti,ab. (0)
49 pamorelin.ti,ab. (0)
50 trelstar.ti,ab. (0)
51 triptodur.ti,ab. (0)
52 ("WY 42422" or WY42422).ti,ab. (0)
53 ("WY 42462" or WY42462).ti,ab. (0)
54 gonapeptyl.ti,ab. (0)
55 decapeptyl.ti,ab. (0)
56 salvacyl.ti,ab. (0)
57 Buserelin/ (0)
58 buserelin.ti,ab. (0)
59 bigonist.ti,ab. (0)
60 ("hoe 766" or hoe-766 or hoe766).ti,ab. (0)
61 profact.ti,ab. (0)
62 receptal.ti,ab. (0)
63 suprecur.ti,ab. (0)
64 suprefact.ti,ab. (0)
65 tiloryth.ti,ab. (0)
66 histrelin.ti,ab. (0)
67 "LHRH-hydrogel implant".ti,ab. (0)
68 ("RL 0903" or RL0903).ti,ab. (0)
69 ("SPD 424" or SPD424).ti,ab. (0)
70 goserelin.ti,ab. (1)
71 Goserelin/ (2)
72 ("ici 118630" or ici118630).ti,ab. (0)
73 ("ZD-9393" or ZD9393).ti,ab. (0)
74 zoladex.ti,ab. (0)
75 leuprorelin.ti,ab. (0)
76 carcinil.ti,ab. (0)
77 enanton*.ti,ab. (0)
78 ginecrin.ti,ab. (0)
79 leuplin.ti,ab. (0)
80 Leuprolide/ (0)
81 leuprolide.ti,ab. (0)

61

82 lucrin.ti,ab. (0)
83 lupron.ti,ab. (0)
84 provren.ti,ab. (0)
85 procrin.ti,ab. (0)
86 ("tap 144" or tap144).ti,ab. (0)
87 (a-43818 or a43818).ti,ab. (0)
88 Trenantone.ti,ab. (0)
89 staladex.ti,ab. (0)
90 prostap.ti,ab. (0)
91 Nafarelin/ (0)
92 nafarelin.ti,ab. (0)
93 ("76932-56-4" or "76932564").ti,ab. (0)
94 ("76932-60-0" or "76932600").ti,ab. (0)
95 ("86220-42-0" or "86220420").ti,ab. (0)
96 ("rs 94991 298" or rs94991298).ti,ab. (0)
97 synarel.ti,ab. (0)
98 deslorelin.ti,ab. (0)
99 gonadorelin.ti,ab. (0)
100 ("33515-09-2" or "33515092").ti,ab. (0)
101 ("51952-41-1" or "51952411").ti,ab. (0)
102 ("52699-48-6" or "52699486").ti,ab. (0)
103 cetrorelix.ti,ab. (0)
104 cetrotide.ti,ab. (0)
105 ("NS 75A" or NS75A).ti,ab. (0)
106 ("NS 75B" or NS75B).ti,ab. (0)
107 ("SB 075" or SB075).ti,ab. (0)
108 ("SB 75" or SB75).ti,ab. (0)
109 gonadoliberin.ti,ab. (0)
110 kryptocur.ti,ab. (0)
111 cetrorelix.ti,ab. (0)
112 cetrotide.ti,ab. (0)
113 antagon.ti,ab. (0)
114 ganirelix.ti,ab. (0)
115 ("ORG 37462" or ORG37462).ti,ab. (0)
116 orgalutran.ti,ab. (0)
117 ("RS 26306" or RS26306).ti,ab. (0)
118 ("AY 24031" or AY24031).ti,ab. (0)
119 factrel.ti,ab. (0)
120 fertagyl.ti,ab. (0)
121 lutrelef.ti,ab. (0)
122 lutrepulse.ti,ab. (0)
123 relefact.ti,ab. (0)
124 fertiral.ti,ab. (0)
125 (hoe471 or "hoe 471").ti,ab. (0)
126 relisorm.ti,ab. (0)
127 cystorelin.ti,ab. (0)
128 dirigestran.ti,ab. (0)
129 or/33-128 (23)

62

130 32 and 129 (1)
131 limit 130 to english language (1)
132 limit 131 to yr="2000 -Current" (1)

Database: Embase
Platform: Ovid
Version: Embase <1974 to 2020 July 22>
Search date: 23/7/2020
Number of results retrieved: 367
Search strategy:

1 exp Gender Dysphoria/ (5399)
2 Gender Identity/ (16820)
3 "Sexual and Gender Disorders"/ (24689)
4 Transsexualism/ (3869)
5 exp Transgender/ (6597)
6 Health Services for Transgender Persons/ (158848)
7 exp Sex Reassignment Procedures/ or sex transformation/ (3058)
8 (gender* adj3 (dysphori* or affirm* or incongru* or identi* or disorder* or confus* or minorit* or queer*)).tw. (13005)
9 (transgend* or transex* or transsex* or transfem* or transwom* or transma* or transmen* or transperson* or transpeopl*).tw. (22509)
10 (trans or crossgender* or cross-gender* or crossex* or cross-sex* or genderqueer*).tw. (154446)
11 ((sex or gender*) adj3 (reassign* or chang* or transform* or transition*)).tw. (10327)
12 (male-to-female or m2f or female-to-male or f2m).tw. (200166)
13 or/1-12 (582812)
14 exp juvenile/ or Child Behavior/ or Child Welfare/ or Child Health/ or infant welfare/ or "minor (person)"/ or elementary student/ (3437324)
15 (prematur* or pre-matur* or preterm* or pre-term* or infan* or newborn* or new-born* or perinat* or peri-nat* or neonat* or neo-nat* or baby* or babies or toddler*).ti,ab,in,jn. (1186161)
16 (child* or minor or minors or boy* or girl* or kid or kids or young*).ti,ab,in,jn. (3586795)
17 exp pediatrics/ (106214)
18 (pediatric* or paediatric* or peadiatric*).ti,ab,in,jn. (1491597)
19 exp adolescence/ or exp adolescent behavior/ or adolescent health/ or high school student/ or middle school student/ (105108)
20 (adolescen* or pubescen* or prepubescen* or pre-pubescen* or pubert* or prepubert* or pre-pubert* or teen* or preteen* or pre-teen* or juvenil* or youth* or under*age*).ti,ab,in,jn. (641660)
21 school/ or high school/ or kindergarten/ or middle school/ or primary school/ or nursery school/ or day care/ (103791)
22 (pre-school* or preschool* or kindergar* or daycare or day-care or nurser* or school* or pupil* or student*).ti,ab,jn. (687437)
23 (("eight" or "nine" or "ten" or "eleven" or "twelve" or "thirteen" or "fourteen" or "fifteen" or "sixteen" or "seventeen" or "eighteen" or "nineteen") adj2 (year or years or age or ages or aged)).ti,ab. (138908)
24 (("8" or "9" or "10" or "11" or "12" or "13" or "14" or "15" or "16" or "17" or "18" or "19") adj2 (year or years or age or ages or aged)).ti,ab. (1562903)

25 or/14-24 (7130881)
26 13 and 25 (182161)
27 (transchild* or transyouth* or transteen* or transadoles* or transgirl* or transboy*).tw. (17)
28 26 or 27 (182161)
29 gonadorelin/ (37580)
30 (pubert* adj3 block*).ti,ab. (142)
31 ((gonadotrophin or gonadotropin) and releasing).ti,ab. (21450)
32 (GnRH adj2 analog*).ti,ab. (4013)
33 GnRH*.ti,ab. (29862)
34 "GnRH agonist*".ti,ab. (6719)
35 exp gonadorelin agonist/ or gonadorelin derivative/ or gonadorelin acetate/ (23304)
36 Triptorelin/ (5427)
37 triptorelin.ti,ab. (1182)
38 arvekap.ti,ab. (3)
39 ("AY 25650" or AY25650).ti,ab. (1)
40 ("BIM 21003" or BIM21003).ti,ab. (0)
41 ("BN 52014" or BN52014).ti,ab. (0)
42 ("CL 118532" or CL118532).ti,ab. (0)
43 Debio.ti,ab. (185)
44 diphereline.ti,ab. (51)
45 moapar.ti,ab. (0)
46 pamorelin.ti,ab. (0)
47 trelstar.ti,ab. (5)
48 triptodur.ti,ab. (1)
49 ("WY 42422" or WY42422).ti,ab. (0)
50 ("WY 42462" or WY42462).ti,ab. (0)
51 gonapeptyl.ti,ab. (10)
52 decapeptyl.ti,ab. (307)
53 salvacyl.ti,ab. (1)
54 buserelin acetate/ or buserelin/ (5164)
55 buserelin.ti,ab. (1604)
56 bigonist.ti,ab. (1)
57 ("hoe 766" or hoe-766 or hoe766).ti,ab. (89)
58 profact.ti,ab. (4)
59 receptal.ti,ab. (37)
60 suprecur.ti,ab. (8)
61 suprefact.ti,ab. (30)
62 tiloryth.ti,ab. (0)
63 histrelin/ (446)
64 histrelin.ti,ab. (107)
65 "LHRH-hydrogel implant".ti,ab. (1)
66 ("RL 0903" or RL0903).ti,ab. (1)
67 ("SPD 424" or SPD424).ti,ab. (1)
68 goserelin.ti,ab. (1487)
69 Goserelin/ (7128)
70 ("ici 118630" or ici118630).ti,ab. (49)
71 ("ZD-9393" or ZD9393).ti,ab. (0)

72 zoladex ti,ab. (501)
73 leuprorelin/ (11312)
74 leuprorelin.ti,ab. (727)
75 carcinil.ti,ab. (0)
76 enanton*.ti,ab. (38)
77 ginecrin.ti,ab. (1)
78 leuplin.ti,ab. (26)
79 leuprolide.ti,ab. (2788)
80 lucrin.ti,ab. (47)
81 lupron.ti,ab. (361)
82 provren.ti,ab. (0)
83 procrin.ti,ab. (11)
84 ("tap 144" or tap144).ti,ab. (63)
85 (a-43818 or a43818).ti,ab. (3)
86 Trenantone.ti,ab. (7)
87 staladex.ti,ab. (0)
88 prostap.ti,ab. (11)
89 nafarelin acetate/ or nafarelin/ (1441)
90 nafarelin.ti,ab. (324)
91 ("76932-56-4" or "76932564").ti,ab. (0)
92 ("76932-60-0" or "76932600").ti,ab. (0)
93 ("86220-42-0" or "86220420").ti,ab. (0)
94 ("rs 94991 298" or rs94991298).ti,ab. (0)
95 synarel.ti,ab. (28)
96 deslorelin/ (452)
97 deslorelin.ti,ab. (324)
98 gonadorelin.ti,ab. (338)
99 ("33515-09-2" or "33515092").ti,ab. (0)
100 ("51952-41-1" or "51952411").ti,ab. (0)
101 ("52699-48-6" or "52699486").ti,ab. (0)
102 cetrorelix/ (2278)
103 cetrorelix.ti,ab. (717)
104 cetrotide.ti,ab. (113)
105 ("NS 75A" or NS75A).ti,ab. (0)
106 ("NS 75B" or NS75B).ti,ab. (0)
107 ("SB 075" or SB075).ti,ab. (1)
108 ("SB 75" or SB75).ti,ab. (76)
109 gonadoliberin.ti,ab. (152)
110 kryptocur.ti,ab. (6)
111 cetrorelix.ti,ab. (717)
112 cetrotide.ti,ab. (113)
113 antagon.ti,ab. (32)
114 ganirelix/ (1284)
115 ganirelix.ti,ab. (293)
116 ("ORG 37462" or ORG37462).ti,ab. (4)
117 orgalutran/ (1284)
118 orgalutran.ti,ab. (68)
119 ("RS 26306" or RS26306).ti,ab. (6)

120 ("AY 24031" or AY24031).ti,ab. (0)
121 factrel.ti,ab. (14)
122 fertagyl.ti,ab. (20)
123 lutrelef.ti,ab. (7)
124 lutrepulse.ti,ab. (6)
125 relefact.ti,ab. (10)
126 fertiral.ti,ab. (0)
127 (hoe471 or "hoe 471").ti,ab. (4)
128 relisorm.ti,ab. (6)
129 cystorelin.ti,ab. (26)
130 dirigestran.ti,ab. (5)
131 or/29-130 (80790)
132 28 and 131 (988)
133 limit 132 to english language (940)
134 133 not (letter or editorial).pt. (924)
135 134 not (conference abstract or conference paper or conference proceeding or "conference review").pt. (683)
136 nonhuman/ not (human/ and nonhuman/) (4649157)
137 135 not 136 (506)
138 limit 137 to yr="2000 -Current" (420)
139 elsevier.cr. (25912990)
140 138 and 139 (372)
141 remove duplicates from 140 (367)

Database: Cochrane Library – incorporating Cochrane Database of Systematic Reviews (CDSR); CENTRAL
Platform: Wiley
Version:
CDSR – Issue 7 of 12, July 2020
CENTRAL – Issue 7 of 12, July 2020
Search date: 23/7/2020
Number of results retrieved: CDSR – 1, CENTRAL - 8

#1 [mh ^"Gender Dysphoria"] 3
#2 [mh ^"gender identity"] 227
#3 [mh ^"sexual and gender disorders"] 2
#4 [mh ^transsexualism] 27
#5 [mh ^"transgender persons"] 36
#6 [mh ^"health services for transgender persons"] 0
#7 [mh "sex reassignment procedures"] 4
#8 (gender* NEAR/3 (dysphori* or affirm* or incongruen* or identi* or disorder* or confus* or minorit* or queer*)):ti,ab 308
#9 (transgend* or transex* or transsex* or transfem* or transwom* or transma* or transmen* or transperson* or transpeopl*):ti,ab 929
#10 (trans or crossgender* or cross-gender* or crossex* or cross-sex* or genderqueer*):ti,ab 3915
#11 ((sex or gender*) NEAR/3 (reassign* or chang* or transform* or transition*)):ti,ab 493
#12 (male-to-female or m2f or female-to-male or f2m).ti,ab 489

#13 {or #1-#12} 6142
#14 [mh infant] or [mh ^"infant health"] or [mh ^"infant welfare"] 27769
#15 (prematur* or pre-matur* or preterm* or pre-term* or infan* or newborn* or new-born* or perinat* or peri-nat* or neonat* or neo-nat* or baby* or babies or toddler*):ti,ab 69476
#16 [mh child] or [mh "child behavior"] or [mh ^"child health"] or [mh ^"child welfare"] 42703
#17 [mh ^minors] 8
#18 (child* or minor or minors or boy* or girl* or kid or kids or young*):ti,ab 175826
#19 [mh pediatrics]661
#20 (pediatric* or paediatric* or peadiatric*):ti,ab 30663
#21 [mh ^adolescent] or [mh ^"adolescent behavior"] or [mh ^"adolescent health"] 102154
#22 [mh ^puberty] 295
#23 (adolescen* or pubescen* or prepubescen* or pre-pubescen* or pubert* or prepubert* or pre-pubert* or teen* or preteen* or pre-teen* or juvenil* or youth* or under*age*):ti,ab 34139
#24 [mh ^schools] 1914
#25 [mh ^"Child Day Care Centers"] or [mh nurseries] or [mh ^"schools, nursery"] 277
#26 (pre-school* or preschool* or kindergar* or daycare or day-care or nurser* or school* or pupil* or student*):ti,ab 54723
#27 (("eight" or "nine" or "ten" or "eleven" or "twelve" or "thirteen" or "fourteen" or "fifteen" or "sixteen" or "seventeen" or "eighteen" or "nineteen") NEAR/2 (year or years or age or ages or aged)):ti,ab 6710
#28 (("8" or "9" or "10" or "11" or "12" or "13" or "14" or "15" or "16" or "17" or "18" or "19") NEAR/2 (year or years or age or ages or aged)):ti,ab 196881
#29 {or #14-#28} 469351
#30 #13 and #29 2146
#31 (transchild* or transyouth* or transteen* or transadoles* or transgirl* or transboy*):ti,ab 0
#32 #30 or #31 2146
#33 [mh ^"Gonadotropin-Releasing Hormone"] 1311
#34 (pubert* NEAR/3 block*):ti,ab 1
#35 ((gonadotrophin or gonadotropin) and releasing):ti,ab 2095
#36 (GnRH NEAR/2 analog*):ti,ab 493
#37 GnRH*:ti,ab 3764
#38 "GnRH agonist*":ti,ab 1399
#39 [mh ^"Triptorelin Pamoate"] 451
#40 triptorelin:ti,ab 451
#41 arvekap:ti,ab 4
#42 ("AY 25650" or AY25650):ti,ab 0
#43 ("BIM 21003" or BIM21003):ti,ab 0
#44 ("BN 52014" or BN52014):ti,ab 0
#45 ("CL 118532" or CL118532):ti,ab 0
#46 Debio:ti,ab 301
#47 diphereline:ti,ab 25
#48 moapar:ti,ab 0
#49 pamorelin:ti,ab 5
#50 trelstar:ti,ab 3

#51 triptodur:ti,ab 0
#52 ("WY 42422" or WY42422):ti,ab 0
#53 ("WY 42462" or WY42462):ti,ab 0
#54 gonapeptyl:ti,ab 11
#55 decapeptyl:ti,ab 135
#56 salvacyl:ti,ab 0
#57 [mh ^Buserelin] 290
#58 Buserelin:ti,ab 339
#59 bigonist:ti,ab 0
#60 ("hoe 766" or hoe-766 or hoe766):ti,ab 11
#61 profact:ti,ab 1
#62 receptal:ti,ab 4
#63 suprecur:ti,ab 0
#64 suprefact:ti,ab 28
#65 tiloryth:ti,ab 0
#66 histrelin:ti,ab 5
#67 "LHRH-hydrogel implant":ti,ab 0
#68 ("RL 0903" or RL0903):ti,ab 0
#69 ("SPD 424" or SPD424):ti,ab 0
#70 goserelin:ti,ab 761
#71 [mh ^goserelin] 568
#72 ("ici 118630" or ici118630):ti,ab 7
#73 ("ZD-9393" or ZD9393):ti,ab 1
#74 zoladex:ti,ab 318
#75 leuprorelin:ti,ab 248
#76 carcinil:ti,ab 0
#77 enanton*:ti,ab 21
#78 ginecrin:ti,ab 1
#79 leuplin:ti,ab 7
#80 [mh ^Leuprolide] 686
#81 leuprolide:ti,ab696
#82 lucrin:ti,ab 21
#83 lupron:ti,ab 77
#84 provren:ti,ab 0
#85 procrin:ti,ab 2
#86 ("tap 144" or tap144):ti,ab 24
#87 (a-43818 or a43818):ti,ab 0
#88 Trenantone:ti,ab 3
#89 staladex:ti,ab 0
#90 prostap:ti,ab 9
#91 [mh ^Nafarelin] 77
#92 nafarelin:ti,ab 114
#93 ("76932-56-4" or "76932564"):ti,ab 0
#94 ("76932-60-0" or "76932600"):ti,ab 2
#95 ("86220-42-0" or "86220420"):ti,ab 0
#96 ("rs 94991 298" or rs94991298):ti,ab 0
#97 synarel:ti,ab 10
#98 deslorelin:ti,ab16

#99 gonadorelin:ti,ab 11
#100 ("33515-09-2" or "33515092"):ti,ab 0
#101 ("51952-41-1" or "51952411"):ti,ab 0
#102 ("52699-48-6" or "52699486"):ti,ab 0
#103 cetrorelix:ti,ab 221
#104 cetrotide:ti,ab 111
#105 ("NS 75A" or NS75A):ti,ab 0
#106 ("NS 75B" or NS75B):ti,ab 0
#107 ("SB 075" or SB075):ti,ab 0
#108 ("SB 75" or SB75):ti,ab 10
#109 gonadoliberin:ti,ab 5
#110 kryptocur:ti,ab 0
#111 cetrorelix:ti,ab 221
#112 cetrotide:ti,ab 111
#113 antagon:ti,ab 12
#114 ganirelix:ti,ab 142
#115 ("ORG 37462" or ORG37462) ti,ab 4
#116 orgalutran:ti,ab 45
#117 ("RS 26306" or RS26306):ti,ab 0
#118 ("AY 24031" or AY24031):ti,ab 0
#119 factrel:ti,ab 1
#120 fertagyl:ti,ab 0
#121 lutrelef:ti,ab 0
#122 lutrepulse:ti,ab1
#123 relefact:ti,ab 1
#124 fertiral:ti,ab 0
#125 (hoe471 or "hoe 471") ti,ab 3
#126 relisorm:ti,ab 0
#127 cystorelin:ti,ab0
#128 dirigestran:ti,ab 0
#129 {or #33-#128} 6844
#130 #32 and #129 27
#131 #130 with Cochrane Library publication date Between Jan 2000 and Jul 2020, in Cochrane Reviews 1
#132 #130 27
#133 "conference":pt or (clinicaltrials or trialsearch) so 492465
#134 #132 not #133 9
#135 #134 with Publication Year from 2000 to 2020, in Trials 8

Database: HTA
Platform: CRD
Version: HTA
Search date: 23/7/2020
Number of results retrieved: 26
Search strategy:

1 MeSH DESCRIPTOR Gender Dysphoria EXPLODE ALL TREES 0
2 MeSH DESCRIPTOR Gender Identity EXPLODE ALL TREES 14

3 MeSH DESCRIPTOR Sexual and Gender Disorders EXPLODE ALL TREES 2

4 MeSH DESCRIPTOR Transsexualism EXPLODE ALL TREES 12
5 MeSH DESCRIPTOR Transgender Persons EXPLODE ALL TREES 3
6 MeSH DESCRIPTOR Health Services for Transgender Persons EXPLODE ALL TREES 0
7 MeSH DESCRIPTOR Sex Reassignment Procedures EXPLODE ALL TREES 1

8 ((gender* adj3 (dysphon* or affirm* or incongruen* or identi* or disorder* or confus* or minorit* or queer*))) 28
9 ((transgend* or transex* or transsex* or transfem* or transwom* or transma* or transmen* or transperson* or transpeopl*)) 76
10 ((trans or crossgender* or cross-gender* or crossex* or cross-sex* or genderqueer*)) 83
11 (((sex or gender*) adj3 (reassign* or chang* or transform* or transition*))) 24
12 (male-to-female or m2f or female-to-male or f2m) 86
13 ((transchild* or transyouth* or transteen* or transadoles* or transgirl* or transboy*)) 0
14 #1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 262
15 (#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13) IN HTA 30

*26 results are from 200 onwards. Downloaded as a set to sift for drug terms rather than continuing with search strategy.

Database: APA PsycInfo
Search date: July 2020 (Week 2)
Search Strategy:
---

1 Gender Dysphoria/ (936)
2 Gender Identity/ (8648)
3 Transsexualism/ (2825)
4 Transgender/ (5257)
5 exp Gender Reassignment/ (568)
6 (gender* adj3 (dysphori* or affirm* or incongruen* or identi* or disorder* or confus* or minorit* or queer*)).tw. (15471)
7 (transgend* or transex* or transsex* or transfem* or transwom* or transma* or transmen* or transperson* or transpeopl*).tw. (13028)
8 (trans or crossgender* or cross-gender* or crossex* or cross-sex* or genderqueer*).tw. (7679)
9 ((sex or gender*) adj3 (reassign* or chang* or transform* or transition*)).tw. (5796)
10 (male-to-female or m2f or female-to-male or f2m).tw. (63688)
11 or/1-10 (99560)
12 exp Infant Development/ (21841)
13 (prematur* or pre-matur* or preterm* or pre-term* or infan* or newborn* or new-born* or perinat* or peri-nat* or neonat* or neo-nat* or baby* or babies or toddler*).ti,ab,in,jn. (150219)

14 Child Characteristics/ or exp Child Behavior/ or Child Psychology/ or exp Child Welfare/ or Child Psychiatry/ (23423)
15 (child* or minor or minors or boy* or girl* or kid or kids or young*).ti,ab,in,jn. (984230)
16 (pediatric* or paediatric* or peadiatric*).ti,ab,in,jn. (78962)
17 Adolescent Psychiatry/ or Adolescent Behavior/ or Adolescent Development/ or Adolescent Psychology/ or Adolescent Characteristics/ or Adolescent Health/ (62142)
18 Puberty/ (2753)
19 (adolescen* or pubescen* or prepubescen* or pre-pubescen* or pubert* or prepubert* or pre-pubert* or teen* or preteen* or pre-teen* or juvenil* or youth* or under*age*).ti,ab,in,jn. (347604)
20 Schools/ or exp elementary school students/ or high school students/ or junior high school students/ or middle school students/ (113053)
21 Child Day Care/ or Nursery Schools/ (2836)
22 (pre-school* or preschool* or kindergar* or daycare or day-care or nurser* or school* or pupil* or student*).ti,ab,jn. (772814)
23 (("eight" or "nine" or "ten" or "eleven" or "twelve" or "thirteen" or "fourteen" or "fifteen" or "sixteen" or "seventeen" or "eighteen" or "nineteen") adj2 (year or years or age or ages or aged)).ti,ab. (21475)
24 (("8" or "9" or "10" or "11" or "12" or "13" or "14" or "15" or "16" or "17" or "18" or "19") adj2 (year or years or age or ages or aged)).ti,ab. (285697)
25 or/12-24 (1772959)
26 11 and 25 (49612)
27 (transchild* or transyouth* or transteen* or transadoles* or transgirl* or transboy*).tw. (14)
28 26 or 27 (49613)
29 exp Gonadotropic Hormones/ (4226)
30 (pubert* adj3 block*).ti,ab. (29)
31 ((gonadotrophin or gonadotropin) and releasing).ti,ab. (1060)
32 (GnRH adj2 analog*).ti,ab. (49)
33 GnRH*.ti,ab. (998)
34 "GnRH agonist*".ti,ab. (72)
35 triptorelin.ti,ab. (25)
36 arvekap.ti,ab. (0)
37 ("AY 25650" or AY25650).ti,ab. (0)
38 ("BIM 21003" or BIM21003).ti,ab. (0)
39 ("BN 52014" or BN52014).ti,ab. (0)
40 ("CL 118532" or CL118532).ti,ab. (0)
41 Debio.ti,ab. (7)
42 diphereline.ti,ab. (0)
43 moapar.ti,ab. (0)
44 pamorelin.ti,ab. (0)
45 trelstar.ti,ab. (0)
46 triptodur.ti,ab. (0)
47 ("WY 42422" or WY42422).ti,ab. (0)
48 ("WY 42462" or WY42462).ti,ab. (0)
49 gonapeptyl.ti,ab. (0)
50 decapeptyl.ti,ab. (3)
51 salvacyl.ti,ab. (1)

71

52 buserelin.ti,ab. (6)
53 bigonist.ti,ab. (0)
54 ("hoe 766" or hoe-766 or hoe766).ti,ab. (0)
55 profact.ti,ab. (0)
56 receptal.ti,ab. (0)
57 suprecur.ti,ab. (0)
58 suprefact.ti,ab. (0)
59 tiloryth.ti,ab. (0)
60 histrelin.ti,ab. (1)
61 "LHRH-hydrogel implant".ti,ab. (0)
62 ("RL 0903" or RL0903).ti,ab. (0)
63 ("SPD 424" or SPD424).ti,ab. (0)
64 goserelin.ti,ab. (30)
65 ("ici 118630" or ici118630).ti,ab. (0)
66 ("ZD-9393" or ZD9393).ti,ab. (0)
67 zoladex.ti,ab. (3)
68 leuprorelin.ti,ab. (12)
69 carcinil.ti,ab. (0)
70 enanton*.ti,ab. (1)
71 ginecrin.ti,ab. (0)
72 leuplin.ti,ab. (0)
73 leuprolide.ti,ab. (79)
74 lucrin.ti,ab. (1)
75 lupron.ti,ab. (18)
76 provren.ti,ab. (0)
77 procrin.ti,ab. (0)
78 ("tap 144" or tap144).ti,ab. (1)
79 (a-43818 or a43818).ti,ab. (0)
80 Trenantone.ti,ab. (0)
81 staladex.ti,ab. (0)
82 prostap.ti,ab. (0)
83 nafarelin.ti,ab. (1)
84 ("76932-56-4" or "76932564").ti,ab. (0)
85 ("76932-60-0" or "76932600").ti,ab. (0)
86 ("86220-42-0" or "86220420").ti,ab. (0)
87 ("rs 94991 298" or rs94991298).ti,ab. (0)
88 synarel.ti,ab. (0)
89 destorelin.ti,ab. (8)
90 gonadorelin.ti,ab. (3)
91 ("33515-09-2" or "33515092").ti,ab. (0)
92 ("51952-41-1" or "51952411").ti,ab. (0)
93 ("52699-48-6" or "52699486").ti,ab. (0)
94 cetrorelix.ti,ab. (9)
95 cetrotide.ti,ab. (0)
96 ("NS 75A" or NS75A).ti,ab. (0)
97 ("NS 75B" or NS75B).ti,ab. (0)
98 ("SB 075" or SB075).ti,ab. (0)
99 ("SB 75" or SB75).ti,ab. (1)

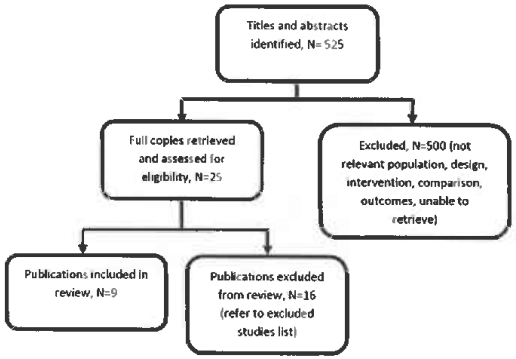
72

100 gonadoliberin.ti,ab. (1)
101 kryptocur.ti,ab. (0)
102 cetrorelix.ti,ab. (9)
103 cetrotide.ti,ab. (0)
104 antagon.ti,ab. (0)
105 ganirelix.ti,ab. (0)
106 ("ORG 37462" or ORG37462).ti,ab. (0)
107 orgalutran.ti,ab. (0)
108 ("RS 26306" or RS26306).ti,ab. (0)
109 ("AY 24031" or AY24031).ti,ab. (0)
110 factrel.ti,ab. (0)
111 fertagyl.ti,ab. (0)
112 lutrelef.ti,ab. (0)
113 lutrepulse.ti,ab. (0)
114 relefact.ti,ab. (0)
115 fertiral.ti,ab. (0)
116 (hoe471 or "hoe 471").ti,ab. (0)
117 relisorm.ti,ab. (0)
118 cystorelin.ti,ab. (0)
119 dirigestran.ti,ab. (0)
120 or/29-119 (4869)
121 28 and 120 (130)
122 limit 121 to english language (120)
123 limit 122 to yr="2000 -Current" (93)

Appendix C Evidence selection

The literature searches identified 525 references. These were screened using their titles and abstracts and 25 references were obtained and assessed for relevance. Of these, 9 references are included in the evidence review. The remaining 16 references were excluded and are listed in appendix D.

Figure 1 – Study selection flow diagram



References submitted with Preliminary Policy Proposal

There is no preliminary policy proposal for this policy.

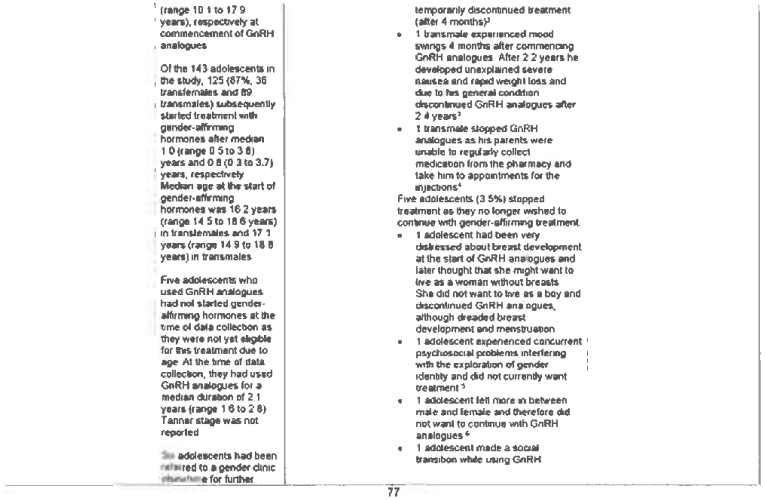
Appendix D Excluded studies table

Study reference	Reason for exclusion
Achille, C., Taggart, T., Eaton, N.R. et al. (2020) Longitudinal impact of gender-affirming endocrine intervention on the mental health and well-being of transgender youths: Preliminary results. International Journal of Pediatric Endocrinology 2020(1): 8	Intervention – data for GnRH analogues not reported separately from other interventions
Bechard, Melanie, Vanderlaan, Doug P, Wood, Hayley et al. (2017) Psychosocial and Psychological Vulnerability in Adolescents with Gender Dysphoria: A "Proof of Principle" Study. Journal of sex & marital therapy 43(7): 678-688	Population – no GnRH analogues at time of study
Chew, Denise, Anderson, Jemma, Williams, Katrina et al. (2018) Hormonal Treatment in Young People With Gender Dysphoria: A Systematic Review. Pediatrics 141(4)	All primary studies included apart from 1 conference abstract
de Vries, Annelou L C, McGuire, Jenifer K et al. (2014) Young adult psychological outcome after puberty suppression and gender reassignment. Pediatrics 134(4) 696-704	Population – relevant population included in de Vries et al. 2011
Ghelani, Rahul, Lim, Cheryl, Brain, Caroline et al. (2020) Sudden sex hormone withdrawal and the effects on body composition in late pubertal adolescents with gender dysphoria. Journal of pediatric endocrinology & metabolism: JPEM 33(1): 107-112	Outcomes – not in the PICO

Study reference	Reason for exclusion
Giovanardi, G, Morales, P, Mirabella, M et al. (2019) Transition memories: experiences of trans adult women with hormone therapy and their beliefs on the usage of hormone blockers to suppress puberty. Journal of endocrinological investigation 42(10): 1231-1240	Population – adults only
Hewitt, Jacqueline K, Paul, Campbell, Kasiannan, Porpavai et al. (2012) Hormone treatment of gender identity disorder in a cohort of children and adolescents. The Medical journal of Australia 196(9): 578-81	Outcomes – no data reported for relevant outcomes
Jensen, R.K., Jensen, J.K., Simons, L.K. et al. (2019) Effect of Concurrent Gonadotropin-Releasing Hormone Agonist Treatment on Dose and Side Effects of Gender-Affirming Hormone Therapy in Adolescent Transgender Patients. Transgender Health 4(1): 300-303	Outcomes – not in the PICO
Klaver, Maartje, de Mutsert, Renee, Wiepjes, Chantal M et al. (2018) Early Hormonal Treatment Affects Body Composition and Body Shape in Young Transgender Adolescents. The journal of sexual medicine 15(2): 251-260	Outcomes – not in the PICO
Klaver, Maartje, de Mutsert, Renee van der Loos, Maria A T C et al. (2020) Hormonal Treatment and Cardiovascular Risk Profile in Transgender Adolescents. Pediatrics 145(3)	Outcomes – not in the PICO
Lopez, Carla Marisa, Solomon, Daniel, Boulware, Susan D et al. (2018) Trends in the use of puberty blockers among transgender children in the United States. Journal of pediatric endocrinology & metabolism : JPEM 31(6): 665-670	Outcomes – not in the PICO
Schagen, Sebastian E E, Lustenhouwer, Paul, Cohen-Kettenis, Peggy T et al. (2018) Changes in Adrenal Androgens During Puberty Suppression and Gender-Affirming Hormone Treatment in Adolescents With Gender Dysphoria. The journal of sexual medicine 15(9): 1357-1363	Outcomes – not in the PICO
Swendiman, Robert A, Vogiatzi, Maria G, Alter, Craig A et al. (2019) Histrelin implantation in the pediatric population: A 10-year institutional experience. Journal of pediatric surgery 54(7): 1457-1461	Population – less than 10% of participants had gender dysphoria; data not reported separately
Turban, Jack L, King, Dana, Carswell, Jeremi M et al. (2020) Pubertal Suppression for Transgender Youth and Risk of Suicidal Ideation. Pediatrics 145(2)	Intervention – data for GnRH analogues not reported separately from other interventions
Vrouenraets, Lieke Josephina Jeanne Johanna, Fredriks, A Miranda, Hannema, Sabine E et al. (2016) Perceptions of Sex, Gender, and Puberty Suppression: A Qualitative Analysis of Transgender Youth. Archives of sexual behavior 45(7): 1697-703	Outcomes – not in the PICO
Zucker, Kenneth J, Bradley, Susan J, Owen-Anderson, Allison et al. (2010) Puberty-blocking hormonal therapy for adolescents with gender identity disorder: A descriptive clinical study. Journal of Gay & Lesbian Mental Health 15(1): 58-82	Intervention – data for GnRH analogues not reported separately from other interventions

Appendix E Evidence tables

Study details	Population	Interventions	Study outcomes	Appraisal and Funding
Ilse T Vrouenraets L, de Vries M, et al (2020) Trajectories of adolescents treated with gonadotropin-releasing hormone analogues for gender dysphoria. Archives of Sexual Behaviour https://doi.org/10.1007/s10508-020-01660-8 Netherlands Retrospective observational single-centre study To document trajectories after the initiation of GnRH analogue and explore reasons for extended use and discontinuation of GnRH analogues Includes participants seen between November 2010 and January 1, 2018	Inclusion criteria were adolescents with gender dysphoria, according to the DSM-5 criteria, seen at the single centre and treated with GnRH analogues between November 2010 and January 1, 2018 The study excluded adolescents without a diagnosis of gender dysphoria, those who had coexisting problems that interfered with the diagnostic process and/or might interfere with successful treatment (not further defined), those adolescents not wanting hormones, those with ongoing diagnostic evaluation and those who did not attend appointments The sample consisted of 143 adolescents meeting the inclusion/exclusion criteria, 38 transfemales 105 transmales, with median ages of 15 0 years (range 11 1 to 18 6 years) and 16 1 years	The study only reports that GnRH analogues were given, no specific drug, dose, route, or frequency of administration are reported No comparator cohort was used in the study Follow-up was at (up to) 9 years (last follow-up July 2018)	**Critical outcomes** No critical outcomes assessed **Important outcomes** ***Psychosocial impact*** Not assessed ***Engagement with health care services*** Not formally assessed but the study reported that out of 214 age and developmentally appropriate adolescents for potential inclusion in the study, 9 were excluded as they stopped attending appointments (4 2%) ***Stopping treatment*** Of the 143 adolescents, 9 (6 2% 1 transfemale and 8 transmales) stopped taking GnRH analogues after a median duration of 0 8 years (range 0 1 to 3 0) Four adolescents (2 8%) discontinued GnRH analogues although they wanted to continue endocrine treatments for gender dysphoria • 1 transmale stopped due to increase in mood problems, suicidal thoughts and confusion attributed to GnRH analogues (later had gender-affirming hormones at an adult gender clinic)[1] • 1 transmale experienced hot flushes increased migraines, had a fear of injections, stress at school and unrelated medical issues, and	This study was appraised using the Newcastle-Ottawa tool for cohort studies **Domain 1: Selection** 1 somewhat representative 2 no-non exposed cohort 3 secure record 4 yes **Domain 2: Comparability** 1 no comparator **Domain 3: Outcome** 1 record linkage 2 yes 3 complete follow-up **Overall quality is assessed as poor.** Other comments Physical and psychological comorbidity was poorly reported, concomitant use of other medicines was not reported Source of funding not reported



(range 10 1 to 17 9 years), respectively at commencement of GnRH analogues

Of the 143 adolescents in the study, 125 (87%, 36 transfemales and 89 transmales) subsequently started treatment with gender-affirming hormones after median 1 0 (range 0 5 to 3 8) years and 0 8 (0 3 to 3.7) years, respectively Median age at the start of gender-affirming hormones was 16 2 years (range 14 5 to 18 6 years) in transfemales and 17 1 years (range 14 9 to 18 8 years) in transmales

Five adolescents who used GnRH analogues had not started gender-affirming hormones at the time of data collection as they were not yet eligible for this treatment due to age At the time of data collection, they had used GnRH analogues for a median duration of 2 1 years (range 1 6 to 2 8) Tanner stage was not reported

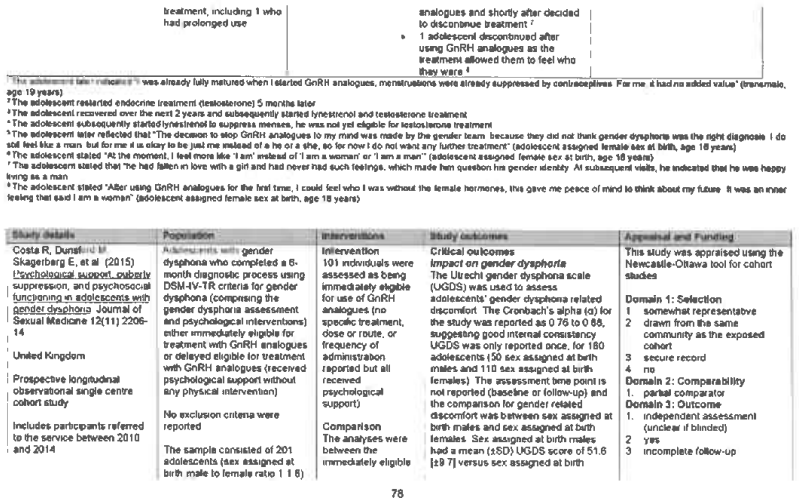
[redacted] adolescents had been referred to a gender clinic elsewhere for further

temporarily discontinued treatment (after 4 months)[2]

- 1 transmale experienced mood swings 4 months after commencing GnRH analogues After 2 2 years he developed unexplained severe nausea and rapid weight loss and due to his general condition discontinued GnRH analogues after 2 4 years[3]
- 1 transmale stopped GnRH analogues as his parents were unable to regularly collect medication from the pharmacy and take him to appointments for the injections[4]

Five adolescents (3 5%) stopped treatment as they no longer wished to continue with gender-affirming treatment.

- 1 adolescent had been very distressed about breast development at the start of GnRH analogues and later thought that she might want to live as a woman without breasts. She did not want to live as a boy and discontinued GnRH ana ogues, although dreaded breast development and menstruation
- 1 adolescent experienced concurrent psychosocial problems interfering with the exploration of gender identity and did not currently want treatment[5]
- 1 adolescent felt more in between male and female and therefore did not want to continue with GnRH analogues[6]
- 1 adolescent made a social transition while using GnRH



treatment, including 1 who had prolonged use

analogues and shortly after decided to discontinue treatment[7]

- 1 adolescent discontinued after using GnRH analogues as the treatment allowed them to feel who they were[8]

[redacted] was already fully matured when I started GnRH analogues, menstruations were already suppressed by contraceptives For me it had no added value" (transmale, age 19 years)

[2] The adolescent restarted endocrine treatment (testosterone) 5 months later

[3] The adolescent recovered over the next 2 years and subsequently started lynestrenol and testosterone treatment

[4] The adolescent subsequently started lynestrenol to suppress menses, he was not yet eligible for testosterone treatment

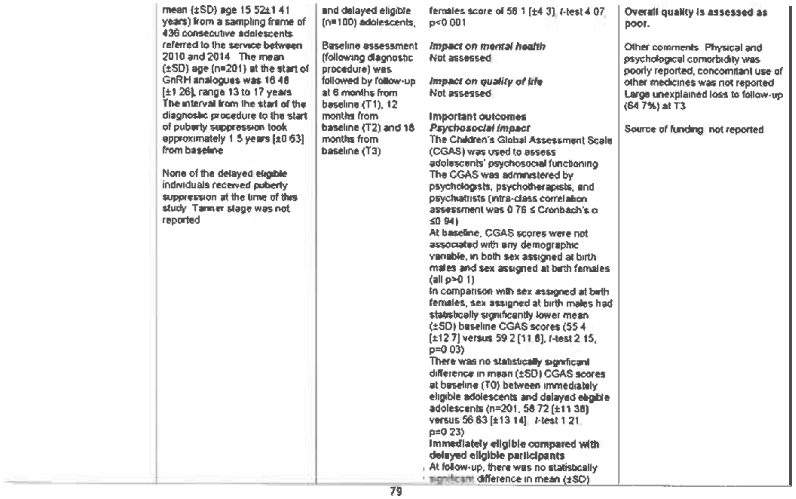
[5] The adolescent later reflected that "The decision to stop GnRH analogues to my mind was made by the gender team because they did not think gender dysphoria was the right diagnosis I do still feel like a man but for me it is okay to be just me instead of a he or a she, so for now I do not want any further treatment" (adolescent assigned female sex at birth, age 18 years)

[6] The adolescent stated "At the moment, I feel more like 'I am' instead of 'I am a woman' or 'I am a man'" (adolescent assigned female sex at birth, age 18 years)

[7] The adolescent stated that "he had fallen in love with a girl and had never had such feelings, which made him question his gender identity At subsequent visits, he indicated that he was happy living as a man

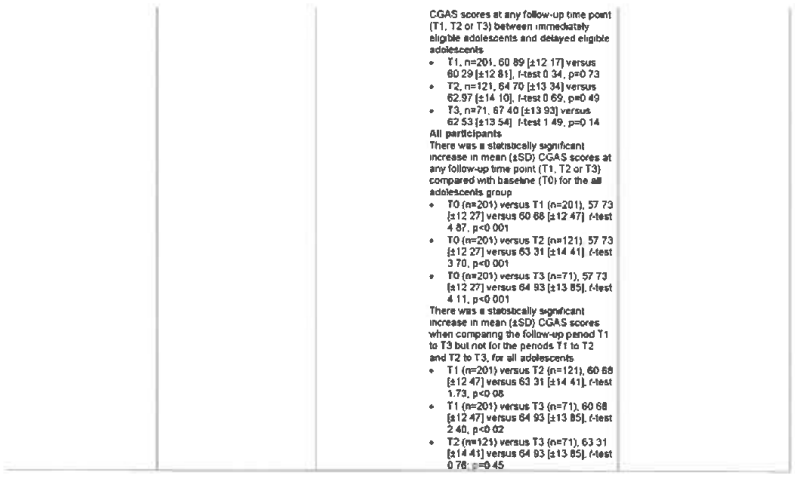
[8] The adolescent stated "After using GnRH analogues for the first time, I could feel who I was without the female hormones, this gave me peace of mind to think about my future It was an inner feeling that said I am a woman" (adolescent assigned female sex at birth, age 18 years)

Study details	Population	Interventions	Study outcomes	Appraisal and Funding
Costa R, Dunsford M, Skagerberg E, et al (2015) Psychological support, puberty suppression, and psychosocial functioning in adolescents with gender dysphoria Journal of Sexual Medicine 12(11) 2206-14 United Kingdom Prospective longitudinal observational single centre cohort study Includes participants referred to the service between 2010 and 2014	[redacted] gender dysphoria who completed a 6-month diagnostic process using DSM-IV-TR criteria for gender dysphoria (comprising the gender dysphoria assessment and psychological interventions) either immediately eligible for treatment with GnRH analogues or delayed eligible for treatment with GnRH analogues (received psychological support without any physical intervention) No exclusion criteria were reported The sample consisted of 201 adolescents (sex assigned at birth male to female ratio 1 1 6)	Intervention 101 individuals were assessed as being immediately eligible for use of GnRH analogues (no specific treatment, dose or route, or frequency of administration reported but all received psychological support) Comparison The analyses were between the immediately eligible	Critical outcomes *Impact on gender dysphoria* The Utrecht gender dysphoria scale (UGDS) was used to assess adolescents' gender dysphoria related discomfort The Cronbach's alpha (α) for the study was reported as 0 76 to 0 88, suggesting good internal consistency UGDS was only reported once, for 180 adolescents (50 sex assigned at birth males and 110 sex assigned at birth females). The assessment time point is not reported (baseline or follow-up) and the comparison for gender related discomfort was between sex assigned at birth males and sex assigned at birth females Sex assigned at birth males had a mean (±SD) UGDS score of 51.6 [±9 7] versus sex assigned at birth	This study was appraised using the Newcastle-Ottawa tool for cohort studies Domain 1: Selection 1 somewhat representative 2 drawn from the same community as the exposed cohort 3 secure record 4 no Domain 2: Comparability 1. partial comparator Domain 3: Outcome 1. independent assessment (unclear if blinded) 2 yes 3 incomplete follow-up



mean (±SD) age 15 52±1 41 years) from a sampling frame of 436 consecutive adolescents referred to the service between 2010 and 2014 The mean (±SD) age (n=201) at the start of GnRH analogues was 16 48 [±1 26], range 13 to 17 years The interval from the start of the diagnostic procedure to the start of puberty suppression took approximately 1 5 years [±0 63] from baseline None of the delayed eligible individuals received puberty suppression at the time of this study Tanner stage was not reported	and delayed eligible (n=100) adolescents, Baseline assessment (following diagnostic procedure) was followed by follow-up at 6 months from baseline (T1), 12 months from baseline (T2) and 18 months from baseline (T3)	females score of 58 1 [±4 3], *t*-test 4 07, p<0 001 ***Impact on mental health*** Not assessed ***Impact on quality of life*** Not assessed **Important outcomes** ***Psychosocial impact*** The Children's Global Assessment Scale (CGAS) was used to assess adolescents' psychosocial functioning The CGAS was administered by psychologists, psychotherapists, and psychiatrists (intra-class correlation assessment was 0 76 ≤ Cronbach's α ≤0 94) At baseline, CGAS scores were not associated with any demographic variable, in both sex assigned at birth males and sex assigned at birth females (all p>0 1) In comparison with sex assigned at birth females, sex assigned at birth males had statistically significantly lower mean (±SD) baseline CGAS scores (55 4 [±12 7] versus 59 2 [±11 8], *t*-test 2 15, p=0 03) There was no statistically significant difference in mean (±SD) CGAS scores at baseline (T0) between immediately eligible adolescents and delayed eligible adolescents (n=201, 58 72 [±13 38] versus 56 83 [±13 14] *t*-test 1 21, p=0 23) **Immediately eligible compared with delayed eligible participants** At follow-up, there was no statistically significant difference in mean (±SD)	**Overall quality is assessed as poor.** Other comments Physical and psychological comorbidity was poorly reported, concomitant use of other medicines was not reported Large unexplained loss to follow-up (64 7%) at T3 Source of funding not reported

79



CGAS scores at any follow-up time point (T1, T2 or T3) between immediately eligible adolescents and delayed eligible adolescents

- T1, n=201, 60 89 [±12 17] versus 60 29 [±12 81], *t*-test 0 34, p=0 73
- T2, n=121, 64 70 [±13 34] versus 62 97 [±14 10], *t*-test 0 69, p=0 49
- T3, n=71, 67 40 [±13 93] versus 62 53 [±13 54] *t*-test 1 49, p=0 14

All participants

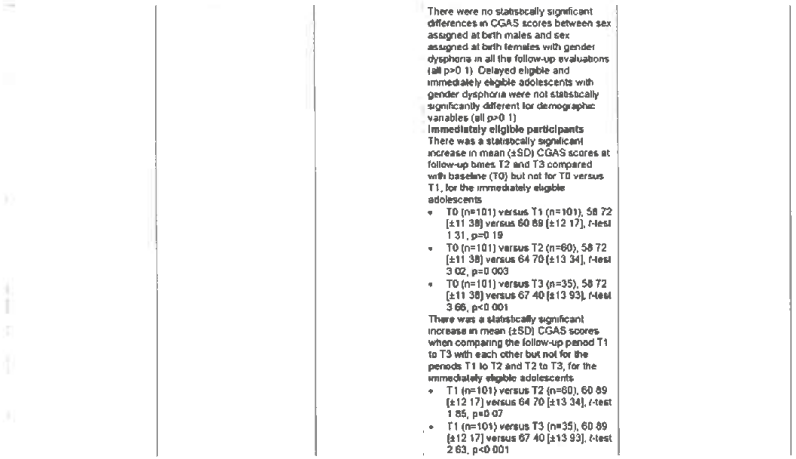
There was a statistically significant increase in mean (±SD) CGAS scores at any follow-up time point (T1, T2 or T3) compared with baseline (T0) for the all adolescents group

- T0 (n=201) versus T1 (n=201), 57 73 [±12 27] versus 60 68 [±12 47] *t*-test 4 87, p<0 001
- T0 (n=201) versus T2 (n=121), 57 73 [±12 27] versus 63 31 [±14 41] *t*-test 3 70, p<0 001
- T0 (n=201) versus T3 (n=71), 57 73 [±12 27] versus 64 93 [±13 85], *t*-test 4 11, p<0 001

There was a statistically significant increase in mean (±SD) CGAS scores when comparing the follow-up period T1 to T3 but not for the periods T1 to T2 and T2 to T3, for all adolescents

- T1 (n=201) versus T2 (n=121), 60 68 [±12 47] versus 63 31 [±14 41], *t*-test 1 73, p<0 05
- T1 (n=201) versus T3 (n=71), 60 68 [±12 47] versus 64 93 [±13 85], *t*-test 2 40, p<0 02
- T2 (n=121) versus T3 (n=71), 63 31 [±14 41] versus 64 93 [±13 85], *t*-test 0 76, p=0 45

80



There were no statistically significant differences in CGAS scores between sex assigned at birth males and sex assigned at birth females with gender dysphoria in all the follow-up evaluations (all p>0 1) Delayed eligible and immediately eligible adolescents with gender dysphoria were not statistically significantly different for demographic variables (all p>0 1)

Immediately eligible participants

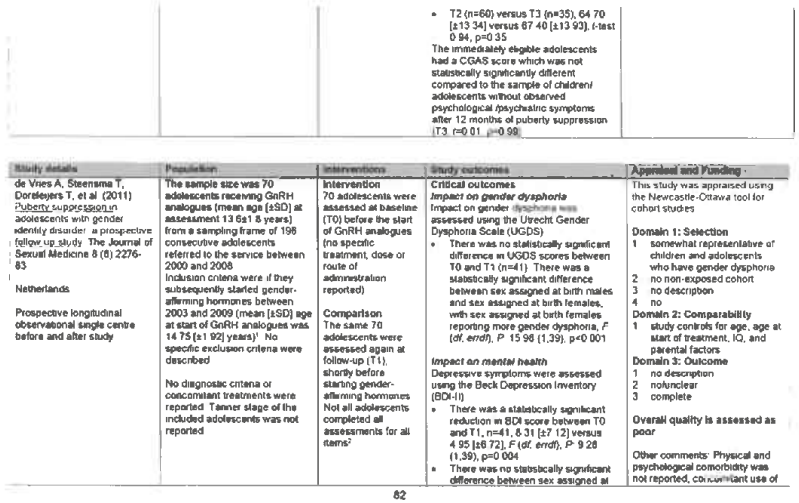
There was a statistically significant increase in mean (±SD) CGAS scores at follow-up times T2 and T3 compared with baseline (T0) but not for T0 versus T1, for the immediately eligible adolescents

- T0 (n=101) versus T1 (n=101), 58 72 [±11 38] versus 60 89 [±12 17], *t*-test 1 31, p=0 19
- T0 (n=101) versus T2 (n=60), 58 72 [±11 38] versus 64 70 [±13 34], *t*-test 3 02, p=0 003
- T0 (n=101) versus T3 (n=35), 58 72 [±11 38] versus 67 40 [±13 93], *t*-test 3 66, p<0 001

There was a statistically significant increase in mean (±SD) CGAS scores when comparing the follow-up period T1 to T3 with each other but not for the periods T1 to T2 and T2 to T3, for the immediately eligible adolescents

- T1 (n=101) versus T2 (n=60), 60 89 [±12 17] versus 64 70 [±13 34], *t*-test 1 85, p=0 07
- T1 (n=101) versus T3 (n=35), 60 89 [±12 17] versus 67 40 [±13 93], *t*-test 2 63, p<0 001

81



- T2 (n=60) versus T3 (n=35), 64 70 [±13 34] versus 67 40 [±13 93], *t*-test 0 94, p=0 35

The immediately eligible adolescents had a CGAS score which was not statistically significantly different compared to the sample of children/ adolescents without observed psychological /psychiatric symptoms after 12 months of puberty suppression (T3 t=0 01, p=0 99)

Study details	Population	Interventions	Study outcomes	Appraisal and Funding
de Vries A, Steensma T, Doreleyers T, et al (2011) Puberty suppression in adolescents with gender identity disorder a prospective follow up study The Journal of Sexual Medicine 8 (8) 2276-83 Netherlands Prospective longitudinal observational single centre before and after study	**The sample size was 70 adolescents receiving GnRH analogues (mean age [±SD] at assessment 13 6±1 8 years)** from a sampling frame of 198 consecutive adolescents referred to the service between 2000 and 2008 Inclusion criteria were if they subsequently started gender-affirming hormones between 2003 and 2009 (mean [±SD] age at start of GnRH analogues was 14 75 [±1 92] years)[1] No specific exclusion criteria were described No diagnostic criteria or concomitant treatments were reported Tanner stage of the included adolescents was not reported	**Intervention** 70 adolescents were assessed at baseline (T0) before the start of GnRH analogues (no specific treatment, dose or route of administration reported) **Comparison** The same 70 adolescents were assessed again at follow-up (T1), shortly before starting gender-affirming hormones Not all adolescents completed all assessments for all items[2]	**Critical outcomes** ***Impact on gender dysphoria*** Impact on gender dysphoria was assessed using the Utrecht Gender Dysphoria Scale (UGDS) • There was no statistically significant difference in UGDS scores between T0 and T1 (n=41) There was a statistically significant difference between sex assigned at birth males and sex assigned at birth females, with sex assigned at birth females reporting more gender dysphoria, *F* (*df, error*), *P* 15 98 (1,39), p<0 001 ***Impact on mental health*** Depressive symptoms were assessed using the Beck Depression Inventory (BDI-I) • There was a statistically significant reduction in BDI score between T0 and T1, n=41, 8 31 [±7 12] versus 4 95 [±6 72], *F* (*df, error*), *P* 9 28 (1,39), p=0 004 • There was no statistically significant difference between sex assigned at	This study was appraised using the Newcastle-Ottawa tool for cohort studies **Domain 1: Selection** 1 somewhat representative of children and adolescents who have gender dysphoria 2 no non-exposed cohort 3 no description 4 no **Domain 2: Comparability** 1 study controls for age, age at start of treatment, IQ, and parental factors **Domain 3: Outcome** 1 no description 2 no/unclear 3 complete **Overall quality is assessed as poor** Other comments: Physical and psychological comorbidity was not reported, concomitant use of

82

birth males and sex assigned at birth females, F (*df, error*), P 3.85 (1,39), p=0.057

Anger and anxiety were assessed using Trait Anger and Anxiety (TPI and STAI, respectively) Scales of the State-Trait Personality Inventory

- There was no statistically significant difference in anger (TPI) scale scores between T0 and T1 (n=41) There was a statistically significant difference between sex assigned at birth males and sex assigned at birth females, with sex assigned at birth females reporting increased anger compared with sex assigned at birth males, F (*df, error*), P 5.70 (1,39), p=0.022
- Similarly, there was no statistically significant difference in anxiety (STAI) scale scores between T0 and T1 (n=41) There was a statistically significant difference between sex assigned at birth males and sex assigned at birth females, with sex assigned at birth females reporting increased anxiety compared with sex assigned at birth males, F (*df, error*), P 16.07 (1,39), p<0.001

Impact on quality of life
Not assessed

Important outcomes
Impact on body image
Impact on body image was assessed using the Body Image Scale to measure body satisfaction (BIS)

other medicines was not reported

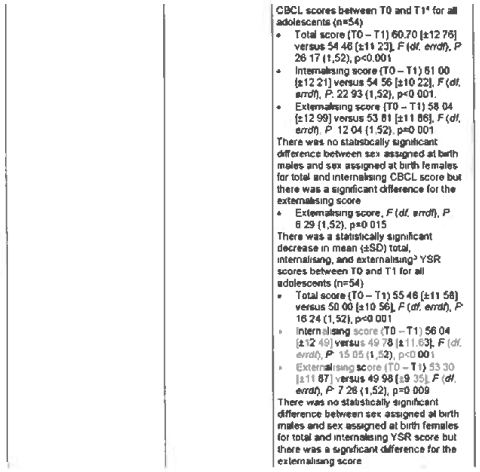
Source of funding This study was supported by a personal grant awarded to the first author by the Netherlands Organization for Health Research and Development

There was no statistically significant difference between T0 and T1 for any of the 3 BIS scores (primary sex characteristics, secondary sex characteristics or neutral characteristics n=57) There were statistically significant differences between sex assigned at birth males and sex assigned at birth females, with sex assigned at birth females reporting more dissatisfaction, for

- primary sexual characteristics, F (*df, error*), P 4.11 (1,55), p=0.047
- secondary sexual characteristics, F (*df, error*), P 11.57 (1,55), p=0.001

But no statistically significant difference between sex assigned at birth males and sex assigned at birth females was found for neutral characteristics However, there was a significant interaction effect between sex assigned at birth sex and the changes of gender dysphoria between T0 and T1, sex assigned at birth females became more dissatisfied with their secondary sex characteristics compared with sex assigned at birth males, F (*df, error*), P 14.59 (1,55), p<0.001) and neutral characteristics, F (*df error*), P 15.26 (1,55), p<0.001)

Psychosocial impact
Psychosocial impact was assessed using both the Child Behaviour Checklist (CBCL) and the Youth Self-Report (YSR) to parents and adolescents, respectively The Children's Global Assessment Scale was also reported
There was a statistically significant decrease in mean (±SD) total, internalising, and externalising[3] parental



CBCL scores between T0 and T1[4] for all adolescents (n=54)

- Total score (T0 – T1) 60.70 [±12.76] versus 54.46 [±11.23], *F (df, errdf), P* 26.17 (1,52), p<0.001
- Internalising score (T0 – T1) 61.00 [±12.21] versus 54.56 [±10.22], *F (df, errdf), P*. 22.93 (1,52), p<0.001.
- Externalising score (T0 – T1) 58.04 [±12.99] versus 53.81 [±11.86], *F (df, errdf), P*. 12.04 (1,52), p=0.001

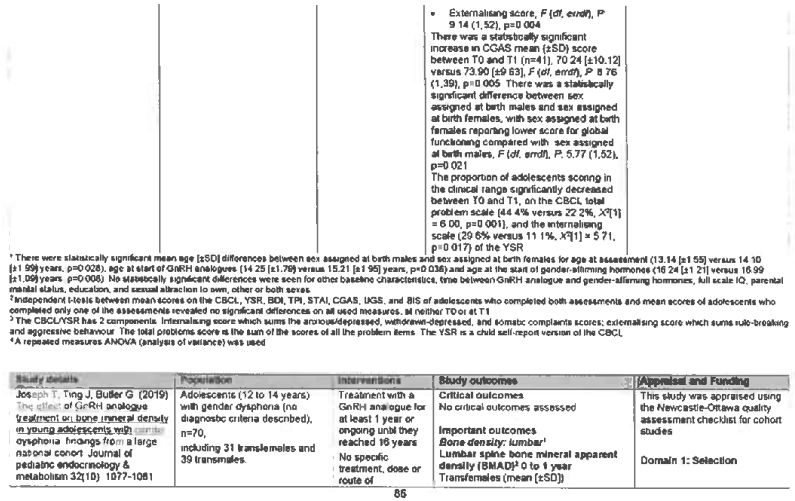
There was no statistically significant difference between sex assigned at birth males and sex assigned at birth females for total and internalising CBCL score but there was a significant difference for the externalising score

- Externalising score, *F (df, errdf), P* 6.29 (1,52), p=0.015

There was a statistically significant decrease in mean (±SD) total, internalising, and externalising[3] YSR scores between T0 and T1 for all adolescents (n=54)

- Total score (T0 – T1) 55.48 [±11.58] versus 50.00 [±10.56], *F (df, errdf), P*. 16.24 (1,52), p<0.001
- Internalising score (T0 – T1) 56.04 [±12.49] versus 49.78 [±11.63], *F (df, errdf), P*. 15.05 (1,52), p<0.001
- Externalising score (T0 – T1) 53.30 [±11.87] versus 49.98 [±9.35], *F (df, errdf), P*. 7.28 (1,52), p=0.009

There was no statistically significant difference between sex assigned at birth males and sex assigned at birth females for total and internalising YSR score but there was a significant difference for the externalising score



- Externalising score, *F (df, errdf), P* 9.14 (1,52), p=0.004

There was a statistically significant increase in CGAS mean (±SD) score between T0 and T1 (n=41), 70.24 [±10.12] versus 73.90 [±9.93], *F (df, errdf), P*. 8.76 (1,39), p=0.005. There was a statistically significant difference between sex assigned at birth males and sex assigned at birth females, with sex assigned at birth females reporting lower score for global functioning compared with sex assigned at birth males, *F (df, errdf), P*. 5.77 (1,52), p=0.021

The proportion of adolescents scoring in the clinical range significantly decreased between T0 and T1, on the CBCL total problem scale (44.4% versus 22.2%, $X^2[1] = 6.00$, p=0.001), and the internalising scale (29.6% versus 11.1%, $X^2[1] = 5.71$, p=0.017) of the YSR

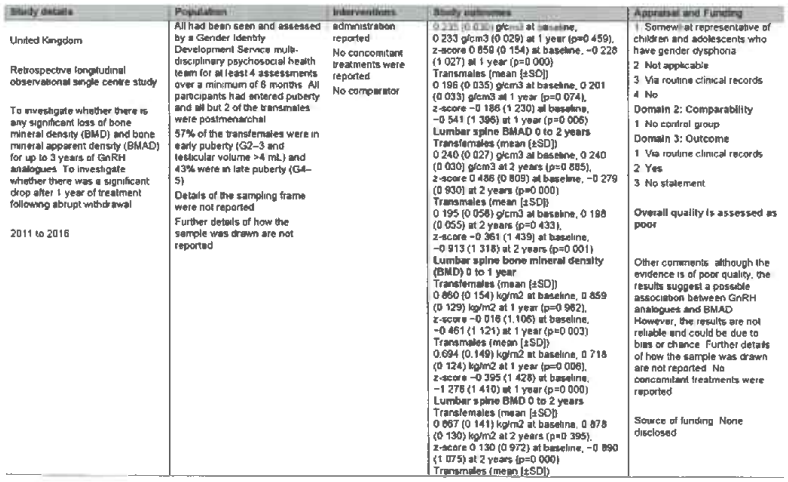
[1] There were statistically significant mean age [±SD] differences between sex assigned at birth males and sex assigned at birth females for age at assessment (13.14 [±1.55] versus 14.10 [±1.99] years, p=0.028), age at start of GnRH analogues (14.25 [±1.78] versus 15.21 [±1.95] years, p=0.036) and age at the start of gender-affirming hormones (16.24 [±1.21] versus 16.99 [±1.09] years p=0.008). No statistically significant differences were seen for other baseline characteristics, time between GnRH analogue and gender-affirming hormones, full scale IQ, parental marital status, education, and sexual attraction to own, other or both sexes.

[2] Independent t-tests between mean scores on the CBCL, YSR, BDI, TPI, STAI, CGAS, UGS, and BIS of adolescents who completed both assessments and mean scores of adolescents who completed only one of the assessments revealed no significant differences on all used measures, at neither T0 or at T1

[3] The CBCL/YSR has 2 components. Internalising score which sums the anxious/depressed, withdrawn/depressed, and somatic complaints scores; externalising score which sums rule-breaking and aggressive behaviour. The total problems score is the sum of the scores of all the problem items. The YSR is a child self-report version of the CBCL

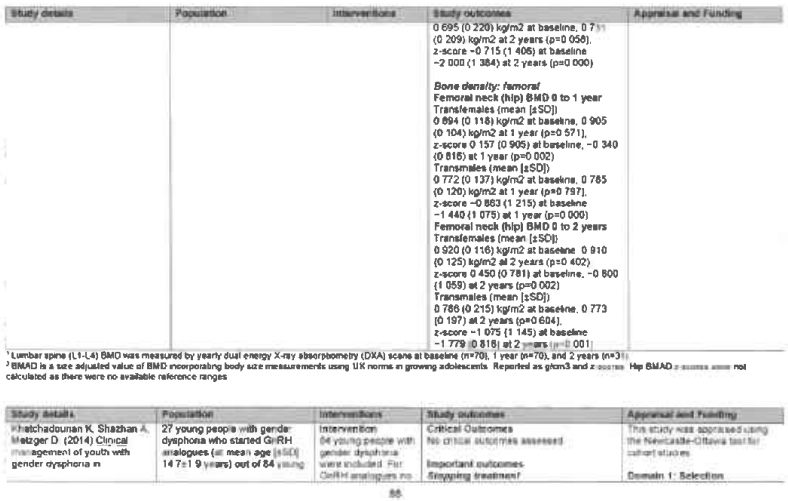
[4] A repeated measures ANOVA (analysis of variance) was used

Study details	Population	Interventions	Study outcomes	Appraisal and Funding
Joseph T, Ting J, Butler G (2019) The effect of GnRH analogue treatment on bone mineral density in young adolescents with gender dysphoria: findings from a large national cohort. Journal of pediatric endocrinology & metabolism 32(10): 1077-1081	Adolescents (12 to 14 years) with gender dysphoria (no diagnostic criteria described), n=70, including 31 transfemales and 39 transmales.	Treatment with a GnRH analogue for at least 1 year or ongoing until they reached 16 years. No specific treatment, dose or route of	Critical outcomes No critical outcomes assessed Important outcomes ***Bone density: lumbar***[1] Lumbar spine bone mineral apparent density (BMAD)[2] 0 to 1 year Transfemales (mean [±SD])	This study was appraised using the Newcastle-Ottawa quality assessment checklist for cohort studies Domain 1: Selection



Study details	Population	Interventions	Study outcomes	Appraisal and Funding
United Kingdom Retrospective longitudinal observational single centre study To investigate whether there is any significant loss of bone mineral density (BMD) and bone mineral apparent density (BMAD) for up to 3 years of GnRH analogues. To investigate whether there was a significant drop after 1 year of treatment following abrupt withdrawal 2011 to 2016	All had been seen and assessed by a Gender Identity Development Service multi-disciplinary psychosocial health team for at least 4 assessments over a minimum of 6 months. All participants had entered puberty and all but 2 of the transmales were postmenarchal 57% of the transfemales were in early puberty (G2–3 and testicular volume >4 mL) and 43% were in late puberty (G4–5) Details of the sampling frame were not reported Further details of how the sample was drawn are not reported	administration reported No concomitant treatments were reported No comparator	[illegible] g/cm3 at baseline, 0 233 g/cm3 (0 029) at 1 year (p=0 459), z-score 0 859 (0 154) at baseline, −0 228 (1 027) at 1 year (p=0 000) Transmales (mean [±SD]) 0 196 (0 035) g/cm3 at baseline, 0 201 (0 033) g/cm3 at 1 year (p=0 074), z-score −0 186 (1 230) at baseline, −0 541 (1 396) at 1 year (p=0 006) **Lumbar spine BMAD 0 to 2 years** **Transfemales (mean [±SD])** 0 240 (0 027) g/cm3 at baseline, 0 240 (0 030) g/cm3 at 2 years (p=0 885), z-score 0 486 (0 809) at baseline, −0 279 (0 930) at 2 years (p=0 000) **Transmales (mean [±SD])** 0 195 (0 058) g/cm3 at baseline, 0 198 (0 055) at 2 years (p=0 433), z-score −0 361 (1 439) at baseline, −0 913 (1 318) at 2 years (p=0 001) **Lumbar spine bone mineral density (BMD) 0 to 1 year** **Transfemales (mean [±SD])** 0 860 (0 154) kg/m2 at baseline, 0 859 (0 129) kg/m2 at 1 year (p=0 982), z-score −0 016 (1 106) at baseline, −0 461 (1 121) at 1 year (p=0 003) **Transmales (mean [±SD])** 0 694 (0 149) kg/m2 at baseline, 0 718 (0 124) kg/m2 at 1 year (p=0 006), z-score −0 395 (1 428) at baseline, −1 278 (1 410) at 1 year (p=0 000) **Lumbar spine BMD 0 to 2 years** **Transfemales (mean [±SD])** 0 967 (0 141) kg/m2 at baseline, 0 878 (0 130) kg/m2 at 2 years (p=0 395), z-score 0 130 (0 972) at baseline, −0 890 (1 075) at 2 years (p=0 000) **Transmales (mean [±SD])**	1 Somewhat representative of children and adolescents who have gender dysphoria 2 Not applicable 3 Via routine clinical records 4 No **Domain 2: Comparability** 1 No control group **Domain 3: Outcome** 1 Via routine clinical records 2 Yes 3 No statement **Overall quality is assessed as poor** Other comments: although the evidence is of poor quality, the results suggest a possible association between GnRH analogues and BMAD. However, the results are not reliable and could be due to bias or chance. Further details of how the sample was drawn are not reported. No concomitant treatments were reported Source of funding: None disclosed

87



Study details	Population	Interventions	Study outcomes	Appraisal and Funding
			0 695 (0 220) kg/m2 at baseline, 0 7[illegible] (0 209) kg/m2 at 2 years (p=0 056), z-score −0 715 (1 406) at baseline −2 000 (1 384) at 2 years (p=0 000) ***Bone density: femoral*** **Femoral neck (hip) BMD 0 to 1 year** **Transfemales (mean [±SD])** 0 894 (0 118) kg/m2 at baseline, 0 905 (0 104) kg/m2 at 1 year (p=0 571), z-score 0 157 (0 905) at baseline, −0 340 (0 816) at 1 year (p=0 002) **Transmales (mean [±SD])** 0 772 (0 137) kg/m2 at baseline, 0 785 (0 120) kg/m2 at 1 year (p=0 797), z-score −0 883 (1 215) at baseline −1 440 (1 075) at 1 year (p=0 000) **Femoral neck (hip) BMD 0 to 2 years** **Transfemales (mean [±SD])** 0 920 (0 116) kg/m2 at baseline 0 910 (0 125) kg/m2 at 2 years (p=0 402) z-score 0 450 (0 781) at baseline, −0 800 (1 059) at 2 years (p=0 002) **Transmales (mean [±SD])** 0 786 (0 215) kg/m2 at baseline, 0 773 (0 197) at 2 years (p=0 604), z-score −1 075 (1 145) at baseline −1 779 (0 818) at 2 years (p=0 001)	

[1] Lumbar spine (L1-L4) BMD was measured by yearly dual energy X-ray absorptiometry (DXA) scans at baseline (n=70), 1 year (n=70), and 2 years (n=3)[illegible]
[2] BMAD is a size adjusted value of BMD incorporating body size measurements using UK norms in growing adolescents. Reported as g/cm3 and z-scores. Hip BMAD z-scores were not calculated as there were no available reference ranges

Study details	Population	Interventions	Study outcomes	Appraisal and Funding
Khatchadourian K, Shazhan A, Metzger D (2014) Clinical management of youth with gender dysphoria in	27 young people with gender dysphoria who started GnRH analogues (at mean age [±SD] 14 7±1 9 years) out of 84 young	Intervention 84 young people with gender dysphoria were included. For GnRH analogues no	Critical Outcomes No critical outcomes assessed Important outcomes Stopping treatment	This study was appraised using the Newcastle-Ottawa tool for cohort studies Domain 1: Selection

88

Vancouver The Journal of Pediatrics 164 (4) 906-11 Canada Retrospective observational chart review single centre study	people seen at the unit between 1998 and 2011 Note the transmale and transfemale subgroups reported in the paper is discrepant, 15 transmales and 11 transfemales (n=26) reported in the outcomes section rather than the n=27 stated in the paper, complete outcome reporting is also incomplete for the transfemale group Inclusion criteria were at least Tanner stage 2 pubertal development, previous assessment by a mental health professional and a confirmed diagnosis of gender dysphoria (diagnostic criteria not specified) No exclusion criteria are specified	specific treatment, dose or route of administration reported **Comparison** No comparator	**The authors report that of 15 transmales taking GnRH analogues** • 14 transitioned to testosterone treatment during the observation period • 7 continued taking GnRH analogues after starting testosterone • 7 discontinued GnRH analogues after a median of 3 0 years (range 0 2 to 9 2 years), of which 5 discontinued after hysterectomy and salpingo-oophorectomy 1 discontinued after 2 2 years (transitioned to gender-affirming hormones) 1 discontinued after <2 months due to mood and emotional lability **The authors report that of 11 transfemales taking GnRH analogues** • 5 received oestrogen treatment during the observation period • 4 continued taking GnRH analogues during oestrogen treatment • 1 discontinued GnRH analogues during oestrogen treatment (no reason reported) • 1 stopped GnRH analogues after a few months due to emotional lability • 1 stopped GnRH analogues before oestrogen treatment (the following year delayed due to heavy smoking) • 1 discontinued GnRH analogues after 13 months due to choosing not to pursue transition ***Safety*** Of the 27 patients treated with GnRH analogues	1 not reported 2 no non-exposed cohort 3 secure record 4 no **Domain 2: Comparability** 1. not applicable **Domain 3: Outcome** 1 record linkage 2 yes 3 incomplete missing data **Overall quality is assessed as poor.** Other comments mental health comorbidity was reported for all participants but not for the GnRH analogue cohort separately Concomitant use of other medicines was not reported Source of funding No source of funding identified

89

			• 1 transmale participant developed sterile abscesses, they were switched from leuprolide acetate to triptorelin, and this was well tolerated • 1 transmale participant developed leg pains and headaches on GnRH an ogues which eventually resolved without treatment • 1 participant gained 19 kg within 9 months of initiating GnRH analogues, although their body mass index was >85 percentile before GnRH analogues	
Study details	**Population**	**Interventions**	**Study outcomes**	**Appraisal and Funding**
Klink D Cams M, Heyboer A et al (2015) Bone mass in young adulthood following gonadotropin releasing hormone analog treatment and cross-sex hormone treatment in adolescents with gender dysphoria The Journal of clinical endocrinology and metabolism 100(2) e270-5 Netherlands Retrospective longitudinal observational single centre study To assess BMD development during GnRH analogues and at age 22 years in adolescents with gender dysphoria who started treatment for gender dysphoria during adolescence	34 adolescents (mean age ±SD 14 9±1 9 for transfemales and 15 0±2 0 for transmales at start of GnRH analogues) Participants were included if they met DSM-IV-TR criteria for gender identity disorder of adolescence and had been treated with GnRH analogues and gender-affirming hormones during their pubertal years No concomitant treatments were reported	The intervention was GnRH analogue monotherapy (triptorelin pamoate 3 75 mg subcutaneously every 4 weeks) followed by gender affirming hormones from 16 years with discontinuation of GnRH analogue after gonadectomy Median duration of GnRH analogue monotherapy in transfemales was 1 3 years (range, 0 5 to 3 8 years), and in transmales was 1 5 years	**Critical outcomes** No critical outcomes assessed **Important outcomes** ***Bone density: lumbar*** **Lumbar spine bone mineral apparent density (BMAD)[1]** Change from starting GnRH analogue (mean age 14 9±1 9) to starting gender-affirming hormones (mean age 16 6±1 4) in transfemales (mean [±SD]) GnRH analogue 0 22 (0 03) g/cm3, gender-affirming hormones 0 22 (0 02) g/cm3 (NS). z-score GnRH analogue −0 44 (1 10), gender-affirming hormones −0 90 (0 80) (p=NS) Change from starting GnRH analogue (mean age 15 0±2 0) to starting gender-affirming hormones (mean age 16 4±2 3) in transmales (mean [±SD] GnRH analogue 0 25 (0 03) g/cm3, gender-affirming hormones 0 24 (0 02) g/cm3 (NS).	This study was appraised using the Newcastle Ottawa quality assessment checklist for cohort studies **Domain 1: Selection** 1. somewhat representative of children and adolescents who have gender dysphoria 2 not applicable 3 via routine clinical records 4 no **Domain 2: Comparability** 1 no control group **Domain 3: Outcome** 1. via routine clinical records 2 yes 3 follow-up rate variable across timepoints and no description of those lost **Overall quality is assessed as poor**

90

Study details	Population	Interventions	Study outcomes	Appraisal and Funding
1998 to 2012		(range, 0 25 to 5 2 years)	z-score GnRH analogue 0 28 (0 90), gender-affirming hormones −0 50 (0 81) (p=0 004) **Lumbar spine bone mineral density (BMD)[1]** Change from starting GnRH analogue (mean age 14 9±1 9) to starting gender-affirming hormones (mean age 16 6±1 4) in transfemales (mean [±SD]) GnRH analogue 0 84 (0 13) g/m2, gender-affirming hormones 0 84 (0 11) g/m2 (NS), z-score GnRH analogue −0 77 (0 89), gender-affirming hormones −1 01 (0 98) (NS) Change from starting GnRH analogue (mean age 15 0±2 0) to starting gender-affirming hormones (mean age 16 4±2 3) in transmales (mean [±SD]) GnRH analogue 0 95 (0 12) g/m2, gender-affirming hormones 0 91 (0 10) g/m2 (p=0 006), z-score GnRH analogue 0 17 (1 18), gender-affirming hormones −0 72 (0 98) (p<0 001). ***Bone density; femoral*** **Femoral area BMAD[1]** Change from starting GnRH analogue (mean age 14 9±1 9) to starting gender-affirming hormones (mean age 16 6±1 4) in transfemales (mean [±SD]), GnRH analogue 0 28 (0 04) g/cm3, gender-affirming hormones 0 28 (0 04) g/cm3 (NS), z-score GnRH analogue −0 93 (1 22) gender-affirming hormones −1 57 (1 74) (p=NS) Change from starting GnRH analogue	Other comments Within person comparison. Small numbers of participants in each subgroup. No concomitant treatments or comorbidities were reported Source of funding None disclosed

91

Study details	Population	Interventions	Study outcomes	Appraisal and Funding
			(mean age 15 0±2 0) to starting gender-affirming hormones (mean age 16 4±2 3) in transmales (mean [±SD]), GnRH analogue 0 32 (0 04) g/cm3 gender-affirming hormones 0 31 (0 04) (NS), z-score GnRH analogue 0 01 (0 70) gender-affirming hormones −0 28 (0 74) (NS) **Femoral area BMD[1]** Change from starting GnRH analogue (mean age 14 9±1 9) to starting gender-affirming hormones (mean age 16 6±1 4) in transfemales (mean [±SD]), GnRH analogue 0 88 (0 12) g/m2, gender-affirming hormones 0 87 (0 08) (NS), z-score GnRH analogue −0 66 (0 77), gender-affirming hormones −0 95 (0 63) (NS) Change from starting GnRH analogue (mean age 15 0±2 0) to starting gender-affirming hormones (mean age 16 4±2 3) in transmales (mean [±SD]), GnRH analogue 0 92 (0 10) g/m2, gender-affirming hormones 0 88 (0 09) (p=0 005) z-score GnRH analogue 0 36 (0 88), gender-affirming hormones −0 35 (0 79) (p=0 001)	

[1] BMD and BMAD of the lumbar spine and femoral region (nondominant side) measured by DXA scans at start of GnRH analogues (n=32), start of gender affirming hormones (n=34), and at 22 years (n=34)

Study details	Population	Interventions	Study outcomes	Appraisal and Funding
Schagen SEE, Cohen-Kettenis PT, Delemarre-van de Waal HA et al (2016)	Adolescents with gender dysphoria (n=116), median age (range) 13 8 years (11 6 to 17 9) in transfemales and 14 2 years (11 1 to	GnRH analogue monotherapy (triptorelin pamoate 3 75 mg at 0 2 and 4	Critical outcomes No critical outcomes assessed Important outcomes	The study was appraised using the Newcastle-Ottawa quality assessment checklist for cohort studies

92

Study details	Population	Interventions	Study outcomes	Appraisal and Funding
[illegible] Puberty in Gender Dysphoric Adolescents. The journal of sexual medicine 13(7) 1125-32 Netherlands Prospective longitudinal study To describe the changes in Tanner stage, testicular volume, gonadotropins, and sex steroids during GnRH analogues of adolescents with gender dysphoria to evaluate the efficacy To report on liver enzymes, renal function and changes in body composition 1998 to 2009	18 6) in transmales during first year of GnRH analogues Participants were included if they met DSM-IV-TR criteria for gender dysphoria, had lifelong extreme gender dysphoria, were psychologically stable and were living in a supportive environment No concomitant treatments were reported	weeks followed by injections every 4 weeks, route of administration not described) for at least 3 months	*Other safety outcomes: liver function* Glutamyl transferase was not elevated at baseline or during treatment in any subject Mild elevations of aspartate aminotransferase (AST) and alanine aminotransferase (ALT) above the reference range were present at baseline but were not more prevalent during treatment than at baseline Glutamyl transferase, AST, and ALT levels did not significantly change from baseline to 12 months of treatment No values or statistical analyses were reported *Other safety outcomes: kidney function* Change in serum creatinine between 0 and 1 year Transfemales (mean [±SD]) 70 (12) micromol/l at baseline, 66 (13) micromol/l at 1 year (p=0 20) Transmales (mean [±SD]) 73 (8) micromol/l at baseline, 68 (13) micromol/l at 1 year (p=0 01)	Domain 1: Selection 1 somewhat representative of children and adolescents who have gender dysphoria 2 not applicable 3 via routine clinical records 4 no Domain 2: Comparability 1 no control group Domain 3: Outcome 1 via routine clinical records 2 yes 3 no statement Overall quality is assessed as poor Other comments Within person comparison No concomitant treatments or comorbidities were reported Source of funding Ferring pharmaceuticals (triptorelin manufacturer)
Staphorsius A, Baudewyntje P, Kreukels P, et al. (2015) Puberty suppression and execu[illegible] functioning: an fMRI stu[illegible]	The inclusion criteria were diagnosed with Gender Identity Disorder according to the DSM-IV-TR and at least 12 years old and Tanner stage of at least B2 or G2 to G3 with	Intervention GnRH analogues (triptorel n pamoate 3 75 mg every 4 weeks	Critical Outcomes No critical outcomes assessed Important outcomes *Psychosocial impact*	This study was appraised using the Newcastle-Ottawa tool for cohort studies Domain 1: Selection domain

Study details	Population	Interventions	Study outcomes	Appraisal and Funding
[illegible] Psychoneuroendocrinology 565 190-9 Netherlands Cross-sectional (single time point) assessment single centre study	measurable oestradiol and testosterone levels in girls and boys, respectively For all group's exclusion criteria were an insufficient command of the Dutch language (how assessed not reported), unadjusted endocrine disorders, neurological or psychiatric disorders that could lead to deviant test results (details not reported) use of psychotropic medication, and contraindications for an MRI scan Additionally, adolescents receiving puberty delaying medication or any form of hormones besides oral contraceptives were excluded as controls The sample size was 85 of whom 41 were adolescents (the numbers are discrepant with the number for whom outcomes are reported n=40) with gender dysphoria (20 of whom were being treated with GnRH analogues), 24 girls and 21 boys without gender dysphoria acted as controls (not further reported here) Details of the sampling frame are not reported The ages at which GnRH analogues were started was not reported The mean duration of treatment was 1 6 years (SD 1 0) Mean (±SD) Tanner stage for each group was reported • Transfemales 3 9 [±1 1] • Transfemales on GnRH [illegible] 4 1 [±1 0	subcutaneously or intramuscularly) Comparison The comparison was between adolescents with gender dysphoria receiving GnRH analogues and those without GnRH [illegible]	The Child Behaviour Checklist (CBCL) was used to assess psychosocial impact The CBCL was administered once during the study The reported outcomes for each group were (n, mean [±SD]) • Transfemales (all, n=18) 57 8 [±9 2] • Transfemales on GnRH analogues (n=8) 57 4 [±9 8] • Transfemales without GnRH analogues (n=10) 58 2 [±9 3] • Transmales (all, n=22) 80 4 [±10 2] • Transmales on GnRH analogues (n=12) 57 5 [±9 4] • Transmales without GnRH analogues (n=10) 63 9 [±10 5] The analysis of the CBCL data is not discussed, and statistical analysis is unclear *Cognitive development or functioning* IQ[1] • Transfemales (mean [±SD]) on GnRH analogues 94 0 (10 3) • Transfemales (mean [±SD]) without GnRH analogues 109 4 (21 2) • Transmales (mean [±SD]) on GnRH analogues 95 8 (15 8) • Transmales (mean [±SD]) without GnRH analogues 98.5 (15 9) Reaction time[2] • Transfemales (mean [±SD]) on GnRH analogues 10 9 (4 1) • Transfemales (mean [±SD]) without GnRH analogues 9 9 (3 1)	1 somewhat representative of children and adolescents who have gender dysphoria 2 drawn from the same community as the exposed cohort 3 via routine clinical records 4 no Domain 2: Comparability 1 study controls for age and diagnosis Domain 3: Outcome 1 via clinical assessment 2 yes 3 unclear Overall quality is assessed as poor. Other comments: Physical and psychological comorbidity was not reported, concomitant use of other medicines was not reported Source of funding This work was supported by an educational grant from the pharmaceutical firm Ferring BV, and by a VICI grant (453-08-003) from the Dutch Science Foundation The authors state that funding sources did not play a role in any component of this study

Study details	Population	Interventions	Study outcomes	Appraisal and Funding
	• Transfemales without GnRH analogues 3 8 [±1 1] • Transmales 4.5 [±0 9] • Transmales on GnRH analogues 4 1 [±1 1] Transmales without GnRH analogues 4 9 [-0 3]		• Transmales (mean [±SD]) on GnRH analogues 9 9 (3 1) • Transmales (mean [±SD]) without GnRH analogues 10.0 (2 0) Accuracy[2] • Transfemales (mean [±SD]) on GnRH analogues 73 9 (9 1) • Transfemales (mean [±SD]) without GnRH analogues 83 4 (9 5) • Transmales (mean [±SD]) on GnRH analogues 85 7 (10 5) • Transmales (mean [±SD]) without GnRH analogues 88 8 (9 7)	

[1] Estimated with 4 subscales (arithmetic, vocabulary, picture arrangement, and block design) of the Wechsler Intelligence Scale for Children, third edition (WISC-III) (Wechsler 1991) or the Wechsler Adult Intelligence Scale, third edition (WAIS-III®, Wechsler 1997), depending on the participant's age
[2] Reaction time in seconds in the Tower of London task
[3] Percentage of correct trials in the Tower of London task

Study details	Population	Interventions	Study outcomes	Appraisal and Funding
Vlot, Mariska C, Klink, Daniel T, den Heijer, Martin et al (2017) Effect of pubertal suppression and cross-sex hormone therapy on bone turnover markers and bone mineral apparent density (BMAD) in transgender adolescents Bone 95 11-19 Netherlands Retrospective observational data analysis study	Adolescents with gender dysphoria, n=70 Median age (range) 15 1 years (11 7 to 18 6) for transmales and 13.5 years (11 5 to 18 3) for transfemales at start of GnRH analogues Participants were included if they had a diagnosis of gender dysphoria according to DSM-IV-TR criteria who were treated with GnRH analogues and then gender-affirming hormones No concomitant treatments were reported The study categorised	GnRH analogues (triptorelin pamoate 3 75 mg every 4 weeks subcutaneously)	Critical outcomes No critical outcomes reported Important outcomes *Bone density: lumbar* **Lumbar spine bone mineral apparent density (BMAD)** Change from starting GnRH analogue to starting gender-affirming hormones in transfemales (bone age of <15 years, median [range]), GnRH analogue 0 21 (0 17 to 0 25) g/cm3, gender-affirming hormones 0 20 (0 16 to 0 24) g/cm3 (NS), z-score GnRH analogue −0 20 (−1 82 to 1 18), gender-affirming hormones −1 52 (−2 36 to 0 42) (p=0 001)	This study was appraised using the Newcastle-Ottawa quality assessment checklist for cohort studies **Domain 1: Selection** 1 Somewhat representative of children and adolescents who have gender dysphoria 2 Not applicable 3 Via routine clinical records 4 No **Domain 2: Comparability** 1 No control group **Domain 3: Outcome** 1 Via routine clinical records 2 Yes

95

Study details	Population	Interventions	Study outcomes	Appraisal and Funding
To investigate the course of 3 bone turnover markers in relation to bone mineral density, in adolescents with gender dysphoria during GnRH analogue and gender-affirming hormones 2001 to 2011	participants into a young and old pubertal group, based on their bone age The young transmales had a bone age of <14 years and the old transmales had a bone age of ≥14 years The young transfemales group had a bone age of <15 years and the old transfemales group ≥15 years		Change from starting GnRH analogue to starting gender-affirming hormones in transfemales (bone age of ≥15, median [range]), GnRH analogue 0 22 (0 18 to 0 25) g/cm3, gender-affirming hormones 0 22 (0 19 to 0 24) g/cm3 (NS), z-score GnRH analogue −1 18 (−1 78 to 1 09), gender-affirming hormones −1 15 (−2 21 to 0 08) (ps0 1) Change from starting GnRH analogue to starting gender-affirming hormones in transmales (bone age of <15 years, median [range]), GnRH analogue 0 23 (0 20 to 0 29) g/cm3, gender-affirming hormones 0 23 (0 19 to 0 28) g/cm3 (NS), z-score GnRH analogue −0 05 (−0 78 to 2 94), gender-affirming hormones −0 84 (−2 20 to 0 87) (p=0 003) Change from starting GnRH analogue to starting gender-affirming hormones in transmales (bone age of ≥15, median [range]), GnRH analogue 0 26 (0 21 to 0 29) g/cm3, gender-affirming hormones 0 24 (0 20 to 0 28) g/cm3 (ps0 01) z-score GnRH analogue 0 27 (−1 60 to 1 80), gender-affirming hormones −0 29 (−2 28 to 0 90) (ps 0 0001) *Bone density: femoral* **Femoral neck BMAD** Change from starting GnRH analogue to starting gender-affirming hormones in transfemales (bone age of <15 years, median [range]), GnRH analogue 0 29 (0 20 to 0 33) g/cm3, gender-affirming hormones 0 27 (0 20 to 0 33) g/cm3 (ps0 1), z-score GnRH analogue −0 71 (−3 35 to	3 Follow-up rate variable across outcomes and no description of those lost Overall quality is assessed as poor Other comments Within person comparison No concomitant treatments were reported Source of funding grant from Abbott diagnostics

96

Study details	Population	Interventions	Study outcomes	Appraisal and Funding
			0 37), gender-affirming hormones −1 32 (−3 39 to 0 21) (p≤0 1) Change from starting GnRH analogue to starting gender-affirming hormones in transfemales (bone age of ≥15, median [range]), GnRH analogue 0 30 (0 26 to 0 36) g/cm3, gender-affirming hormones 0 30 (0 26 to 0 34) g/cm3 (NS), z-score GnRH analogue −0 44 (−1 37 to 0 93), gender-affirming hormones −0 36 (−1 50 to 0 46) (NS) Change from starting GnRH analogue to starting gender-affirming hormones in transmales (bone age of <15 years, median [range]), GnRH analogue 0 31 (0 26 to 0 36) g/cm3, gender-affirming hormones 0 30 (0 22 to 0 35) g/cm3 (NS), z-score GnRH analogue −0 01 (−1 38 to 0 91), gender-affirming hormones −0 37 (−2 28 to 0 47) (NS) Change from starting GnRH analogue to starting gender-affirming hormones in transmales (bone age of ≥15, median [range]), GnRH analogue 0 33 (0 25 to 0 39) g/cm3, gender-affirming hormones 0 30 (0 23 to 0 41) g/cm3 (p≤0 01), z-score GnRH analogue 0 27 (−1 39 to 1 32), gender-affirming hormones −0 27 (−1 91 to 1 29) (p=0 002)	

97

Appendix F Quality appraisal checklists

Newcastle-Ottawa tool for cohort studies

Question	
Domain: Selection	
1. Representativeness of the exposed cohort	Truly representative of the average [describe] in the community Somewhat representative of the average [describe] in the community Selected group of users e.g. nurses, volunteers No description of the derivation of the cohort
2. Selection of the non-exposed cohort	Drawn from the same community as the exposed cohort Drawn from a different source No description of the derivation of the non-exposed cohort
3. Ascertainment of exposure	Secure record (e.g. surgical records) Structured interview Written self-report No description
4. Demonstration that outcome of interest was not present at start of study	Yes / No
Domain: Comparability	
1. Comparability of cohorts on the basis of the design or analysis	Study controls for [select most important factor] Study controls for any additional factor [this criteria could be modified to indicate specific control for a second important factor]
Domain: Outcome	
1. Assessment of outcome	Independent blind assessment Record linkage Self-report No description
2. Was follow-up long enough for outcomes to occur	Yes [select and adequate follow up period for outcome of interest] No
3. Adequacy of follow up of cohorts	Complete follow up (all subjects accounted for) Subjects lost to follow up unlikely to introduce bias (small number lost to follow up [select an adequate %] follow up or description provided of those lost) Follow up rate [select an adequate %] and no description of those lost No statement

98

Appendix G Grade profiles

Table 2: Question 1. For children and adolescents with gender dysphoria, what is the clinical effectiveness of treatment with GnRH analogues compared with one or a combination of psychological support, social transitioning to the desired gender or no intervention? – gender dysphoria

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of events/No of patients (n/N%)		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result		
Impact on gender dysphoria									
Mean±SD Utrecht Gender Dysphoria Scale[1] (version(s) not reported), time point at baseline (before GnRH analogues) versus follow-up (before gender-affirming hormones, higher scores indicate more gender dysphoria)									
1 cohort study de Vries et al 2011	Serious limitations[2]	No serious indirectness	Not applicable	Not calculable	N=41	None	Baseline 53.20±7.91 GnRH analogue 53.9±17.42 P=0.393	Critical	VERY LOW

Abbreviations: GnRH, gonadotrophin releasing hormone; P, P-value; SD, Standard deviation

1 The UGDS is a validated screening tool for both adolescents and adults to assess gender dysphoria. It consists of 12 items, to be answered on a 1- to 5-point scale, resulting in a sum score between 12 and 60. The higher the UGDS score the greater the gender dysphoria
2 Downgraded 1 level - the cohort study by de Vries et al (2011) was assessed as at high risk of bias (poor quality overall, lack of blinding and no control group)

Table 3: Question 1. For children and adolescents with gender dysphoria, what is the clinical effectiveness of treatment with GnRH analogues compared with one or a combination of psychological support, social transitioning to the desired gender or no intervention? – mental health

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of events/No of patients (n/N%)		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result		
Impact on mental health									

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of events/No of patients (n/N%)		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result		
Mean±SD Beck Depression Inventory-II, time point at baseline (before GnRH analogues) versus follow-up (just before gender-affirming hormones), (Lower scores indicate benefit)									
1 cohort study de Vries et al 2011	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=41	None	Baseline 8.31±7.12 GnRH analogue 4.95±6.72 P=0.004	Critical	VERY LOW
Mean±SD Trait Anger (TPI), time point at baseline (before GnRH analogues) versus follow-up (just before gender-affirming hormones, lower scores indicate benefit)									
1 cohort study de Vries et al 2011	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=41	None	Baseline 18.29±5.54 GnRH analogue 17.98±5.24 P=0.803	Critical	VERY LOW
Mean±SD Trait Anxiety (STAI), time point at baseline (before GnRH analogues) versus follow-up (just before gender-affirming hormones, lower scores indicate benefit)									
1 cohort study de Vries et al 2011	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=41	None	Baseline 38.43±10.07 GnRH analogue 37.95±9.38 P=0.276	Critical	VERY LOW

Abbreviations: GnRH, gonadotrophin releasing hormone; P, P-value; SD, Standard deviation

1 Downgraded 1 level - the cohort study by de Vries et al (2011) was assessed as at high risk of bias (poor quality overall, lack of blinding and no control group)

Table 4: Question 1. For children and adolescents with gender dysphoria, what is the clinical effectiveness of treatment with GnRH analogues compared with one or a combination of psychological support, social transitioning to the desired gender or no intervention? – body image

QUALITY						Summary of findings			IMPORTANCE	CERTAINTY
						No of participants (studies)		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision		Intervention	Comparator	Result		
Impact on body image										
Mean±SD Body Image Scale (primary sexual characteristics), time point at baseline (before GnRH analogues) versus follow-up (just before gender-affirming hormones, lower scores indicate benefit)										
1 cohort study de Vries et al 2011	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=57	None		Baseline 4 10±0 56 GnRH analogue 3 96±0 71 P=0 145	Important	VERY LOW
Mean±SD Body Image Scale (secondary sexual characteristics), time point at baseline (before GnRH analogues) versus follow-up (just before gender-affirming hormones, lower scores indicate benefit)										
1 cohort study de Vries et al 2011	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=57	None		Baseline 2 74±0 65 GnRH analogue 2 82±0 68 P=0 569	Important	VERY LOW
Mean±SD Body Image Scale (neutral characteristics), time point at baseline (before GnRH analogues) versus follow-up (just before gender-affirming hormones, lower scores indicate benefit)										
1 cohort study de Vries et al 2011	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=57	None		Baseline 2 41±0 63 GnRH analogue 2 47±0 56 P=0 620	Important	VERY LOW

Abbreviations: GnRH gonadotrophin releasing hormone P P-value SD Standard deviation

[1] Downgraded 1 level - the cohort study by de Vries et al (2011) was assessed as at high risk of bias (poor quality overall lack of blinding and no control group)

101

Table 5: Question 1. For children and adolescents with gender dysphoria, what is the clinical effectiveness of treatment with GnRH analogues compared with one or a combination of psychological support, social transitioning to the desired gender or no intervention? – psychosocial impact

QUALITY						Summary of findings			IMPORTANCE	CERTAINTY
						No of participants (studies)		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision		Intervention	Comparator	Result		
Psychosocial impact										
Mean (±SD) Children's Global Assessment Scale score, at baseline, higher scores indicate benefit)										
1 cohort study Costa et al 2015	Serious limitations[1]	No serious indirectness	No serious inconsistency	Not calculable		n=101 58 72 [±11 38]	n=100 56 63 [±13 14]	P=0 23	Important	VERY LOW
Mean (±SD) Children's Global Assessment Scale score, at 6 months[2] (higher scores indicate benefit).										
1 cohort study Costa et al 2015	Serious limitations[1]	No serious indirectness	No serious inconsistency	Not calculable		n=101 60 89 [±12 17]	n=100 60 29 [±12 81]	P=0 73	Important	VERY LOW
Mean (±SD) Children's Global Assessment Scale score, at 12 months[3] (higher scores indicate benefit).										
1 cohort study Costa et al 2015	Serious limitations[1]	No serious indirectness	No serious inconsistency	Not calculable		n=60 64 70 [±13 34]	n=61 62 92 [±14 10]	P=0 49	Important	VERY LOW
Mean (±SD) Children's Global Assessment Scale score, at 18 months[4] (higher scores indicate benefit).										
1 cohort study Costa et al 2015	Serious limitations[1]	No serious indirectness	No serious inconsistency	Not calculable		n=30 67 40 [±13 93]	n=30 62 53 [±13 54]	P=0 14	Important	VERY LOW
Mean (±SD) Children's Global Assessment Scale score, participants at 6 months compared to baseline (higher scores indicate benefit).										
1 cohort study Costa et al 2015	Serious limitations[1]	No serious indirectness	No serious inconsistency	Not calculable		N=101 N=101	None	Baseline 58 72±11 38 6 months 60 60±12 17 P=0 19	Important	VERY LOW
Mean (±SD) Children's Global Assessment Scale score, participants at 12 months compared to baseline (higher scores indicate benefit).										

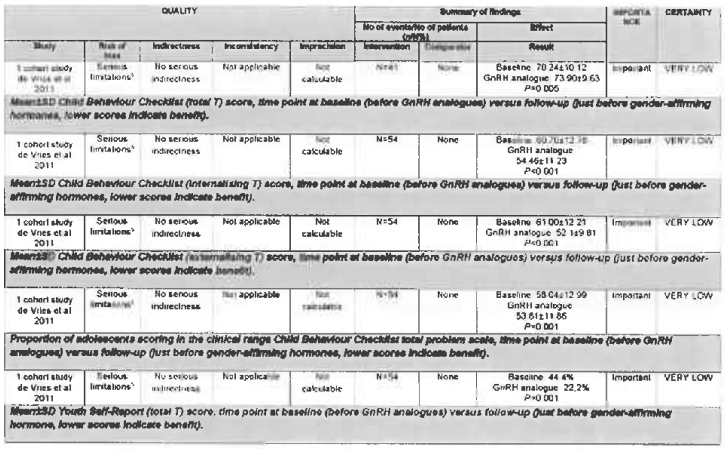
102

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of events/No of patients (n/N%)		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result		
1 cohort study Costa et al 2015	Serious limitations[1]	No serious indirectness	No serious inconsistency	Not calculable	N=101 N=60	None	Baseline 58.72±11.38 12 months 64.70±13.34 P=0.003	Important	VERY LOW
Mean [±SD] Children's Global Assessment Scale score, participants at 18 months compared to baseline (higher scores indicate benefit).									
1 cohort study Costa et al 2015	Serious limitations[1]	No serious indirectness	No serious inconsistency	Not calculable	N=101 N=35	None	Baseline 58.72±11.38 18 months 67.40±13.93 P<0.001	Important	VERY LOW
Mean [±SD] Children's Global Assessment Scale score, participants at 12 months compared to 6 months (higher scores indicate benefit).									
1 cohort study Costa et al 2015	Serious limitations[1]	No serious indirectness	No serious inconsistency	Not calculable	N=101 N=60	None	6 months 60.89±12.17 12 months 64.70±13.34 P=0.07	Important	VERY LOW
Mean [±SD] Children's Global Assessment Scale score, participants at 18 months compared to 6 months (higher scores indicate benefit).									
1 cohort study Costa et al 2015	Serious limitations[1]	No serious indirectness	No serious inconsistency	Not calculable	N=101 N=35	None	6 months 60.89±12.17 18 months 67.40±13.93 P<0.001	Important	VERY LOW
Mean [±SD] Children's Global Assessment Scale score, participants at 18 months compared to 12 months (higher scores indicate benefit).									
1 cohort study Costa et al 2015	Serious limitations[1]	No serious indirectness	No serious inconsistency	Not calculable	N=60 N=35	None	12 months 64.70±13.34 18 months 67.40±13.93 P=0.35	Important	VERY LOW
Mean [±SD] Children's Global Assessment Scale score, in all participants (including those not treated with GnRH analogues) at 6 months[2] compared to baseline (higher scores indicate benefit).									
1 cohort study Costa et al 2015	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=201	None	Baseline 57.73±12.27 6 months 60.68±12.47 P<0.001	Important	VERY LOW
Mean [±SD] Children's Global Assessment Scale score, in all participants (including those not treated with GnRH analogues) at 12 months[2] compared to baseline (higher scores indicate benefit).									

103

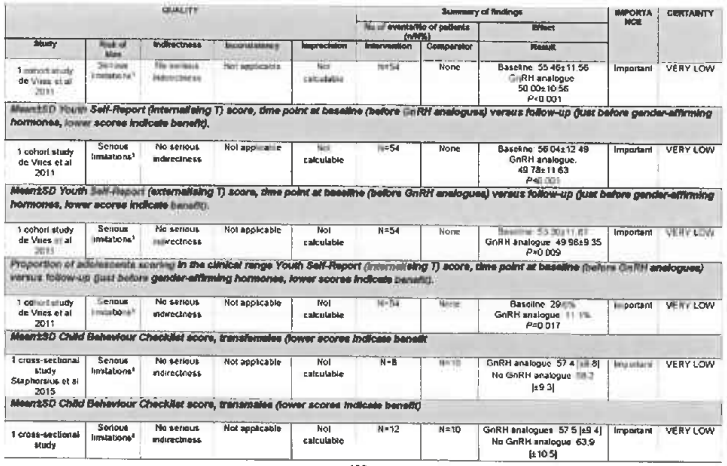
QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of events/No of patients (n/N%)		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result		
1 cohort study Costa et al 2015	Serious limitations[1]	No serious indirectness	No serious inconsistency	Not calculable	N=201 N=121	None	Baseline 57.73±12.27 12 months 63.31±14.41 P<0.001	Important	VERY LOW
Mean±SD Children's Global Assessment Scale score, in all participants (including those not treated with GnRH analogues) at 18 months[2] compared to baseline (higher scores indicate benefit).									
1 cohort study Costa et al 2015	Serious limitations[1]	No serious indirectness	No serious inconsistency	Not calculable	N=201 N=71	None	Baseline 57.73±12.27 18 months 64.93±13.85 P<0.001	Important	VERY LOW
Mean±SD Children's Global Assessment Scale score, in all participants (including those not treated with GnRH analogues) at 12 months compared to 6 months (higher scores indicate benefit).									
1 cohort study Costa et al 2015	Serious limitations[1]	No serious indirectness	No serious inconsistency	Not calculable	N=201 N=121	None	6 months 60.68±12.47 12 months 63.31±14.41 P<0.06	Important	VERY LOW
Mean±SD Children's Global Assessment Scale score, in all participants (including those not treated with GnRH analogues) at 18 months compared to 6 months (higher scores indicate benefit).									
1 cohort study Costa et al 2015	Serious limitations[1]	No serious indirectness	No serious inconsistency	Not calculable	N=201 N=71	None	6 months 60.68±12.47 18 months 64.93±13.85 P<0.02	Important	VERY LOW
Mean±SD Children's Global Assessment Scale score, in all participants (including those not treated with GnRH analogues) at 18 months compared to 12 months (higher scores indicate benefit).									
1 cohort study Costa et al 2015	Serious limitations[1]	No serious indirectness	No serious inconsistency	Not calculable	N=121 N=71	None	12 months 63.31±14.41 18 months 64.93±13.85 P<0.45	Important	VERY LOW
Mean±SD Children's Global Assessment Scale score, time point at baseline (before GnRH analogues) versus follow-up (just before gender-affirming hormones, higher scores indicate benefit).									

104



QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of events/No of patients (n/N)		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result		
1 cohort study de Vries et al 2011	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=41	None	Baseline 70.24±10.12 GnRH analogue 73.90±9.63 P=0.005	Important	VERY LOW
Mean±SD Child Behaviour Checklist (total T) score, time point at baseline (before GnRH analogues) versus follow-up (just before gender-affirming hormones, lower scores indicate benefit).									
1 cohort study de Vries et al 2011	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=54	None	Baseline 60.70±12.76 GnRH analogue 54.46±11.23 P<0.001	Important	VERY LOW
Mean±SD Child Behaviour Checklist (internalising T) score, time point at baseline (before GnRH analogues) versus follow-up (just before gender-affirming hormones, lower scores indicate benefit).									
1 cohort study de Vries et al 2011	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=54	None	Baseline 61.00±12.21 GnRH analogue 52.1±9.81 P<0.001	Important	VERY LOW
Mean±SD Child Behaviour Checklist (externalising T) score, time point at baseline (before GnRH analogues) versus follow-up (just before gender-affirming hormones, lower scores indicate benefit).									
1 cohort study de Vries et al 2011	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=54	None	Baseline 58.04±12.99 GnRH analogue 53.61±11.86 P=0.001	Important	VERY LOW
Proportion of adolescents scoring in the clinical range Child Behaviour Checklist total problem scale, time point at baseline (before GnRH analogues) versus follow-up (just before gender-affirming hormones, lower scores indicate benefit).									
1 cohort study de Vries et al 2011	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=54	None	Baseline 44.4% GnRH analogue 22.2% P=0.001	Important	VERY LOW
Mean±SD Youth Self-Report (total T) score, time point at baseline (before GnRH analogues) versus follow-up (just before gender-affirming hormone, lower scores indicate benefit).									

105



QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of events/No of patients (n/N)		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result		
1 cohort study de Vries et al 2011	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=54	None	Baseline 55.46±11.56 GnRH analogue 50.00±10.56 P<0.001	Important	VERY LOW
Mean±SD Youth Self-Report (internalising T) score, time point at baseline (before GnRH analogues) versus follow-up (just before gender-affirming hormones, lower scores indicate benefit).									
1 cohort study de Vries et al 2011	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=54	None	Baseline 56.04±12.49 GnRH analogue 49.78±11.63 P<0.001	Important	VERY LOW
Mean±SD Youth Self-Report (externalising T) score, time point at baseline (before GnRH analogues) versus follow-up (just before gender-affirming hormones, lower scores indicate benefit).									
1 cohort study de Vries et al 2011	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=54	None	Baseline 53.30±11.61 GnRH analogue 49.98±9.35 P=0.009	Important	VERY LOW
Proportion of adolescents scoring in the clinical range Youth Self-Report (internalising T) score, time point at baseline (before GnRH analogues) versus follow-up (just before gender-affirming hormones, lower scores indicate benefit).									
1 cohort study de Vries et al 2011	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=54	None	Baseline 29.6% GnRH analogue 11.1% P=0.017	Important	VERY LOW
Mean±SD Child Behaviour Checklist score, transfemales (lower scores indicate benefit									
1 cross-sectional study Staphorsius et al 2015	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=8	N=10	GnRH analogue 57.4 [±8.8] No GnRH analogue 58.2 [±9.3]	Important	VERY LOW
Mean±SD Child Behaviour Checklist score, transmales (lower scores indicate benefit)									
1 cross-sectional study	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=12	N=10	GnRH analogues 57.5 [±9.4] No GnRH analogue 63.9 [±10.5]	Important	VERY LOW

106

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of events/No of patients (n/N%)		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result		
Stephorsius et al 2015									

Abbreviations: GnRH, gonadotrophin releasing hormone, *P*, P-value, SD, Standard deviation

1 Downgraded 1 level - the cohort study by Costa et al (2015) was assessed as at high risk of bias (poor quality overall, lack of blinding and no control group)
2 6 months from baseline (after 6 months of psychological support – both groups)
3 12 months from baseline (delayed eligible gender dysphoria (GD) adolescents, after 12 months of psychological support, immediately eligible GD adolescents, after 12 months of psychological support + 6 months of puberty suppression)
4 18 months from baseline (delayed eligible gender dysphoria (GD) adolescents, after 12 months of psychological support, immediately eligible GD adolescents, after 12 months of psychological support + 6 months of puberty suppression)
5 Downgraded 1 level - the cohort study by de Vries et al (2011) was assessed as at high risk of bias (poor quality overall, lack of blinding and no control group)
6 Downgraded 1 level - the cohort study by Staphorsius et al (2015) was assessed as at high risk of bias (poor quality overall, lack of blinding and no randomisation)

Table 6: Question 1. For children and adolescents with gender dysphoria, what is the clinical effectiveness of treatment with GnRH analogues compared with one or a combination of psychological support, social transitioning to the desired gender or no intervention? – engagement with healthcare services

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of events/No of patients (n/N%)		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result		
Engagement with healthcare services									
Number (proportion) failing to engage with health care services (did not attend clinic), at (up to) 9 years follow-up									
1 cohort study Bak et al 2018	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	9/214 (4.2%)	None	9 adolescents out of 214 failed to attend clinic and were excluded from the study (4.2%)	Important	VERY LOW
Loss to follow-up									
1 cohort study	Serious limitations[2]	No serious indirectness	Not applicable		201	None	The sample size at baseline and 6 months was 201, which dropped by 39.8% to 121 after	Important	VERY LOW

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of events/No of patients (n/N%)		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result		
Costa et al 2015				Not calculable			12 months and by 64.7% to 71 at 18 months follow-up. No explanation of the reasons for loss to follow-up are reported		

Abbreviations: GnRH, gonadotrophin releasing hormone

1 Downgraded 1 level - the cohort study by Bak et al (2018) was assessed as at high risk of bias (poor quality overall, lack of blinding and no control group)
2 Downgraded 1 level - the cohort study by Costa et al (2015) was assessed as at high risk of bias (poor quality overall, lack of blinding and no control group)

Table 7: Question 1. For children and adolescents with gender dysphoria, what is the clinical effectiveness of treatment with GnRH analogues compared with one or a combination of psychological support, social transitioning to the desired gender or no intervention? – stopping treatment

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of events/No of patients (n/N%)		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result		
Stopping treatment									
Number (proportion) stopping GnRH analogues, at (up to) 9 years follow-up									
1 cohort study Bak et al 2018	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	9/143 (6.2%)	None	9/143 adolescents stopped GnRH analogues (6.2%)[2]	Important	VERY LOW
Number (proportion) stopping from GnRH analogues, at (up to) 13 years follow-up									
1 cohort study Khatchadourian et al 2014	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	11/27 (42%)	None	11/26 stopped GnRH analogues (42%)[2]	Important	VERY LOW
Number (proportion) stopping GnRH analogues but who wished to continue endocrine treatment, at (up to) 9 years follow-up									

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of events/No of patients/% (n/N%)		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result		
1 cohort study Bnk et al 2018	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	4/143 (2.8%)	None	4/143 adolescents stopped GnRH analogues but wished to continue treatment (2.8%)	Important	VERY LOW
Number (proportion) stopping GnRH analogues who no longer wished gender-affirming treatment, at (up to) 9 years follow-up									
1 cohort study Bnk et al 2018	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	5/143 (3.5%)	None	5/143 adolescents stopped GnRH analogues and no longer wished to continue gender-affirming treatment (3.5%)	Important	VERY LOW

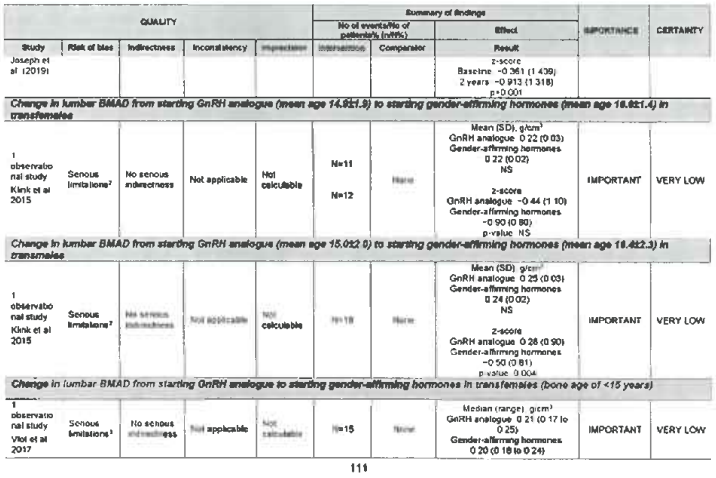
Abbreviations: GnRH, gonadotrophin releasing hormone

1 Downgraded 1 level - the cohort study by Bnk et al (2018) was assessed as at high risk of bias (poor quality overall, lack of blinding and no control group)
2 Median duration of 9.8 years (range 0.1 to 20). Five adolescents stopped treatment because they no longer wished to receive gender-affirming treatment for various reasons. In 4 adolescents (all transmales), although they wanted to continue treatments for gender dysphoria, GnRH analogues were stopped mainly because of adverse effects (such as mood and emotional lability)
3 Downgraded 1 level - the cohort study by Khatchadourian et al (2014) was assessed as at high risk of bias (poor quality overall, lack of blinding, no control group and high number of participants lost to follow-up)
4 Because of transitioning to gender-affirming hormones or gender-affirming surgery, adverse effects (such as mood and emotional lability) or no longer wishing to pursue transition

Table 8. Question 2. For children and adolescents with gender dysphoria, what is the short-term and long-term safety of GnRH analogues compared with one or a combination of psychological support, social transitioning to the desired gender or no intervention? – bone density

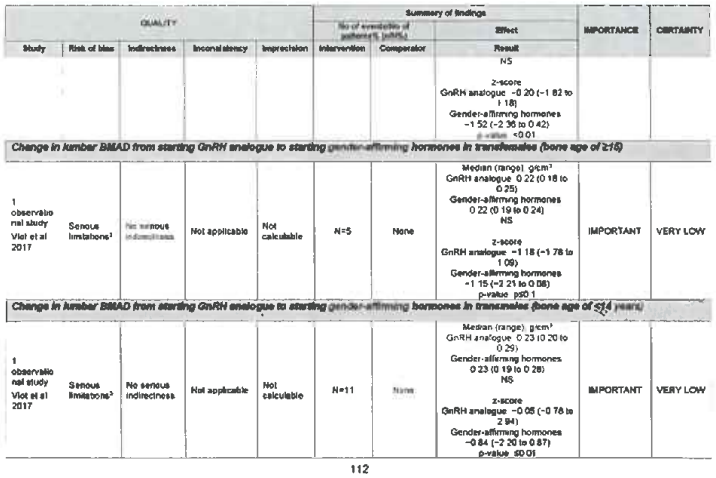
QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of events/No of patients/% (n/N%)		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result		
Bone density: change in lumbar BMAD									
Change in lumbar spine BMAD from baseline to 1 year in transfemales									

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of events/No of patients/% (n/N%)		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result		
1 observational study Joseph et al (2019)	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=21	None	Mean (SD), g/cm² Baseline 0.225 (0.030) 1 year 0.233 (0.029) p=0.459 z-score Baseline 0.859 (0.154) 1 year -0.228 (1.027) p=0.000	IMPORTANT	VERY LOW
Change in lumbar spine BMAD from baseline to 1 year in transmales									
1 observational study Joseph et al (2019)	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=39	None	Mean (SD), g/cm² Baseline 0.196 (0.035) 1 year 0.201 (0.033) p=0.074 z-score Baseline -0.186 (1.210) 1 year -0.541 (1.206) p=0.006	IMPORTANT	VERY LOW
Change in lumbar spine BMAD from baseline to 2 years in transfemales									
1 observational study Joseph et al (2019)	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=10	None	Mean (SD), g/cm² Baseline 0.240 (0.027) 2 years 0.240 (0.030) p=0.865 z-score Baseline 0.486 (0.806) 2 years -0.279 (0.930) p=0.000	IMPORTANT	VERY LOW
Change in lumbar spine BMAD from baseline to 2 years in transmales									
1 observational study	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=21	None	Mean (SD), g/cm² Baseline 0.185 (0.058) 2 years 0.198 (0.055) p=0.433	IMPORTANT	VERY LOW



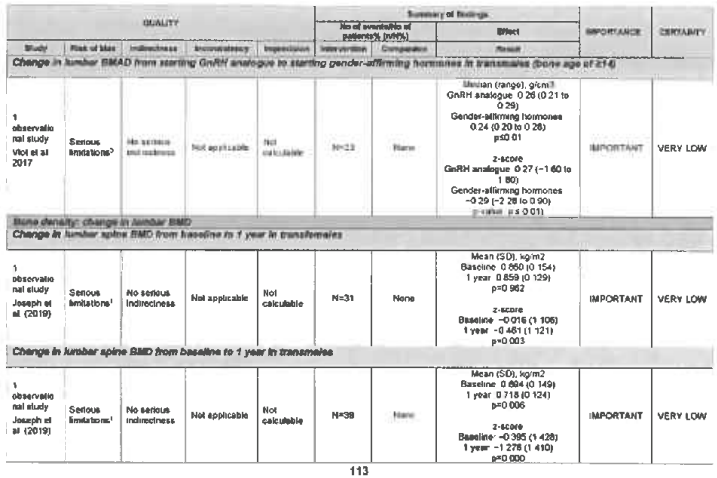
QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of events/No of patients (n/N%)		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result		
Joseph et al (2019)[1]							z-score Baseline −0 361 (1 439) 2 years −0 913 (1 318) p<0 001		
Change in lumbar BMAD from starting GnRH analogue (mean age 14.9±1.8) to starting gender-affirming hormones (mean age 16.8±1.4) in transfemales									
1 observational study Klink et al 2015	Serious limitations[2]	No serious indirectness	Not applicable	Not calculable	N=11 N=12	None	Mean (SD), g/cm² GnRH analogue 0 22 (0 03) Gender-affirming hormones 0 22 (0 02) NS z-score GnRH analogue −0 44 (1 10) Gender-affirming hormones −0 90 (0 80) p-value NS	IMPORTANT	VERY LOW
Change in lumbar BMAD from starting GnRH analogue (mean age 15.0±2 6) to starting gender-affirming hormones (mean age 18.4±2.3) in transmales									
1 observational study Klink et al 2015	Serious limitations[2]	No serious indirectness	Not applicable	Not calculable	N=18	None	Mean (SD), g/cm² GnRH analogue 0 25 (0 03) Gender-affirming hormones 0 24 (0 02) NS z-score GnRH analogue 0 26 (0 90) Gender-affirming hormones −0 50 (0 81) p-value 0 004	IMPORTANT	VERY LOW
Change in lumbar BMAD from starting GnRH analogue to starting gender-affirming hormones in transfemales (bone age of <15 years)									
1 observational study Viot et al 2017	Serious limitations[3]	No serious indirectness	Not applicable	Not calculable	N=15	None	Median (range), g/cm² GnRH analogue 0 21 (0 17 to 0 25) Gender-affirming hormones 0 20 (0 18 to 0 24)	IMPORTANT	VERY LOW

111



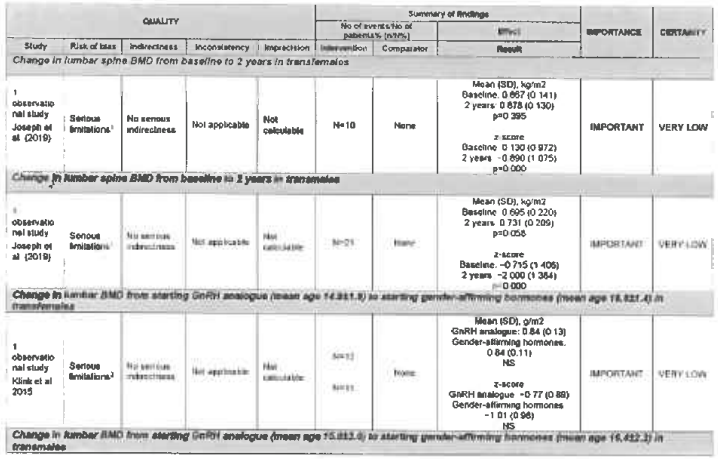
QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of events/No of patients (n/N%)		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result		
							NS z-score GnRH analogue −0 20 (−1 82 to 1 18) Gender-affirming hormones −1 52 (−2 36 to 0 42) p-value <0 01		
Change in lumbar BMAD from starting GnRH analogue to starting gender-affirming hormones in transfemales (bone age of ≥15)									
1 observational study Viot et al 2017	Serious limitations[3]	No serious indirectness	Not applicable	Not calculable	N=5	None	Median (range), g/cm² GnRH analogue 0 22 (0 18 to 0 25) Gender-affirming hormones 0 22 (0 19 to 0 24) NS z-score GnRH analogue −1 18 (−1 78 to 1 09) Gender-affirming hormones −1 15 (−2 21 to 0 08) p-value p≤0 1	IMPORTANT	VERY LOW
Change in lumbar BMAD from starting GnRH analogue to starting gender-affirming hormones in transmales (bone age of ≤14 years)									
1 observational study Viot et al 2017	Serious limitations[3]	No serious indirectness	Not applicable	Not calculable	N=11	None	Median (range), g/cm² GnRH analogue 0 23 (0 20 to 0 29) Gender-affirming hormones 0 23 (0 19 to 0 28) NS z-score GnRH analogue −0 05 (−0 78 to 2 84) Gender-affirming hormones −0 84 (−2 20 to 0 87) p-value ≤0 01	IMPORTANT	VERY LOW

112



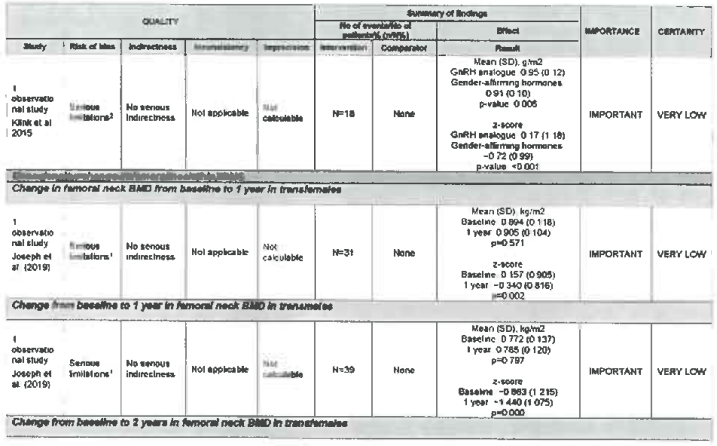
QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	No of events/No of patients (n/N%) Intervention	Comparator	Effect Result		
Change in lumbar BMAD from starting GnRH analogue to starting gender-affirming hormones in transitalss (bone age ≥12.0)									
1 observational study Vlot et al 2017	Serious limitations[2]	No serious indirectness	Not applicable	Not calculable	N=33	None	Median (range), g/cm3 GnRH analogue 0.26 (0.21 to 0.29) Gender-affirming hormones 0.24 (0.20 to 0.28) p≤0.01 z-score GnRH analogue 0.27 (−1.60 to 1.80) Gender-affirming hormones −0.29 (−2.28 to 0.90) p value p ≤ 0.01	IMPORTANT	VERY LOW
Bone density: change in lumbar BMD									
Change in lumbar spine BMD from baseline to 1 year in transfemales									
1 observational study Joseph et al (2019)	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=31	None	Mean (SD), kg/m2 Baseline 0.860 (0.154) 1 year 0.859 (0.129) p=0.962 z-score Baseline −0.016 (1.106) 1 year −0.461 (1.121) p=0.003	IMPORTANT	VERY LOW
Change in lumbar spine BMD from baseline to 1 year in transmales									
1 observational study Joseph et al (2019)	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=38	None	Mean (SD), kg/m2 Baseline 0.894 (0.149) 1 year 0.718 (0.124) p=0.006 z-score Baseline −0.395 (1.428) 1 year −1.275 (1.410) p=0.000	IMPORTANT	VERY LOW

113



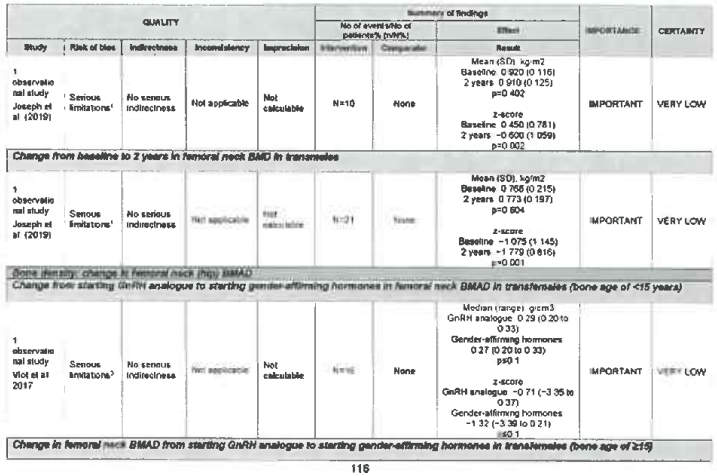
QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	No of events/No of patients (n/N%) Intervention	Comparator	Effect Result		
Change in lumbar spine BMD from baseline to 2 years in transfemales									
1 observational study Joseph et al (2019)	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=10	None	Mean (SD), kg/m2 Baseline 0.867 (0.141) 2 years 0.878 (0.130) p=0.395 z-score Baseline 0.130 (0.972) 2 years 0.890 (1.075) p=0.000	IMPORTANT	VERY LOW
Change in lumbar spine BMD from baseline to 2 years in transmales									
1 observational study Joseph et al (2019)	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=21	None	Mean (SD), kg/m2 Baseline 0.695 (0.220) 2 years 0.731 (0.209) p=0.058 z-score Baseline. −0.715 (1.406) 2 years −2.000 (1.364) p=0.000	IMPORTANT	VERY LOW
Change in lumbar BMD from starting GnRH analogue (mean age 14.3±1.9) to starting gender-affirming hormones (mean age 16.8±1.4) in transfemales									
1 observational study Klink et al 2015	Serious limitations[2]	No serious indirectness	Not applicable	Not calculable	N=11 N=11	None	Mean (SD), g/m2 GnRH analogue: 0.84 (0.13) Gender-affirming hormones. 0.84 (0.11) NS z-score GnRH analogue +0.77 (0.89) Gender-affirming hormones −1.01 (0.96) NS	IMPORTANT	VERY LOW
Change in lumbar BMD from starting GnRH analogue (mean age 15.8±2.0) to starting gender-affirming hormones (mean age 16.4±2.2) in transmales									

114



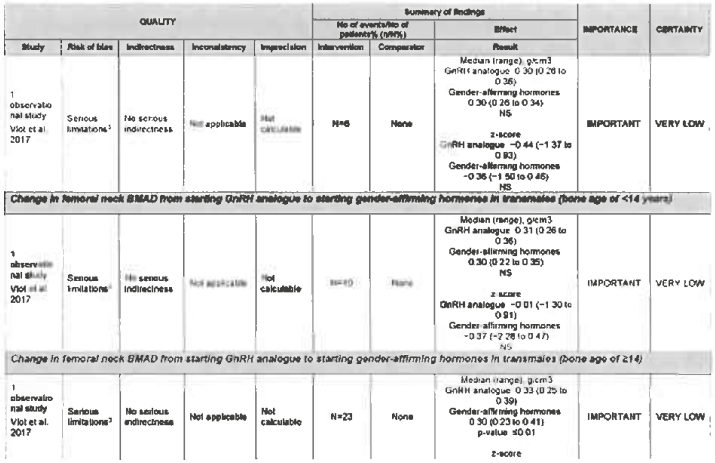
QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of events/No of patients (n/N%)		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result		
1 observational study Klink et al 2015	Serious limitations[2]	No serious indirectness	Not applicable	Not calculable	N=18	None	Mean (SD), g/m2 GnRH analogue 0 95 (0 12) Gender-affirming hormones 0 93 (0 10) p-value 0 005 z-score GnRH analogue 0 17 (1 18) Gender-affirming hormones −0 72 (0 99) p-value <0 001	IMPORTANT	VERY LOW
Change in femoral neck BMD from baseline to 1 year in transfemales									
1 observational study Joseph et al (2019)	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=31	None	Mean (SD), kg/m2 Baseline 0 894 (0 118) 1 year 0 905 (0 104) p=0 571 z-score Baseline 0 157 (0 905) 1 year −0 340 (0 816) p=0 002	IMPORTANT	VERY LOW
Change from baseline to 1 year in femoral neck BMD in transmales									
1 observational study Joseph et al (2019)	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=39	None	Mean (SD), kg/m2 Baseline 0 772 (0 137) 1 year 0 765 (0 120) p=0 797 z-score Baseline −0 863 (1 215) 1 year −1 440 (1 075) p=0 000	IMPORTANT	VERY LOW
Change from baseline to 2 years in femoral neck BMD in transfemales									

115

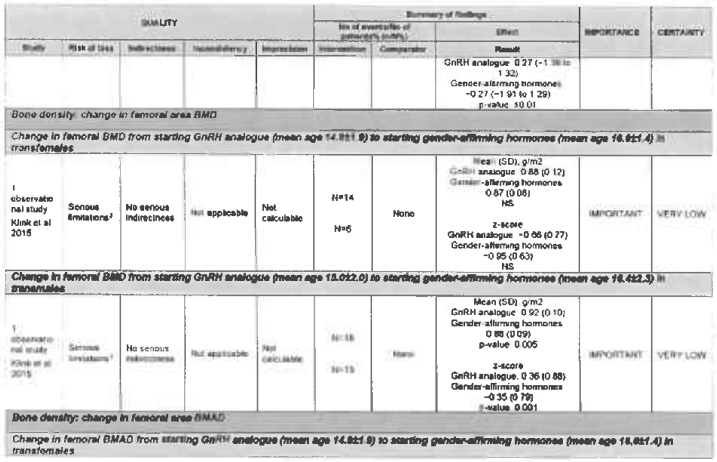


QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of events/No of patients (n/N%)		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result		
1 observational study Joseph et al (2019)	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=10	None	Mean (SD), kg/m2 Baseline 0 920 (0 116) 2 years 0 910 (0 125) p=0 402 z-score Baseline 0 450 (0 781) 2 years −0 600 (1 059) p=0 002	IMPORTANT	VERY LOW
Change from baseline to 2 years in femoral neck BMD in transmales									
1 observational study Joseph et al (2019)	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=21	None	Mean (SD), kg/m2 Baseline 0 766 (0 215) 2 years 0 773 (0 197) p=0 804 z-score Baseline −1 075 (1 145) 2 years −1 779 (0 816) p=0 001	IMPORTANT	VERY LOW
Bone density: change in femoral neck (hip) BMAD									
Change from starting GnRH analogue to starting gender-affirming hormones in femoral neck BMAD in transfemales (bone age of <15 years)									
1 observational study Vlot et al 2017	Serious limitations[3]	No serious indirectness	Not applicable	Not calculable	N=16	None	Median (range) g/cm3 GnRH analogue 0 29 (0 20 to 0 33) Gender-affirming hormones 0 27 (0 20 to 0 33) p≤0 1 z-score GnRH analogue −0 71 (−3 35 to 0 37) Gender-affirming hormones −1 32 (−3 39 to 0 21) p≤0 1	IMPORTANT	VERY LOW
Change in femoral neck BMAD from starting GnRH analogue to starting gender-affirming hormones in transfemales (bone age of ≥15)									

116



QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of events/No of patients (n/N)		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result		
1 observational study Viot et al. 2017	Serious limitations[2]	No serious indirectness	Not applicable	Not calculable	N=6	None	Median (range), g/cm3 GnRH analogue 0.30 (0.26 to 0.36) Gender-affirming hormones 0.30 (0.26 to 0.34) NS z-score GnRH analogue −0.44 (−1.37 to 0.93) Gender-affirming hormones −0.36 (−1.50 to 0.46) NS	IMPORTANT	VERY LOW
Change in femoral neck BMAD from starting GnRH analogue to starting gender-affirming hormones in transmales (bone age of <14 years)									
1 observational study Viot et al. 2017	Serious limitations[2]	No serious indirectness	Not applicable	Not calculable	N=10	None	Median (range), g/cm3 GnRH analogue 0.31 (0.26 to 0.36) Gender-affirming hormones 0.30 (0.22 to 0.35) NS z-score GnRH analogue −0.01 (−1.30 to 0.91) Gender-affirming hormones −0.37 (−2.26 to 0.47) NS	IMPORTANT	VERY LOW
Change in femoral neck BMAD from starting GnRH analogue to starting gender-affirming hormones in transmales (bone age of ≥14)									
1 observational study Viot et al. 2017	Serious limitations[2]	No serious indirectness	Not applicable	Not calculable	N=23	None	Median (range), g/cm3 GnRH analogue 0.33 (0.25 to 0.39) Gender-affirming hormones 0.30 (0.23 to 0.41) p-value ≤0.01 z-score	IMPORTANT	VERY LOW



QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of events/No of patients (n/N)		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result		
							GnRH analogue 0.27 (−1.36 to 1.32) Gender-affirming hormones −0.27 (−1.91 to 1.29) p-value ≤0.01		
Bone density: change in femoral area BMD									
Change in femoral BMD from starting GnRH analogue (mean age 14.8±1.9) to starting gender-affirming hormones (mean age 16.6±1.4) in transfemales									
1 observational study Klink et al 2018	Serious limitations[2]	No serious indirectness	Not applicable	Not calculable	N=14 N=6	None	Mean (SD), g/m2 GnRH analogue 0.88 (0.12) Gender-affirming hormones 0.87 (0.08) NS z-score GnRH analogue −0.06 (0.77) Gender-affirming hormones −0.95 (0.63) NS	IMPORTANT	VERY LOW
Change in femoral BMD from starting GnRH analogue (mean age 13.0±2.0) to starting gender-affirming hormones (mean age 16.4±2.3) in transmales									
1 observational study Klink et al 2018	Serious limitations[2]	No serious indirectness	Not applicable	Not calculable	N=15 N=15	None	Mean (SD), g/m2 GnRH analogue 0.92 (0.10) Gender-affirming hormones 0.88 (0.09) p-value 0.005 z-score GnRH analogue 0.36 (0.88) Gender-affirming hormones −0.35 (0.79) p-value 0.001	IMPORTANT	VERY LOW
Bone density: change in femoral area BMAD									
Change in femoral BMAD from starting GnRH analogue (mean age 14.8±1.9) to starting gender-affirming hormones (mean age 16.6±1.4) in transfemales									

QUALITY					Summary of findings: No of events/No of patients (n/N%)		Effect	IMPORTANCE	CERTAINTY
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result		
1 observational study Klink et al 2015	Serious limitations[2]	No serious indirectness	Not applicable	Not calculable	N=12 N=10	None	Mean (SD), g/cm3 GnRH analogue 0.28 (0.04) Gender-affirming hormones 0.28 (0.04) NS z-score GnRH analogue −0.93 (1.22) Gender-affirming hormones −1.57 (1.74) p value NS	IMPORTANT	VERY LOW
Change in femoral BMAD from starting GnRH analogue (mean age 15.0±2.0) to starting gender-affirming hormones (mean age 16.4±2.3) in transmales									
1 observational study Klink et al 2015	Serious limitations[2]	No serious indirectness	Not applicable	Not calculable	N=16 N=16	None	Mean (SD), g/cm3 GnRH analogue 0.32 (0.04) Gender-affirming hormones 0.31 (0.04) NS z-score GnRH analogue 0.01 (0.70) Gender-affirming hormones −0.26 (0.74) NS	IMPORTANT	VERY LOW

Abbreviations: BMAD, bone mineral apparent density; BMD, bone mineral density; GnRH, gonadotrophin releasing hormone; NS, not significant; SD, standard deviation

1 Downgraded 1 level - the cohort study by Joseph et al (2019) was assessed as at high risk of bias (poor quality overall, lack of blinding and no control group)
2 Downgraded 1 level - the cohort study by Klink et al (2015) was assessed as at high risk of bias (poor quality overall, lack of blinding, no randomisation, no control group and high number of participants lost to follow-up)
3 Downgraded 1 level - the cohort study by Vlot et al (2017) was assessed as at high risk of bias (poor quality overall, lack of blinding and no control)

Table 9 Question 2: For children and adolescents with gender dysphoria, what is the short-term and long-term safety of GnRH analogues compared with one or a combination of psychological support, social transitioning to the desired gender or no intervention? – cognitive development or functioning

QUALITY					Summary of findings: No of events/No of patients (n/N%)		Effect	IMPORTANCE	CERTAINTY
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result		
Cognitive development or functioning (1 cross-sectional study)									
IQ (4 subscales: arithmetic, vocabulary, picture arrangement, and block design) at a single time point between GnRH analogue treated and untreated transmales									
1 Cross-sectional study Staphorsius et al 2015	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=8 Mean (SD) 94.0 (10.3)	N=10 Mean (SD) 109.4 (21.2)	NR	IMPORTANT	VERY LOW
IQ (4 subscales: arithmetic, vocabulary, picture arrangement, and block design) at a single time point between GnRH analogue treated and untreated transmales									
1 Cross-sectional study Staphorsius et al 2015	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=12 Mean (SD) 95.8 (15.6)	N=10 Mean (SD) 98.5 (15.9)	NR	IMPORTANT	VERY LOW
Reaction time at a single time point between GnRH analogue treated and untreated transmales									
1 Cross-sectional study Staphorsius et al 2015	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=8 Mean (SD) 10.9 (4.1)	N=10 Mean (SD) 9.9 (3.1)	NR	IMPORTANT	VERY LOW
Reaction time at a single time point between GnRH analogue treated and untreated transmales									
1 Cross-sectional study	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=12 Mean (SD) 9.9 (3.1)	N=10 Mean (SD) 10.0 (2.0)	NR	IMPORTANT	VERY LOW

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of participants (studies)		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result		
Staphorsius et al 2015									
Accuracy at a single time point between GnRH analogue treated and untreated transfemales									
1 cohort study Staphorsius et al 2015	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=8 Mean (SD) 73 9 (9 1)	N=10 Mean (SD) 83 4 (9 5)	NR	IMPORTANT	VERY LOW
Accuracy at a single time point between GnRH analogue treated and untreated transmales									
1 cohort study Staphorsius et al 2015	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=12 Mean (SD) 85.7 (10.5)	N=10 Mean (SD) 88 8 (9 7)	NR	IMPORTANT	VERY LOW

Abbreviations: GnRH, gonadotrophin releasing hormone, NR, not reported, P, P-value, SD, Standard deviation

[1] Downgraded 1 level - the cohort study by Staphorsius et al (2015) was assessed as at high risk of bias (poor quality overall: lack of blinding and no randomisation)

Table 10: Question 2: In children and adolescents with gender dysphoria, what is the short-term and long-term safety of GnRH analogues compared with one or a combination of psychological support, social transitioning to the desired gender or no intervention? – other safety outcomes

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of participants (studies)		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result		
Other safety outcomes: change in serum creatinine									
Change in serum creatinine (micromol/l) between baseline and 1 year in transfemales									

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of participants (studies)		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result		
1 observational study Schagen et al 2016	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=28	None	Mean (SD) Baseline 70 (17) 1 year 66 (13) p-value 0 20	IMPORTANT	VERY LOW
Change in serum creatinine (umol/l) between baseline and 1 year in transmales									
1 observational study Schagen et al 2016	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=29	None	Mean (SD) Baseline 73 (8) 1 year 68 (13) p-value 0 01	IMPORTANT	VERY LOW
Other safety outcomes: liver enzymes									
Presence of elevated liver enzymes (AST, ALT, and glutamyl transferase) between baseline and during treatment									
1 observational study Schagen et al 2016	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	39	None	Glutamyl transferase was not elevated at baseline or during treatment in any subject Mild elevations of AST and ALT above the reference range were present at baseline but were not more prevalent during treatment than at baseline Glutamyl transferase, AST, and ALT levels did not significantly change from baseline to 12 months of treatment	IMPORTANT	VERY LOW
Proportion of patients reporting adverse effects									
1 cohort study Khatchadourian et al 2014	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable[2]	27	None	3/27 adolescents[3]	Important	VERY LOW

Abbreviations: ALT, alanine aminotransferase, AST, aspartate aminotransferase, GnRH, gonadotrophin releasing hormone, P, P-value, SD, standard deviation

1 Downgraded 1 level - the cohort study by Schagen et al (2016) was assessed as at high risk of bias (poor quality overall, lack of blinding and no control)
2 Downgraded 1 level - the cohort study by Khatchadourian et al (2014) was assessed as at high risk of bias (poor quality overall, lack of blinding, no control group and high number of participants lost to follow-up)
3 1 transmale developed sterile abscesses, they were switched from leuprolide acetate to triptorelin, and this was well tolerated. 1 transmale developed leg pains and headaches, which eventually resolved without treatment. 1 participant gained 15 kg within 9 months of initiating GnRH analogues

Table 11: Question 4. From the evidence selected, are there any subgroups of children and adolescents with gender dysphoria that may derive more (or less) advantage from treatment with GnRH analogues than the wider population of children and adolescents with gender dysphoria? – critical outcomes

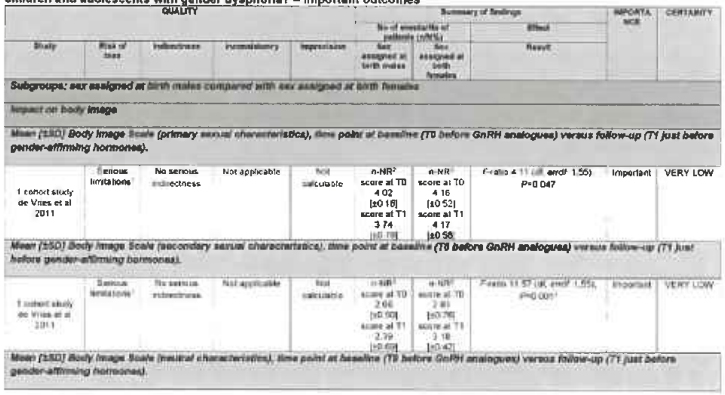
QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of events/No of patients (n/N%)		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Sex assigned at birth males	Sex assigned at birth females	Result		
Subgroups: sex assigned at birth males compared with sex assigned at birth females									
Impact on gender dysphoria									
Mean [±SD] Utrecht Gender Dysphoria Scale (version(s) not reported), time point at baseline (before GnRHa) versus follow-up (just before gender-affirming hormones)									
1 cohort study de Vries et al 2011	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	n-NR[2] score at T0 47.95 [±9.70] score at T1 49.67 [±9.47]	n-NR[2] score at T0 58.57 [±3.09] score at T1 58.62 [±4.0]	F-ratio 19.06 (df, error 1,39), P=0.001	Critical	VERY LOW
Impact on mental health									
Mean [±SD] Beck Depression Inventory-II, time point at baseline (T0 before GnRH analogues) versus follow-up (T1 just before gender-affirming hormones).									

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of events/No of patients (n/N%)		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Sex assigned at birth males	Sex assigned at birth females	Result		
1 cohort study de Vries et al 2011	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	n-NR[2] score at T0 5.71 [±4.31] score at T1 3.50 [±4.58]	n-NR[2] score at T0 10.34 [±8.24] score at T1 6.09 [±7.93]	F-ratio 3.85 (df, error 1,39), P=0.057	Critical	VERY LOW
Mean [±SD] Trait Anger (TP), time point at baseline (T0 before GnRH analogues) versus follow-up (T1 just before gender-affirming hormones).									
1 cohort study de Vries et al 2011	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	n-NR[2] score at T0 5.22 [±2.76] score at T1 5.00 [±3.07]	n-NR[2] score at T0 6.43 [±2.78] score at T1 6.39 [±2.59]	F-ratio 5.70 (df, error 1,39), P=0.022	Critical	VERY LOW
Mean [±SD] Trait Anxiety (STAI), time point at baseline (T0 before GnRH analogues) versus follow-up (T1 just before gender-affirming hormones).									
1 cohort study de Vries et al 2011	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	n-NR[2] score at T0 4.33 [±2.66] score at T1 4.39 [±2.64]	n-NR[2] score at T0 7.00 [±2.36] score at T1 6.17 [±2.69]	F-ratio 16.07 (df, error 1,39), P<0.001	Critical	VERY LOW

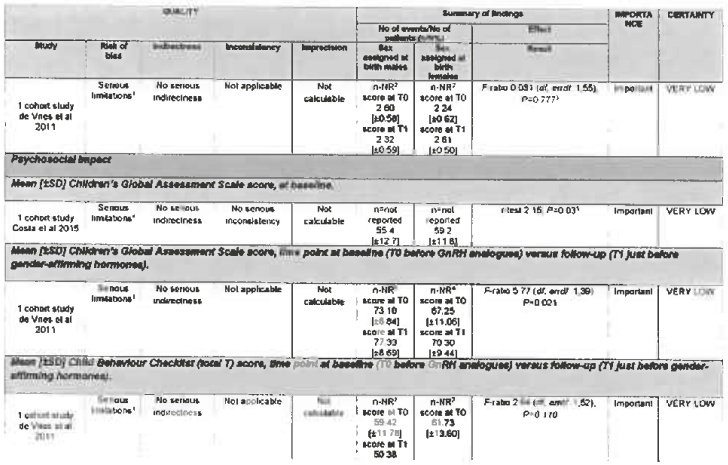
Abbreviations: GnRH, gonadotrophin releasing hormone, NR, not reported, P, P-value, SD, Standard deviation

1 Downgraded 1 level - the cohort study by de Vries et al (2011) was assessed as at high risk of bias (poor quality overall, lack of blinding and no control group)
2 The overall sample size completing the outcome at both time points was 41

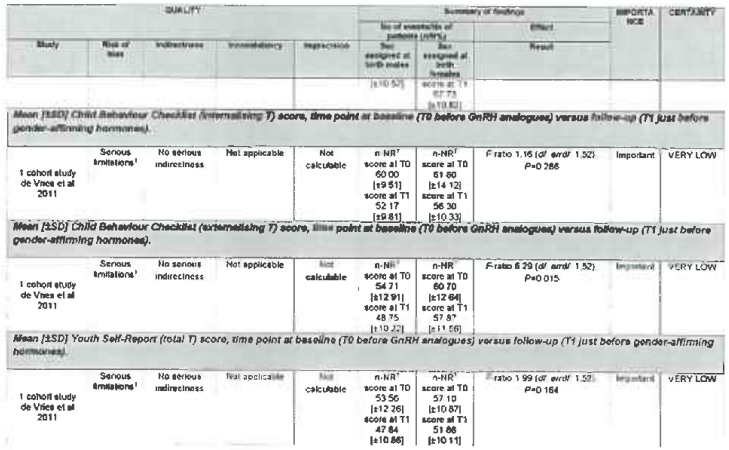
Table 11: Question: 4. From the evidence selected, are there any subgroups of children and adolescents with gender dysphoria that may derive more (or less) advantage from treatment with GnRH analogues than the wider population of children and adolescents with gender dysphoria? – Important outcomes



QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	No of events/No of patients (n/N%): Sex assigned at birth males	No of events/No of patients (n/N%): Sex assigned at birth females	Effect: Result		
Subgroups: sex assigned at birth males compared with sex assigned at birth females									
Impact on body image									
Mean [±SD] Body Image Scale (primary sexual characteristics), time point at baseline (T0 before GnRH analogues) versus follow-up (T1 just before gender-affirming hormones).									
1 cohort study de Vries et al 2011	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	n-NR[2] score at T0 4.02 [±0.18] score at T1 3.74 [±0.78]	n-NR[2] score at T0 4.16 [±0.52] score at T1 4.17 [±0.58]	F-ratio 4.11 (df, error df 1,55) P=0.047	Important	VERY LOW
Mean [±SD] Body Image Scale (secondary sexual characteristics), time point at baseline (T0 before GnRH analogues) versus follow-up (T1 just before gender-affirming hormones).									
1 cohort study de Vries et al 2011	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	n-NR[2] score at T0 2.66 [±0.90] score at T1 2.79 [±0.69]	n-NR[2] score at T0 2.81 [±0.76] score at T1 3.18 [±0.42]	F-ratio 11.57 (df, error df 1,55), P=0.001[3]	Important	VERY LOW
Mean [±SD] Body Image Scale (neutral characteristics), time point at baseline (T0 before GnRH analogues) versus follow-up (T1 just before gender-affirming hormones).									

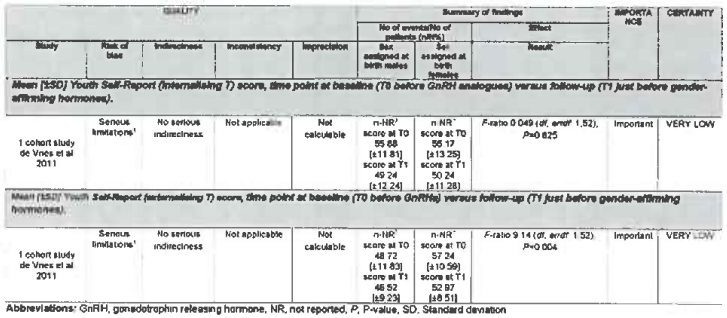


QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	No of events/No of patients (n/N%): Sex assigned at birth males	No of events/No of patients (n/N%): Sex assigned at birth females	Effect: Result		
1 cohort study de Vries et al 2011	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	n-NR[2] score at T0 2.60 [±0.58] score at T1 2.32 [±0.59]	n-NR[2] score at T0 2.24 [±0.62] score at T1 2.61 [±0.50]	F-ratio 0.031 (df, error df 1,55), P=0.777[3]	Important	VERY LOW
Psychosocial impact									
Mean [±SD] Children's Global Assessment Scale score, at baseline.									
1 cohort study Costa et al 2015	Serious limitations[1]	No serious indirectness	No serious inconsistency	Not calculable	n=not reported 55.4 [±12.7]	n=not reported 59.2 [±11.8]	ttest 2.15, P=0.03[3]	Important	VERY LOW
Mean [±SD] Children's Global Assessment Scale score, time point at baseline (T0 before GnRH analogues) versus follow-up (T1 just before gender-affirming hormones).									
1 cohort study de Vries et al 2011	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	n-NR[2] score at T0 73.10 [±8.64] score at T1 77.39 [±8.69]	n-NR[2] score at T0 67.25 [±11.06] score at T1 79.30 [±9.44]	F-ratio 5.77 (df, error df 1,39) P=0.021[3]	Important	VERY LOW
Mean [±SD] Child Behaviour Checklist (total T) score, time point at baseline (T0 before GnRH analogues) versus follow-up (T1 just before gender-affirming hormones).									
1 cohort study de Vries et al 2011	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	n-NR[2] score at T0 59.42 [±11.78] score at T1 50.38	n-NR[2] score at T0 61.73 [±13.60]	F-ratio 2.84 (df, error df 1,52), P=0.110	Important	VERY LOW



QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of events/No of patients (n/N*)		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Sex assigned at birth males	Sex assigned at birth females	Result		
						(±10.52)	score at T1 67.73 (±10.62)		
Mean [±SD] Child Behaviour Checklist (internalising T) score, time point at baseline (T0 before GnRH analogues) versus follow-up (T1 just before gender-affirming hormones).									
1 cohort study de Vries et al 2011	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	n-NR[2] score at T0 60.00 (±9.51) score at T1 52.17 (±9.81)	n-NR[2] score at T0 61.60 (±14.12) score at T1 56.30 (±10.33)	F-ratio 1.16 (df, erdf 1,52), P=0.286	Important	VERY LOW
Mean [±SD] Child Behaviour Checklist (externalising T) score, time point at baseline (T0 before GnRH analogues) versus follow-up (T1 just before gender-affirming hormones).									
1 cohort study de Vries et al 2011	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	n-NR[2] score at T0 54.21 (±12.91) score at T1 48.75 (±10.22)	n-NR[2] score at T0 60.70 (±12.64) score at T1 57.87 (±11.56)	F-ratio 6.29 (df, erdf 1,52), P=0.015	Important	VERY LOW
Mean [±SD] Youth Self-Report (total T) score, time point at baseline (T0 before GnRH analogues) versus follow-up (T1 just before gender-affirming hormones).									
1 cohort study de Vries et al 2011	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	n-NR[2] score at T0 53.56 (±12.26) score at T1 47.84 (±10.86)	n-NR[2] score at T0 57.10 (±10.87) score at T1 51.86 (±10.11)	F-ratio 1.99 (df, erdf 1,52), P=0.164	Important	VERY LOW

127



QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of events/No of patients (n/N*)		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Sex assigned at birth males	Sex assigned at birth females	Result		
Mean [±SD] Youth Self-Report (internalising T) score, time point at baseline (T0 before GnRH analogues) versus follow-up (T1 just before gender-affirming hormones).									
1 cohort study de Vries et al 2011	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	n-NR[2] score at T0 55.68 (±11.81) score at T1 49.24 (±12.24)	n-NR[2] score at T0 55.17 (±13.25) score at T1 50.24 (±11.28)	F-ratio 0.049 (df, erdf 1,52), P=0.825	Important	VERY LOW
Mean [±SD] Youth Self-Report (externalising T) score, time point at baseline (T0 before GnRHa) versus follow-up (T1 just before gender-affirming hormones).									
1 cohort study de Vries et al 2011	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	n-NR[2] score at T0 48.72 (±11.83) score at T1 46.52 (±9.23)	n-NR[2] score at T0 57.24 (±10.59) score at T1 52.97 (±8.51)	F-ratio 9.14 (df, erdf 1,52), P=0.004	Important	VERY LOW

Abbreviations: GnRH, gonadotrophin releasing hormone, NR, not reported, P, P-value, SD, Standard deviation

1 Downgraded 1 level - the cohort study by de Vries et al (2011) was assessed as at high risk of bias (poor quality overall, lack of blinding and no control group)
2 The overall sample size completing the outcome at both time points was 57
3 There was a significant interaction effect between sex assigned at birth and BDI between T0 and T1, sex assigned at birth females became more dissatisfied with their secondary F (df, erdf), P 14.59 (1,55), P<0.001) and neutral F (df, erdf), P 15.26 (1,55), P<0.001) sex characteristics compared with sex assigned at birth males
4 Serious limitations – the cohort study by Costa et al 2015 was assessed as at high risk of bias (poor quality)
5 At baseline, CGAS scores were not associated with any demographic variable, in both sex assigned at birth males and females. There were no statistically significant differences in CGAS scores between gender dysphoric sex assigned at birth males and females in all follow-up evaluations (P>0.1; full data not reported)
6 The overall sample size completing the outcome at both time points was 41
7 The overall sample size completing the outcome at both time points was 54

128

Glossary

Beck Depression Inventory-II (BDI-II)	The BDI-II is a tool for assessing depressive symptoms. There are no specific scores to categorise depression severity, but it is suggested that 0 to 13 is minimal symptoms, 14 to 19 is mild depression, 20 to 28 is moderate depression, and severe depression is 29 to 63.
Body Image Scale (BIS)	The BIS is used to measure body satisfaction. The scale consists of 30 body features, which the person rates on a 5-point scale. Each of the 30 items falls into one of 3 basic groups based on its relative importance as a gender-defining body feature: primary sex characteristics, secondary sex characteristics, and neutral body characteristics. A higher score indicates more dissatisfaction.
Bone mineral apparent density (BMAD)	BMAD is a size adjusted value of bone mineral density (BMD) incorporating body size measurements using UK norms in growing adolescents.
Child Behaviour Checklist (CBCL)	CBCL is a checklist parents complete to detect emotional and behavioural problems in children and adolescents.
Children's Global Assessment Scale (CGAS)	The CGAS tool is a validated measure of global functioning on a single rating scale from 1 to 100. Lower scores indicate poorer functioning
Gender	The roles, behaviours, activities, attributes, and opportunities that any society considers appropriate for girls and boys, and women and men.
Gender dysphoria	Discomfort or distress that is caused by a discrepancy between a person's gender identity (how they see themselves regarding their gender) and that person's sex assigned at birth (and the associated gender role, and/or primary and secondary sex characteristics)
Gonadotrophin releasing hormone (GnRH) analogues	GnRH analogues competitively block GnRH receptors to prevent the spontaneous release of 2 gonadotropin hormones, Follicular Stimulating Hormone (FSH) and Luteinising Hormone (LH) from the pituitary gland. The reduction in FSH and LH secretion reduces oestradiol secretion from the ovaries in those whose sex assigned at birth was female and testosterone secretion from the testes in those whose sex assigned at birth was male.
Sex assigned at birth	Sex assigned at birth (male or female) is a biological term and is based on genes and how external and internal sex and reproductive organs work and respond to hormones. Sex is the label that is recorded when a baby's birth is registered.
Tanner stage	Tanner staging is a scale of physical development
Trait Anger Spielberger scales of the State-Trait Personality Inventory (TPI)	The TPI is a validated 20-item inventory tool which measures the intensity of anger as the disposition to experience angry feelings as a personality trait. Higher scores indicate greater anger.
Transgender (including transmale and transfemale)	Transgender is a term for someone whose gender identity is not congruent with their birth-registered sex. A transmale is a person who identifies as male and a transfemale is a person who identifies as female.

Utrecht Gender Dysphoria Scale (UGDS)	The UGDS is a validated screening tool for both adolescents and adults to assess gender dysphoria. It consists of 12 items, to be answered on a 1- to 5-point scale, resulting in a sum score between 12 and 60. The higher the UGDS score the greater the impact on gender dysphoria.
Youth Self-Report (YSR)	The self-administered YSR is a checklist to detect emotional and behavioural problems in children and adolescents. It is self-completed by the child or adolescent. The scales consist of a Total problems score, which is the sum of the scores of all the problem items. An internalising problem scale sums the anxious/depressed, withdrawn-depressed, and somatic complaints scores while the externalising problem scale combines rule-breaking and aggressive behaviour.

References

Included studies

- Bnk T, Vrouenraets L, de Vries M et al. (2020) Trajectories of Adolescents Treated with Gonadotropin-Releasing Hormone Analogues for Gender Dysphoria. Archives of Sexual Behaviour. [Accessed 6 August 2020]
- Costa R, Dunsford M, Skagerberg E et al. (2015) Psychological Support, Puberty Suppression, and Psychosocial Functioning in Adolescents with Gender Dysphoria. Journal of Sexual Medicine. [online] Volume 12(11), Pages 2206-2214. Available at: https://doi.org/10.1111/jsm.13034 [Accessed 7 August 2020]
- de Vries A, Steensma T, Doreleijers T et al. (2011) Puberty Suppression in Adolescents with Gender Identity Disorder: A Prospective Follow-Up Study. The Journal of Sexual Medicine Volume 8, Issue 8, August, Pages 2276-2283. [Accessed 11 August 2020].
- Joseph T, Ting J, Butler G (2019) The effect of GnRH analogue treatment on bone mineral density in young adolescents with gender dysphoria: findings from a large national cohort. Journal of pediatric endocrinology & metabolism 32(10): 1077-1081
- Khatchadourian K, Shazhan A, Metzger D. (2014) Clinical Management of Youth with Gender Dysphoria in Vancouver. The Journal of Pediatrics. Volume 164, Issue 4, April, Pages 906-911. [Accessed 14 August 2020]
- Klink D, Caris M, Heijboer A et al. (2015) Bone mass in young adulthood following gonadotropin-releasing hormone analog treatment and cross-sex hormone treatment in adolescents with gender dysphoria. The Journal of clinical endocrinology and metabolism 100(2): e270-5
- Schagen SEE, Cohen-Kettenis PT, Delemarre-van de Waal HA et al. (2016) Efficacy and Safety of Gonadotropin-Releasing Hormone Agonist Treatment to Suppress Puberty in Gender Dysphoric Adolescents. The journal of sexual medicine 13(7): 1125-32
- Staphorsius A, Baudewijntje P, Kreukels P, et al. (2015) Puberty suppression and executive functioning: An fMRI-study in adolescents with gender dysphoria. Psychoneuroendocrinology Volume 565. Pages 190-199. [Accessed 10 August 2020]

- Vlot, Mariska C, Klink, Daniel T, den Heijer, Martin et al. (2017) Effect of pubertal suppression and cross-sex hormone therapy on bone turnover markers and bone mineral apparent density (BMAD) in transgender adolescents. Bone 95: 11-19

Other references

- World Health Organisation (2018) International Classification of Diseases 11. Available from https://icd.who.int/ [online; accessed 20 August 2020]
- American Psychiatric Association (2013) Diagnostic and statistical Manual of Mental Disorders (DSM-5) (5th ed). Washington, DC and London: American Psychiatric Publishing pp.451-460. [accessed 20 August 2020]
- NHS England (2013) NHS Standard contract for gender identity development service for children and adolescents [accessed 20 August 2020]

The NEW ENGLAND JOURNAL of MEDICINE

ORIGINAL ARTICLE

Psychosocial Functioning in Transgender Youth after 2 Years of Hormones

Diane Chen, Ph.D., Johnny Berona, Ph.D., Yee-Ming Chan, M.D., Ph.D., Diane Ehrensaft, Ph.D., Robert Garofalo, M.D., M.P.H., Marco A. Hidalgo, Ph.D., Stephen M. Rosenthal, M.D., Amy C. Tishelman, Ph.D., and Johanna Olson-Kennedy, M.D.

ABSTRACT

BACKGROUND

Limited prospective outcome data exist regarding transgender and nonbinary youth receiving gender-affirming hormones (GAH; testosterone or estradiol).

METHODS

We characterized the longitudinal course of psychosocial functioning during the 2 years after GAH initiation in a prospective cohort of transgender and nonbinary youth in the United States. Participants were enrolled in a four-site prospective, observational study of physical and psychosocial outcomes. Participants completed the Transgender Congruence Scale, the Beck Depression Inventory–II, the Revised Children's Manifest Anxiety Scale (Second Edition), and the Positive Affect and Life Satisfaction measures from the NIH (National Institutes of Health) Toolbox Emotion Battery at baseline and at 6, 12, 18, and 24 months after GAH initiation. We used latent growth curve modeling to examine individual trajectories of appearance congruence, depression, anxiety, positive affect, and life satisfaction over a period of 2 years. We also examined how initial levels of and rates of change in appearance congruence correlated with those of each psychosocial outcome.

RESULTS

A total of 315 transgender and nonbinary participants 12 to 20 years of age (mean [±SD], 16±1.9) were enrolled in the study. A total of 190 participants (60.3%) were transmasculine (i.e., persons designated female at birth who identify along the masculine spectrum), 185 (58.7%) were non-Latinx or non-Latine White, and 25 (7.9%) had received previous pubertal suppression treatment. During the study period, appearance congruence, positive affect, and life satisfaction increased, and depression and anxiety symptoms decreased. Increases in appearance congruence were associated with concurrent increases in positive affect and life satisfaction and decreases in depression and anxiety symptoms. The most common adverse event was suicidal ideation (in 11 participants [3.5%]); death by suicide occurred in 2 participants.

CONCLUSIONS

In this 2-year study involving transgender and nonbinary youth, GAH improved appearance congruence and psychosocial functioning. (Funded by the Eunice Kennedy Shriver National Institute of Child Health and Human Development.)

The authors' affiliations are listed in the Appendix. Dr. Chen can be contacted at dichen@luriechildrens.org or at the Ann and Robert H. Lurie Children's Hospital of Chicago, Potocsnak Family Division of Adolescent and Young Adult Medicine, 225 E. Chicago Ave., Box 161B, Chicago, IL 60611.

N Engl J Med 2023;388:240-50.
DOI: 10.1056/NEJMoa2206297

TRANSGENDER AND NONBINARY YOUTH comprise 2 to 9% of high-school–aged persons in the United States.[1-3] Many transgender and nonbinary youth have gender dysphoria, the persistent distress arising from incongruence between gender identity and external phenotype. Increasingly, transgender and nonbinary youth receive medical care to alleviate gender dysphoria, including gonadotropin-releasing hormone (GnRH) agonists to suppress gender-incongruent puberty and gender-affirming hormones (GAH; testosterone or estradiol) to foster gender-congruent secondary sex characteristics. An important goal of such treatment is to attenuate gender dysphoria by increasing appearance congruence — that is, the degree to which youth experience alignment between their gender and their physical appearance.

The available prospective research indicates that gender-affirming medical care is associated with improvements in psychosocial functioning.[4-9] Previously published studies with modest sample sizes[5,6,9] have examined outcomes for relatively short follow-up periods (approximately 1 year on average),[5,6,9] focused exclusively on outcomes of GnRH agonists,[7,8] or examined outcomes for mixed samples of youth initiating GnRH agonists or GAH,[4,6,9] despite evidence that such cohorts have distinct psychosocial profiles.[10] Evidence has been lacking from longitudinal studies that explore potential mechanisms by which gender-affirming medical care affects gender dysphoria and subsequent well-being.

We characterized the longitudinal course of psychosocial functioning over a period of 2 years after GAH initiation in a prospective cohort of more than 300 transgender and nonbinary young people in the United States. We hypothesized that appearance congruence, positive affect, and life satisfaction would increase and that depression and anxiety symptoms would decrease. We also hypothesized that improvements would be secondary to treatment for gender dysphoria, such that increasing appearance congruence would be associated with concurrent improvements in psychosocial outcomes. We also explored the potential moderating effects of demographic and clinical characteristics, including age, designated sex at birth, racial and ethnic identity, and the initiation of GAH in early as compared with later stages of puberty.

METHODS

STUDY DESIGN AND PARTICIPANT RECRUITMENT

Participants were recruited from gender clinics at the Ann and Robert H. Lurie Children's Hospital of Chicago, UCSF Benioff Children's Hospitals, Boston Children's Hospital, and Children's Hospital Los Angeles from July 2016 through June 2019 for the Trans Youth Care–United States (TYCUS) Study,[11] a prospective, observational study evaluating the physical and psychosocial outcomes of medical treatment for gender dysphoria in two distinct cohorts of transgender and nonbinary youth — those initiating GnRH agonists and those initiating GAH as part of their clinical care. All participating clinics employ a multidisciplinary team that includes medical and mental health providers and that collaboratively determines whether gender dysphoria is present and whether gender-affirming medical care is appropriate. For minors, parental consent is required to initiate medical treatment. Publications by individual study teams provide details on site-specific approaches to care.[12-15]

Study visits occurred at baseline and at 6, 12, 18, and 24 months after treatment initiation. Details on study procedures have been published previously,[11] and the protocol is available with the full text of this article at NEJM.org. The present analyses focus on the GAH cohort; outcomes for the cohort initiating GnRH agonists are being analyzed separately, given differences in baseline functioning between the two cohorts[10] and distinct outcomes of GnRH agonists[8] as compared with GAH treatment.[4] Participants provided written informed consent or assent; parents provided permission for minors to participate. Procedures were approved by the institutional review board at each study site.

The first and second authors analyzed the data and wrote the initial draft of the manuscript. All the authors critically reviewed the manuscript. The authors vouch for the accuracy and completeness of the data and for the fidelity of the study to the protocol. There were no agreements regarding confidentiality of the data among the sponsor (Eunice Kennedy Shriver National Institute of Child Health and Human Development), the authors, and the participating institutions. The sponsor had no role in the design of the study; the collection, analysis, or in-

terpretation of data; the writing of the manuscipt; or the decision to submit the manuscript for publication.

MEASURES

Participants reported age, racial and ethnic identity, gender identity, and designated sex at birth (details are provided in the Supplementary Appendix, available at NEJM.org). A small subgroup had been treated with GnRH agonists in early puberty (Tanner stage 2 or 3) (20 participants) or had a relatively late age at onset of endogenous puberty, such that they began receiving GAH in Tanner stage 3 (at 13 to 15 years of age) even without previous treatment with GnRH agonists (4 participants). These 24 participants comprise a subcohort in that they did not undergo extensive gender-incongruent puberty. Participants with a history of GnRH agonist treatment that was initiated in Tanner stage 4 (5 participants) were not included in this subcohort, because their experience of substantial gender-incongruent puberty is more similar to that of youth initiating GAH in Tanner stage 4 or 5.

With respect to longitudinal outcomes, participants completed the Transgender Congruence Scale,[16] the Beck Depression Inventory–II,[17] the Revised Children's Manifest Anxiety Scale (Second Edition),[18] and the Positive Affect and Life Satisfaction measures from the NIH (National Institutes of Health) Toolbox Emotion Battery[19] at each study visit. Scoring information and sample items from each scale are provided in the Supplementary Appendix. Higher scores on these measures reflect greater appearance congruence, depression, anxiety, positive affect, and life satisfaction, respectively.

STATISTICAL ANALYSIS

Trajectories of psychosocial functioning were examined with the use of repeated-measures multivariate analysis of variance and mixed-effects models. Multivariate analysis of variance provided a preliminary omnibus test for significant within-person change over time. Owing to listwise deletion, 150 participants were excluded from the multivariate analysis of variance (the analysis involved 141 participants). Mixed-effects modeling was therefore selected owing to greater flexibility in accommodating missing data and nonnormal distributions and examining parallel processes. Specifically, we used latent growth curve modeling, which uses a structural equation modeling framework to examine changes in mean scores over time.[20] Repeated measures are treated as indicators of latent factors: an intercept factor (estimates of initial levels) and a slope factor (rate of change). Intercept and slope factors can be regressed on covariates in adjusted models to explore moderation effects. In addition, growth curves for two different outcomes can be combined to examine how intercepts and slopes of those constructs correlate with each other. Data were Winsorized at the 95th percentile to reduce the influence of outliers.

Analyses involving latent growth curve modeling proceeded in three steps. First, we modeled trajectories of appearance congruence and psychosocial outcomes (i.e., effects of time only). Second, we adjusted models to estimate the effects of covariates on baseline scores and rates of change over time. Third, because changes in appearance congruence and psychosocial outcomes occur as parallel, simultaneous processes during GAH treatment, we examined how initial levels and rates of change in appearance congruence correlated with those of each psychosocial outcome. Standardized β levels were used as indicators of effect sizes for longitudinal models using conventional ranges (small, 0.20; medium, 0.50; and large, 0.80). Our conceptual model is shown in Figure S1 in the Supplementary Appendix. All statistical analyses were conducted with the use of SPSS software, version 27, and Mplus software, version 8.8.

RESULTS

ANALYTIC SAMPLE

There were a total of 6114 observations from 315 participants, who were assessed up to five times over a period of 2 years (data were available for 81% of all possible observations). Most participants (238 [75.6%]) completed either four study visits (76 participants) or five visits (162 participants). Tables S1 and S2 show the number of completed visits by time point and data coverage for key variables. The analytic sample for longitudinal models included 291 participants with follow-up data on primary outcome variables (Fig. S2). The analytic sample did not differ substantially from the overall sample with respect to age, designated sex at birth, racial and ethnic

identity, initiation of GAH in early puberty, or baseline scores on psychosocial measures (Table S3).

SAMPLE CHARACTERISTICS

We enrolled 315 eligible participants 12 to 20 years of age (mean [±SD], 16±1.9 years) (Table 1). Most were transmasculine (i.e., persons designated female at birth who identify along the masculine spectrum; 60.3%), designated female at birth (64.8%), and non-Latinx or non-Latine White (58.7%). Transmasculine, non-Latinx or non-Latine White, and multiracial participants were overrepresented and nonbinary and Black participants were underrepresented as compared with the study sample in the Williams Institute Executive Report[21] (Table S4); however, the study sample was representative of transgender and nonbinary youth presenting to pediatric subspecialty gender programs[22] and generalizable to this population. Two participants died by suicide during the study (one after 6 months of follow-up and the other after 12 months of follow-up), and 6 participants withdrew from the study. For these eight participants, data that had been collected before death or study withdrawal were included in the analyses. Data on adverse events are provided in Table 2.

APPEARANCE CONGRUENCE AND PSYCHOSOCIAL OUTCOMES OVER TIME

Table S5 depicts mean scores for appearance congruence, depression, anxiety, positive affect, and life satisfaction at baseline and 24 months. Results for multivariate analysis of variance indicated that there were significant within-participant changes over time for all psychosocial outcomes in hypothesized directions (Wilk's lambda, 0.32; F statistic with 20 and 122 degrees of freedom; 12.86; P<0.001). Specifically, scores for appearance congruence, positive affect, and life satisfaction increased significantly, and scores for depression and anxiety decreased significantly.

Means and variances of the variables for latent growth curve modeling, with estimated baseline levels and change over time for both time-only and adjusted models, are provided in Table 3. Scores for appearance congruence increased (annual increase on a 5-point scale, 0.48 points; 95% confidence interval [CI], 0.42 to 0.54; standardized $\beta=1.47$), as did T scores for positive affect (annual increase on a 100-point scale, 0.80 points; 95% CI, 0.08 to 1.54; $\beta=0.19$) and life satisfaction (annual increase on a 100-point scale, 2.32 points; 95% CI, 1.64 to 3.00; $\beta=0.52$). We observed decreased scores for depression (annual change on a 63-point scale, −1.27 points; 95% CI, −1.98 to −0.57; standardized $\beta=-0.29$) and decreased T scores for anxiety (annual change on a 100-point scale, −1.46 points; 95% CI, −2.13 to −0.79; $\beta=-0.35$) over a period of 2 years of GAH treatment.

Unadjusted models can be interpreted on their original scale. For instance, depression scores range from 0 to 63 (ranges of severity, minimal, 0 to 13; mild, 14 to 19; moderate, 20 to 28; and severe, 29 to 63). The model had an intercept (baseline mean) of 15.46 and estimated slope (change per year) of −1.27. Thus, on average, depression started in the mild range and decreased to the subclinical level by 24 months. Table S6 shows the percentages of youth scoring in the clinical range for depression and anxiety at each time point. Of 27 participants with depression scores in the severe range at baseline, 18 (67%) reported a depression score in the minimal or moderate ranges at 24 months. Similarly, 21 of 33 participants (64%) with depression scores in the moderate range at baseline reported a depression score in the minimal or moderate ranges at 24 months (chi-square statistic with 9 degrees of freedom, 49.85; P<0.001). With respect to anxiety, 47 of 122 participants (38.5%) with baseline scores in the clinical range (T scores, >60) were in the nonclinical range at 24 months (chi-square statistic with 1 degree of freedom, 22.05; P<0.001).

ASSOCIATIONS BETWEEN APPEARANCE CONGRUENCE AND PSYCHOSOCIAL OUTCOMES

Figure 1 depicts parallel processes between appearance congruence and each psychosocial outcome as analyzed by means of latent growth curve modeling. As described above, we used linear latent growth curve modeling to estimate baseline scores (intercepts) and linear rates of change (slopes) of each outcome (see Table 3 for details of each model). In parallel-process models, we examined how the components for latent growth curve modeling for appearance congruence related to those for scores for depression (Fig. 1A) and T scores for anxiety (Fig. 1B), positive affect (Fig. 1C), and life satisfaction

Table 1. **Demographic and Clinical Characteristics of the Participants.***

Characteristic	Participants (N=315)
	no. (%)
Gender identity†	
Transmasculine	190 (60.3)
Transfeminine	106 (33.7)
Nonbinary	19 (6.0)
Designated sex at birth	
Female	204 (64.8)
Male	111 (35.2)
Racial and ethnic identity	
Non-Latinx or non-Latine White	185 (58.7)
Latinx or Latine non-White	50 (15.9)
Latinx or Latine White	25 (7.9)
Black	11 (3.5)
Asian or Pacific Islander	10 (3.2)
Multiracial	32 (10.2)
Other	1 (0.3)
Unknown	1 (0.3)
Age at baseline	
12 yr	6 (1.9)
13 yr	23 (7.3)
14 yr	38 (12.1)
15 yr	67 (21.3)
16 yr	55 (17.5)
17 yr	51 (16.2)
18 yr	48 (15.2)
19 yr	15 (4.8)
20 yr	12 (3.8)
Tanner stage at GAH initiation‡	
1	2 (0.6)
2	13 (4.1)
3	9 (2.9)
4	29 (9.2)
5	262 (83.2)
Past use of GnRH agonist	
No	290 (92.1)
Yes	25 (7.9)
Tanner stage at initiation of GnRH agonist	
2	12 (3.8)
3	8 (2.5)
4	5 (1.6)
Not applicable	290 (92.1)
Initiation of GAH in early puberty subcohort§	
No	291 (92.4)
Yes	24 (7.6)

* The table does not include demographic and clinical characteristics for one participant who was accidentally enrolled and did not meet criteria for study eligibility. Percentages may not total 100 because of rounding. GAH denotes gender-affirming hormones, and GnRH gonadotropin-releasing hormone.

† Transmasculine refers to persons designated female at birth who identify along the masculine spectrum. Transfeminine refers to persons designated male at birth who identify along the feminine spectrum.

‡ Three participants began receiving GnRH agonists in either Tanner stage 2 or 3 and subsequently had pubertal regression to Tanner stage 1 or 2 by the time of GAH initiation.

§ This subcohort includes 20 participants who began receiving GnRH agonists at Tanner stage 2 or 3 and 4 participants who had not previously received GnRH agonists but had begun receiving GAH in Tanner stage 3 owing to a relatively late onset of puberty (13 to 15 years of age) and thus did not have physical changes associated with later stages of endogenous puberty. This subcohort does not include 5 participants with a history of initiation of GnRH agonists in Tanner stage 4 and who thus did undergo substantial gender-incongruent puberty.

(Fig. 1D). Higher appearance congruence at baseline was associated with lower baseline scores for depression (r=−0.60) and T scores for anxiety (r=−0.40), and increases in appearance congruence were associated with decreases in scores for depression (r=−0.68) and T scores for anxiety (r=−0.52) over time. In addition, higher appearance congruence at baseline was associated with higher baseline T scores for positive affect (r=0.46) and life satisfaction (r=0.72), and increases in appearance congruence were associated with increases in T scores for positive affect (r=0.74) and life satisfaction (r=0.84) over time.

MODERATING EFFECTS OF DEMOGRAPHIC AND CLINICAL COVARIATES

Table 3 shows the effects of covariates on scores for appearance congruence and depression and T scores for anxiety, positive affect, and life satisfaction. Age was not associated with any outcomes at baseline or over time.

Designated Sex at Birth

Depression and anxiety scores decreased among youth designated female at birth but not among those designated male at birth. Similarly, T scores for life satisfaction increased among youth designated female at birth but not among those designated male at birth (Fig. S3). Designated sex at birth was not associated with any other outcomes at baseline or over time.

Table 2. **Adverse Events.**

Event	No. of Events in Sample
Any event	15
Death by suicide	2
Suicidal ideation reported during study visit	11
Severe anxiety triggered by study visit	2

Effects of Racial and Ethnic Identity

At baseline, youth of color had higher scores for appearance congruence, lower scores for depression, and higher scores for positive affect than non-Latinx or non-Latine White youth. With respect to change over time, non-Latinx or non-Latine White youth had greater decreases in depression scores than youth of color (Fig. S4). Racial and ethnic identity were not associated with any other outcomes at baseline or over time.

Initiation of GAH in Early Puberty

Youth who had initiated GAH in early puberty had higher scores for appearance congruence, positive affect, and life satisfaction at baseline and lower scores for depression and anxiety at baseline than those who had initiated GAH in later puberty. Tables S7, S8, and S9 provide more information regarding differences between youth initiating GAH in early puberty and those initiating GAH in late puberty. With respect to change over time, youth initiating GAH in later puberty had greater improvements in appearance congruence than those initiating GAH in early puberty (Fig. 2).

DISCUSSION

Understanding the effect of GAH on the psychosocial outcomes of transgender and nonbinary youth would appear crucial, given the documented mental health disparities observed in this population,[10,15,23,24] particularly in the context of increasing politicization of gender-affirming medical care.[25] In our U.S.-based cohort of transgender and nonbinary youth treated with GAH, we found decreases in depression and anxiety symptoms and increases in positive affect and life satisfaction as assessed through validated instruments. Our findings are consistent with those of other longitudinal studies involving transgender and nonbinary youth receiving GAH, which showed reductions in depression[6,9] and anxiety[6] and increases in overall well-being[5] with small-to-moderate effects over a follow-up period of up to 1 year. We replicated these findings in a larger sample of racially and ethnically diverse transgender and nonbinary youth recruited from four geographically distinct regions in the United States and found sustained improvements over a period of 2 years.

Increasing appearance congruence is a primary goal of GAH, and we observed appearance congruence improve over 2 years of treatment. This was a moderate effect, and the strongest effect observed across our outcomes, consistent with the effect seen in research involving other samples, which has noted large effects of GAH on body image and small-to-moderate effects on mental health.[6] Appearance congruence was also associated with each psychosocial outcome assessed at baseline and during the follow-up period, such that increases in appearance congruence were associated with decreases in depression and anxiety symptoms and increases in positive affect and life satisfaction. These findings suggest that appearance congruence is a candidate mechanism by which GAH influences psychosocial functioning.

The importance of appearance congruence for psychosocial well-being is further highlighted by the effect of avoiding gender-incongruent pubertal changes. Youth who had not undergone substantial gender-incongruent puberty had higher scores for appearance congruence, positive affect, and life satisfaction and lower scores for depression and anxiety at baseline than youth who had undergone substantial endogenous puberty. These observations align with other published reports that earlier access to gender-affirming medical care is associated with more positive psychosocial functioning.[10,26] Alternatively, youth who first recognize their gender incongruence in adolescence may represent a distinct subgroup of transgender and nonbinary youth who have more psychosocial complexities than youth recognizing gender incongruence in childhood.[27]

The effects of GAH on some psychosocial outcomes varied on the basis of designated sex

Table 3. **Variable Estimates for Individual Latent Growth Curve Models of 2-Year Outcomes.***

Model	Appearance Congruence†	Depression‡	Anxiety§	Positive Affect¶	Life Satisfaction‖
			value (95% confidence interval)		
Unconditional model: time					
Intercept mean	2.99 (2.90 to 3.08)	15.46 (14.27 to 16.70)	59.58 (58.22 to 60.68)	42.93 (41.82 to 44.03)	40.12 (38.99 to 41.26)
Intercept variance	0.35 (0.27 to 0.50)	86.23 (68.13 to 106.85)	17.84 (11.38 to 24.54)	63.50 (46.23 to 81.79)	75.21 (59.76 to 93.98)
Slope mean	0.48 (0.42 to 0.54)	−1.27 (−1.98 to −0.57)	−1.46 (−2.13 to −0.79)	0.80 (0.08 to 1.54)	2.32 (1.64 to 3.00)
Slope variance	0.11 (0.07 to 0.15)	19.44 (12.23 to 27.14)	17.84 (11.38 to 24.54)	17.98 (9.25 to 27.57)	20.33 (14.12 to 27.70)
Conditional model					
Time					
Intercept mean	2.59 (1.91 to 3.27)	20.01 (10.79 to 29.48)	60.82 (53.56 to 67.95)	47.27 (38.93 to 55.81)	38.86 (29.90 to 47.75)
Intercept variance	0.32 (0.25 to 0.42)	80.92 (63.35 to 100.47)	114.74 (91.96 to 138.23)	56.96 (41.19 to 74.75)	71.93 (57.15 to 90.22)
Slope mean	0.51 (0.07 to 0.96)	−0.92 (−3.82 to −0.06)	−1.95 (−3.81 to −0.09)	1.79 (0.14 to 3.43)	4.54 (2.66 to 6.43)
Slope variance	0.10 (0.06 to 0.14)	18.81 (11.71 to 26.34)	18.37 (11.78 to 25.63)	17.97 (9.29 to 27.66)	19.74 (13.61 to 27.06)
Time-invariant effects on intercept					
Baseline age	0.02 (−002 to 0.06)	−0.23 (−0.08 to 0.36)	−0.20 (−0.78 to 0.38)	−0.32 (−0.84 to 0.21)	0.06 (−0.49 to 0.62)
Designated sex at birth**	−0.12 (−0.31 to 0.06)	1.74 (−0.69 to 4.09)	0.05 (−2.37 to 2.49)	−1.26 (−3.53 to 0.91)	−2.36 (−4.89 to 0.18)
Racial and ethnic identity††	0.19 (0.03 to 0.36)	−2.60 (−4.82 to −0.32)	−2.22 (−4.48 to 0.06)	2.30 (0.22 to 4.38)	1.70 (−0.58 to 3.98)
Early gender-affirming care‡‡	0.70 (0.35 to 1.04)	−5.88 (−9.67 to −1.96)	−7.41 (−11.30 to −3.52)	5.34 (1.70 to 8.98)	7.55 (2.82 to 12.28)
Time-invariant effects on slope					
Baseline age	0.00 (−0.03 to 0.03)	−0.04 (−0.18 to 0.10)	−0.02 (−0.15 to 0.12)	−0.03 (−0.15 to 0.10)	−0.09 (−0.22 to 0.05)
Designated sex at birth**	0.03 (−0.09 to 0.15)	1.91 (0.33 to 3.50)	1.56 (0.01 to 3.10)	−0.43 (−2.10 to 1.31)	−1.86 (−3.49 to −0.24)
Racial and ethnic identity††	−0.10 (−0.20 to 0.01)	1.70 (0.23 to 3.15)	0.62 (−0.77 to 1.98)	−1.42 (−2.98 to 0.13)	−1.08 (−2.52 to 0.36)
Early gender-affirming care‡‡	−0.42 (−0.66 to −0.19)	−0.73 (−3.41 to 1.93)	0.04 (−2.53 to 2.59)	−0.78 (−3.56 to 2.06)	−1.08 (−4.01 to 1.86)

* Shown are unstandardized variable estimates with 95% confidence intervals. Slope means indicate change over time, and slope variances indicate heterogeneity within the sample.
† Scores on the Appearance Congruence subscale of the Transgender Congruence Scale range from 1 to 5, with higher scores indicating greater appearance congruence.
‡ Scores on the Beck Depression Inventory–II range from 0 to 63, with scores of 20 to 28 indicating moderate depression and scores of 29 to 63 indicating severe depression.
§ T scores on the Revised Children's Manifest Anxiety Scale (Second Edition) have a mean of 50 and a standard deviation of 10, with scores of 60 or more indicating clinical levels of anxiety.
¶ T scores for the Positive Affect measure from the NIH (National Institutes of Health) Toolbox Emotion Battery have a mean of 50 and a standard deviation of 10, with higher scores indicating greater positive affect.
‖ T scores for the Life Satisfaction measure from the NIH Toolbox Emotion Battery have a mean of 50 and a standard deviation of 10, with higher scores indicating greater life satisfaction.
** Coding for designated sex at birth was as follows: 0=assigned female at birth (reference) and 1=assigned male at birth.
†† Coding for racial and ethnic identity was as follows: 0=non-Latinx or non-Latine White (reference) and 1=other racial and ethnic identities.
‡‡ Coding for early gender-affirming care was as follows: 0=initiated GAH in later puberty (Tanner stage 4 or 5) (reference) and 1=initiated GAH in early puberty (Tanner stage 2 or 3).

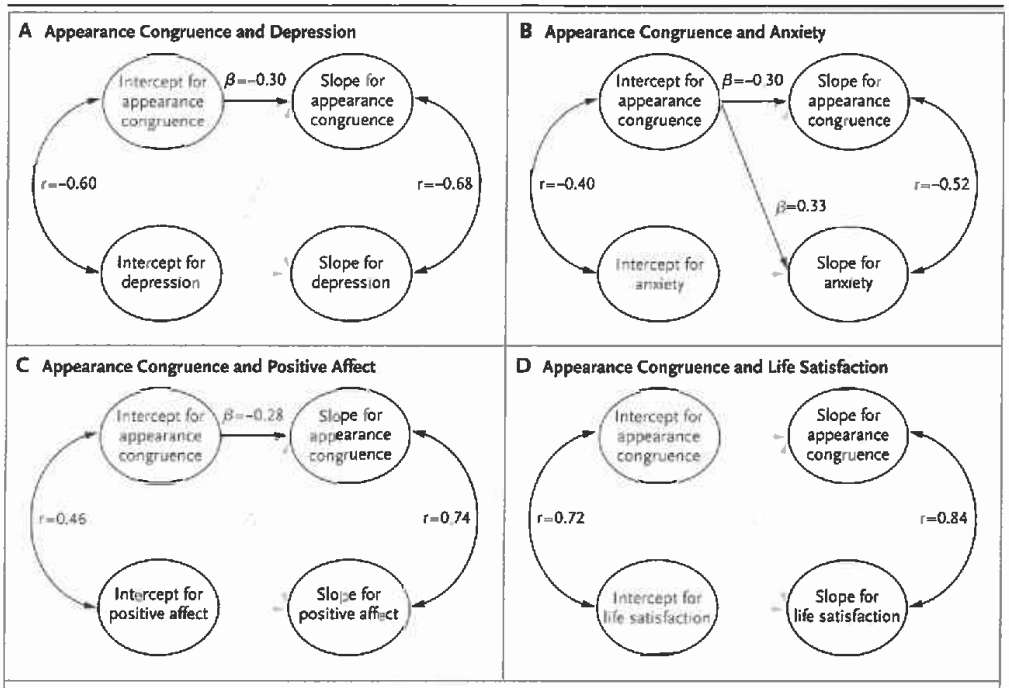


Figure 1. **Appearance Congruence and Depression, Anxiety, Positive Affect, and Life Satisfaction.**
Parallel-process latent growth curve models are depicted. A linear latent growth curve model was fitted for each outcome, with model-based estimates of baseline scores (intercept) and rates of linear change over time (slope). Parallel-process models can provide tests of how aspects of trajectories relate to each other. Each panel provides estimates for correlations between baseline scores of appearance congruence and each outcome (intercept correlations, arcs displayed on the left side of each panel), correlations between rate of change of appearance congruence and rate of change of each outcome (slope correlations, arcs displayed on the right side of each panel), and effects of baseline scores on slopes (straight lines in the middle of each panel). Solid black lines and arcs indicate significant effects (confidence intervals for variable estimates do not contain 0); nonsignificant effects are shown with dashed gray lines. All models were controlled for age, designated sex at birth, racial and ethnic identity, and early gender-affirming care (not shown for ease of interpretation).

at birth. Depression and anxiety symptoms decreased significantly, and life satisfaction increased significantly, among youth designated female at birth but not among those designated male at birth. Given that some key estrogen-mediated phenotypic changes can take between 2 and 5 years to reach their maximum effect (e.g., breast growth),[28] we speculate that a longer follow-up period may be necessary to see an effect on depression, anxiety, and life satisfaction. Furthermore, changes that are associated with an endogenous testosterone-mediated puberty (e.g., deeper voice) may be more pronounced and observable than those associated with an endogenous estrogen-mediated puberty. Thus, we hypothesize that observed differences in depression, anxiety, and life satisfaction among youth designated female at birth as compared with those designated male at birth may be related to differential experiences of gender minority stress, which could arise from differences in societal acceptance of transfeminine (i.e., persons designated male at birth who identity along the feminine spectrum) as compared with transmasculine persons. Indeed, gender minority stress is consistently associated with more negative mental health outcomes,[29] and research suggests that transfeminine youth may experience more minority stress than transmasculine youth.[30]

Our study has certain limitations. Because participants were recruited from four urban pediatric gender centers, the findings may not be generalizable to youth without access to comprehensive interdisciplinary services or to transgen-

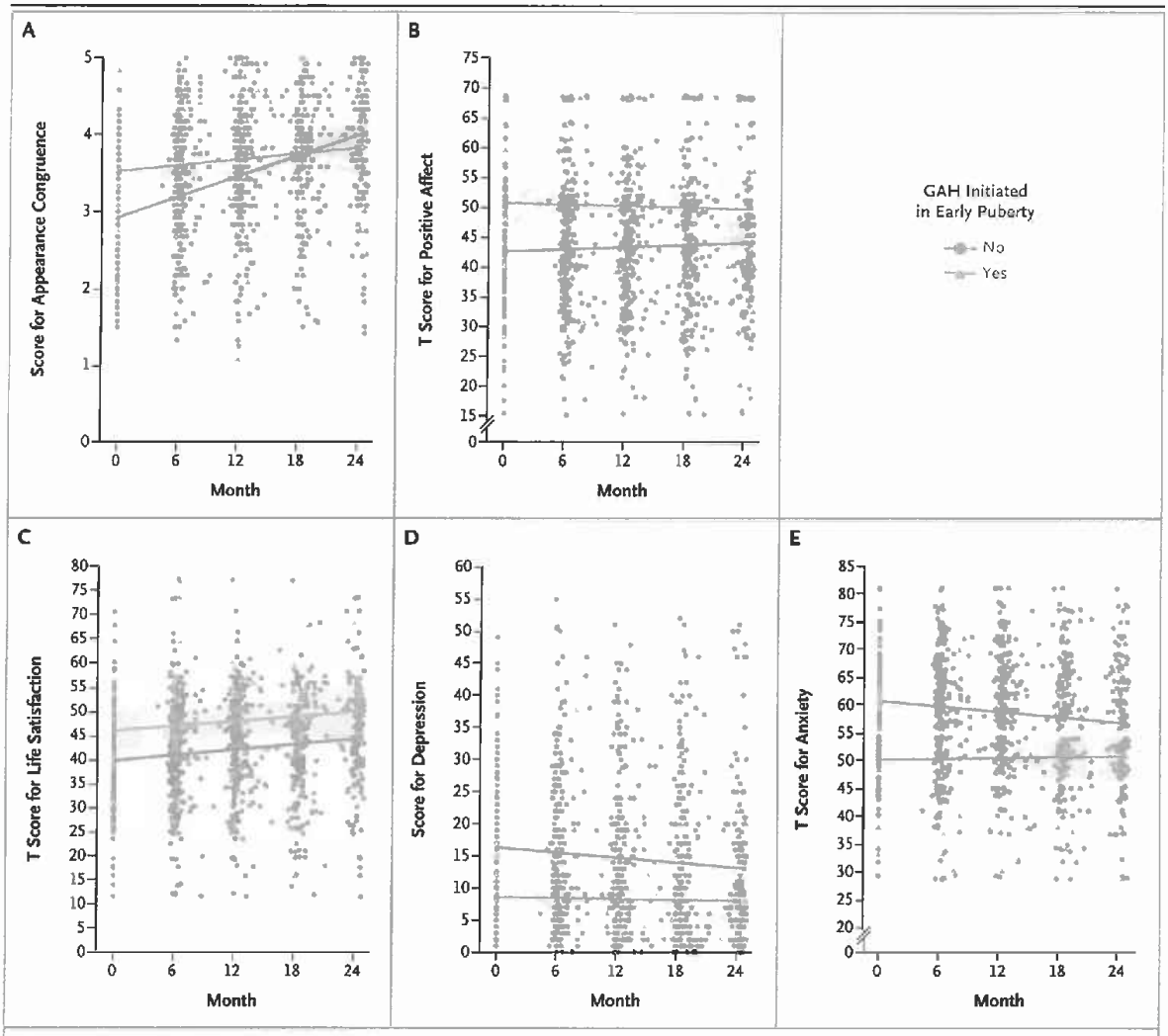


Figure 2. **Psychosocial Outcomes during 2 Years of GAH.**

Shown are changes in participant-reported measures over a period of 2 years of treatment with gender-affirming hormones (GAH). Scores on the Appearance Congruence subscale of the Transgender Congruence Scale (Panel A) range from 1 to 5, with higher scores indicating greater appearance congruence. T scores for the Positive Affect measure from the NIH (National Institutes of Health) Toolbox Emotion Battery (Panel B) range from 0 to 100, with higher scores indicating greater positive affect. T scores for the Life Satisfaction measure from the NIH Toolbox Emotion Battery (Panel C) range from 0 to 100, with higher scores indicating greater life satisfaction. Scores on the Beck Depression Inventory–II (Panel D) range from 0 to 63, with higher scores indicating greater depression. T scores on the Revised Children's Manifest Anxiety Scale (Second Edition) (Panel D), range from 0 to 100, with higher scores indicating greater anxiety. Individual scores are depicted with orange triangles for youth initiating GAH in early puberty ("Yes") and with blue circles for youth who did not initiate GAH in early puberty ("No"). Lines indicate mean scores for each group, with gray shaded bands for 95% confidence intervals.

der and nonbinary youth who are self-medicating with GAH. In addition, despite improvement across psychosocial outcomes on average, there was substantial variability around the mean trajectory of change. Some participants continued to report high levels of depression and anxiety and low positive affect and life satisfaction, despite the use of GAH. We plan to examine other factors that are known to contribute to psychosocial functioning among transgender and non-

binary youth and may not be affected by GAH, such as parental support,[31,32] in this cohort. Finally, our study lacked a comparison group, which limits our ability to establish causality. However, the large effects in parallel-process models examining associations between improvements in appearance congruence and improvements in psychosocial outcomes provide support for the concept that GAH may affect psychosocial outcomes through increasing gender congruence.

Despite these limitations, our findings showed improvements in psychosocial functioning across 2 years of GAH treatment, which supports the use of GAH as effective treatment for transgender and nonbinary youth. We are now following this cohort to see whether gains in functioning are sustained over a longer follow-up period, and — given substantial variability in outcomes even after controlling for a number of factors — we hope to discover additional predictors of change to identify youth for whom GAH alone is not adequate to address mental health challenges. We intend to initiate further work with this cohort to focus on understanding reasons for discontinuing GAH among the small subgroup of youth who stopped medical treatment. Overall, our results provide evidence that GAH improved appearance congruence and psychosocial functioning in transgender and nonbinary youth.

Supported by a grant (R01 HD082554) from the Eunice Kennedy Shriver National Institute of Child Health and Human Development.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

We thank the participants, their families, their referring clinicians, and the many research staff for their contributions in conducting this study, and Norman Spack, one of the original principal investigators, for his contributions to the study.

APPENDIX

The authors' affiliations are as follows: the Gender and Sex Development Program, Potocsnak Family Division of Adolescent and Young Adult Medicine (D.C., R.G.), and the Pritzker Department of Psychiatry and Behavioral Health (D.C.), Ann and Robert H. Lurie Children's Hospital of Chicago, the Departments of Pediatrics (D.C., R.G.) and Psychiatry and Behavioral Sciences (D.C., J.B.), Northwestern University Feinberg School of Medicine, and the Institute for Sexual and Gender Minority Health and Wellbeing, Northwestern University (J.B.) — all in Chicago; the Division of Endocrinology, Department of Pediatrics, Boston Children's Hospital (Y.-M.C.), and the Department of Pediatrics, Harvard Medical School (Y.-M.C.), Boston, and the Department of Psychology and Neuroscience, Boston College, Newton (A.C.T.) — all in Massachusetts; the Department of Pediatrics, Division of Pediatric Endocrinology (D.E., S.M.R.), and the Child and Adolescent Gender Center, Benioff Children's Hospital (D.E., S.M.R.), University of California, San Francisco, San Francisco, and the Gender Health Program, UCLA Health (M.A.H.), and the Division of General Internal Medicine and Health Services Research, Medicine–Pediatrics Section, Department of Medicine, David Geffen School of Medicine (M.A.H.), University of California, Los Angeles, the Center for Transyouth Health and Development, Division of Adolescent and Young Adult Medicine, Children's Hospital Los Angeles (J.O.-K.), and the Department of Pediatrics, Keck School of Medicine, University of Southern California (J.O.-K.), Los Angeles — all in California.

REFERENCES

1. Johns MM, Lowry R, Andrzejewski J, et al. Transgender identity and experiences of violence victimization, substance use, suicide risk, and sexual risk behaivors among high school students — 19 states and large urban school districts, 2017. MMWR Morb Mortal Wkly Rep 2019;68:67-71.
2. Rider GN, McMorris BJ, Gower AL, Coleman E, Eisenberg ME. Health and care utilization of transgender and gender nonconforming youth: a population-based study. Pediatrics 2018;141(3):e20171683.
3. Kidd KM, Sequeira GM, Douglas C, et al. Prevalence of gender-diverse youth in an urban school district. Pediatrics 2021; 147(6):e2020049823.
4. de Vries AL, McGuire JK, Steensma TD, Wagenaar ECF, Doreleijers TA, Cohen-Kettenis PT. Young adult psychological outcome after puberty suppression and gender reassignment. Pediatrics 2014; 134:696-704.
5. Allen LR, Watson LB, Egan AM, Moser CN. Well-being and suicidality among transgender youth after gender-affirming hormones. Clin Pract Pediatr Psychol 2019;7:302-11.
6. Kuper LE, Stewart S, Preston S, Lau M, Lopez X. Body dissatisfaction and mental health outcomes of youth on gender-affirming hormone therapy. Pediatrics 2020;145(4):e20193006.
7. Costa R, Dunsford M, Skagerberg E, Holt V, Carmichael P, Colizzi M. Psychological support, puberty suppression, and psychosocial functioning in adolescents with gender dysphoria. J Sex Med 2015; 12:2206-14.
8. de Vries AL, Steensma TD, Doreleijers TA, Cohen-Kettenis PT. Puberty suppression in adolescents with gender identity disorder: a prospective follow-up study. J Sex Med 2011;8:2276-83.
9. Achille C, Taggart T, Eaton NR, et al. Longitudinal impact of gender-affirming endocrine intervention on the mental health and well-being of transgender youths: preliminary results. Int J Pediatr Endocrinol 2020;2020:8.
10. Chen D, Abrams M, Clark L, et al. Psychosocial characteristics of transgender youth seeking gender-affirming medical treatment: baseline findings from the trans youth care study. J Adolesc Health 2021;68:1104-11.
11. Olson-Kennedy J, Chan Y-M, Garofalo R, et al. Impact of early medical treatment for transgender youth: protocol for the longitudinal, observational trans youth care study. JMIR Res Protoc 2019;8(7):e14434.
12. Chen D, Hidalgo MA, Leibowitz S, et al. Multidisciplinary care for gender-diverse youth: a narrative review and unique model of gender-affirming care. Transgend Health 2016;1:117-23.
13. Sherer I, Rosenthal SM, Ehrensaft D, Baum J. Child and Adolescent Gender Center: a multidisciplinary collaboration

to improve the lives of gender nonconforming children and teens. Pediatr Rev 2012;33:273-5.
14. Spack NP, Edwards-Leeper L, Feldman HA, et al. Children and adolescents with gender identity disorder referred to a pediatric medical center. Pediatrics 2012;129:418-25.
15. Olson J, Schrager SM, Belzer M, Simons LK, Clark LF. Baseline physiologic and psychosocial characteristics of transgender youth seeking care for gender dysphoria. J Adolesc Health 2015;57:374-80.
16. Kozee HB, Tylka TL, Bauerband LA. Measuring transgender individuals' comfort with gender identity and appearance: development and validation of the Transgender Congruence Scale. Psychol Women Q 2012;36:179-96.
17. Beck AT, Steer RA, Brown GK. BDI-II, Beck depression inventory. San Antonio, TX: Psychological Corporation, 1996.
18. Reynolds CR, Richmond BO. Revised children's manifest anxiety scale (RCMAS-2), second edition. Torrance, CA: Western Psychological Services, 2008.
19. Slotkin J, Nowinski CJ, Hays RD, et al. NIH Toolbox Scoring and Interpretation Guide. Washington, DC: National Institutes of Health, September 18, 2012 (https://repository.niddk.nih.gov/media/studies/look-ahead/Forms/Look_AHEAD_Cognitive_Function/NIH%20Toolbox%20Scoring%20and%20Interpretation%20Manual%209-27-12.pdf).
20. Muthén LK, Muthén BO. Mplus user's guide, version 8. 2017 (https://www.statmodel.com/download/usersguide/MplusUserGuideVer_8.pdf).
21. Herman JL, Flores AR, O'Neill KK. How many adults and youth identify as transgender in the United States? UCLA School of Law, June 2022 (https://williamsinstitute.law.ucla.edu/publications/trans-adults-united-states/).
22. Chen D, Lash B, Kim E, et al. A comparison of demographic and psychosocial characteristics between transgender youth enrolling versus not enrolling in a multisite study. Transgend Health 2020;6:229-34.
23. Reisner SL, Vetters R, Leclerc M, et al. Mental health of transgender youth in care at an adolescent urban community health center: a matched retrospective cohort study. J Adolesc Health 2015;56:274-9.
24. Toomey RB, Syvertsen AK, Shramko M. Transgender adolescent suicide behavior. Pediatrics 2018;142(4):e20174218.
25. American Civil Liberties Union. Legislation affecting LGBT rights across the country. December 17, 2021 (https://www.aclu.org/legislation-affecting-lgbt-rights-across-country).
26. Sorbara JC, Chiniara LN, Thompson S, Palmert MR. Mental health and timing of gender-affirming care. Pediatrics 2020;146(4):e20193600.
27. de Vries ALC. Challenges in timing puberty suppression for gender-nonconforming adolescents. Pediatrics 2020;146(4):e2020010611.
28. Coleman E, Bockting W, Botzer M, et al. Standards of care for the health of transsexual, transgender, and gender-nonconforming people, version 7. Int J Transgenderism 2012;13:165-232.
29. Delozier AM, Kamody RC, Rodgers S, Chen D. Health disparities in transgender and gender expansive adolescents: a topical review from a minority stress framework. J Pediatr Psychol 2020;45:842-7.
30. Poquiz JL, Coyne CA, Garofalo R, Chen D. Comparison of gender minority stress and resilience among transmasculine, transfeminine, and nonbinary adolescents and young adults. J Adolesc Health 2021;68:615-8.
31. Simons L, Schrager SM, Clark LF, Belzer M, Olson J. Parental support and mental health among transgender adolescents. J Adolesc Health 2013;53:791-3.
32. Pariseau EM, Chevalier L, Long KA, Clapham R, Edwards-Leeper L, Tishelman AC. The relationship between family acceptance-rejection and transgender youth psychosocial functioning. Clin Pract Pediatr Psychol 2019;7:267-77.

Supplementary Appendix

Supplement to: Chen D, Berona J, Chan Y-M, et al. Psychosocial functioning in transgender youth after 2 years of hormones. N Engl J Med 2023;388:240-50. DOI: 10.1056/NEJMoa2206297

This appendix has been provided by the authors to give readers additional information about the work.

Table of Contents

METHODS

Measures

Demographic and Clinical Characteristics

Participants self-reported age, race/ethnicity, gender identity, and designated sex at birth. For age, participants were asked "How old are you?" For race/ethnicity, between the start of the study and May 2018, participants were asked "With which racial or ethnic group do you most closely identify? (Choose one) and provided with the following options: (a) American Indian or Alaska Native; (b) Asian; (c) Black or African American; (d) Hispanic or Latino; (e) Native Hawaiian or Other Pacific Islander; (f) White; (g) Other. After May 2018, participants were asked "What race or ethnicity are you? Check all that apply" and provided with the following options: (a) American Indian or Alaska Native; (b) Asian; (c) Black or African American; (d) Hispanic or Latino; (e) Native Hawaiian or other Pacific Islander; (f) White; (g) other. Those selecting "other" were asked to specify race or ethnicity in free text form. Participant responses were recoded into the following: (a) non-Latinx/Latine White; (b) Latinx/Latine, non-White; (c) Latinx/Latine, White; (d) Black/African American; (e) Asian/Pacific Islander; (f) Multiracial; (g) other; and (h) Unknown.

For gender identity, youth either selected from eight response options [male, female, transgender female (male-to-female), transgender male (female-to-male), gender fluid, gender queer, bigender, or nonbinary] or indicated "other" and specified. Responses were recoded into three categories: transmasculine, transfeminine, and nonbinary. For designated sex at birth, participants were asked "What was your assigned sex at birth?" with male and female as response options.

Longitudinal Outcomes

Appearance Congruence. Appearance congruence was captured through the 9-item appearance congruence subscale of the Transgender Congruence Scale.[1] Each item was rated on a 5-point scale from "strongly disagree" to "strongly agree" and averaged. Example items include: "My outward appearance represents my gender identity" and "I am happy with the way my appearance expresses my gender identity". Higher scores reflect greater appearance congruence.

Depression Symptoms. Depression symptoms were assessed using the 21-item Beck Depression Inventory-II (BDI-II).[2] Each item was rated on a 4-point scale, summed and compared to standardized cutoffs reflecting minimal (0-13), mild (14-19), moderate (20-28), or severe depression symptoms (29-63).

Anxiety Symptoms. Anxiety symptoms were assessed by the Revised Children's Manifest Anxiety Scale, Second Edition (RCMAS2).[3] Forty-nine items were rated "yes"/ "no". "Yes" responses were tallied and transformed into a *T* score; for this scale *T* scores >60 are considered clinically significant.

Positive Affect. Positive affect was assessed using the 10-item Positive Affect measure from the National Institutes of Health (NIH) Toolbox—Emotion Battery.[4] Participants were asked to rate how frequently they experienced a variety of positive feelings over the past seven days. Example items include "I felt joyful" and "I felt content". Each item was rated on a 5-point scale from 1 = "not at all" to 5 = "very much". Raw scores were summed and converted to *T* scores; higher scores indicate greater positive affect.

Life Satisfaction. Life satisfaction was assessed using the 10-item General Life Satisfaction measure from the NIH Toolbox—Emotion Battery.[4] Participants were asked to rate how much they agree or disagree with statements about their personal well-being. Example items

include "If I could live my life over, I would change almost nothing," "I have what I want in life," and "My life is going well." Each item was rated on a 5-point scale from "strongly disagree" to "strongly agree". Raw scores were summed and converted to *T* scores; higher scores indicate greater life satisfaction.

Rationale for Selecting Primary Mental Health Outcome Measures

The Trans Youth Care—United States (TYCUS) study used various measures to assess different domains of mental health and psychosocial functioning,[1] including the Youth Self-Report (YSR),[2] a widely used child-report measure that assesses problem behaviors along two "broadband scales" (Internalizing, Externalizing) and eight empirically-based syndrome and DSM-oriented scales and provides a Total Problems score, and the age-appropriate version of the MINI International Neuropsychiatric Interview (MINI)[3] or the MINI International Neuropsychiatric Interview for Children and Adolescents (MINI-KID).[4] We chose to use the BDI-II and RCMAS2 as our primary mental health outcome measures in this paper as they are more granular than the YSR and have clinical thresholds that aid in interpretation of findings. Furthermore, the YSR and MINI/MINI-KID were administered annually (baseline, 12-month, and 24-month) versus the BDI-II and RCMAS2 which were administered every 6 months. Having more datapoints to model change across time allowed us to explore whether change in these outcomes were non-linear in nature. Future work using the YSR and MINI/MINI-KID data will allow for comparison across samples, as these measures are widely used among other study teams.[5,6]

Statistical Analysis Plan

Missing Data

At least four out of five total time points were available for 75% of participants (Table S1). As a result, there was high covariance coverage with data available for the majority of the sample for each variable of interest at all time points (range of data present: 0.66-0.99; Table S2). Within our sample, data exhibited skew and were determined to be missing at random (Little's MCAR test: χ^2 [751] = 803.25, p = 0.09).[5,6] This type of missing data can be appropriately handled using maximum likelihood estimation methods (described below).

Longitudinal Modeling Approach

Analyses were conducted in a latent growth curve modeling (LGCM) framework using Mplus 8.8.[7] This approach provides a unified modeling framework with several pertinent computational techniques including specification of hierarchical data structure, accommodation of missing data, and integration of both maximum likelihood and Bayesian estimation techniques. Consistent with NEJM recommendations, we handled missing data using model-based methods.[8] More specifically, LGCM was conducted with a two-stage estimation process in which starting values were generated for parameter estimates using full-information maximum likelihood estimation (FIML) followed by optimization using the Bayes estimator. The Bayes estimator was used in the second stage optimization as it is recommended for use when variables of interest exhibit non-normal distributions.[9,10] Bayesian estimation uses Markov chain Monte Carlo (MCMC) resampling algorithms and do not require large sample sizes.[11,12] These methods accommodate multilevel models that would otherwise be computationally intractable due to small sample sizes, modest effect sizes, and skewed response distributions.[13]

Model Specifications

Latent growth curves were generated for each variable of interest. Linear and quadratic effects of time were explored for inclusion. In all cases, quadratic effects were either non-significant (i.e., confidence intervals included 0) or had small parameter estimates that did not alter interpretation of results. For parsimony, all growth curves included intercepts and linear slopes. Intercept priors were estimated based on median values from observed data. Models employed MCMC algorithms to generate a series of 50,000 random draws from 4 stationary Markov chains to approximate the multivariate posterior distribution of our sample, with a burn-in period of 2,500 iterations. Model convergence was determined by the Gelman-Rubin potential scale reduction factor (PSR) values, with values close to 1 indicating convergence.[14] Trace plots were also inspected to evaluate model fit. All PSR values (range: 1.01-1.03) and trace plots indicated that the models converged and fit the data well.

Table S1. Count of Visits Completed

Visits	n	Proportion present
1	12	0.04
2	27	0.09
3	38	0.11
4	76	0.24
5	162	0.51

Proportion present is out of N=315 eligible participants.

Table S2. Data Coverage for Key Variables

	Baseline		Month 6		Month 12		Month 18		Month 24	
Variable	n	present*	n	present	n	present	n	present	n	present
AC	310	0.98	283	0.90	249	0.79	212	0.67	221	0.70
BDI	307	0.97	281	0.89	248	0.79	210	0.67	219	0.70
RCMAS	308	0.98	282	0.90	248	0.79	209	0.66	216	0.69
NPA	311	0.99	284	0.90	250	0.79	211	0.67	223	0.71
NLS	312	0.99	282	0.90	250	0.79	210	0.67	224	0.71

Note. Proportion present is out of N=315 eligible participants. AC = appearance congruence. BDI = Beck Depressive Inventory. RCMAS = Revised Children's Manifest Anxiety Scale. NPA = NIH Toolbox Positive Affect. NLS = NIH Toolbox Life Satisfaction
*present= proportion present.

Table S3. Comparison of Analytic Sample (n=291) and Participants Excluded from Longitudinal Analysis (n=24)

	t	df	*p*	Cohen's *d*
Baseline Age	0.28	26.27	0.78	0.06
Appearance Congruence	-0.63	25.58	0.54	-0.13
Depression	1.99	22.17	0.06	0.48
Anxiety	1.02	21.42	0.32	0.24
Positive Affect	-0.09	23.07	0.93	-0.02
Life Satisfaction	-1.56	24.03	0.13	-0.35
	c^2	df	*p*	f
Designated sex	0.47	1	0.49	0.04
Early gender-affirming care	0.44	1	0.51	0.04
Racial/ethnic identity	0.002	1	0.97	0.002

Note. For continuous variables, negative *t*-scores and Cohen's *d* indicate higher scores among participants excluded from longitudinal analysis.

Table S4. Representativeness of Study Participants

Category	Example
Disease, problem, or condition under investigation	People who identify as transgender in the U.S.
Special considerations related to:	
Sex and gender	Of the estimated 1.3 million transgender adults, 38.5% are transgender women, 35.9% are transgender men, and 25.6% are nonbinary.
Age	Youth ages 13 to 17 comprise 7.6% of the U.S. population and represent 18% of the transgender population in the U.S. Youth ages 18 to 24 comprise 11% of the U.S. population and represent 24.4% of the transgender population in the U.S. Approximately 1.4% of youth ages 13 to 17 and 1.3% of youth ages 18 to 24 identify as transgender.
Race or ethnic group	The racial/ethnic distribution of youth and adults who identify as transgender appears generally similar to the U.S. population, though transgender youth and adults are more likely to report being Latinx and less likely to report being White compared to the U.S. population. Among youth ages 13 to 17, white youth represent 51.3% of the U.S. population and 46.3% of transgender youth are white. Black youth represent 13.4% of the U.S. population and 13.2% of transgender youth are Black. Asian youth represent 5% of the U.S. population and 3.6% of transgender youth are Asian. American Indian or Alaska Native (AIAN) youth represent 0.8% of the U.S. population and 1% of transgender youth are AIAN. Latinx youth represent 24.8% of the U.S. population and 31% of transgender youth are Latinx. Multiracial youth represent 4.7% of the U.S. population and 5% of transgender youth are multiracial.
Geography	Percentage of residents in U.S. regions who identify as transgender range from 1.8% in the Northeast to 1.2% in the Midwest for youth ages 13 to 17. At the state level, estimates range from 3% of youth ages 13-17 identifying as transgender in New York to 0.6% in Wyoming.
Other considerations	In the last decade, the number of youth presenting for gender-affirming medical care has increased exponentially. In addition, the number of youth reporting a nonbinary identity also has increased significantly in recent years.
Overall representativeness of this trial	Transmasculine participants are over-represented in our study and non-binary participants are under-represented. Non-Latinx white and multiracial participants are over-represented in our sample, whereas Black participants are vastly under-represented in our sample. The proportion of Latinx and Asian participants are

	comparable to population estimates. Because study recruitment occurred at 4 study sites in the Northeast, Midwest, and California, youth in the Southeastern and Southwestern United States are not represented in the sample.

Note. Numbers are predominately pulled from the most recent Williams Institute Executive Summary "How many adults and youth identify as transgender in the United States" published in June 2022 by Jody L. Herman, Andrew R. Flores, and Kathryn K. O'Neill.

Table S5. Paired Samples *t*-tests Comparing Scores at Baseline and 24 Months

	n	baseline	24 Months	*p*-value	effect size
Appearance congruence	213	2.86 (0.74)	3.86 (0.76)	<0.001	-1.12
Depression	211	16.39 (11.88)	13.95 (12.76)	<0.001	0.20
Anxiety	208	60.25 (11.18)	57.38 12.00)	<0.001	0.25
Positive affect	215	42.90 (10.05)	43.72 (12.03)	0.37	-0.05
Life satisfaction	217	39.92 (10.55)	44.61 (12.29)	<0.001	-0.39

Note. Variables are presented as mean (SD). Results are based on *t*-tests (baseline minus 24-months). Negative *t*-test values indicate increases in appearance congruence, positive affect, and life satisfaction. Effect sizes are Cohen's *d* (ranges: 0.20, small; 0.50, medium; 0.80, large).

Table S6. Proportions of Youth Scoring in the Clinical Range for Depression and Anxiety at Each Timepoint

	Baseline	**6-month**	**12-month**	**18-month**	**24-month**
Beck Depression Inventory-II n (%)	*n*=307	*n*=281	*n*=248	*n*=210	*n*=219
Minimal Depression	149 (48.5)	152 (54.1)	143 (57.7)	125 (59.5)	126 (57.5)
Mild Depression	53 (17.3)	46 (16.4)	41 (16.5)	25 (11.9)	41 (18.7)
Moderate Depression	57 (18.6)	43 (15.3)	24 (9.7)	30 (14.3)	22 (10)
Severe Depression	48 (15.6)	40 (14.2)	40 (16.1)	30 (14.3)	30 (13.7)
Revised Children's Manifest Anxiety Scale 2	*n*=308	*n*=282	*n*=248	*n*=209	*n*=216
M (*SD*)	60.0 (11.5)	58.6 (11.6)	58.6 (11.3)	56.8 (11.4)	57.4 (12.1)
n (%) in Clinical range (*T*>60)	181 (58.8)	145 (51.4)	115 (46.4)	90 (43.1)	103 (47.7)

Note. % calculated as valid percent using the n for each timepoint as the denominator.

Table S7. Independent Samples *t*-tests Comparing Baseline Scores between Youth Initiating GAH in Early versus Late Puberty

	Total sample	Early gender-affirming care			
		Yes	No		
	N=315	n = 24	n = 291	*p*-value	effect size
Appearance congruence	2.36 (0.88)	3.08 (0.95)	2.31 (0.85)	<0.001	0.86
Depression	16.44 (12.11)	9.57 (8.26)	17.00 (12.21)	<0.001	0.71
Anxiety	60.03 (11.48)	51.54 (12.20)	60.75 (11.15)	<0.001	0.79
Positive affect	43.05 (10.78)	50.27 (12.08)	42.47 (10.49)	<0.001	0.69
Life satisfaction	39.76 (10.85)	44.90 (14.13)	39.35 (10.46)	0.08	0.45

Note. Variables are presented as mean (SD). Results are based on *t*-tests. Effect sizes are Cohen's *d* (ranges: 0.20, small; 0.50, medium; 0.80, large).

Table S8. Independent Samples *t*-tests Comparing Baseline Scores between Youth Initiating GAH in Early versus Late Puberty Among Youth Designated Male at Birth

	DMAB	Early gender-affirming care			
		Yes	No		
	n=111	n = 20	n = 91	*p*-value	Effect Size
Appearance congruence	2.27 (1.03)	3.09 (1.02)	2.10 (0.95)	<0.001	1.00
Depression	17.52 (13.35)	9.41 (8.70)	19.23 (13.56)	<0.001	0.86
Anxiety	59.12 (11.47)	52.30 (11.94)	60.67 (10.85)	0.008	0.73
Positive affect	42.06 (12.68)	51.24 (12.70)	40.14 (11.87)	0.002	0.90
Life satisfaction	38.82 (13.47)	45.71 (15.20)	37.38 (12.71)	0.04	0.59

Note. DMAB = designated male at birth. Variables are presented as mean (SD). Results are based on *t*-tests. Effect sizes are Cohen's *d* (ranges: 0.20, small; 0.50, medium; 0.80, large).

Table S9. Independent Samples *t*-tests Comparing Baseline Scores between Youth Initiating GAH in Early versus Late Puberty among Youth Designated Female at Birth

	DFAB	Early gender-affirming care			
		Yes	No		
	n=204	n = 4	n = 200	*p*-value	Effect Size
Appearance congruence	2.42 (0.78)	3.04 (0.56)	2.40 (0.77)	0.11	0.94
Depression	15.85 (11.36)	10.32 (6.69)	15.96 (11.42)	0.19	0.60
Anxiety	60.52 (11.48)	47.75 (14.66)	60.78 (11.30)	0.17	1.00
Positive affect	43.59 (9.59)	45.65 (8.19)	43.55 (9.62)	0.65	0.24
Life satisfaction	40.27 (9.10)	41.08 (7.43)	40.25 (9.14)	0.84	0.10

Note. DFAB = designated female at birth. Variables are presented as mean (SD). Results are based on *t*-tests. Effect sizes are Cohen's *d* (ranges: 0.20, small; 0.50, medium; 0.80, large).

Figure S1 Conceptual Model of Parallel Process Latent Growth Curve Models

Conceptual model of parallel process latent growth curve models. Rectangles indicate measured variables. Ovals represent model-based estimates of baseline scores (intercepts) and linear rates of change (slopes). Straight arrows indicate regression paths to model (1) moderating effects of baseline covariates on growth curve intercepts and slopes and (2) effects of intercepts on slopes. Curved arrows represent correlations between intercepts and slopes.

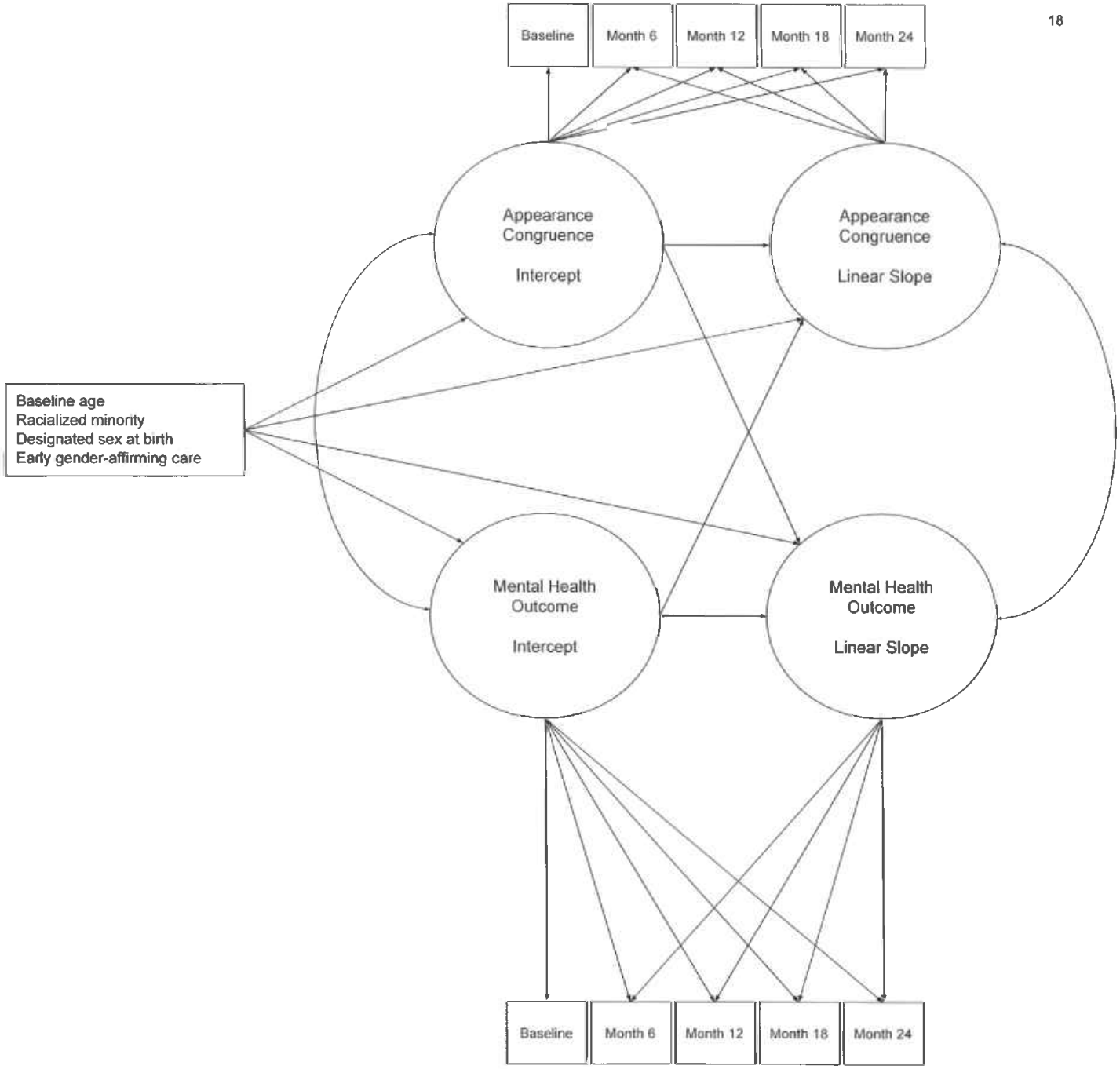

Baseline
Month 6
Month 12
Month 18
Month 24
Appearance Congruence Intercept
Appearance Congruence Linear Slope
Baseline age
Racialized minority
Designated sex at birth
Early gender-affirming care
Mental Health Outcome Intercept
Mental Health Outcome Linear Slope
Baseline
Month 6
Month 12
Month 18
Month 24

Figure S2 Consort Diagram

Flow diagram of the progress through the phases of a prospective, observational study, including enrollment, follow-up, and data analysis for latent growth curve models.

Figure S2. Consort diagram

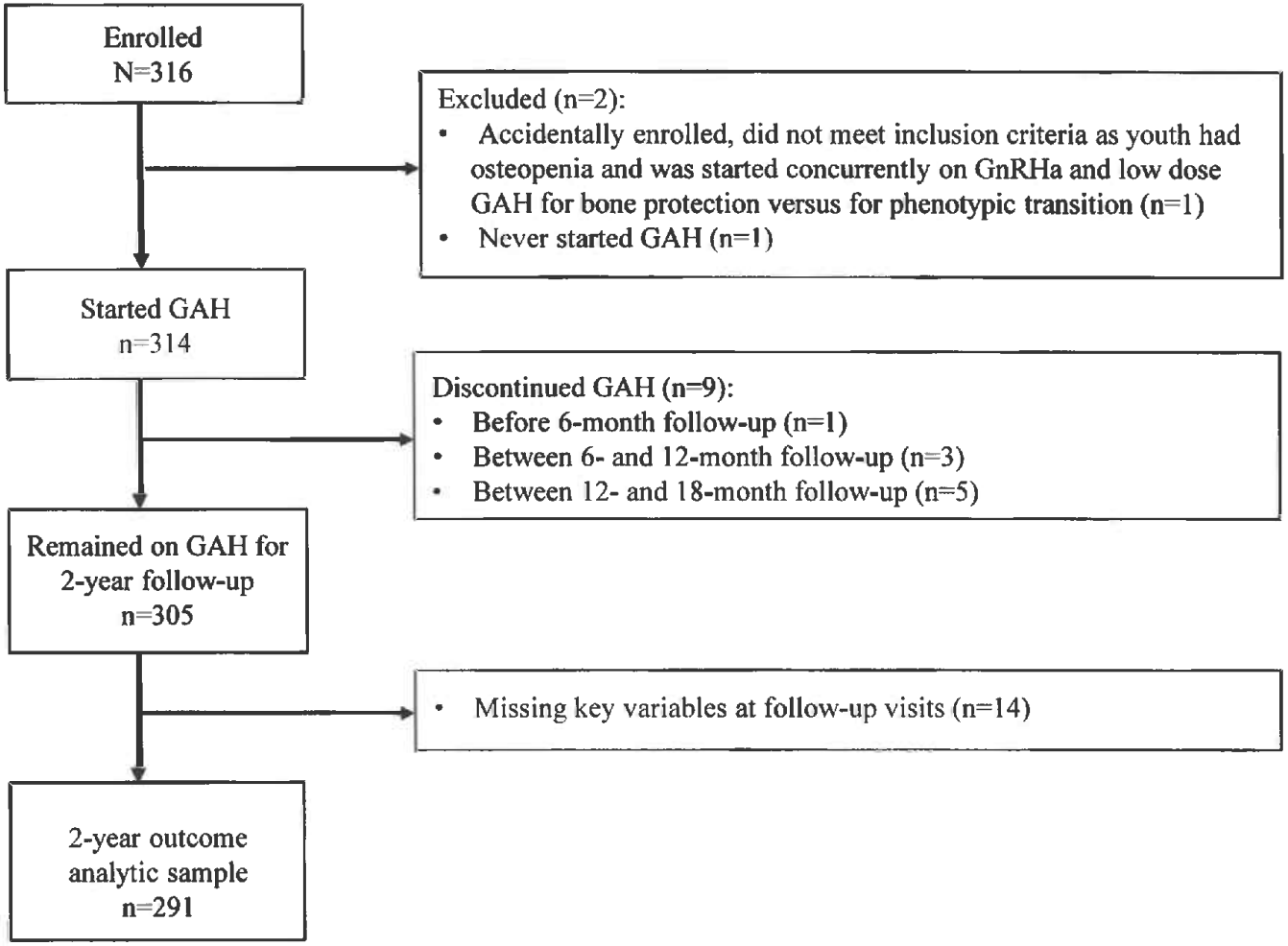


Figure S3 Change in Psychosocial Outcomes by Designated Sex at Birth

Figure panels display changes in psychosocial outcomes over two years of GAH by designated sex at birth (designated female at birth: blue circles; designated male at birth: orange triangles). Lines indicate mean scores for each group with gray shaded bands for 95% confidence intervals. Outcomes shown are as follows: (S3-A) Transgender Congruence Scale, range: 1-5; (S3-B) Positive Affect Scale T-Score (NIH Toolbox), range: 0-100; (S3-C) Life Satisfaction T-Score (NIH Toolbox), range 0-100); (S3-D) Beck Depression Inventory-II, range: 0-63; (S3-E) Revised Children's Manifest Anxiety Scale, Second Edition T-Score, range: 0-100.

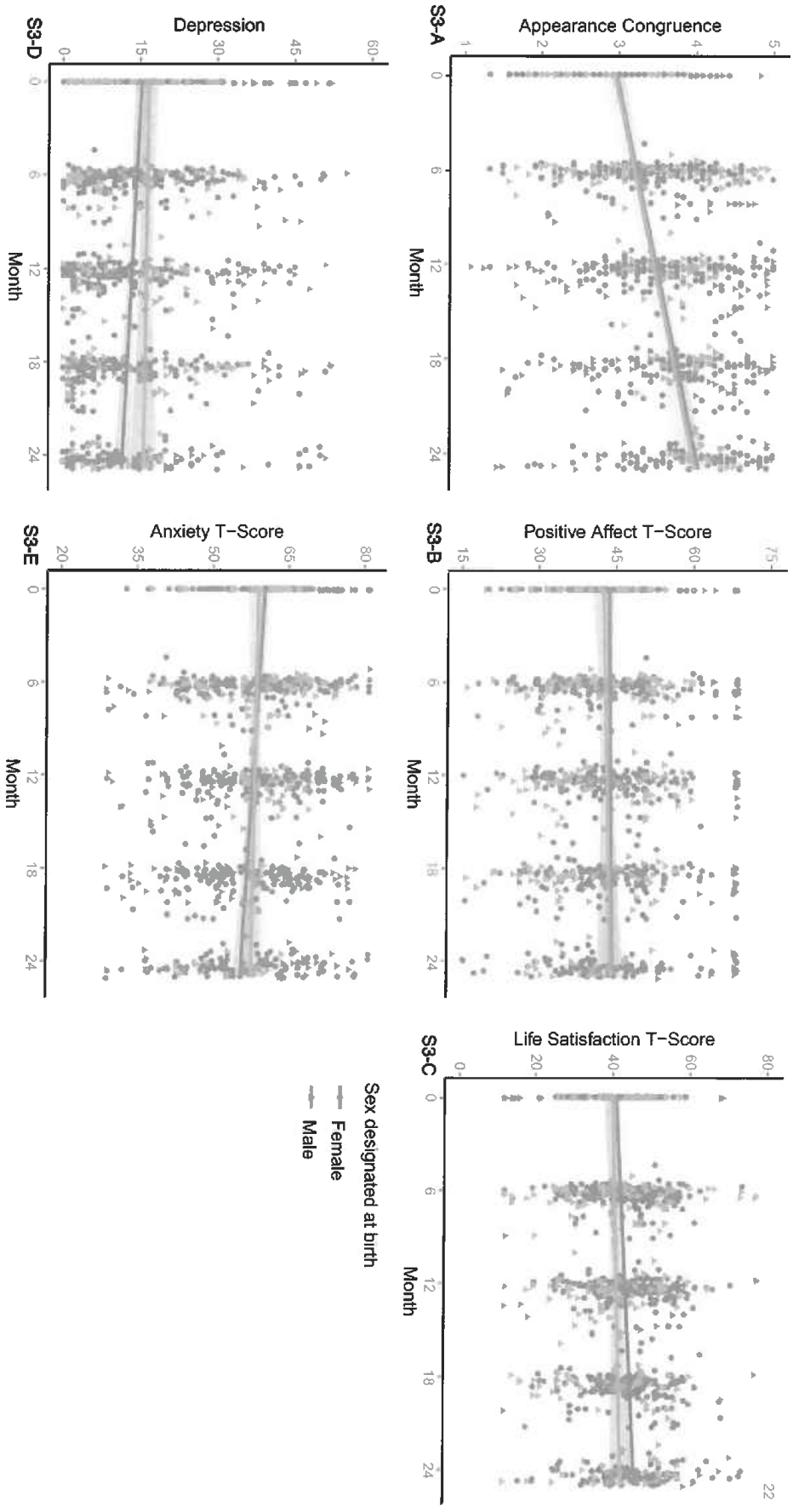
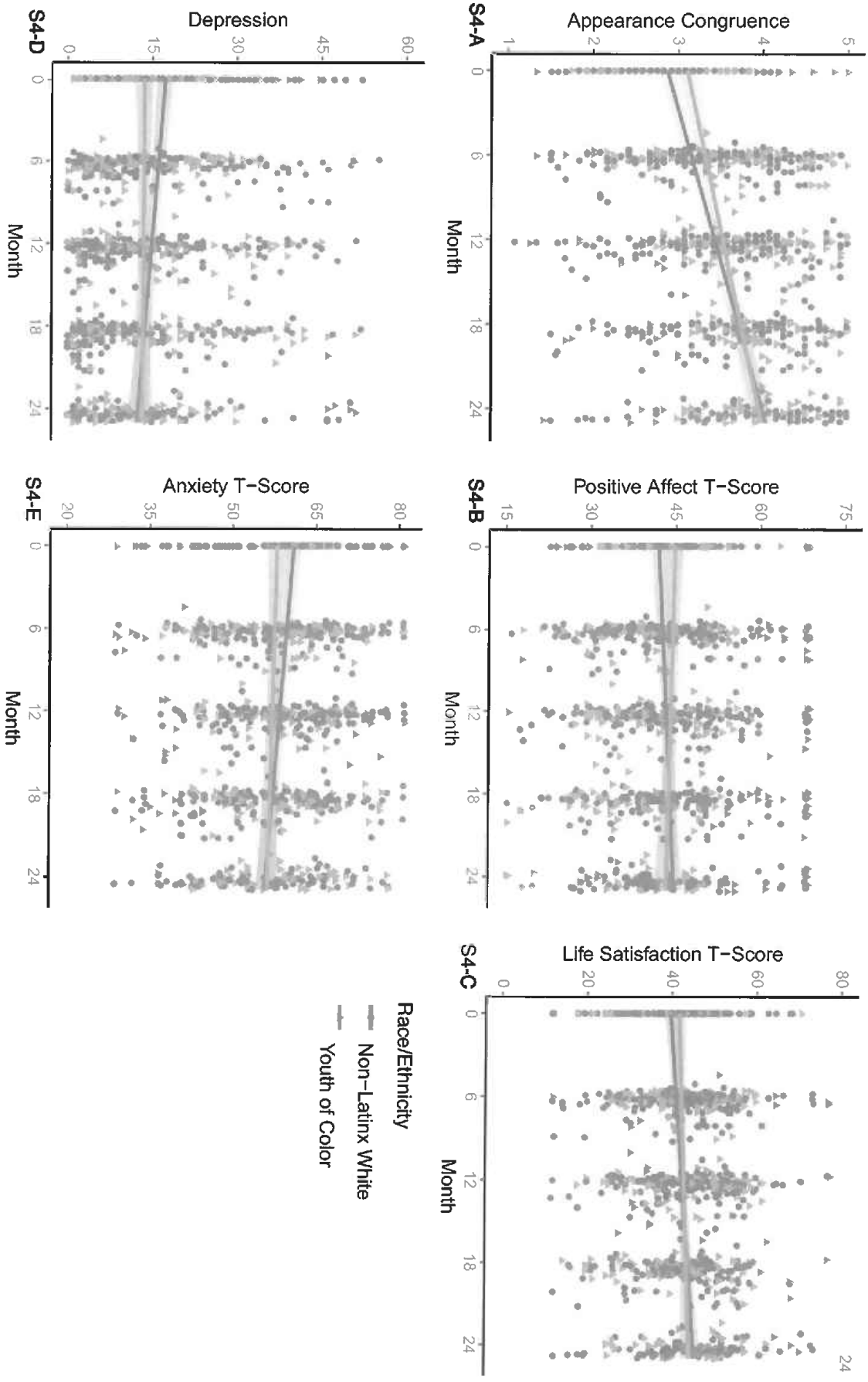

S3-A
Appearance Congruence
S3-B
Positive Affect T−Score
S3-C
Life Satisfaction T−Score
S3-D
Depression
S3-E
Anxiety T−Score
Month
Sex designated at birth
Female
Male

Figure S4 Change in Psychosocial Outcomes by Racial/Ethnic Identity

Figure panels display changes in psychosocial outcomes over two years of GAH by racial/ethnic identity (Non-Latinx White: blue circles; youth of color: orange triangles). Lines indicate mean scores for each group with gray shaded bands for 95% confidence intervals. Outcomes shown are as follows: (S4-A) Transgender Congruence Scale, range: 1-5; (S4-B) Positive Affect Scale T-Score (NIH Toolbox), range: 0-100; (S4-C) Life Satisfaction T-Score (NIH Toolbox), range 0-100); (S4-D) Beck Depression Inventory-II, range: 0-63; (S4-E) Revised Children's Manifest Anxiety Scale, Second Edition T-Score, range: 0-100.


S4-A
Appearance Congruence
S4-B
Positive Affect T-Score
S4-C
Life Satisfaction T-Score
S4-D
Depression
S4-E
Anxiety T-Score
Month
Race/Ethnicity
Non-Latinx White
Youth of Color

Published Manuscripts Using TYC Data

1. Olson-Kennedy, J., Chan, Y.M., Garofalo, R., Spack, N., Chen, D., Clark, L., Ehrensaft, D., Hidalgo, M.A., Tishelman, A.C., & Rosenthal, S.M., (2019). Impact of early medical treatment for transgender youth: Protocol for the longitudinal, observational Trans Youth Care study. JMIR Research Protocols, 8(7): e14434.

2. Lee, J.Y., Finlayson, C., Olson-Kennedy, J., Garofalo, R., Chan, Y.M., Glidden, D.V., & Rosenthal, S.M. (2020). Low bone mineral density in early pubertal transgender/gender diverse youth: Findings from the Trans Youth Care Study. *Journal of the Endocrine Society, 4(9):* bvaa065. DOI: 10.1210/jendso/bvaa065.

3. Millington, K., Schulmeister, C., Finlayson, C., Grabert, R., Olson-Kennedy, J., Garofalo, R., Rosenthal, S.M., & Chan, Y.M. (2020). Physiological and metabolic characteristics of a cohort of transgender and gender-diverse youth in the United States. *Journal of Adolescent Health, 67(3):* 376-383. DOI: 10.1016/j.jadohealth.2020.03.028.

4. Chen, D., Abrams, M., Clark, L., Ehrensaft, D., Tishelman, A.C., Chan, Y.M., Garofalo, R., Olson-Kennedy, J., Rosenthal, S.M., & Hidalgo, M.A. (2021). Psychosocial characteristics of transgender youth seeking gender-affirming medical treatment: Baseline findings from the Trans Youth Care Study. *Journal of Adolescent Health, 68(6):* 1104-1111. DOI: 10.1016/j.jadohealth.2020.07.033.

5. Millington, K., Finlayson, C., Olson-Kennedy, J., Garofalo, R., Rosenthal, S.M., & Chan, Y.M. (2021). Association of high-density lipoprotein cholesterol with sex steroid treatment in transgender and gender-diverse youth. *JAMA Pediatrics, 175(5):* 520-521. DOI: 10.1001/jamapediatrics.2020.5620. DOI: 10.1089/trgh.2020.0055.

6. Olson-Kennedy, J., Streeter, L.H., Garofalo, R., Chan, Y.M., & Rosenthal, S.M. (2021). Histrelin implants for suppression of puberty in youth with gender dysphoria: A comparison of 50 mcg/day (Vantas) and 65 mcg/day (SupprelinLA). *Transgender Health, 6(1):* 36-42.

Supplemental Appendix References

1. Kozee HB, Tylka TL, Bauerband LA. Measuring transgender individuals' comfort with gender identity and appearance: Development and validation of the transgender congruence scale. *Psychol Women Q*. 2012;36(2):179-196.

2. Beck AT, Steer RA, Brown GK, others. Manual for the beck depression inventory-II. *San Antonio TX Psychol Corp*. 1996;1(82):10-1037.

3. Reynolds CR, Richmond BO. *Revised Children's Manifest Anxiety Scale*. Western Psychological Services Los Angeles; 1985.

4. Slotkin J, Nowinski C, Hays R, et al. NIH toolbox scoring and interpretation guide. *Wash DC Natl Inst Health*. Published online 2012:6-7.

5. Little RJ. A test of missing completely at random for multivariate data with missing values. *J Am Stat Assoc*. 1988;83(404):1198-1202.

6. Li C. Little's Test of Missing Completely at Random. *Stata J*. 13(4):795-809.

7. Muthén B, Muthén L. Mplus User's Guide. Eighth Edition. Published online 2017.

8. Ware JH, Harrington D, Hunter DJ, D'Agostino RB. Missing Data. *N Engl J Med*. 2012;367(14):1353-1354. doi:10.1056/NEJMsm1210043

9. McNeish D. On Using Bayesian Methods to Address Small Sample Problems. *Struct Equ Model Multidiscip J*. 2016;23(5):750-773. doi:10.1080/10705511.2016.1186549

10. Oravecz Z, Muth C. Fitting growth curve models in the Bayesian framework. *Psychon Bull Rev*. 2018;25(1):235-255. doi:10.3758/s13423-017-1281-0

11. Stefan AM, Oertzen T. Bayesian power equivalence in latent growth curve models. *Br J Math Stat Psychol*. 2020;73(S1):180-193. doi:10.1111/bmsp.12193

12. van de Schoot R, Depaoli S, King R, et al. Bayesian statistics and modelling. *Nat Rev Methods Primer*. 2021;1(1):1. doi:10.1038/s43586-020-00001-2

13. Zondervan-Zwijnenburg M, Depaoli S, Peeters M, van de Schoot R. Pushing the Limits: The Performance of Maximum Likelihood and Bayesian Estimation With Small and Unbalanced Samples in a Latent Growth Model. *Methodology*. 2019;15(1):31-43. doi:10.1027/1614-2241/a000162

14. Depaoli S, Clifton JP. A Bayesian Approach to Multilevel Structural Equation Modeling With Continuous and Dichotomous Outcomes. *Struct Equ Model Multidiscip J*. 2015;22(3):327-351. doi:10.1080/10705511.2014.937849

1. Olson-Kennedy J, Chan Y-M, Garofalo R, et al. Impact of Early Medical Treatment for Transgender Youth: Protocol for the Longitudinal, Observational Trans Youth Care Study. *JMIR Res Protoc.* 2019;8(7):e14434.
2. Achenbach TM. *Manual for the Youth Self-Report/4-18 and 1991 Profiles.* Burlington, VT: Department of Psychiatry: University of Vermont;1991.
3. Sheehan DV, Lecrubier Y, Sheehan KH, et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J Clin Psychiatry.* 1998;59 Suppl 20:22-33;quiz 34-57.
4. Sheehan DV, Sheehan KH, Shytle RD, et al. Reliability and validity of the Mini International Neuropsychiatric Interview for Children and Adolescents (MINI-KID). *J Clin Psychiatry.* 2010;71(3):313-326.
5. de Vries AL, McGuire JK, Steensma TD, Wagenaar EC, Doreleijers TA, Cohen-Kettenis PT. Young adult psychological outcome after puberty suppression and gender reassignment. *Pediatrics.* 2014;134(4):696-704.
6. Berg D, Edwards-Leeper L, eds. *Child and Family Assessment.* Washington, DC: American Psychological Association; 2018. The Gender Affirmative Model: An Interdisciplinary Approach to Supporting Transgender and Gender Expansive Children.

Acta Psychiatrica Scandinavica

Acta Psychiatr Scand 2020: 141: 486–491

DOI: 10.1111/acps.13164

ACTA PSYCHIATRICA SCANDINAVICA

Trends in suicide death risk in transgender people: results from the Amsterdam Cohort of Gender Dysphoria study (1972–2017)

Wiepjes CM, den Heijer M, Bremmer MA, Nota NM, de Blok CJM, Coumou BJG, Steensma TD. Trends in suicide death risk in transgender people: results from the Amsterdam Cohort of Gender Dysphoria study (1972–2017).

Objective: This study explored the overall suicide death rate, the incidence over time, and the stage in transition where suicide deaths were observed in transgender people.
Methods: A chart study, including all 8263 referrals to our clinic since 1972. Information on death occurrence, time, and cause of death was obtained from multiple sources.
Results: Out of 5107 trans women (median age at first visit 28 years, median follow-up time 10 years) and 3156 trans men (median age at first visit 20 years, median follow-up time 5 years), 41 trans women and 8 trans men died by suicide. In trans women, suicide deaths decreased over time, while it did not change in trans men. Of all suicide deaths, 14 people were no longer in treatment, 35 were in treatment in the previous two years. The mean number of suicides in the years 2013–2017 was higher in the trans population compared with the Dutch population.
Conclusions: We observed no increase in suicide death risk over time and even a decrease in suicide death risk in trans women. However, the suicide risk in transgender people is higher than in the general population and seems to occur during every stage of transitioning. It is important to have specific attention for suicide risk in the counseling of this population and in providing suicide prevention programs.

C. M. Wiepjes[1,2], M. den Heijer[1,2], M. A. Bremmer[2,3], N. M. Nota[1,2], C. J. M. de Blok[1,2], B. J. G. Coumou[2,4], T. D. Steensma[2,4]

[1]Department of Endocrinology, Amsterdam UMC, VU University Medical Center, Amsterdam, the Netherlands, [2]Center of Expertise on Gender Dysphoria, Amsterdam UMC, VU University Medical Center, Amsterdam, the Netherlands, [3]Department of Psychiatry, Amsterdam UMC, VU University Medical Center, Amsterdam, the Netherlands and [4]Department of Medical Psychology, Amsterdam UMC, VU University Medical Center, Amsterdam, the Netherlands

Key words. gender dysphoria; transgender; suicide

Thomas D. Steensma, Department of Medical Psychology, Amsterdam UMC, VU University Medical Center, PO Box 7057, 1007 MB Amsterdam, the Netherlands. E-mail: t.steensma@amsterdamumc.nl

Accepted for publication February 16, 2020

Significant outcomes

- Suicide death risk in trans people did not increase over time.
- Suicide deaths occurred during every stage of transitioning.
- Suicide death risk is higher in trans people than in the general population.

Limitations

- Psychological comorbidity was not known.
- No data were available for people on the waiting list for their first appointment.

Introduction

Gender dysphoria (GD) refers to the distress related to a marked incongruence between one's assigned gender at birth and the experienced gender (1). Trans people are diverse in the intensity of experienced GD (2), their needs for medical

transition (3), and the impairment that GD can have on their life. Studies focusing on the wellbeing of trans people show a greater vulnerability for experiencing mental health problems compared with the non-trans (cis) population (4). Most prevalent are affective and anxiety problems (5-7), often accompanied by feelings, thoughts, or behaviours linked to suicidality (8,9).

The prevalence of suicidality in trans people in suicidal ideation, suicidal attempts, and suicide death rates is studied in varying degrees and shows high variability in findings. A systematic review by McNeil et al (9). reported suicidal ideation rates across 17 identified studies, ranging from 37% (10) up to 83% (11). Prevalence rates on suicidal attempts in trans people, which are generally observed to be lower than suicidal ideation, showed to be lower but also with a wide variation in reported rates, ranging from 9.8% (12) up to 44% (13). Since structured prevalence studies on suicide deaths are lacking in the transgender literature, an estimation comes from a limited number of studies reporting on suicide death rates in small study samples. Derived from a systematic review on suicidality in trans people by Marshall et al. (8), suicide death rates varied from 0% (14) to 4.2% in a sample of 24 post-treatment trans people from Sweden (15). Six of these studies only included postsurgical people (14-19), whereas two studies also included trans people who were only using hormones without surgery (20,21). However, studies differentiating the treatment stage during which death by suicide occurred are lacking. In addition, studies differentiating between suicide in trans women and trans men are scarce. While some studies found that trans men have a higher risk of suicide attempts than trans women (22,23), other studies reported no differences in suicide attempts between trans women and trans men (24,25). Only one cohort distinguished suicide death risk in trans women and trans men and found that trans women had an increased risk of suicide death compared with trans men (20,21).

Aims of the study

The aim of the current study is to explore the overall suicide death rate in trans women and trans men in the largest clinical cohort of gender-referred people seen at the Center of Expertise on Gender Dysphoria of the Amsterdam University Medical Centers between 1972 and 2017 the Netherlands (26). In addition, the change in incidence of suicide death rate over time and at what stage in transition (pretreatment, during hormonal treatment and/or surgical phase, or post-treatment) suicide deaths were observed was explored. The relevance of such information is to get a greater understanding of how large the risk is in clinically referred transgender people and whether suicide prevention interventions should focus on specific stages in transition or not.

Material and methods

Study design

A retrospective chart study was performed, including all people who once visited the Center of Expertise on Gender Dysphoria of the Amsterdam UMC, Vrije Universiteit Amsterdam, the Netherlands, between 1972 and 2017. The selection of the study population is described previously (26). A total of 8263 adults, adolescents, and children were included, with a median age at first visit of 25 years (range 4 to 81 years) and a median follow-up time of 7.5 years (range 0.0 to 45.5 years). Information on death occurrence, time, and cause of death was obtained by cross-checking multiple sources: the National Civil Record Registry (21), which contains date of birth and date of death of all inhabitants of the Netherlands, and the hospital registration system, medical, and psychological files for cause of death.

The Medical Ethics Review Committee of the Amsterdam UMC, Vrije Universiteit Amsterdam, reviewed this study and determined that the Medical Research Involving Human Subjects Act (WMO) did not apply to this study. Therefore, and because of the retrospective design, necessity for informed consent was waived.

Treatment

After an initial visit to the endocrinologist (for adults) or child psychiatrist (for children and adolescents), all people were referred to the psychology department for the diagnostic phase. In this phase, people were seen to gain insight into their experienced gender identity, to verify whether they fulfill the diagnosis gender dysphoria, to explore their treatment desires, and to prepare them for possible medical interventions. After this phase, people may start with hormonal treatment. Trans women received treatment with anti-androgens and estrogens. Trans men were treated with testosterone. In adolescents, treatment first started with a period of puberty suppression, followed by estrogens of testosterone around the age of 16 years (27).

Surgical interventions can be offered to people aged 18 years or older. Depending on the desired

treatment, the surgery is preceded after at least one year of hormonal treatment (genital surgery) or can be offered after the diagnostic phase (e.g., breast removal). After surgery, all people were usually seen every 2 years for medical check-up.

Statistical analyses

Characteristics of the population were shown as median with range due to the non-normal distribution. The total number of people seen at our center and the total number of suicide deaths were counted and were expressed as percentages as well as incidence per 100 000 person years. For each year, the number of people at risk and the number of people who died by suicide were calculated. Cox regression analyses were performed to calculate hazard ratios (HR) with corresponding 95% confidence intervals (95% CI). Date of first visit was used as start date of follow-up. The end date of follow-up was either date of death or date of closing the database (December 31, 2017). Suicide death was analyzed as event. To analyze whether the incidence of suicide deaths changed over time, the year of first visit was added as determinant to the analyses. Analyses were adjusted for age at first visit as age might be related to suicide death risk. Time between date of suicide death and first visit, and between date of suicide death and start of hormonal treatment, if applicable, were calculated. All analyses were performed for the total population and were stratified for trans women and trans men.

All analyses were performed using STATA Statistical software (Statacorp, College Station, TX, USA), version 15.1.

Results

The characteristics of the study population are shown in Table 1. In total, 8263 people attended the gender identity clinic, of which 5107 were trans women (median age at first visit 28 years, range 4 to 81 years) and 3156 were trans men (median age at first visit 20 years, range 4 to 73 years). The median follow-up time was 7.5 years (range 0.0–45.5 years), which was longer in trans women (10.2 years, range 0.0–45.5 years) than in trans men (4.8 years, range 0.0–45.5 years). The total follow-up time was 92 227 person years (64 287 in trans women and 27 940 in trans men).

Table 1. Characteristics of the study population (A) and the people who died by suicide (B)

	Total	Trans women	Trans men
(A)			
Number of people	8263	5107	3156
Age at first visit, year	25 (4–81)	28 (4–81)	20 (4–73)
Follow-up time, year	7.5 (0.0–45.5)	10.2 (0.0–45.5)	4.8 (0.0–45.5)
(B)			
Number of suicides	49 (0.6%)	41 (0.8%)	8 (0.3%)
Age at first visit, year	31 (15–59)	31 (15–58)	21 (16–59)
Age at time of suicide, year	41 (18–66)	41 (18–66)	36 (21–60)
Follow-up time, year	6.7 (0.6–32.7)	6.7 (0.6–32.7)	6.7 (0.6–23.1)
Time between start hormones and suicide, year	6.4 (0.4–32.5) *n = 42*	6.1 (0.4–32.5) *n = 35*	6.9 (3.7–23.1) *n = 7*

Data are shown as number or median (range).

Forty-nine people died by suicide: 41 trans women (0.8%) and 8 trans men (0.3%), which is 64 per 100 000 person years in trans women and 29 per 100 000 person years in trans men. The median follow-up time between first visit and suicide death was 6.7 years (range 0.6 to 32.7 years) in trans women and 6.7 years (range 0.6 to 23.1 years) in trans men. Trans women had a higher overall suicide death risk than trans men (per year: HR 2.26, 95% CI 1.06–4.82). Four suicide deaths occurred in individuals who were referred to the clinic before the age of 18 (0.2%), which is a lower risk than in adults (0.7%, $P = 0.010$).

The course of number of people at risk and the number of people who died by suicide over the years is shown in Fig. 1. Overall suicide deaths did not increase over the years: HR per year 0.97 (95% CI 0.94–1.00). In trans women, suicide death rates decreased slightly over time (per year: HR 0.96, 95% CI 0.93–0.99), while it did not change in trans men (per year: HR 1.10, 95% CI 0.97–1.25). Adjustment for age at the first visit did not change these numbers.

As the median follow-up time between first visit and suicide death was 6.7 years, subgroup analyses were performed in those who had their first visit before 2011. This did not change the outcomes: trans women ($n = 3115$) HR 0.94, 95% CI 0.91–0.98; trans men ($n = 1269$) HR 1.02, 95% CI 0.90–1.16).

Of the 49 people who died by suicide, 35 had a face-to-face contact with the endocrinologist or psychologist of the gender identity clinic in the previous two years, while the other 14 people were no longer in active counseling with the clinic. Sixteen of the 35 people who recently had visited the clinic, only came for a medical check-up, as they were postsurgery (vaginoplasty or phalloplasty). Two people were in the surgery trajectory, and 17 were still in the diagnostic or hormonal phase at time of suicide. The transition phases separately for trans women and trans men who died by suicide are shown in Table 2.

The mean number of suicides in the years 2013–2017 was higher in the trans population (40 per 100 000 person years; 43 per 100 000 trans women

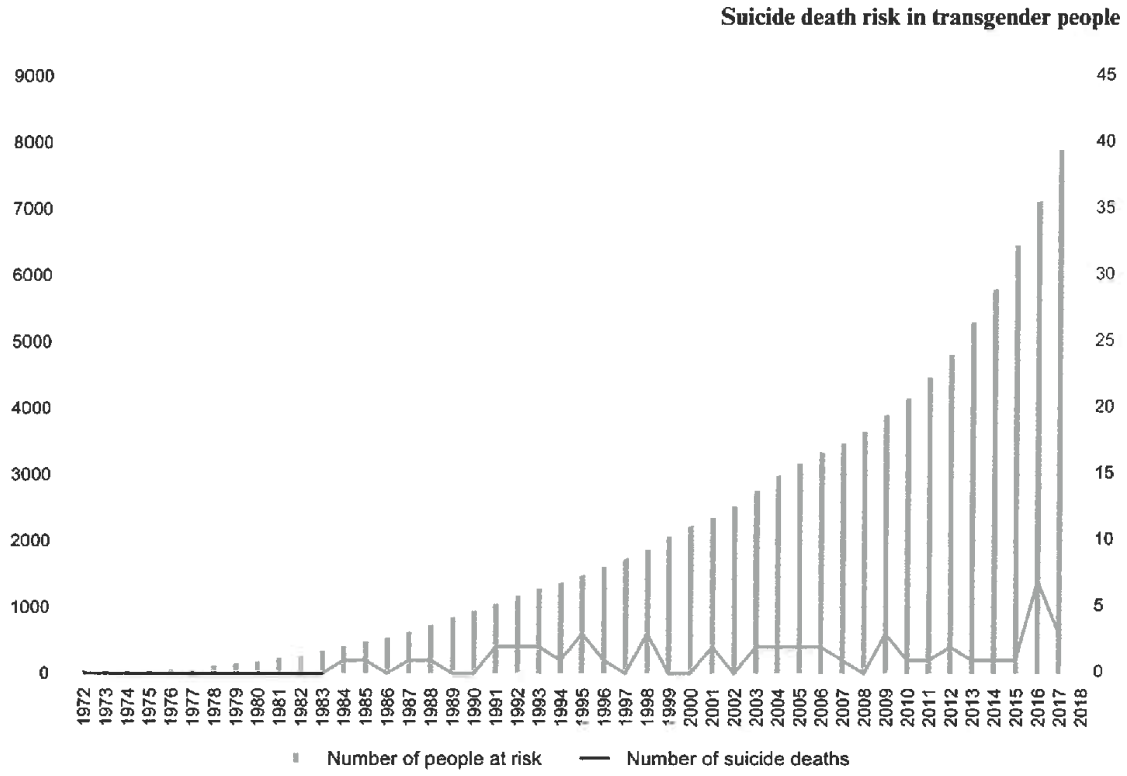


Fig. 1. Number of people at risk (left *y*-axis) and the number of suicides (right *y*-axis), between 1972 and 2017.

Table 2. The occurrence of suicide deaths distinguished for transition stage, and trans women or trans men

	Total (n = 49)	Trans women (n = 41)	Trans men (n = 8)
In active counseling	35	29	6
In diagnostic or hormonal phase	17	16	1
In surgical phase	2	0	2
Only medical follow-up care	16	13	3
No active counseling	14	12	2

Data are shown as number. In active counseling is defined as a face-to-face contact with the endocrinologist or psychologist of the gender identity clinic in the previous two years.

and 34 per 100 000 trans men) compared with the Dutch population in this time frame (11 per 100 000 person years; 15 per 100 00 registered men and 7 per 100 000 registered women) (28).

Discussion

The current study investigated the suicide death risk in the largest clinical cohort of gender-referred individuals to the Center of Expertise on Gender Dysphoria at the Amsterdam UMC, the Netherlands, between 1972 and 2017. Findings from the chart reviews showed us a decrease in suicide death risk over time in trans women and no change in suicide death risk in trans men. Trans women, however, showed a higher suicide death risk than trans men. Between 2013 and 2017, the suicide risk in Dutch referred transgender people (40 per 100 000 person years) showed to be three to four times higher than the general Dutch population (11 per 100 000 person years) (28). Evaluation of transition stage in relation to suicide deaths showed that approximately two-third of the observed suicides occurred in those who were still in active treatment (diagnostic, hormonal, or surgical phase). The incidence of suicide deaths and transition stage was similar in trans women and trans men.

Suicidal behaviour is a complex phenomenon that is a result of many individual (age, male sex assigned at birth, previous suicide attempts, mental health history, substance abuse) as well as more distant environmental factors. A recent literature review clearly demonstrates the specific risk factors for suicide in sexual minority youth, which includes negative social environments, inadequate support within the closest social network, and an absence of lesbian, gay, bisexual, and transgender (LGBT) movements in communities (29). In our cohort, both trans women and trans men show a three- to four-fold elevated

risk of suicide compared with the population rate in the Netherlands and can therefore be considered a high-risk group. Although the Netherlands is known for its tolerance toward sexual minority groups in comparison to most countries in the world (30), the societal position of trans people is generally less favorable compared with the lesbian, gay, bisexual, and cisgender population. Furthermore, compared with trans men, the societal position of trans women is lower (31,32).

In the Netherlands, between 1972 and 2017 suicide rates showed a fluctuating course. Our finding of a slightly decreasing suicide risk in Dutch trans women may confer some hope. Recent studies showed an increase in societal acceptance toward lesbian, gay, bisexual, and transgender people (31), and indications of an increase in social-economic status over the years (33). Although specific information on trans men and trans women is unavailable, it is conceivable that the improvement of societal position may have effect on the psychological functioning and the prevention of suicidal risk in trans women. The cause of this increase in tolerance seems largely to be the effect of a national and international increase in visibility and attention for trans people in media and society. Another explanation may be that, with the increase in attention and acceptance, the threshold for transgender people to seek treatment or professional help has become lower over the years. This is also reflected by the increase in referrals each year (26). Lastly, with the increase of knowledge in this field and the literature about the vulnerability of the transgender population for suicidal ideation, suicidal attempts, and suicide death rates, it is conceivable to assume that the attention to these risks has increased in clinical counseling and may have its effect on prevention of suicide deaths over the years.

Although the incidence of suicide deaths in trans women decreased over the years, the overall incidence still showed to be higher in trans women compared with trans men. Conflicting results in literature are reported about the risk of suicide attempts between trans women and trans men. Some studies reported that trans men had a higher risk of suicide attempts than trans women (22,23), while in other studies no differences in suicide attempts between trans women and trans men were found (24,25). Only two studies looked at the differences in the risk of death by suicide between trans women and trans men and found that trans women had an increased risk compared with trans men (20,21). However, these two studies were earlier studies performed in our center and therefore include a smaller part of our current study population.

An important finding was that the incidence for observed suicide deaths was almost equally distributed over the different stages of treatment. Although the distribution showed that one-third of the suicides occurred in people who were no longer in active treatment in our center, the other two-third of the people who died by suicide still visited our center in the previous two years. About half of these last two-third people were still in active diagnostic or medical treatment, while the other half completed their transition and only came for a medical check-up. This indicates that vulnerability for suicide occurs similarly in the different stages of transition. Although the literature on suicide risk factors is comprehensive, and particular suicidal risk factors like verbal victimization, physical and sexual violence, and the absence of social support (9,34), may apply for transgender people in all stages of transitioning, it seems clinically highly relevant to understand and explore possible differences in motives and risk factors in the different stages of treatment. Therefore, future research on suicide deaths and suicide risk factors in transgender people should have a greater focus on transition status in relation to these motives and risk factors.

This study is performed in the largest cohort of gender-referred people from the Netherlands, consisting of a large population of both adult and adolescent trans women and trans men at different stages of their transition with a long follow-up time. However, this study has also some limitations. First, this study is a retrospective chart study. Although we used multiple strategies to obtain data about date of death, it is possible that we missed some data. Second, we did not have information about psychological comorbidities or other psychological information, such as social support. Third, we only had information about people who actually visited our gender identity clinic. Information about people on the waiting list for their first appointment was lacking.

To conclude, in our clinic we observed no increase in suicide death risk over time and even a decrease over time in suicide death risk in trans women was found. Since the suicide risk in the transgender population is higher than the general population and seems to occur during every stage of transitioning, it is important that (mental) health practitioners pay attention to this risk and create a safe environment in which these feelings can be discussed at all stages of treatment and counseling. Further research is necessary to investigate the motives behind the suicides, as input in

the development of adequate suicide prevention programs.

Conflicts of interests

None.

Data availability statement

Author elects to not share data.

References

1. American Psychiatry Association. Diagnostic and Statistical Manual of Mental Disorders (DSM-5), 5th ed. Arlington, VA: American Psychiatric Publishing; 2013.
2. Meyer-Bahlburg HF. From mental disorder to iatrogenic hypogonadism: dilemmas in conceptualizing gender identity variants as psychiatric conditions. Arch Sex Behav 2010;**39**(2):461–476.
3. Beek TF, Kreukels BP, Cohen-Kettenis PT, Steensma TD. Partial treatment requests and underlying motives of applicants for gender affirming interventions. J Sex Med 2015;**12**(11):2201–2205.
4. Dhejne C, Van Vlerken R, Heylens G, Arcelus J. Mental health and gender dysphoria: A review of the literature. Int Rev Psychiatry 2016;**28**(1):44–57.
5. Haraldsen IR, Dahl AA. Symptom profiles of gender dysphoric patients of transsexual type compared with patients with personality disorders and healthy adults. Acta Psychiatr Scand 2000;**102**:276–281.
6. Fisher AD, Bandini E, Casale H et al. Sociodemographic and clinical features of gender identity disorder: an Italian multicentric evaluation. J Sex Med 2013;**10**(2):408–419.
7. Heylens G, Elaut E, Kreukels BP et al. Psychiatric characteristics in transsexual individuals: multicentre study in four European countries. Br J Psychiatry 2014;**204**(2):151–156.
8. Marshall E, Claes L, Bouman WP, Witcomb GL, Arcelus J. Non-suicidal self-injury and suicidality in trans people: A systematic review of the literature. Int Rev Psychiatry 2016;**28**(1):58–69.
9. McNeil J, Ellis SJ, Eccles FJR. Suicide in trans populations: A systematic review of prevalence and correlates. Psychol Sex Orient Gender Div 2017;**4**(3):341–353.
10. Mathy RM. Transgender Identity and Suicidality in a Nonclinical Sample. J Psychol Human Sexual. 2003;**14**(4):47–65.
11. Testa RJ, Sciacca LM, Wang F et al. Effects of Violence on Transgender People. Profess Psychol: Res Pract 2012;**43**(5):452–459.
12. Heylens G, Verroken C, De Cock S, T'Sjoen G, De Cuypere G. Effects of different steps in gender reassignment therapy on psychopathology: a prospective study of persons with a gender identity disorder. J Sex Med 2014;**11**(1):119–126.
13. Miller LR, Grollman EA. The social costs of gender nonconformity for transgender adults: implications for discrimination and health. Sociol Forum 2015;**30**(3):809–831.
14. Johansson A, Sundbom E, Hojerback T, Bodlund O. A five-year follow-up study of Swedish adults with gender identity disorder. Archiv Sex Behav 2010;**39**(6):1429–1437.
15. Walinder J, Law ITA. Concerning Sex Reassignment of Transsexuals in Sweden. Archiv Sex Behav 1976;**5**(3):255–258.
16. Dhejne C, Lichtenstein P, Boman M, Johansson AL, Langstrom N, Landen M. Long-term follow-up of transsexual persons undergoing sex reassignment surgery: cohort study in Sweden. PLoS One 2011;**6**(2):e16885.
17. Berg JEA, Gustafsson M. Long term follow up after sex reassignment surgery. Scand J Plast Reconstruct Surg Hand Surg 2009;**31**(1):39–45.
18. De Cuypere G, T'Sjoen G, Beerten R et al. Sexual and physical health after sex reassignment surgery. Arch Sex Behav. 2005;**34**(6):679–690.
19. Sorensen T, Hertoft P. Male and female transsexualism: the danish experience with 37 patients. Archiv Sex Behav 1982;**11**(2):133–155.
20. Van Kesteren P, Asscheman H, Megens JAJ, Gooren LJG. Mortality and morbidity in transsexual subjects treated with cross-sex hormones. Clin Endocrinol 1997;**47**:337–342.
21. Asscheman H, Giltay EJ, Megens JA, de Ronde WP, van Trotsenburg MA, Gooren LJ. A long-term follow-up study of mortality in transsexuals receiving treatment with cross-sex hormones. Eur J Endocrinol 2011;**164**(4):635–642.
22. Marshall BD, Socias ME, Kerr T, Zalazar V, Sued O, Aristegui I. Prevalence and correlates of lifetime suicide attempts among transgender persons in Argentina. J Homosex 2016;**63**(7):955–967.
23. Maguen S, Shipherd JC. Suicide risk among transgender individuals. Psychol Sexual 2010;**1**(1):34–43.
24. Reisner SL, White JM, Bradford JB, Mimiaga MJ. Transgender Health disparities: COMPARING full cohort and nested matched-pair study designs in a community health center. LGBT Health 2014;**1**(3):177–184.
25. Skagerberg E, Parkinson R, Carmichael P. Self-harming thoughts and behaviors in a group of children and adolescents with gender dysphoria. Int J Transgend 2013;**14**(2):86–92.
26. Wiepjes CM, Nota NM, de Blok CJM et al. The amsterdam cohort of gender Dysphoria study (1972–2015): Trends in prevalence, treatment, and regrets. J Sex Med 2018;**15**(4):582–590.
27. Kreukels BP, Cohen-Kettenis PT. Puberty suppression in gender identity disorder: the Amsterdam experience. Nat Rev Endocrinol 2011;**7**(8):466–472.
28. StatLine CBSOverledenen; belangrijke doodsoorzaken (korte lijst), leeftijd, geslacht; 2018.
29. Poštuvan V, Podlogar T, Zadravec Šedivy N, De Leo D. Suicidal behaviour among sexual-minority youth: a review of the role of acceptance and support. Lancet Child Adol Health 2019;**3**(3):190–198.
30. Veenhoven R. Wat bracht de seksuele revolutie (What caused the sexual revolution?). In: Couwenberg SW ed. Seksuele Revolutie Ter Discussie. Van Phil Bloom tot Sex and the Citiy. Vol 1. Eindhoven, the Netherlands: Damon, 2005:93-105.
31. Kuyper L. Opvattingen over seksuele en genderdiversiteit in Nederland en Europa. Den Haag: Sociaal en Cultureel Planbureau, 2018.
32. Keuzenkamp S. Worden wie je bent. Het leven van transgenders in Nederland. Den Haag: Sociaal en Cultureel Planbureau, 2012.
33. Beusekom G, Kuyper L. LHBT-monitor 2018. De leefsituatie van lesbische, homoseksuele, biseksuele en transgender personen in Nederland. Den Haag: Sociaal en Cultureel Planbureau, 2018.
34. Haas AP, Rodgers PL, Herman JL. Suicide attempts among transgender and gender non-conforming adults. findings of the national transgender discrimination survey. New York, NY: American Foundation for Suicide Prevention and The Williams Institute, 2014.

European Journal of Endocrinology (2011) 164 635–642 ISSN 0804-4643

CLINICAL STUDY

A long-term follow-up study of mortality in transsexuals receiving treatment with cross-sex hormones

Henk Asscheman[1], Erik J Giltay[3], Jos A J Megens[2], W (Pim) de Ronde[1], Michael A A van Trotsenburg[2] and Louis J G Gooren[1]

[1]*Endocrine Unit, Department of Internal Medicine and* [2]*Center of Expertise on Gender Dysphoria, VU University Medical Center, PO Box 7057, NL-1007 MB Amsterdam, The Netherlands and* [3]*Department of Psychiatry, Leiden University Medical Center, NL-2300 RC Leiden, The Netherlands*

(Correspondence should be addressed to H Asscheman who is now at HAJAP, Valeriusstraat 4hs, 1071 MH Amsterdam, The Netherlands; Email: h.ajap@worldonline.nl)

Abstract

Objective: Adverse effects of long-term cross-sex hormone administration to transsexuals are not well documented. We assessed mortality rates in transsexual subjects receiving long-term cross-sex hormones.

Design: A cohort study with a median follow-up of 18.5 years at a university gender clinic.

Methods: Mortality data and the standardized mortality rate were compared with the general population in 966 male-to-female (MtF) and 365 female-to-male (FtM) transsexuals, who started cross-sex hormones before July 1, 1997. Follow-up was at least 1 year. MtF transsexuals received treatment with different high-dose estrogen regimens and cyproterone acetate 100 mg/day. FtM transsexuals received parenteral/oral testosterone esters or testosterone gel. After surgical sex reassignment, hormonal treatment was continued with lower doses.

Results: In the MtF group, total mortality was 51% higher than in the general population, mainly from increased mortality rates due to suicide, acquired immunodeficiency syndrome, cardiovascular disease, drug abuse, and unknown cause. No increase was observed in total cancer mortality, but lung and hematological cancer mortality rates were elevated. Current, but not past ethinyl estradiol use was associated with an independent threefold increased risk of cardiovascular death. In FtM transsexuals, total mortality and cause-specific mortality were not significantly different from those of the general population.

Conclusions: The increased mortality in hormone-treated MtF transsexuals was mainly due to non-hormone-related causes, but ethinyl estradiol may increase the risk of cardiovascular death. In the FtM transsexuals, use of testosterone in doses used for hypogonadal men seemed safe.

European Journal of Endocrinology **164** 635–642

Introduction

Psychological evaluation has shown that sex reassignment increases the well-being of transsexual subjects (1–3). Cross-sex hormone treatment has an important role in acquiring the secondary sex characteristics of the desired sex (4). Transsexuals often start taking sex hormones at young to middle age and in higher than recommended dosages. Fearing loss of secondary characteristics of the reassigned sex, transsexual subjects usually continue hormones lifelong. Previous reports from our clinic, in 1989 (5) and 1997 (6), assessed clinical endpoints, such as morbidity and mortality, in transsexuals receiving cross-sex hormones. Both these studies found no increase in mortality rates in subjects receiving cross-sex hormones compared with the general population, but reported higher than expected rates of completed suicide and death due to acquired immunodeficiency syndrome (AIDS) in male-to-female (MtF) transsexual subjects, while no increased morbidity/mortality was observed in female-to-male (FtM) transsexual subjects.

Several studies have looked at the effects of cross-sex hormone administration on laboratory variables related to cardiovascular risk in transsexuals, finding partially favorable and partially unfavorable effects (7–10). The skewed sex ratio in cardiovascular disease favoring women at all ages and the increasing incidence of cardiovascular disease after menopause were previously interpreted to indicate that estrogens are cardioprotective. By contrast, hyperandrogenemia in women is associated with increases in cardiovascular risk factors (11), which has led to the belief that androgens are detrimental to cardiovascular health (12). However, the large randomized trials (Heart and Estrogen/Progestin Replacement Study (13) and Women's Health Initiative

DOI: 10.1530/EJE-10-1038
Online version via www.eje-online.org

(14)) refuted the cardioprotective effects of exogenous estrogens, in its generality, leading to revision of the practice of hormone replacement therapy in (post)menopausal women.

Another factor to be considered is the route of administration of estrogens, possibly having relevance for their adverse effects. Oral versus transdermal administration of 17β-estradiol (E_2) may impact differently on variables such as inflammation markers (15), lipoproteins (16), and coagulation markers (17). The pharmacological nature of the estrogen compound may be of significance too: oral administration of the synthetic compound ethinyl estradiol may have more negative effects on hemostasis than oral or transdermal E_2 (18).

This study aims primarily to describe all-cause and cause-specific mortality rates in subjects receiving cross-sex hormone treatment. This analysis extends our previous reports by assessing the effects of longer term use of cross-sex hormones in subjects treated at this clinic, increasing the accrued person-years of follow-up data from 10 152 (6) to 25 544. In addition, the effects of aging and co-morbidity have likely increased the number of endpoints, which will increase the impact and precision of the effect size measures with smaller confidence intervals (CIs); associations previously not detected due to the smaller sample size and lower power may now become apparent. We report the observed mortality rates in 1331 transsexuals followed-up for a median period of more than 18 years, and we compare the observed number of deaths with the expected number as found in the general population. In a subanalysis, the type of estrogen (i.e. oral ethinyl estradiol versus other estrogen compound and routes of administration) is analyzed in relation to the risk of cardiovascular mortality.

Subjects and methods

Baseline and follow-up data of all transsexual subjects referred to our outpatient department since 1975 were entered into a cumulative database. In the present analysis on mortality aiming to measure longer term effects, we included only subjects who had started cross-sex hormone treatment before July 1, 1997, followed-up for at least 1 year and included 2 MtF who had died the first year of hormone administration.

In total, 1331 subjects met the above inclusion criteria, 966 (72.6%) MtF transsexuals, with a mean age of 31.4 years at the start of cross-sex hormones (range: 16–76 years), with 18 678 patient-years of follow-up, and 365 (27.4%) FtM transsexuals, with a mean age 26.1 years (range: 16–57 years) at the start of hormone therapy with 6866 patient-years of follow-up. Subjects were followed-up until July 1, 2007, or until the date of death. In 2009, we could cross check our database against the National Civil Record Registry (Gemeentelijke Basis Administratie) which registers all residents in the Netherlands and, if deceased, their date of death (but not cause of death). We identified another 45 MtF and 3 FtM subjects included in our database who had died before July 1, 2007, but were unknown to us in our initial analysis on mortality based on hospital records (19). Of these additional deaths, the cause of death could be ascertained in two out of three FtM (66%), and in 27 out of 45 (60%) MtF transsexual subjects. The mean follow-up period of subjects receiving cross-sex hormones was 19.3±7.7 years (median 18.6, range 0.7–44.5 years) in the MtF group. In the FtM group, the follow-up was 18.8±6.3 years (median 18.4, range 4.7–42.6 years; Table 1).

The cause of death was ascertained by medical report or information from the family physician and was coded according to the International Classification of Disease (ICD-10, 10th revision 2007; www.who.int/classification/icd10online). When initiating sex reassignment treatment, all subjects had agreed that their data could potentially be used in future scientific analysis with the provision that data could not be related to an individual person.

Hormone treatment

In MtF transsexuals, hormone treatment before sex reassignment surgery consisted of estrogens combined with anti-androgens. Until 1989, mainly ethinyl estradiol was prescribed in a dose of 100 μg/day, and

Table 1 Baseline and treatment-related characteristics of 1331 male-to-female and female-to-male transsexuals who underwent cross-sex hormone treatment.

	Male-to-female transsexuals	Female-to-male transsexuals
n	966	365
Age at start (mean±s.d.)	31.4±11.4	26.1±7.6
Range (years of age)	16–76	16–56
Age groups (*n* (%))		
15–24	329 (34.1)	204 (55.9)
25–39	429 (44.4)	145 (39.7)
40–64	199 (20.6)	16 (4.4)
65–80	9 (0.9)	0
Smoking status (*n* (%))		
Never	254 (26.3)	94 (25.8)
Current	373 (38.6)	131 (35.9)
Former or unknown	339 (35.1)	140 (38.3)
Starting date before 1990 (*n* (%))	619 (64.2)	197 (54.0)
Sex reassignment surgery (*n* (%))	834 (86.7)	343 (94.0)
Follow-up on hormone treatment (years±s.d.)	19.4±7.7	18.8±6.3
<5 years (*n* (%))	22 (2.2)	1 (0.3)
5–10 years (*n* (%))	50 (5.2)	6 (1.6)
10–15 years (*n* (%))	229 (23.7)	111 (30.4)
15–20 years (*n* (%))	252 (26.1)	99 (27.2)
20–25 years (*n* (%))	190 (19.7)	86 (23.5)
25–30 years (*n* (%))	131 (13.6)	43 (11.8)
>30 years (*n* (%))	92 (9.5)	19 (5.2)

only small numbers of patients used estrogen injections or other oral estrogen compounds, such as conjugated estrogens. But after publication of an elevated risk of venous thrombosis associated with ethinyl estradiol use (5), particularly in patients over 40 years of age, we started to recommend transdermal E_2 to all MtF, particularly to those over 40 years of age. In those MtF who did not tolerate or refused transdermal estrogens, oral estradiol valerate 2–4 mg/day was prescribed. However, some subjects were reluctant to change their previous estrogen therapy and continued with ethinyl estradiol.

Before surgical sex reassignment in MtF transsexuals (which includes orchiectomy), estrogens were always combined with anti-androgen treatment (usually cyproterone acetate 100 mg/day and spironolactone 100–200 mg/day in <5% of MtF) to decrease testosterone levels and/or block androgen action. In the period before we started to advice not to use ethinyl estradiol, the standard practice was to reduce the dose of ethinyl estradiol to 50 μg/day after surgery, or estrogen treatment was changed to transdermal or oral E_2. Furthermore, anti-androgens were discontinued, but about 30% of the MtF subjects experienced regrowth of undesired (facial) hair to some extent and asked for continuation of anti-androgens, though in significant lower doses.

FtM transsexuals were prescribed testosterone as esters intramuscularly 250 mg/2 weeks, reduced postoperatively to every 3 weeks, oral testosterone undecanoate 160–240 mg/day (Andriol, not available in the USA) and, more recently, transdermal testosterone 50 mg/day. If uterine bleeding persisted, a progestin was added until hysterectomy, usually lynestrenol.

It is of note that the Dutch health care system fully covers sex reassignment treatment, with the result that almost all transsexual subjects have undergone sex reassignment surgery 2 years after starting cross-sex hormones. Consequently, the observed effects of sex hormones on biological systems in this study are largely attributable to exogenous hormones.

Statistical analysis

The observed number of deaths in the study population was set against the expected numbers of deaths (except from AIDS and drug abuse) derived from the 2001 mortality data of the general population provided by the Central Bureau of Statistics of the Netherlands (Centraal Bureau voor Statistiek on www.statline.cbs.nl) stratified per age group (i.e. 15–24; 25–39; 40–64; and 65–80 years of age) and biological sex. Numbers of deaths were adjusted for the years of follow-up on cross-sex hormone treatment. Expected number of deaths from AIDS and drug abuse, which varied largely from year to year, was calculated from specific reports by Statistics Netherlands. The risk was expressed as standardized mortality ratio (SMR), and the 95% CIs were calculated by regarding the observed number as a Poisson variable with tables based on Poisson distribution (20).

In a subanalysis, the association of use of ethinyl estradiol to mortality was analyzed. The use of ethinyl estradiol (dichotomized into i) never or former users during hormone administration, and ii) ongoing users) was analyzed in relation to all-cause mortality, cardiovascular mortality, mortality due to external causes, cancer mortality, and non-cardiovascular mortality in 964 MtF transsexuals. The never/former users were combined into one reference group, as the risk of cardiovascular death was not increased in former versus never users of ethinyl estradiol. The potentially mediating or confounding variables such as age, smoking status, and starting date before 1990 were adjusted for Cox proportional hazards models by incrementally including them as covariates. The associations of different groups of ethinyl estradiol use and mortality were explored by selecting the first group (i.e. never or former users of ethinyl estradiol) as the reference category (i.e. a hazard ratio of 1). The software used was SPSS version 17.0 (SPSS, Inc., Chicago, IL, USA).

Results

Baseline characteristics

Baseline data and duration of follow-up in the patient groups are shown in Table 1. MtF transsexual subjects were older when they started cross-sex hormones (31.4 ± 11.4 years) than FtM (26.1 ± 7.4 years; $P<0.001$). In the MtF group, 207 subjects (21.4%) were over 40 years of age, and nine subjects (0.9%) were even over 65 years of age, whereas only few FtM ($n=16$, 4.4%) were over 40 years of age at the start of cross-sex hormone treatment. The mean duration of follow-up was not significantly different between MtF and FtM subjects (19.4 ± 7.7 vs 18.8 ± 6.3 years; $P=0.12$). The rate of sex reassignment surgery (defined as orchiectomy/penectomy + vaginoplasty in MtF and extirpation of the internal genitalia with both ovaries in FtM) was significantly lower in MtF compared to FtM subjects (86.7 vs 94.0%, $P<0.001$).

Mortality rates in MtF transsexuals

In the MtF group, 122 (12.6%) out of 966 subjects had died during follow-up. When compared with the adjusted expected mortality in the general population, MtF had a significantly increased mortality with a SMR of 1.51 (95% CI: 1.47–1.55; Table 2). The increased mortality in MtF in the 25–39 years of age group (SMR 4.47; 95% CI: 4.04–4.92) was mainly due to the relatively high numbers of suicides (in six), drugs-related death (in four), and death due to AIDS (in 13 subjects).

In 40–64 year age group, the SMR of total mortality was increased with 1.42 (95% CI: 1.35–1.48).

Table 2 SMR adjusted for age and period of follow-up on hormone treatment by biological sex in 1331 male-to-female and female-to-male transsexual subjects.

	Male-to-female transsexuals		Female-to-male transsexuals	
Cause of death	Observed cases	SMR (95% CI)	Observed cases	SMR (95% CI)
Malignant neoplasm	28	0.98 (0.68–1.08)	5	0.99 (0.65–1.44)
Lung	13	1.35 (1.14–1.58)	1	1.06 (0.26–3.19)
Digestive tract	3	0.42 (0.28–0.60)	2	2.41 (0.90–5.18)
Hematological	6	2.58 (1.97–3.30)	1	2.86 (0.69–8.57)
Brain	2	1.59 (0.95–2.46)	0	–
Other: kidney, melanoma, bone, and prostate in MtF. In FtM: leiomyosarcoma	4	0.79 (0.57–1.07)	1	0.77 (0.25–1.77)
Ischemic heart disease	18	1.64 (1.43–1.87)	1	1.19 (0.39–2.74)
Cerebrovascular accidents	5	1.26 (0.93–1.64)	0	–
AIDS	16	30.20 (26.0–34.7)	0	–
Endocrine/diabetes	2	0.85 (0.41–1.32)	0	–
Respiratory system diseases	4	0.85 (0.61–1.14)	0	–
Digestive system diseases	3	1.01 (0.68–1.45)	1	2.56 (0.62–7.69)
Genitourinary system disease (ESRD)	1	1.21 (0.58–2.17)	0	–
Nervous system disease (MS)	0		1	3.57 (0.86–10.7)
External causes	24	7.67 (6.84–8.56)	2	2.22 (1.07–5.44)
Illicit drugs use	5	13.20 (9.70–17.6)	1	25.00 (6.00–32.5)
Suicide	17	5.70 (4.93–6.54)	1	2.22 (0.53–6.18)
Unknown/ill-defined symptoms	21	4.00 (3.52–4.51)	2	2.08 (0.69–4.79)
Total	122	1.51 (1.47–1.55)	12	1.12 (0.89–1.59)

ESRD, end-stage renal disease; MS, multiple sclerosis.

The higher rate as compared with the general population was largely explained by eight suicides (where only one was expected on the basis of mortality data in the general population) and 17 deaths from cardiovascular diseases (where only eight were expected). In the relatively small MtF group over 65 years of age, total mortality was not increased (SMR 0.95, 95% CI: 0.86–1.06) as compared to the general population.

In MtF, ischemic heart disease was the cause of death in 18 subjects (SMR 1.64; 95% CI: 1.43–1.87). The mean age of occurrence of the lethal ischemic cardiac event was 59.7 years (range: 42–79 years). The mean duration of estrogen use was 13.2 years (range: 2–42 years). Eleven of them (61%) had been using ethinyl estradiol during a mean period of 9.7 years (range: 2–16 years), whereas the other seven had used transdermal estrogen ($n=2$), stilbestrol ($n=1$), tibolon ($n=1$), or conjugated estrogens ($n=3$) for a mean period of 16.9 years (range: 5–42 years). The mean age at the start of estrogen treatment was 45.9 years (range: 18–70 years), 46.5 years in ethinyl estradiol users, and 44.9 years in users of other estrogens. Nine (50%) of the deceased subjects were current smokers, two non-smokers, and seven former smokers or unknown. Four (22%) had hypercholesterolemia (>6.5 mmol/l or >250 mg/dl). Four (22%) had been diagnosed earlier with venous thrombosis, and five (28%) had suffered a previous myocardial infarction.

Five MtF subjects died from stroke (SMR 1.26; 95% CI: 0.93–1.64). Two subjects died before the age of 60, and the other three subjects died when they were 60, 62, and 75 years old; therefore, in 40–64 years of age, the SMR for fatal stroke was 2.11 (95% CI: 1.32–3.21). All had been using ethinyl estradiol, and in only one of the two who had suffered a previous transient ischemic attack, the treatment regimen had been changed to transdermal estrogen.

In the Cox proportional hazard analysis of the type of estrogen treatment in MtF, current use of ethinyl estradiol was significantly associated with cardiovascular mortality, but not with an increased risk of all-cause mortality or mortality due to other causes. The threefold increased hazard ratio of cardiovascular mortality in current users compared with never and former users of ethinyl estradiol remained significant after adjustment for covariates (Table 3).

In the MtF group, the observed total number of deaths due to malignant neoplasm ($n=28$) was not increased compared with the general population, but lung cancer ($n=13$) showed a statistically significant increased SMR of 1.35 (95% CI: 1.14–1.58). The risk of leukemia/lymphoma with six deaths (one acute myeloid leukemia, one chronic lymphoid leukemia, one unclassified leukemia, and three non-Hodgkin lymphomas) was significantly increased with a SMR of 2.66 (95% CI: 1.93–3.60).

External causes of death were increased almost eightfold due to suicide and illicit drug use. The suicide rate in MtF was increased sixfold. Thirteen out of the seventeen (76%) had received psychiatric treatment in the past. No suicides occurred within the first 2 years of hormone treatment, while there were six suicides after 2–5 years, seven after 5–10 years, and four after more than 10 years of cross-sex hormone treatment at a mean age of 41.5 years (range 21–73 years).

Table 3 Hazard ratios (95% CIs) of mortality according to the use of ethinyl estradiol in 964 male-to-female transsexuals during a median of 18.6 years of follow-up. Two deaths within the first year of follow-up were excluded to reduce the chance of reverse causation. Cardiovascular mortality was defined as death due to myocardial infarction or stroke.

	Use of ethinyl estradiol		
	Never or former use	Continuous use	*P* value
No. of male-to-female transsexuals	596	368	
All-cause mortality	69 (11.6%)	51 (13.9%)	
Crude	1.00	1.13 (0.78–1.62)	0.53
Adjusted for age and smoking	1.00	1.33 (0.92–1.92)	0.13
Fully adjusted[a]	1.00	1.28 (0.88–1.86)	0.20
Cardiovascular mortality	8 (1.3%)	15 (4.1%)	
Crude	1.00	2.82 (1.19–6.65)	0.02
Adjusted for age and smoking	1.00	3.64 (1.52–8.73)	0.004
Fully adjusted[a]	1.00	3.12 (1.28–7.63)	0.01
Mortality due to external causes[b]	12 (2.0%)	11 (3.0%)	
Crude	1.00	1.40 (0.62–3.17)	0.43
Adjusted for age and smoking	1.00	1.44 (0.63–3.30)	0.38
Fully adjusted[a]	1.00	1.36 (0.60–3.10)	0.46
Cancer mortality	17 (2.9%)	11 (3.0%)	
Crude	1.00	0.99 (0.46–2.12)	0.98
Adjusted for age and smoking	1.00	1.24 (0.57–2.67)	0.59
Fully adjusted[a]	1.00	1.35 (0.61–3.00)	0.46
Non-cardiovascular mortality	46 (7.7%)	30 (8.2%)	
Crude	1.00	1.00 (0.63–1.59)	0.99
Adjusted for age and smoking	1.00	1.16 (0.73–1.84)	0.54
Fully adjusted[a]	1.00	1.15 (0.71–1.83)	0.58

P values using Cox proportional hazards models.
[a]Adjusted for age, smoking status, and a starting date before 1990 (because before 1990, ethinyl estradiol was the standard estrogen prescribed).
[b]Deaths due to accidents, intentional self-harm and suicide, assault, drugs, and adverse effects.

Five (1.6%) suicides were observed among the 304 MtF who were still using cyproterone acetate and 12 (1.8%) in the group of MtF no longer using cyproterone acetate. Six MtF subjects who committed suicide (35%) had not undergone sex reassignment surgery because there had been doubts about their mental stability. In the whole group of MtF subjects, 87.6% underwent sex reassignment surgery.

Also death due to illicit drug use ($n=5$) was relatively increased (SMR 13.2; 95% CI: 9.7–17.6). All had been past or current substance abusers before the start of hormone treatment but had been evaluated as sufficiently mentally stable to undergo hormone treatment. Sixteen MtF transsexual subjects died from AIDS between 1986 and 2006 (SMR 30.2; 95% CI: 26.0–34.7). The underlying cause of death could not be ascertained in 21 (17.2%) of the 122 subjects who had died.

Mortality rates in FtM transsexuals

In the FtM group, 12 out of 365 (3.4%) died during follow-up. When compared with the adjusted expected mortality in the general population, in FtM, the SMR of 1.12 (95% CI: 0.87–1.42) was not significantly increased (Table 2). Compared with the MtF population, actual numbers were lower in the FtM group, which resulted in large CIs of the point estimates. The FtM group was also on average of younger age (only six over 65 years of age) compared to the MtF group. Only one myocardial infarction was observed in a 72-year-old FtM subject after 42 years of testosterone treatment. External causes of death were increased due to one death by illicit drug abuse, a cause of death extremely rare in the reference group of the female general population. Total number of cancer deaths was not different from the expected number. No deaths due to breast cancer was observed, and other cancer death categories were not statistically significantly different from those expected, but again this has to be set against larger CIs.

Discussion

In this large cohort with a median follow-up of more than 18 years, we observed in MtF transsexual subjects a 51% relatively increased mortality rate compared with the general male population, mainly due to increased rates of death from suicide, illicit drugs, AIDS, cardiovascular disease, and unknown causes. In FtM transsexuals, the observed mortality rate was not significantly increased compared with women in the general population. However, it should be taken into consideration that FtM started cross-sex hormones (testosterone) at a younger age than MtF (at mean age 26.1 years compared with MtF at age 31.1 years), and rarely started treatment after the age of 40. The effects of aging may thus carry less weight in the FtM group (also smaller) than the MtF group. Follow-up of FtM in

the 65–79 years of age group was only 35.4 patient years, implying that no firm conclusions can be drawn in this FtM age group nor in the FtM group as a whole.

The increased mortality risk in MtF in our cohort was characterized by a high SMR of suicide (of 5.70), AIDS (of 30.2), and illicit drug-related deaths (of 13.2). In our previous publication, the increased mortality rates due to suicide and AIDS had already been noted (6). Depressive mood changes have been reported in cyproterone acetate use but these are usually transient occurring during the first 6 months of use. No association of suicide with actual use of cyproterone acetate could be established. The main benefit of 50–100 mg/day cyproterone acetate before surgery is the effective suppression of testosterone levels and counteracting the biological action of androgens, allowing the use of estrogens in a lower dosage and a more potent biological action, particularly on breast tissue. Psychological evaluation has shown that sex reassignment increases the well-being of transsexuals, but it should not be considered as a cure-all; it is rehabilitative relieving gender dysphoria, but some transsexual subjects may still experience other problems (e.g. comorbid psychiatric problems, social isolation, troubled relationships, prejudice, and discrimination).

Our present analysis, as compared to our earlier reports, comprises a larger study population and a longer follow-up, resulting in a more apparent increased mortality rate of cardiovascular disease in MtF. This may be partially explained by heavier smoking and a higher incidence of hypercholesterolemia in MtF than in the male general population. Moreover, long-term ethinyl estradiol use was independently associated with a threefold increased risk of cardiovascular death. Our findings in 1989 (5) of an increased incidence of venous thrombosis associated with the use of ethinyl estradiol had already led to a change in type of estrogen prescription for new patients starting cross-sex hormones, and nowadays only few MtF use ethinyl estradiol or other oral estrogens in high dose. The increased risk of cardiovascular mortality was observed only in those who were still using ethinyl estradiol. No increased risk was found in former users who had changed to other formulations and lower doses of estradiol. This is reassuring for those who have changed to other estrogen preparations. The increased risk with ethinyl estradiol can possibly be explained by the thrombogenic hemostatic changes: a large increase in APC resistance and a decrease in plasma protein S that have been previously described by our group (13). The high prevalence of previous venous thrombosis (22%) in those who died from cardiovascular causes supports this hypothesis. The favorable serum lipid changes associated with ethinyl estradiol (7) – an increase in high-density lipoprotein cholesterol (+20%) and a decrease in low-density lipoprotein cholesterol (−12%) – did apparently not translate into a reduced risk of cardiovascular death. Recently, raised levels of circulatory inflammatory markers in transsexuals treated with high dose of oral estrogens have been reported, which could further contribute to the increased cardiovascular risk (10). An increased risk of cardiovascular disease was also reported in women using oral contraceptives (OC), in particular if they used OC pills with a higher ethinyl estradiol content (50 μg), even more so when they were smokers (21–23). The increased risk, however, disappeared after discontinuing OC use (24). In those MtF who had continued using ethinyl estradiol, subjects had used equally relatively high doses of about 50 μg/day up to advanced age, which could explain our finding of an increased rate of cardiovascular death.

The total cancer mortality rate was not increased. The statistically significant increase in mortality rate of lung cancer may be related to heavier smoking in the transsexual population. The increased mortality rate due to a variety of hematological cancers is puzzling. There are no reports of associations of hematological cancers with use of sex hormones. This may be a chance finding, or may be explained by the association of non-Hodgkin lymphomas with HIV, the latter might have gone unreported. The decreased mortality rate for colon cancer, also observed in the Women's Health Initiative (14), is similarly remarkable, but also this needs confirmation in further studies. We did not observe any cases of breast cancer in the population studied, neither in MtF nor in FtM, in agreement with the low prevalence of breast cancer in the literature.

Our study has a number of limitations inherent in a cohort study. Firstly, it was not randomized nor placebo-controlled which would have been difficult, if not impossible given the nature of the study population. Comparing our cohort with the general population was probably the best available option for this research but it should be noted that this comparison is potentially biased and confounded by lifestyle factors, prone to associated pathology and other factors specific for the transsexual population besides cross-sex hormone treatment. Transsexual subjects, in particular MtF, differed in a number of regards with the general population. Before they presented themselves for sex reassignment, they have an increased history of suicide attempts, more psychopathology, and substance abuse, probably associated with the psychological burden of gender dysphoria, as well as an increased prevalence of HIV infection. Secondly, the data have been collected over a 30-year period, and follow-up was not entirely complete, 40% of the subjects had their last visit to the clinic before 2007. Our cross check with the Dutch civil registry in 2009 confirmed this assumption.

In summary, increased mortality in hormone-treated MtF transsexuals was mainly due to non-hormone-related causes, such as suicide, AIDS, and drug abuse, but current use of ethinyl estradiol was associated with an increased risk of cardiovascular death. In FtM transsexuals, the use of testosterone in doses similar

to those used for replacement for hypogonadal men seemed safe, but our data in over 65-year-old FtM were limited. In line with the Endocrine Society's Clinical practice guidelines on Endocrine Treatment of Transsexual Persons (25), we strongly recommend not to prescribe ethinyl estradiol (or OC, often self-administered in higher dosages) to MtF transsexuals. Transdermal and low dose oral estradiol combined with anti-androgens are effective with fewer side effects in our experience and as published by others (26, 27). Consequently, since ethinyl estradiol is no longer used in our clinic since 2001, there is no indication to routinely test asymptomatic MtF before initiation of cross-sex therapy for (inherited) forms of thrombophilia (27), as long as the subject's history does not suggest any additional risk (25).

Lifestyle behaviors, which include healthy diets, smoking cessation, and regular exercise, may help to reduce cardiovascular risk especially in the group of MtF. Furthermore, intense preventive action may help reduce the mortality from suicide, AIDS, and drug abuse.

Declaration of interest

The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

Funding

This research did not receive any specific grant from a funding agency in the public, commercial, or not-for-profit sector.

Acknowledgements

We thank Prof. Paul Lips (Endocrine Unit) for facilitating the study, and all family physicians who helped gathering the data on causes of death.

References

1 Kuiper B & Cohen-Kettenis P. Sex reassignment surgery: a study of 141 Dutch transsexuals. *Archives of Sexual Behaviour* 1988 **17** 439–457. (doi:10.1007/BF01542484)

2 Weyers S, Elaut E, De Sutter P, Gerris J, T'Sjoen G, Heylens G, De Cuypere G & Verstraelen H. Long-term assessment of the physical, mental, and sexual health among transsexual women. *Journal of Sexual Medicine* 2009 **6** 752–760. (doi:10.1111/j.1743-6109.2008.01082.x)

3 Murad MH, Elamin MB, Garcia MZ, Mullan RJ, Murad A, Erwin PJ & Montori VM. Hormonal therapy and sex reassignment: a systemic review and meta-analysis of quality of life and psychosocial outcomes. *Clinical Endocrinology* 2010 **72** 214–231. (doi:10.1111/j.1365-2265.2009.03625.x)

4 Gooren LJ, Giltay EJ & Bunck MC. Long-term treatment of transsexuals with cross-sex hormones: extensive personal experience. *Journal of Clinical Endocrinology and Metabolism* 2008 **93** 19–25. (doi:10.1210/jc.2007-1809)

5 Asscheman H, Gooren LJG & Eklund PL. Mortality and morbidity in transsexual patients with cross-gender treatment. *Metabolism* 1989 **38** 869–873. (doi:10.1016/0026-0495(89)90233-3)

6 van Kesteren PJM, Asscheman H, Megens JAJ & Gooren LJG. Mortality and morbidity in transsexual subjects treated with cross-sex hormones. *Clinical Endocrinology* 1997 **47** 337–342. (doi:10.1046/j.1365-2265.1997.2601068.x)

7 Elbers JM, Giltay EJ, Teerlink T, Scheffer PG, Asscheman H, Seidell JC & Gooren LJ. Effects of sex steroids on components of the insulin resistance syndrome in transsexual subjects. *Clinical Endocrinology* 2003 **58** 562–571. (doi:10.1046/j.1365-2265.2003.01753.x)

8 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. *American Journal of Physiology* 1999 **276** E317–E325.

9 Polderman KH, Gooren LJ, Asscheman H, Bakker A & Heine RJ. Induction of insulin resistance by androgens and estrogens. *Journal of Clinical Endocrinology and Metabolism* 1994 **79** 265–271. (doi:10.1210/jc.79.1.265)

10 Giltay EJ, Lambert J, Gooren LJ, Elbers JM, Steyn M & Stehouwer CD. Sex steroids, insulin, and arterial stiffness in women and men. *Hypertension* 1999 **34** 590–597.

11 Elting MW, Korsen TJ, Bezemer PD & Schoemaker J. Prevalence of diabetes mellitus, hypertension and cardiac complaints in a follow-up study of a Dutch PCOS population. *Human Reproduction* 2001 **16** 556–560. (doi:10.1093/humrep/16.3.556)

12 Shaw LJ, Bairey Merz CN, Azziz R, Stanczyk FZ, Sopko G, Braunstein GD, Kelsey FZ, Klp KE, Cooper-DeHoff RM, Johnson BD, Vaccarino V, Reis SE, Bittner V, Hodgson K, Rogers W & Pepine CJ. Postmenopausal women with a history of irregular menses and elevated androgen measurements at high risk for worsening cardiovascular event-free survival: results from the National Institute of Health- National Heart, Lung, and Blood Institute sponsored women's ischemia syndrome evaluation. *Journal of Clinical Endocrinology and Metabolism* 2008 **93** 1276–1284. (doi:10.1210/jc.2007-0425)

13 Hulley S, Grady D, Bush T, Furbert C, Herrington D, Riggs B & Vittinghoff E. Randomized trial of estrogen plus progestin for secondary prevention of coronary heart disease in postmenopausal women Heart and Estrogen/progestin Replacement Study (HERS) Research Group. *Journal of the American Medical Association* 1998 **288** 605–613. (doi:10.1001/jama.280.7.605)

14 Rossouw JE, Anderson GL, Prentice RL, LaCroix AZ, Koopersberg C, Stefanick ML, Jackson RD, Beresford SA, Howard BV, Johnson KC, Kotchen JM & Ockene J. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results from the Women's Health Initiative randomized controlled trial. *Journal of the American Medical Association* 2002 **288** 321–333. (doi:10.1001/jama.288.3.321)

15 Wilson R, Spiers A, Ewan J, Johnson P, Jenkins C & Carr S. Effects of high dose oestrogen therapy on circulating inflammatory markers. *Maturitas* 2009 **62** 281–286. (doi:10.1016/j.maturitas.2009.01.009)

16 Giltay EJ, Verhoef P, Gooren LJ, Geleijnse JM, Schouten EG & Stehouwer CD. Oral and transdermal estrogens both lower plasma total homocysteine in male-to-female transsexuals. *Atherosclerosis* 2003 **168** 139–146. (doi:10.1016/S0021-9150(03)00090-X)

17 Post MS, Christella M, Thomassen LG, van der Mooren MJ, Van Baal WM, Rosing J, Kenemans P & Stehouwer CD. Effect of oral and transdermal estrogen replacement therapy on hemostatic variables associated with venous thrombosis: a randomized, placebo-controlled study in postmenopausal women. *Arteriosclerosis, Thrombosis, and Vascular Biology* 2003 **23** 1116–1121. (doi:10.1161/01.ATV.0000074146.36646.C8)

18 Toorians AWFT, Thomassen MCLGD, Zweegman S, Magdeleyns EJP, Tans G, Gooren LJG & Rosing J. Venous thrombosis and changes of hemostatic variables during cross-sex hormone treatment in transsexual people. *Journal of Clinical Endocrinology and Metabolism* 2003 **88** 5723–5729. (doi:10.1210/jc.2003-030520)

19 Asscheman H, Giltay EJ, Megens JAJ & Gooren LJG. Long term mortality in hormone-treated transsexuals. *Program of the 91st Annual meeting of The Endocrine Society, Washington, DC*, 2009, P3–314 (Abstract).

20 Gardner MJ & Altman DFG. *Statistics with Confidence*. London: British Medical Journal, 1989.
21 Khader YS, Rice J, John L & Abueita O. Oral contraceptive use and the risk of myocardial infarction: a meta-analysis. *Contraception* 2003 **68** 11–17. (doi:10.1016/S0010-7824(03)00073-8)
22 Tanis BC, van den Bosch MA, Kemmeren JM, Cats VM, Helmerhorst FM, Algra A, van der Graaf Y & Rosendaal FR. Oral contraceptives and the risk of myocardial infarction. *New England Journal of Medicine* 2001 **345** 1787–1793. (doi:10.1056/NEJMoa003216)
23 Shufelt CL & Bairey Merz CN. Contraceptive hormone use and cardiovascular disease. *Journal of the American College of Cardiology* 2009 **53** 221–231. (doi:10.1016/j.jacc.2008.09.042)
24 Stampfer MJ, Willett WC, Colditz GA, Speizer FE & Hennekens CH. A prospective study of past use of contraceptive agents and the risk of cardiovascular diseases. *New England Journal of Medicine* 1988 **319** 1313–1317. (doi:10.1056/NEJM198811173192004)
25 Hembree WC, Cohen-Kettenis P, Delemarre-van de Waal HA, Gooren LJ, Meyer WJ III, Spack NP, Tangpricha V & Montori VM. Endocrine treatment of transsexual persons: an Endocrine Society clinical practice guideline. *Journal of Clinical Endocrinology and Metabolism* 2009 **94** 3132–3154. (doi:10.1210/jc.2009-0345)
26 Dittrich R, Binder H, Cupisti S, Hoffman I, Beckmann MW & Mueller A. Endocrine treatment of male-to-female transsexuals using gonadotropin-releasing hormone agonist. *Experimental and Clinical Endocrinology & Diabetes* 2005 **113** 586–592. (doi:10.1055/s-2005-865900)
27 Ott J, Kaufmann U, Bentz EK, Huber JC & Tempfer CB. Incidence of thrombophilia and venous thrombosis in transsexuals under cross-sex hormone therapy. *Fertility and Sterility* 2010 **93** 1267–1272. (doi:10.1016/j.fertnstert.2008.12.017)

Received 15 December 2010
Accepted 25 January 2011

Archives of Sexual Behavior
https://doi.org/10.1007/s10508-022-02287-7

LETTER TO THE EDITOR

Suicide by Clinic-Referred Transgender Adolescents in the United Kingdom

Michael Biggs[1]

Received: 24 April 2021 / Revised: 5 January 2022 / Accepted: 5 January 2022

Introduction

Surveys show that adolescents who identify as transgender are vulnerable to suicidal thoughts and self-harming behaviors (dickey & Budge, 2020; Hatchel et al., 2021; Mann et al., 2019). Little is known about death by suicide. This Letter presents data from the Gender Identity Development Service (GIDS), the publicly funded clinic for children and adolescents aged under 18 from England, Wales, and Northern Ireland. From 2010 to 2020, four patients were known or suspected to have died by suicide, out of about 15,000 patients (including those on the waiting list). To calculate the annual suicide rate, the total number of years spent by patients under the clinic's care is estimated at about 30,000. This yields an annual suicide rate of 13 per 100,000 (95% confidence interval: 4–34). Compared to the United Kingdom population of similar age and sexual composition, the suicide rate for patients at the GIDS was 5.5 times higher. The proportion of patients dying by suicide was far lower than in the only pediatric gender clinic which has published data, in Belgium (Van Cauwenberg et al., 2021).

Suicidality in Transgender Adolescents

"About half of young trans people. . .attempt suicide," declared the United Kingdom Parliament's Women and Equalities Committee (2015). Similar figures are cited by news media and campaigning organizations. The *Guardian* reported Stonewall's statistic that "almost half" of transgender young people "have attempted to kill themselves" (Weale, 2017). "Fifty percent of transgender youth attempt suicide before they are at age 21" stated the mother of the most famous transgender youth in the English-speaking world (Jennings & Jennings, 2016). As a transgender theologian has observed, "the statistic about suicide attempts has, in essence, developed a life of its own" (Tanis, 2016).

Representative surveys of students in high schools provide one source of evidence for this statistic. In New Zealand, 20% of transgender students reported attempting suicide in the past 12 months, compared to 4% of all students (Clark et al., 2014). In the United States, 15% of transgender students reported a suicide attempt requiring medical treatment in the last 12 months, compared to 3% of all students (Centers for Disease Control & Prevention, 2018; Jackman et al., 2021; Johns et al., 2019). In another American survey, 41% of transgender students reported having attempted suicide during their lifetime, compared to 14% of all students (Toomey et al., 2018).

To what extent are self-reported suicide attempts reflected in fatalities? The connection is not straightforward. Respondents who report suicide attempts are not necessarily indicating an intent to die. One survey of the American population found that almost half the respondents who reported attempting suicide subsequently stated that their action was a cry for help and not intended to be fatal (Nock & Kessler, 2006). In two small samples of non-heterosexual youth, half the respondents who initially reported attempting suicide subsequently clarified that they went no further than imagining or planning it; for the remainder who did actually attempt suicide, their actions were usually not life-threatening. To an extent, then, "the reports were attempts to communicate the hardships of lives or to identify with a gay community" (Savin-Williams, 2001). Although such elaborate survey methods have not been used to study transgender populations, there is anecdotal evidence for a similar disjuncture. The pediatric endocrinologist who established the first clinic for transgender children in the United States stated that "the majority of self-harmful actions that I see in my clinic are not real suicide attempts and are not usually life threatening" (Spack, 2009).

Suicide mortality has been studied in the transgender population using registry data. The annual suicide rate is calculated by dividing the number of suicides by the total number of years each person was at risk. An individual who was observed for 20 years, for instance, contributes 20 person-years to the denominator. The

✉ Michael Biggs
michael.biggs@sociology.ox.ac.uk

[1] Department of Sociology and St Cross College, University of Oxford, 42 Park End Street, Oxford OX1 1JD, UK

Published online: 18 January 2022

largest study covers over 8,000 patients who visited the gender clinic in Amsterdam from 1972 to 2017 (Wiepjes et al., 2020). The annual suicide rate was 29 per 100,000 for transmen, quadruple the rate for the female population, and 64 for transwomen, quadruple the rate for the male population. A Swedish study of 324 individuals who had undergone genital surgery between 1973 and 2003 found much higher annual suicide rates: 250 per 100,000 for transmen, 43 times the rate for matched female controls, and 285 for transwomen, 16 times the rate for matched male controls (M. Boman, personal communication, 12 April 2021; Dhejne et al., 2011). Only one published study has reported suicide fatalities among transgender adolescents. Belgium's pediatric gender clinic provided counseling to 177 youth aged from 12 to 18 years, who had been referred between 2007 and 2016: five of them (2.8%) committed suicide (Van Cauwenberg et al., 2021). The mean age of referral was 15, implying a mean duration of 3 years before transition to an adult clinic, which translates to an annual suicide rate of 942 per 100,000. This is the highest suicide mortality recorded for any transgender population.

Method

This Letter estimates the suicide rate at the world's largest pediatric gender clinic. Based in London, the GIDS is part of the Tavistock and Portman NHS Foundation Trust, and serves youth under 18 from England, Wales, and Northern Ireland who are "experiencing difficulties with their gender identity development" (Carmichael & Davidson, 2009). Like all such services throughout Western Europe and North America, it has experienced enormous growth; referrals increased from 100 in 2009 to a peak of 2700 in 2019. The waiting list in April 2021 exceeded 5300.

The GIDS patients manifest typically high rates of self-harming behavior. In a sample of 900 adolescents (aged from 13 to 17) admitted to the clinic from 2009 to 2017 and given the Youth Self-Report questionnaire, 44% answered that they sometimes or very often "deliberately try to hurt or kill myself" (de Graaf et al., 2020). Unfortunately, both behaviors are combined in this question. In a different sample of over 700 children and adolescents (aged from 4 to 17) assessed by the GIDS in 2012 and 2015, 10% were flagged by clinicians as having attempted suicide (Morandini et al., 2021).

Suicides

Since the early 2000s, the National Health Service has implemented mandatory reporting of "serious incidents" (Department of Health, 2001, 2010). The death of any patient—including those on the waiting list—suspected to be suicide is reported to the Tavistock's Board of Directors. The Tavistock cooperates with a comprehensive surveillance system for every death classified as suicide or (after an open verdict by the coroner) probable suicide in the United Kingdom (National Confidential Inquiry into Suicide & Homicide by People with Mental Illness, 1999; National Confidential Inquiry into Suicide and Safety in Mental Health, 2019). Papers for the Tavistock's Board meetings are available from April 2007 onwards; those not on the Trust's website were acquired by a Freedom of Information request. The pdf files of the *Agenda and Papers* (through September 2021) were searched for the keyword "suicid"; all 442 instances were inspected. From 2007 to 2020, four patients of the GIDS died by suspected suicide: two on the waiting list, in 2016 and 2017; and two after having been seen, in 2017 and 2020. The last case was described as "likely" to be suicide, because the inquest has not yet been held. These figures were confirmed by Freedom of Information requests to the Tavistock.

Triangulation is possible from two sources. Comprehensive mortality data on all adolescents aged from 10 to 19 who committed suicide in the United Kingdom from 2014 to 2016 include five transgender individuals (Rodway et al., 2020). Due to confidentiality restrictions, it is not possible to disaggregate these further by age or by country. Presumably, one of these is the patient of GIDS who died in 2016. The remaining four might have been 18 or 19—the risk of suicide increases significantly in the late teens—or might have lived in Scotland. Alternatively, they might have been eligible for the GIDS but had not sought a clinical referral (made by the local Child and Adolescent Mental Health Service, the child's general practitioner, social worker, or teacher) or had not obtained it.

Another source is the Transgender Day of Remembrance website, which aims to record all deaths by suicide or violence (Metcalfe, 2021). For the United Kingdom between 2007 and 2020, the website names 3 adolescents under the age of 18 who committed suicide. One was one of the GIDS patients (the match is certain because they were named in the *Agenda and Papers*). The other two had no involvement with the GIDS (or any other gender clinician), as was evident from their inquests, though one was under the psychiatric care of another NHS Trust (BBC News, 2021; Bunyan, 2008). In addition, the website lists suicides by two "young" transgender people, sourced from Twitter, without information on their name or age. In one case, it is not clear whether the person lived in the United Kingdom.

Patients

With suicides as numerator, two denominators are relevant. Because comprehensive data on patient numbers became available from 2010, the period will be the 11 years from 2010 to 2020. (These are financial years; thus, 2020 runs from April 2020 to March 2021.) The first denominator is the total number of individual patients, estimated by summing the annual number of referrals to the GIDS from 2010 to 2020—excluding those aged 18 or over, as they are not accepted. The total number is 15,032. This sum omits patients at the clinic who had been referred before

2010, and so is a slight underestimate. (The Online Supplement provides full details.)

The second denominator is the total number of patient-years: the sum of the number of years spent by each individual as a patient of the GIDS. The number of patients seen by the GIDS each year was available from 2014 to 2020. Before 2014 only the number of patients first seen was available. From 2014 to 2016, the number of patients seen was consistently double the number first seen, and so the former number for 2010 to 2013 was estimated by doubling the latter. All these numbers exclude patients on the waiting list. The number waiting at the beginning of each year from 2016 to 2020 was obtained by Freedom of Information request. Before then the number was not available, and so must be treated as zero. This leads to an underestimate, of course, but the waiting list became appreciable only from 2015. The total number of patient-years over this period is estimated as 30,080. In other words, patients spent on average 2 years at the GIDS (including time on the waiting list). Time on the waiting list contributed 41% of the total patient-years.

Results

From 2010 to 2020, the four suicide deaths equate to 0.03% of the 15,032 patients. Taking the denominator as 30,080 patient-years, the annual suicide rate is calculated as 13 per 100,000 (95% confidence interval: 4 to 34 per 100,000). For comparison, the annual suicide rate in England and Wales between 2010 and 2020 for adolescents aged from 15 to 19 years averaged 4.7 (Office for National Statistics, 2021). This does not quite correspond to the age range of the GIDS patients, however. At referral, the patients ranged in age from 3 to 17 years; only 7% were younger than 10. The mean was 14 years and the median 15. Most patients stay with the GIDS until transitioning to an adult service. Therefore, the average age of patients at any point in time will lie somewhere between 14 and 17. A better comparison is therefore the overall suicide rate for adolescents aged from 14 to 17 (available only for the entire United Kingdom for 2015–2017), which was 2.7 per 100,000 (Office for National Statistics, 2018; Rodway et al., 2020). Comparison should also account for the difference between the sexes, because males are more likely to commit suicide than females. Of the GIDS patients, 69% were female. Adjusting for sex, the GIDS patients were 5.5 times more likely to commit suicide than the overall population of adolescents aged 14 to 17.

Discussion

How reliable are these estimates? The chief uncertainty about the numerator is whether the fourth death will be ruled as suicide when the inquest is eventually held. It could be speculated that there were further suicides unknown to the Tavistock and to the National Confidential Inquiry into Suicide and Safety in Mental Health. All that can be said is that the single suicide by a GIDS patient from 2014 to 2016 is not out of line with comprehensive mortality data on suicides by transgender adolescents in the United Kingdom which counted five suicides in a longer age range and wider geographical area. The denominator for the annual suicide rate, however, is pieced together from various series and so is inevitably approximate. Statistics from the early 2010s are less reliable, though they make only a small contribution to the grand total; the last three years contribute more than half of the total number of patient-years. The most significant limitation is the lack of information on the age and sex of all the patients who committed suicide.

Direct comparison can be made with the Belgian pediatric gender clinic (Van Cauwenberg et al., 2021). Its annual suicide rate was about 70 times greater than the rate at the GIDS. This is especially puzzling because patients at the Belgian clinic scored better, on average, than those at the GIDS on tests of psychological functioning (de Graaf et al., 2018). The explanation for the huge disparity in suicide is not clear. The Amsterdam's clinic annual suicide rate was four times greater than the rate at the GIDS. The higher rate is not surprising, however, because the Dutch clinical population was dominated by older adults: the median age at first visit was 25 (Wiepjes et al., 2020). Suicide rates peak in middle age, and so a population of older adults would be at higher risk than a population of adolescents.

The suicide rate of the GIDS patients is not necessarily indicative of the rate among all adolescents who identify as transgender. On the one hand, individuals with more serious problems (and their families) would be particularly motivated to seek referral and more likely to obtain it, and so the clinical subset would be more prone to suicide. One study suggests that a child who frequently attempted suicide was more readily referred to the GIDS (Carlile et al., 2021). On the other hand, young people facing hostility from their families would be less able to seek referral, and this hostility could make them especially vulnerable to suicide.

Taking into account these limitations, the estimated suicide rate at the GIDS provides the strongest evidence yet published that transgender adolescents are more likely to commit suicide than the overall adolescent population. The higher risk could have various causes: gender dysphoria, accompanying psychological conditions, and ensuing social disadvantages such as bullying. Studies of young people referred to the GIDS in 2012 and 2015 found a high prevalence of eating disorders, depression, and autism spectrum conditions (ASC) (Holt et al., 2016; Morandini et al., 2021)—all known to increase the probability of suicide (Simon & VonKorff, 1998; Smith et al., 2018). Eating disorders and depression could be consequences of transgender identity and its ensuing social repercussions, but this is implausible for ASC insofar as it originates in genes or the prenatal environment. From a sample of over 700 referrals to the GIDS in 2012 and 2015, 14–15% were diagnosed with ASC (Morandini

et al., 2021). This compared to 0.8–1.1% of students in England (Department for Education, 2012, 2015). The association between autism and gender dysphoria is found in many populations (Socialstyrelsen, 2020; Warrier et al., 2020). Autism is known to increase the risk of suicide mortality, especially in females (Hirvikoski et al., 2016; Kirby et al., 2019; Socialstyrelsen, 2020). To some extent, therefore, the elevated suicide rate for transgender youth compared to their peers reflects the higher incidence of ASC. The same holds for other psychiatric disorders associated with gender dysphoria (Dhejne et al., 2016). Ideally, the suicide rate for patients of the GIDS would be compared to the suicide rate for patients in contact with other NHS mental health services, but the latter rate is not available.

One final caveat is that these data shed no light on the question of whether counseling or endocrinological interventions—gonadotropin-releasing hormone agonist or cross-sex hormones—affect the risk of suicide (Biggs, 2020; Turban et al., 2020). Although two out of the four suicides were of patients on the waiting list, and thus would not have obtained treatment, this is not disproportionate: the waiting list contributed nearly half of the total patient-years.

Conclusion

Data from the world's largest clinic for transgender youth over 11 years yield an estimated annual suicide rate of 13 per 100,000. This rate was 5.5 times greater than the overall suicide rate of adolescents of similar age, adjusting for sex composition. The estimate demonstrates the elevated risk of suicide among adolescents who identify as transgender, albeit without adjusting for accompanying psychological conditions such as autism. The proportion of individual patients who died by suicide was 0.03%, which is orders of magnitude smaller than the proportion of transgender adolescents who report attempting suicide when surveyed. The fact that deaths were so rare should provide some reassurance to transgender youth and their families, though of course this does not detract from the distress caused by self-harming behaviors that are non-fatal. It is irresponsible to exaggerate the prevalence of suicide. Aside from anything else, this trope might exacerbate the vulnerability of transgender adolescents. As the former lead psychologist at the Tavistock has warned, "when inaccurate data and alarmist opinion are conveyed very authoritatively to families we have to wonder what the impact would be on children's understanding of the kind of person they are…and their likely fate" (Wren, 2015).

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10508-022-02287-7.

Acknowledgements The author thanks Cathryn Rodway for information on the National Confidential Inquiry into Suicide and Safety in Mental Health, Marcus Boman for data on Swedish suicides, and Susan Bewley, Marcus Evans, Susan Evans, Susan Matthews, Avi Ring, and James Thornhill for criticisms and suggestions.

Declarations

Conflict of interest I acted as an expert witness (without payment) for the claimant in the case of Bell v Tavistock and Portman NHS Foundation Trust [2020] EWHC 3274.

References

BBC News. (2021, June 4). Ellis Murphy-Richards: NHS care questioned over suicide. *BBC News*. https://www.bbc.com/news/uk-england-kent-57358063

Biggs, M. (2020). Puberty blockers and suicidality in adolescents suffering from gender dysphoria [Letter to the Editor]. *Archives of Sexual Behavior, 49*, 2227–2229. https://doi.org/10.1007/s10508-020-01743-6

Bunyan, N. (2008, February 15). Boy, 10, hangs himself after talking to his mother about craze. *Daily Telegraph*.

Carlile, A., Butteriss, E., & Sansfaçon, A. P. (2021). "It's like my kid came back overnight": Experiences of trans and non-binary young people and their families seeking, finding and engaging with clinical care in England. *International Journal of Transgender Health, 22*, 412–424. https://doi.org/10.1080/26895269.2020.1870188

Carmichael, P., & Davidson, S. (2009). A gender identity development service. *Psychologist, 22*, 916–917.

Centers for Disease Control and Prevention. (2018). *2017 Youth Risk Behavior Survey*. https://www.cdc.gov/healthyyouth/data/yrbs/data.htm

Clark, T. C., Lucassen, M. F. G., Bullen, P., Denny, S. J., Fleming, T. M., Robinson, E. M., & Rossen, F. V. (2014). The health and well-being of transgender high school students: Results from the New Zealand Adolescent Health Survey (Youth'12). *Journal of Adolescent Health, 55*, 93–99. https://doi.org/10.1016/j.jadohealth.2013.11.008

de Graaf, N. M., Cohen-Kettenis, P. T., Carmichael, P., de Vries, A. L. C., Dhondt, K., Laridaen, J., Pauli, D., Ball, J., & Steensma, T. D. (2018). Psychological functioning in adolescents referred to specialist gender identity clinics across Europe: A clinical comparison study between four clinics. *European Child & Adolescent Psychiatry, 27*, 909–919. https://doi.org/10.1007/s00787-017-1098-4

de Graaf, N. M., Steensma, T. D., Carmichael, P., VanderLaan, D. P., Aitken, M., Cohen-Kettenis, P. T., de Vries, A. L. C., Kreukels, B. P. C., Wasserman, L., Wood, H., & Zucker, K. J. (2020). Suicidality in clinic-referred transgender adolescents. *European Child & Adolescent Psychiatry*. https://doi.org/10.1007/s00787-020-01663-9

Department for Education. (2012). *Special educational needs in England, January 2012*. https://www.gov.uk/government/statistics/special-educational-needs-in-england-january-2012

Department for Education. (2015). *Special educational needs in England, January 2015*. https://www.gov.uk/government/statistics/special-educational-needs-in-england-january-2015

Department of Health. (2001). *Building a safer NHS for patients: Implementing 'an organisation with a memory'*.

Department of Health. (2010). *National framework for reporting and learning from serious incidents requiring investigation*. https://web.archive.org/web/20101126102908/http://www.nrls.npsa.nhs.uk/EasySiteWeb/getresource.axd?AssetID=68464&type=full&servicetype=Attachment

Dhejne, C., Lichtenstein, P., Boman, M., Johansson, A. L. V., Långström, N., & Landén, M. (2011). Long-term follow-up of transsexual persons undergoing sex reassignment surgery: Cohort study in Sweden. *PLoS ONE, 6*, e16885. https://doi.org/10.1371/journal.pone.0016885

Dhejne, C., Van Vlerken, R., Heylens, G., & Arcelus, J. (2016). Mental health and gender dysphoria: A review of the literature. *International Review of Psychiatry, 28*, 44–57. https://doi.org/10.3109/09540261.2015.1115753

dickey, l. m., & Budge, S. L. (2020). Suicide and the transgender experience: A public health crisis. *American Psychologist, 75*, 380–390. https://doi.org/10.1037/amp0000619

Hatchel, T., Polanin, J. R., & Espelage, D. L. (2021). Suicidal thoughts and behaviors among LGBTQ youth: Meta-analyses and a systematic review. *Archives of Suicide Research, 25*, 1–37. https://doi.org/10.1080/13811118.2019.1663329

Hirvikoski, T., Mittendorfer-Rutz, E., Boman, M., Larsson, H., Lichtenstein, P., & Bölte, S. (2016). Premature mortality in autism spectrum disorder. *British Journal of Psychiatry, 208*, 232–238. https://doi.org/10.1192/bjp.bp.114.160192

Holt, V., Skagerberg, E., & Dunsford, M. (2016). Young people with features of gender dysphoria: Demographics and associated difficulties. *Clinical Child Psychology & Psychiatry, 21*, 108–118. https://doi.org/10.1177/1359104514558431

Jackman, K. B., Caceres, B. A., Kreuze, E. J., & Bockting, W. O. (2021). Suicidality among gender minority youth: Analysis of 2017 Youth Risk Behavior Survey data. *Archives of Suicide Research, 25*, 208–223. https://doi.org/10.1080/13811118.2019.1678539

Jennings, J., & Jennings, J. (2016). Trans teen shares her story. *Pediatrics in Review, 37*, 99–100. https://doi.org/10.1542/pir.2016-002

Johns, M. M., Lowry, R., Andrzejewski, J., Barrios, L. C., Demissie, Z., McManus, T., Rasberry, C. N., Robin, L., & Underwood, J. M. (2019). Transgender identity and experiences of violence victimization, substance use, suicide risk, and sexual risk behaviors among high school students: 19 states and large urban school districts, 2017. *Morbidity and Mortality Weekly Report, 68*, 69–71.

Kirby, A. V., Bakian, A. V., Zhang, Y., Bilder, D. A., Keeshin, B. R., & Coon, H. (2019). A 20-year study of suicide death in a statewide autism population. *Autism Research, 12*, 658–666. https://doi.org/10.1002/aur.2076

Mann, G. E., Taylor, A., Wren, B., & de Graaf, N. (2019). Review of the literature on self-injurious thoughts and behaviours in gender-diverse children and young people in the United Kingdom. *Clinical Child Psychology & Psychiatry, 24*, 304–321. https://doi.org/10.1177/1359104518812724

Metcalfe, A.-J. (2021). *Remembering our dead: Reports*. https://tdor.translivesmatter.info/reports

Morandini, J. S., Kelly, A., de Graaf, N. M., Carmichael, P., & Dar-Nimrod, I. (2021). Shifts in demographics and mental health co-morbidities among gender dysphoric youth referred to a specialist gender dysphoria service. *Clinical Child Psychology & Psychiatry*. https://doi.org/10.1177/13591045211046813

National Confidential Inquiry into Suicide and Homicide by People with Mental Illness. (1999). *Safer services: National confidential inquiry into suicide and homicide by people with mental illness*.

National Confidential Inquiry into Suicide and Safety in Mental Health. (2019). *Annual Report: England, Northern Ireland, Scotland and Wales*. Healthcare Quality Improvement Partnership. https://www.hqip.org.uk/resource/national-confidential-inquiry-into-suicide-and-safety-in-mental-health-annual-report-2019/#.YdQiOy-l3KY

Nock, M. K., & Kessler, R. C. (2006). Prevalence of and risk factors for suicide attempts versus suicide gestures: Analysis of the National Comorbidity Survey. *Journal of Abnormal Psychology, 115*, 616–623. https://doi.org/10.1037/0021-843X.115.3.616

Office for National Statistics. (2018). *Population estimates for UK, England and Wales, Scotland and Northern Ireland: Mid-2012 to mid-2016*. https://www.ons.gov.uk/file?uri=%2fpeoplepopulationandcommunity%2fpopulationandmigration%2fpopulationestimates%2fdatasets%2fpopulationestimatesforukenglandandwalesscotlandandnorthernireland%2fmid2012tomid2016

Office for National Statistics. (2021). *Suicides in England and Wales*. https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/datasets/suicidesintheunitedkingdomreferencetables

Rodway, C., Tham, S.-G., Ibrahim, S., Turnbull, P., Kapur, N., & Appleby, L. (2020). Children and young people who die by suicide: Childhood-related antecedents, gender differences and service contact. *Bjpsych Open, 6*, e49. https://doi.org/10.1192/bjo.2020.33

Savin-Williams, R. C. (2001). Suicide attempts among sexual-minority youths: Population and measurement issues. *Journal of Consulting and Clinical Psychology, 69*, 983–991. https://doi.org/10.1037/0022-006X.69.6.983

Simon, G. E., & VonKorff, M. (1998). Suicide mortality among patients treated for depression in an insured population. *American Journal of Epidemiology, 147*, 155–160. https://doi.org/10.1093/oxfordjournals.aje.a009428

Smith, A. R., Zuromski, K. L., & Dodd, D. R. (2018). Eating disorders and suicidality: What we know, what we don't know, and suggestions for future research. *Current Opinion in Psychology, 22*, 63–67. https://doi.org/10.1016/j.copsyc.2017.08.023

Socialstyrelsen. (2020). *Utvecklingen av diagnosen könsdysfori* [The evolution of the diagnosis of gender dysphoria]. https://www.socialstyrelsen.se/globalassets/sharepoint-dokument/artikelkatalog/ovrigt/2020-2-6600.pdf

Spack, N. P. (2009). An endocrine perspective on the care of transgender adolescents. *Journal of Gay & Lesbian Mental Health, 13*, 309–319. https://doi.org/10.1080/19359700903165381

Tanis, J. (2016). The power of 41%: A glimpse into the life of a statistic. *American Journal of Orthopsychiatry, 86*, 373–377. https://doi.org/10.1037/ort0000200

Toomey, R. B., Syvertsen, A. K., & Shramko, M. (2018). Transgender adolescent suicide behavior. *Pediatrics, 142*, e20174218. https://doi.org/10.1542/peds.2017-4218

Turban, J. L., King, D., Carswell, J. M., & Keuroghlian, A. S. (2020). Pubertal suppression for transgender youth and risk of suicidal ideation. *Pediatrics, 145*, e20191725. https://doi.org/10.1542/peds.2019-1725

Van Cauwenberg, G., Dhondt, K., & Motmans, J. (2021). Ten years of experience in counseling gender diverse youth in Flanders, Belgium: A clinical overview. *International Journal of Impotence Research., 33*, 671–678. https://doi.org/10.1038/s41443-021-00441-8

Warrier, V., Greenberg, D. M., Weir, E., Buckingham, C., Smith, P., Lai, M.-C., Allison, C., & Baron-Cohen, S. (2020). Elevated rates of autism, other neurodevelopmental and psychiatric diagnoses, and autistic traits in transgender and gender-diverse individuals. *Nature Communications, 11*(3959). https://doi.org/10.1038/s41467-020-17794-1

Weale, S. (2017, June 27). Almost half of trans pupils in UK have attempted suicide, survey finds. *Guardian*. Retrieved from https://www.theguardian.com/education/2017/jun/27/half-of-trans-pupils-in-the-uk-tried-to-take-their-own-livessurvey-finds

Wiepjes, C. M., den Heijer, M., Bremmer, M. A., Nota, N. M., Blok, C. J. M., Coumou, B. J. G., & Steensma, T. D. (2020). Trends in

suicide death risk in transgender people: Results from the Amsterdam Cohort of Gender Dysphoria study (1972–2017). *Acta Psychiatrica Scandinavica, 141*, 486–491. https://doi.org/10.1111/acps.13164

Women and Equalities Committee, UK Parliament. (2015). *Transgender equality: First report of session 2015–16*. https://publications.parliament.uk/pa/cm201516/cmselect/cmwomeq/390/390.pdf

Wren, B. (2015). *Making up people*. Presented at the meeting of the European Professional Association for Transgender Health, Ghent, Belgium.

the choice of aspirin or heparin for venous thromboembolism prophylaxis among patients with operatively treated extremity fractures (or any pelvic or acetabular fracture), this is by far the largest trial to date and provides compelling evidence that a readily available, inexpensive drug, taken orally, is a viable alternative to an injectable pharmacologic prophylaxis.

Are there any caveats to this message? The trial shows several secondary outcomes that support the main conclusion of the trial, including a similar risk of pulmonary embolism in the two groups and, in terms of safety outcomes, no evidence of a difference in the incidence of bleeding events, which occurred in 13.72% of patients in the aspirin group and 14.27% in the low-molecular-weight–heparin group. However, in keeping with previous trials, the authors noted that deep-vein thrombosis was more frequent in patients who had received aspirin than in those who had received heparin (2.51% vs. 1.71%), although the absolute difference was small (0.80 percentage points). Although deep-vein thrombosis is clearly not as serious as a fatal pulmonary embolism, it is not an inconsequential problem. Post-thrombotic syndrome affects some people who have had a deep-vein thrombosis of the leg, and this condition can cause chronic pain and swelling.[9]

The findings in this trial clearly indicate that guidelines for the prevention of hospital-acquired venous thromboembolism will need to be rewritten to include the option of aspirin in patients with traumatic injuries. More work is needed to determine whether aspirin should also be considered for venous thromboembolism prophylaxis after other types of surgeries and for nonsurgical patients who have risk factors for venous thromboembolism.

Disclosure forms provided by the author are available with the full text of this editorial at NEJM.org.

From Oxford Trauma and Emergency Care, Nuffield Department of Orthopedics, Rheumatology, and Musculoskeletal Sciences, University of Oxford, Oxford, United Kingdom.

1. National Institute for Health and Care Excellence. Venous thromboembolism in over 16s: reducing the risk of hospital-acquired deep vein thrombosis or pulmonary embolism. August 13, 2019 (https://www.nice.org.uk/guidance/ng89/chapter/Recommendations).
2. Hegsted D, Gritsiouk Y, Schlesinger P, Gardiner S, Gubler KD. Utility of the risk assessment profile for risk stratification of venous thrombotic events for trauma patients. Am J Surg 2013;205:517-20.
3. Geerts WH, Code KI, Jay RM, Chen E, Szalai JP. A prospective study of venous thromboembolism after major trauma. N Engl J Med 1994;331:1601-6.
4. Barrera LM, Perel P, Ker K, Cirocchi R, Farinella E, Morales Uribe CH. Thromboprophylaxis for trauma patients. Cochrane Database Syst Rev 2013;3:CD008303.
5. Rogers FB, Cipolle MD, Velmahos G, Rozycki G, Luchette FA. Practice management guidelines for the prevention of venous thromboembolism in trauma patients: the EAST Practice Management Guidelines Work Group. J Trauma 2002;53:142-64.
6. Colwell CW Jr, Pulido P, Hardwick ME, Morris BA. Patient compliance with outpatient prophylaxis: an observational study. Orthopedics 2005;28:143-7.
7. Horner D, Goodacre S, Pandor A, et al. Thromboprophylaxis in lower limb immobilisation after injury (TiLLI). Emerg Med J 2020;37:36-41.
8. Major Extremity Trauma Research Consortium (METRC). Aspirin or low-molecular-weight heparin for thromboprophylaxis after a fracture. N Engl J Med 2023;388:203-13.
9. Makedonov I, Kahn SR, Abdulrehman J, et al. Prevention of the postthrombotic syndrome with anticoagulation: a narrative review. Thromb Haemost 2022;122:1255-64.

DOI: 10.1056/NEJMe2214045

Growing Evidence and Remaining Questions in Adolescent Transgender Care

Annelou L.C. de Vries, M.D., Ph.D., and Sabine E. Hannema, M.D., Ph.D.

This week in the *Journal*, a much-awaited primary report from Chen et al.[1] on 2 years of gender-affirming hormones (GAH) in transgender adolescents appears. The approach to adolescent transgender care with early treatment with puberty blockers, and GAH in youth from 16 years of age, originated in the Netherlands ("the Dutch model") and became the dominant medical care model for transgender adolescents.[2] Especially over the past decade, marked increases in referrals but limited evidence as to long-term outcomes have led to controversies and debate regarding this approach. Indeed, some European countries are adapting their guidelines and re-

stricting access to care for transgender youth, and some states in the United States have introduced laws to ban such care.[3] Therefore, rigorous longitudinal outcome studies that provide evidence about whether this approach is effective and safe are needed.

The results of the current study — involving a large, multisite sample of 315 participants — provide such evidence. During 24 months of GAH treatment, participant-reported appearance congruence (alignment between gender identity and physical appearance), positive affect, and life satisfaction increased and depression and anxiety decreased. In addition, initial levels and rates of change in appearance congruence correlated with the psychosocial outcomes. These results corroborate the positive effects in several earlier studies of smaller samples of adolescents and add to the evidence base that GAH can have a positive effect on mental health.[4]

Yet the study leaves some concerns unanswered. Although overall psychological functioning in the study participants improved, there was substantial variation among participants; a considerable number still had depression, anxiety, or both at 24 months, and two died by suicide. The correlation between appearance congruence and various psychological-outcome variables suggests an important mediating role of GAH and consequent bodily changes. However, other possible determinants of outcomes were not reported, particularly the extent of mental health care provided throughout GAH treatment. To date, international guidelines for transgender adolescent care recommend a psychosocial assessment and involvement of mental health professionals in a multidisciplinary care model.[5] Whether participating centers in the current study followed that approach is unfortunately unclear. Future studies that compare outcomes with different care models are needed, preferably using similar measures.

In addition, some are concerned that young persons may not be capable of making decisions regarding medical treatments that have irreversible effects that they might regret later in life. In the 2-year study by Chen et al., 9 of 314 adolescents (2.9%) stopped GAH, but it is unclear whether they detransitioned or regretted their treatment or whether they stopped because they were satisfied with treatment-related changes. Despite concerns about detransitioning, few studies have provided data on the incidence of detransitioning, and available results are inconsistent. Although one U.S. study showed that 74% of adolescents who started GAH treatment were still receiving it 4 years later, 98% of 720 Dutch adolescents who began such therapy were receiving it after a median of 2.7 years (range, 0.0 to 20.0).[6,7] Similar studies in other centers, regions, and countries are necessary to learn whether the incidence of detransitioning differs between settings and what factors are associated with these differences. It will be especially important to evaluate outcomes in adolescents starting GAH before 16 years of age, the age limit in the initial Dutch protocol.[2]

Furthermore, although Chen et al. investigated relevant psychological and gender outcome measures (e.g., depression, appearance congruence, and life satisfaction), additional factors such as autism spectrum disorder and the quality of peer relations and family support are also of interest. Social support has been hypothesized as explaining why Dutch transgender adolescents have better psychological function than those in other countries.[8] Understanding additional factors that influence outcomes should help to determine which components of care and support other than GAH might improve the lives of transgender adolescents.

Finally, benefits of early medical intervention, including puberty suppression, need to be weighed against possible adverse effects — for example, with regard to bone and brain development and fertility. At present, studies involving young adults from the Dutch adolescent transgender cohort show that accrual of bone mineral decelerates during puberty suppression but increases during GAH treatment and also that adolescents' educational achievements are as expected given their pretreatment status, which is reassuring.[9,10] However, those results from a single Dutch center should be replicated and validated in other contexts, as in a sample followed in the current study.

Despite uncertainties that call for further study, current information shows that mental health improves with GAH, whereas withholding treatment may lead to increased gender dysphoria and adversely affect psychological functioning. The study by Chen et al. adds to the

evidence of the effectiveness of the current care model that includes hormonal treatment for transgender adolescents.

Disclosure forms provided by the authors are available with the full text of this editorial at NEJM.org.

From the Departments of Child and Adolescent Psychiatry (A.L.C.V.) and Pediatrics (S.E.H.), Center of Expertise on Gender Dysphoria, Amsterdam University Medical Centers, Location Vrije Universiteit, Amsterdam.

1. Chen D, Berona J, Chan Y-M, et al. Psychosocial functioning in transgender youth after 2 years of hormones. N Engl J Med 2023;388:240-50.
2. Delemarre-van de Waal HA, Cohen Kettenis PT. Clinical management of gender identity disorder in adolescents: a protocol on psychological and paediatric endocrinology aspects. Eur J Endocrinol 2006;155:Suppl 1:S131-S137.
3. Cass Review. Independent review of gender identity services for children and young people: interim report. February 2022 (https://cass.independent-review.uk/publications/interim-report/).
4. Kuper LE, Stewart S, Preston S, Lau M, Lopez X. Body dissatisfaction and mental health outcomes of youth on gender-affirming hormone therapy. Pediatrics 2020;145(4):e20193006.
5. Coleman E, Radix AE, Bouman WP, et al. Standards of care for the health of transgender and gender diverse people, version 8. Int J Transgend Health 2022;23:Suppl 1:S1-S259.
6. Roberts CM, Klein DA, Adirim TA, Schvey NA, Hisle-Gorman E. Continuation of gender-affirming hormones among transgender adolescents and adults. J Clin Endocrinol Metab 2022;107(9):e3937-e3943.
7. van der Loos MATC, Hannema SE, Klink DT, den Heijer M, Wiepjes CM. Continuation of gender-affirming hormones in transgender people starting puberty suppression in adolescence: a cohort study in the Netherlands. Lancet Child Adolesc Health 2022;6:869-75.
8. de Graaf NM, Steensma TD, Carmichael P, et al. Suicidality in clinic-referred transgender adolescents. Eur Child Adolesc Psychiatry 2022;31:67-83.
9. Schagen SEE, Schagen SEE, Wouters FM, Cohen-Kettenis PT, Gooren LJ, Hannema SE. Bone development in transgender adolescents treated with GnRH analogues and subsequent gender-affirming hormones. J Clin Endocrinol Metab 2020;105(12):e4252-e4263.
10. Arnoldussen M, Hooijman EC, Kreukels BP, de Vries AL. Association between pre-treatment IQ and educational achievement after gender-affirming treatment including puberty suppression in transgender adolescents. Clin Child Psychol Psychiatry 2022;27:1069-76.

DOI: 10.1056/NEJMe2216191