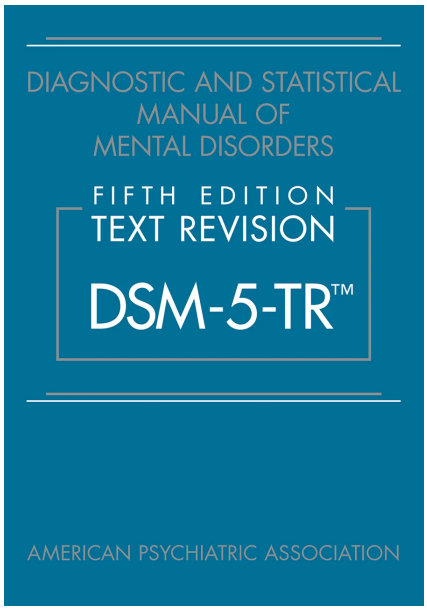
EXHIBIT 21

PSYCHIATRY online

DSM® Library

Sections



Diagnostic And Statistical Manual Of Mental Disorders, Fifth Edition, Text Revision (DSM-5-TR)

Full Access

https://doi.org/10.1176/appi.books.9780890425787

Expand All

Section I: **DSM-5 Basics**

Section II: **Diagnostic Criteria and Codes**

AMERICAN PSYCHIATRIC ASSOCIATION PUBLISHING

American Psychiatric Association Publishing
Powered by Atypon Literatum

Full Access

Gender Dysphoria

https://doi.org/10.1176/appi.books.9780890425787.x14_Gender_Dysophoria

In this chapter, there is one overarching diagnosis of gender dysphoria, with separate developmentally appropriate criteria sets for children and for adolescents and adults. The area of sex and gender is highly controversial and has led to a proliferation of terms whose meanings vary over time and within and between disciplines. An additional source of confusion is that in English “sex” connotes both male/female and sexuality. This chapter employs constructs and terms as they are widely used by clinicians from various disciplines with specialization in treating gender dysphoria (Bouman et al. 2017; Hembree et al. 2017). In this chapter, *sex* and *sexual* refer to the biological indicators of male and female (understood in the context of reproductive capacity), such as in sex chromosomes, gonads, sex hormones, and nonambiguous internal and external genitalia. Disorders of sex development or differences of sex development (DSDs) included the historical terms *hermaphroditism* and *pseudohermaphroditism*. DSDs include somatic intersex conditions such as congenital development of ambiguous genitalia (e.g., clitoromegaly, micropenis), congenital disjunction of internal and external sex anatomy (e.g., complete androgen insensitivity syndrome), incomplete development of sex anatomy (e.g., gonadal agenesis), sex chromosome anomalies (e.g., Turner syndrome; Klinefelter syndrome), or disorders of gonadal development (e.g., ovotestes) (Lee et al. 2016).

Gender is used to denote the public, sociocultural (and usually legally recognized) lived role as boy or girl, man or woman, or other gender. Biological factors are seen as contributing, in interaction with social and psychological factors, to gender development. *Gender assignment* refers to the assignment as male or female. This occurs usually at birth based on phenotypic sex and , thereby, yields the *birth-assigned gender*, historically referred to as "biological sex" or, more recently, "natal gender." *Birth-assigned sex* is often used interchangeably with birth-assigned gender. The terms *assigned sex* and *assigned gender* encompass birth-assigned sex/gender but also include gender/sex assignments and reassignments made after birth but during infancy or early childhood, usually in the case of intersex conditions. *Gender-atypical* refers to somatic features or behaviors that are not typical (in a statistical sense) of individuals with the same assigned gender in a given society and historical era; *gender-nonconforming*, *gender variant*, and *gender diverse* are alternative nondiagnostic terms. *Gender reassignment* denotes an official (and sometimes legal) change of gender. *Gender-affirming treatments* are medical procedures (hormones or surgeries or both) that aim to align an individual's physical characteristics with their *experienced gender*. *Gender identity* is a category of social identity and refers to an individual's identification as male, female, some category in between (i.e., *gender fluid*), or a category other than male or female (i.e., *gender neutral*). There has been a proliferation of gender identities in recent years (Bragg et al. 2018; White et al. 2018). *Gender dysphoria* as a general descriptive term refers to the distress that may accompany the incongruence between one's experienced or expressed gender and one's assigned gender. However, it is more specifically defined when used as a diagnostic category. It does not refer to distress related to stigma, a distinct although possibly co-occurring source of distress. *Transgender* refers to the broad spectrum of individuals whose gender identity is different from their birth-assigned gender. *Cisgender* describes individuals whose gender expression is congruent with their birth-assigned gender (also *nontransgender*). *Transsexual*, a historic term, denotes an individual who seeks, is undergoing, or has undergone a social transition from male to female or female to male, which in many, but not all, cases also involves a somatic transition by gender-affirming hormone treatment and genital, breast, or other gender-affirming surgery (historically referred to as *sex reassignment surgery*).

Although not all individuals will experience distress from incongruence, many are distressed if the desired physical interventions using hormones and/or surgery are not available. The current term is more descriptive than the previous DSM-IV term *gender identity disorder* and focuses on dysphoria as the clinical problem, not identity per se.

Gender Dysphoria

Diagnostic Criteria

Gender Dysphoria in Children (F64.2)

A. A marked incongruence between one's experienced/expressed gender and assigned gender, of at least 6 months' duration, as manifested by at least six of the following (one of which must be Criterion A1):

1. A strong desire to be of the other gender or an insistence that one is the other gender (or some alternative gender different from one's assigned gender).

2. In boys (assigned gender), a strong preference for cross-dressing or simulating female attire; or in girls (assigned gender), a strong preference for wearing only typical masculine clothing and a strong resistance to the wearing of typical feminine clothing.

3. A strong preference for cross-gender roles in make-believe play or fantasy play.

4. A strong preference for the toys, games, or activities stereotypically used or engaged in by the other gender.

5. A strong preference for playmates of the other gender.

6. In boys (assigned gender), a strong rejection of typically masculine toys, games, and activities and a strong avoidance of rough-and-tumble play; or in girls (assigned gender), a strong rejection of typically feminine toys, games, and activities.

7. A strong dislike of one's sexual anatomy.

8. A strong desire for the primary and/or secondary sex characteristics that match one's experienced gender.

B. The condition is associated with clinically significant distress or impairment in social, school, or other important areas of functioning.

Specify if:

With a disorder/difference of sex development (e.g., a congenital adrenogenital disorder such as E25.0 congenital adrenal hyperplasia or E34.50 androgen insensitivity syndrome).

Coding note: Code the disorder/difference of sex development as well as gender dysphoria.

Gender Dysphoria in Adolescents and Adults (F64.0)

A. A marked incongruence between one's experienced/expressed gender and assigned gender, of at least 6 months' duration, as manifested by at least two of the following:

1. 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).

2. 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).

3. A strong desire for the primary and/or secondary sex characteristics of the other gender.

4. A strong desire to be of the other gender (or some alternative gender different from one's assigned gender).

5. A strong desire to be treated as the other gender (or some alternative gender different from one's assigned gender).

6. 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).

B. The condition is associated with clinically significant distress or impairment in social, occupational, or other important areas of functioning.

Specify if:

With a disorder/difference of sex development (e.g., a congenital adrenogenital disorder such as E25.0 congenital adrenal hyperplasia or E34.50 androgen insensitivity syndrome).

Coding note: Code the disorder/difference of sex development as well as gender dysphoria.

Specify if:

Posttransition: The individual has transitioned to full-time living in the experienced gender (with or without legalization of gender change) and has undergone (or is preparing to have) at least one gender-affirming medical procedure or treatment regimen—namely, regular gender-affirming hormone treatment or gender reassignment surgery confirming the experienced gender (e.g., breast augmentation surgery and/or vulvovaginoplasty in an individual assigned male at birth; transmasculine chest surgery and/or phalloplasty or metoidioplasty in an individual assigned female at birth).

Specifiers

The specifier "with a disorder/difference of sex development" should be used in the context of individuals who have a specific and codable disorder/difference of sex development documented in their medical record.

The “posttransition” specifier may be used in the context of continuing treatment procedures that serve to support the new gender assignment.

Diagnostic Features

Individuals with gender dysphoria have a marked incongruence between the gender to which they have been assigned (usually based on phenotypic sex at birth, referred to as *birth-assigned gender*) and their experienced/expressed gender. This discrepancy is the core component of the diagnosis. There must also be evidence of distress about this incongruence. Experienced gender may include alternative gender identities beyond binary stereotypes. Consequently, distress may involve not only the experience that the individual is a male or female gender other than the one assigned at birth but also an experience that the individual is an intermediate or alternative gender that differs from the individual’s birth-assigned gender.

Gender dysphoria manifests itself differently in different age groups. The following examples may be less prominent in children raised in surroundings with fewer gender stereotypes.

Prepubertal individuals assigned female at birth with gender dysphoria may express a marked, persistent feeling or conviction that they are a boy, express aversion to the idea of being a girl, or assert they will grow up to be a man. They often prefer boys’ clothing and hairstyles, may be perceived by strangers as boys, and may ask to be called by a boy’s name. Sometimes they display intense negative reactions to parental attempts to have them wear dresses or other feminine attire. Some may refuse to attend school or social events where such clothes are required. These children may demonstrate marked gender nonconformity in role-playing, dreams, gender-typed play and toy preferences, styles, mannerisms, fantasies, and peer preferences. Contact sports, rough-and-tumble play, traditional boyhood games, and boys as playmates are most often preferred. They show little interest in stereotypically feminine toys (e.g., dolls) or activities (e.g., feminine dress-up or role-play). Occasionally, they refuse to urinate in a sitting position. Some may express a desire to have a penis or claim to have a penis or that they will grow one when older. They may also state that they do not want to develop breasts or menstruate.

Prepubertal individuals assigned male at birth with gender dysphoria may express a

marked, persistent feeling or conviction that they are a girl or assert that they will grow up to be a woman. They may express aversion to the idea of being a boy. They often prefer dressing in girls' or women's clothes or may improvise clothing from available materials (e.g., using towels, aprons, and scarves for long hair or skirts). These children may demonstrate marked gender nonconformity in gender-typed play and toy preferences, styles, mannerisms, and peer preferences. They may role-play female figures (e.g., playing "mother") and may be intensely interested in female fantasy figures. Traditional feminine activities, stereotypical games, and pastimes (e.g., "playing house"; drawing feminine pictures; watching television or videos of favorite female characters) may be preferred. Stereotypical female-type dolls (e.g., Barbie) may be favorite toys, and girls are their preferred playmates. They avoid rough-and-tumble play and have little interest in stereotypically masculine toys (e.g., cars, trucks). They may state that they find their penis or testes disgusting, that they wish them removed, or that they have, or wish to have, a vagina.

Increasingly, parents are presenting to specialized clinics after their child with gender dysphoria has already socially transitioned (Aitken et al. 2015; Ristori and Steensma 2016; Steensma and Cohen-Kettenis 2011).

As the onset of puberty for individuals assigned female at birth is somewhere between ages 9 and 13, and between 11 and 14 for individuals assigned male at birth, their symptoms and concerns may arise in a developmental phase somewhere between childhood and adolescence. As secondary sex characteristics of younger adolescents are not yet fully developed, these individuals may not state dislike of them, but they may be markedly distressed by imminent physical changes.

In adolescents and adults with gender dysphoria, the discrepancy between experienced gender and physical sex characteristics is often, but not always, accompanied by a desire to be rid of primary and/or secondary sex characteristics and/or a strong desire to acquire some primary and/or secondary sex characteristics of another gender. To varying degrees, older adolescents and adults with gender dysphoria may adopt the behavior, clothing, and mannerisms of their experienced gender. They feel uncomfortable being regarded by others, or functioning in society, as members of their assigned gender. Some adults and adolescents may have a strong desire to be of a

different gender and treated as such, and they may have an inner certainty to feel and respond as their experienced gender without seeking medical treatment to alter body characteristics. They may find other ways to resolve the incongruence between experienced/expressed and assigned gender by partially living in the desired role or by adopting a gender role neither conventionally male nor conventionally female (Aitken et al. 2015; Costa et al. 2015; de Graaf et al. 2018b; de Graaf et al. 2018c).

Associated Features

When visible signs of puberty develop, individuals assigned male at birth may shave their facial, body, and leg hair at the first signs of growth. They sometimes bind their genitals to make erections less visible. Individuals assigned female at birth may bind their breasts, walk with a stoop, or use loose sweaters to make breasts less visible. Increasingly, adolescents request, or may obtain without medical prescription and supervision, drugs that suppress production of gonadal steroids (e.g., gonadotropin-releasing hormone [GnRH] agonists) or that block gonadal hormone actions (e.g., spironolactone). Clinically referred adolescents often want hormone treatment and many also wish for gender-affirming surgery. Adolescents living in an accepting environment may openly express the desire to be and be treated as their experienced gender and dress partly or completely as their experienced gender, have a hairstyle typical of their experienced gender, preferentially seek friendships with peers of another gender, and/or adopt a new first name consistent with their experienced gender. Older adolescents, when sexually active, often do not show or allow partners to touch their sexual organs. For adults with an aversion toward their genitals, sexual activity is constrained by the preference that their genitals not be seen or touched by their partners. Not infrequently, adults may seek hormone treatment (sometimes without medical prescription and supervision) and gender-affirming surgery. Others are satisfied with either hormone treatment or surgery alone, or without any gender-affirming medical treatment.

In children, adolescents, and adults with gender dysphoria, an overrepresentation of autism spectrum traits has been observed (Jones et al. 2012; van der Miesen et al. 2016). Also, individuals with autism spectrum disorder are more likely to exhibit gender diversity (Glidden et al. 2016; Thrower et al. 2020; van der Miesen et al. 2018).

Adolescents and adults with gender dysphoria before gender-affirming treatment and legal gender change are at increased risk for mental health problems including suicidal ideation, suicide attempts, and suicides. After gender reassignment, adjustment may vary, and suicide risk and mental health problems may persist.

In prepubertal children, increasing age is associated with having more behavioral or emotional problems; this is related to the increasing nonacceptance of gender-nonconforming behavior by others (Steensma et al. 2014). Children and adolescents who feel supported and accepted in their gender nonconformity may show less or even no psychological problems (de Vries et al. 2014; Olson et al. 2016).

Prevalence

There are no large-scale population studies of gender dysphoria. Based on gender-affirming treatment–seeking populations, the prevalence for gender dysphoria diagnosis across populations has been assessed to be less than 1/1,000 (i.e., < 0.1%) for both individuals assigned male at birth and individuals assigned female at birth (Goodman et al. 2019). Because many adults with gender dysphoria do not seek care at specialty treatment programs, prevalence rates are likely underestimates. Prevalence estimates based on surveys of self-reporting general population samples in the United States and Europe suggest higher numbers, although varied methods of assessment make comparisons difficult across studies (Zhang et al. 2020). Self-identification as transgender ranges from 0.5% to 0.6% (Conron et al. 2012; Crissman et al. 2017; Flores et al. 2016); experiencing oneself as having an incongruent gender identity ranges from 0.6% to 1.1% (Kuyper and Wijsen 2014; Van Caenegem et al. 2015); feeling that one is a person of a different sex ranges from 2.1% to 2.6% (Åhs et al. 2018); and the desire to undergo medical treatment ranges from 0.2% to 0.6% (Åhs et al. 2018; Kuyper and Wijsen 2014).

Development and Course

Because expression of gender dysphoria varies with age, there are separate criteria sets for children versus those for adolescents and adults (Steensma et al. 2011). Criteria for children are defined in a more concrete, behavioral manner than those for adolescents and adults. Young children are less likely than older children, adolescents, and adults to

express extreme and persistent anatomic dysphoria (Ristori and Steensma 2016). In adolescents and adults, incongruence between experienced gender and assigned gender is a central feature of the diagnosis (Leibowitz and de Vries 2016). Factors related to distress and impairment also vary with age. A very young child may show signs of distress (e.g., intense crying) only when parents tell the child that he or she is "really" not a member of another gender but only "desires" to be. Distress may not be manifest in social environments supportive of the child's gender nonconformity and may emerge only if there is parental/social interference with the child's gender variance (Cohen-Kettenis et al. 2003). In adolescents and adults, distress may manifest because of strong incongruence between experienced gender and birth-assigned gender (Leibowitz and de Vries 2016). Such distress may, however, be mitigated by supportive environments and knowledge that biomedical treatments exist to reduce incongruence (de Vries et al. 2011; Olson et al. 2016). Impairment (e.g., school refusal, development of depression, anxiety, peer and behavioral problems, and substance abuse) may be a correlate of gender dysphoria.

Gender dysphoria without a disorder of sex development

For clinic-referred children studied in Canada and the Netherlands (Cohen-Kettenis et al. 2003), onset of gender-nonconforming behaviors is usually between ages 2 and 4 years. This corresponds to the developmental time period in which most children begin expressing gendered behaviors and interests. For some preschool-age children, both marked, persistent gender-atypical behaviors and the expressed desire to be another gender may be present, or labeling themselves as a member of another gender may occur. In other cases, the gender expression appears later, usually at entry into elementary school. Children may sometimes express discomfort with their sexual anatomy or will state the desire to have a sexual anatomy corresponding to their experienced gender ("anatomic dysphoria"). Expressions of anatomic dysphoria become more common as children with gender dysphoria approach and anticipate puberty.

No general population studies exist of adolescent or adult outcomes of childhood gender variance. Some prepubescent children expressing a desire to be another gender will not seek gender-affirming somatic treatments when they reach puberty. They frequently report nonheterosexual orientations and frequently marked gender-nonconforming

behavior, although not necessarily a transgender identity in adolescence/young adulthood (Singh et al. 2021; Steensma and Cohen-Kettenis 2018; Temple Newhook et al. 2018; Zucker 2018). Some children with gender dysphoria in childhood that remits in adolescence may experience a recurrence in adulthood (Steensma and Cohen-Kettenis 2015).

In individuals assigned male at birth, studies from North America and the Netherlands found persistence ranged from 2% to 39% (Ristori and Steensma 2016). In individuals assigned female at birth, persistence ranged from 12% to 50% (Ristori and Steensma 2016). Persistence of gender dysphoria is modestly correlated with dimensional measures of severity ascertained at the time of a childhood baseline assessment. Early social transition may also be a factor in persistence of gender dysphoria in adolescence (Ristori and Steensma 2016; Steensma et al. 2013).

Studies have shown a high incidence of sexual attraction to those of the individual's birth-assigned gender, regardless of the trajectory of the prepubescent child's gender dysphoria. For individuals whose gender dysphoria continues into adolescence and beyond, most self-identify as heterosexual. In those who no longer have gender dysphoria by the time of adolescence, a majority self-identify as gay, lesbian, or bisexual (Ristori and Steensma 2016; Singh et al. 2021; Steensma et al. 2013; Wallien and Cohen-Kettenis 2008).

Two broad trajectories have been described for development of gender dysphoria in individuals who identify as either male or female.

As opposed to gender-nonconforming children, individuals with prepubertal-onset gender dysphoria have symptoms that meet diagnostic criteria for gender dysphoria in childhood. The dysphoria can continue into adolescence and adulthood; alternatively, some individuals go through a period in which the gender dysphoria either desists or is denied. At such times, these individuals may self-identify as being gay or lesbian. Some may identify as heterosexual and cisgender. However, it is possible that some of these individuals may experience a recurrence of gender dysphoria later in life (Nieder et al. 2011; Steensma and Cohen-Kettenis 2015).

Regardless of whether the individual's gender dysphoria persists or desists at a later

date, either the onset of puberty or the realization that puberty will begin with development of secondary sex characteristics can prompt distressing feelings of gender incongruence that can exacerbate the individual's gender dysphoria.

The early/prepubertal-onset group often present for clinical, gender-affirming care during childhood, during adolescence, or in young adulthood (Johansson et al. 2010; Nieder et al. 2011). This may reflect a more intense gender dysphoria compared with individuals with late/postpubertal-onset gender dysphoria, whose distress may be more variable and less intense (Schneider et al. 2016).

Late-onset or pubertal/postpubertal-onset gender dysphoria occurs around puberty or even much later in life. Some of these individuals report having had a desire to be of another gender in childhood that was not expressed verbally to others or had gender-nonconforming behavior that did not meet full criteria for gender dysphoria in childhood. Others have no recollection of any signs of childhood gender dysphoria. Parents of individuals with gender dysphoria of pubertal/postpubertal-onset often report surprise, as they saw no signs of gender dysphoria during childhood (Littman 2019).

Gender dysphoria in association with a disorder of sex development

Individuals with DSDs who require early medical intervention or decisions about gender assignment come to clinical attention at an early age. Depending on the condition, they may have been gonadectomized (often because of risk of future malignancy) before puberty so that administration of exogenous hormones is part of routine care to induce puberty. Infertility is common whether due to the condition itself or to gonadectomy, and genital surgery may have been done in infancy or childhood with the intent of affirming the assigned gender to both the affected individual and caregivers.

Affected individuals may exhibit gender-nonconforming behavior starting in early childhood in a manner that is predictable depending on the specific DSD syndrome and the gender assignment, and thresholds for supporting social and medical gender transition in minors have traditionally been much lower for those with compared to those without DSDs. As individuals with some DSD syndromes become aware of their condition and medical history, many experience uncertainty about their gender, as

opposed to developing a firm conviction that they are of another gender. The proportion who develop gender dysphoria and progress to gender transition varies markedly depending on the particular syndrome and gender assignment (Byne et al. 2012; de Vries et al. 2007; Kreukels et al. 2018; Lee et al. 2016).

Risk and Prognostic Factors

Temperamental

Gender-variant behavior among individuals with prepubertal-onset gender dysphoria can develop in early preschool age. Studies suggest that a greater intensity of gender nonconformity and an older age at presentation make persistence of gender dysphoria into adolescence and adulthood more likely (Steensma et al. 2011; Steensma et al. 2013). A predisposing factor under consideration, especially in individuals with postpubertal-onset gender dysphoria (adolescence, adulthood), includes history of transvestism that may develop into autogynephilia (i.e., sexual arousal associated with the thought or image of oneself as a woman).

Environmental

Individuals assigned male at birth with gender dysphoria without a DSD (in both childhood and adolescence) more commonly have older brothers when compared with cisgender males.

Genetic and physiological

For individuals with gender dysphoria without a DSD, some genetic contribution is suggested by evidence for (weak) familiality of gender dysphoria among nontwin siblings, increased concordance for gender dysphoria in monozygotic compared with dizygotic same-sex twins, and some degree of heritability of gender dysphoria. Research suggests that gender dysphoria has a polygenetic basis involving interactions of several genes and polymorphisms that may affect in utero sexual differentiation of the brain, contributing to gender dysphoria in individuals assigned male at birth (Foreman et al. 2020).

As to endocrine findings in individuals with gender dysphoria, no endogenous systemic

abnormalities in sex-hormone levels have been found in 46,XY individuals, whereas there appear to be increased androgen levels (in the range found in hirsute women but far below normal male levels) in 46,XX individuals. Overall, current evidence is insufficient to label gender dysphoria without a DSD as a form of intersexuality limited to the central nervous system.

In gender dysphoria associated with a DSD, the likelihood of later gender dysphoria is increased if prenatal production and utilization (via receptor sensitivity) of androgens are grossly variant relative to what is usually seen in individuals with the same assigned gender. Examples include 46,XY individuals with a history of normal male prenatal hormone milieu but inborn nonhormonal genital defects (as in cloacal bladder exstrophy or penile agenesis) and who have been assigned to the female gender. The likelihood of gender dysphoria is further enhanced by additional, prolonged, highly gender-variant postnatal androgen exposure with somatic virilization as may occur in female-raised and noncastrated 46,XY individuals with 5-alpha reductase-2 deficiency or 17-beta-hydroxysteroid dehydrogenase-3 deficiency or in female-raised 46,XX individuals with classical congenital adrenal hyperplasia with prolonged periods of nonadherence to glucocorticoid replacement therapy. However, the prenatal androgen milieu is more closely related to gendered behavior than to gender identity. Many individuals with DSDs and markedly gender-variant behavior do not develop gender dysphoria. Thus, gender-nonconforming behavior by itself should not be interpreted as an indicator of current or future gender dysphoria. There appears to be a higher rate of gender dysphoria and patient-initiated gender change from assigned female to male than from assigned male to female in individuals prenatally exposed to a full complement of masculinizing hormonal influences.

Culture-Related Diagnostic Issues

Individuals with gender dysphoria have been reported across many countries and cultural contexts around the world. The equivalent of gender dysphoria has also been reported in individuals living in cultural contexts with institutionalized gender identity categories other than men/boys or women/girls that sanction gender nonconforming development. These include India, Sri Lanka, Myanmar, Oman, Samoa, Thailand, and Indigenous Peoples of North America. It is unclear however, in such cultural contexts,

whether the diagnostic criteria for gender dysphoria would be met with these individuals.

The prevalence of coexisting mental health problems differs among cultures; these differences may also be related to differences in attitudes toward gender nonconformity in children, adolescents, and adults. However, also in some non-Western cultures, anxiety has been found to be relatively common in individuals with gender dysphoria, even in cultures with accepting attitudes toward gender-variant behavior (Campbell et al. 2018; Khoury et al. 2020; Lobato et al. 2019).

Sex- and Gender-Related Diagnostic Issues

Sex differences in rate of referrals to specialty clinics vary by age group. In children, sex ratios of individuals assigned male at birth to individuals assigned female at birth range from 1.25:1 to 4.3:1. Studies show increasing numbers of children and adolescents presenting to specialty clinics, presentation at younger ages, more frequent early social transition, and a shift to a greater number of individuals assigned female at birth in adolescents and young adults than individuals assigned male at birth (Aitken et al. 2015; de Graaf et al. 2018a; de Graaf et al. 2018c; Steensma 2013; Steensma et al. 2018; Wood et al. 2013). In adults, estimates generally suggest more individuals assigned male at birth seek gender-affirming treatment, with ratios ranging from 1:1 to 6.1:1 in most studies in the United States and Europe (Arcelus et al. 2015).

Association With Suicidal Thoughts or Behavior

Rates of suicidality and suicide attempts for transgender individuals are reported to range from 30% to 80%, with risk factors including past maltreatment, gender victimization, depression, substance abuse, and younger age (Mueller et al. 2017). Transgender adolescents referred to gender clinics have substantially higher rates of suicidal thoughts and behaviors when compared with nonreferred adolescents (de Graaf et al. 2020). Prior to receiving gender-affirming treatment and legal gender reassignment, adolescents and adults with gender dysphoria are at increased risk for suicidal thoughts and suicide attempts (Dhejne et al. 2016; Marshall et al. 2016). After gender-affirming treatment, adjustment varies, and while improvement in coexisting symptoms is often seen (Bränström and Pachankis 2020), some individuals continue to

experience prominent anxiety and affective symptoms and remain at increased risk for suicide (Dhejne et al. 2011; Dhejne et al. 2016).

A study of 572 children referred for gender identity concerns in Canada and several comparison groups (siblings, other referred children, and nonreferred children) largely from other high-income countries found that gender-referred children were 8.6 times more likely to self-harm or attempt suicide than comparison children, even after adjustment for overall behavior and peer relationship problems, and particularly in the second half of childhood (Aitken et al. 2016). Among adolescents, the highest rate of suicide attempt is among transgender young men, followed by those defining themselves as neither male nor female (Jackman et al. 2019; Toomey et al. 2018).

Functional Consequences of Gender Dysphoria

Gender nonconformity may appear at all ages after the first 2–3 years of childhood and may interfere with daily activities. In older children, gender nonconformity may affect peer relationships and may lead to isolation from peer groups and to distress. Many children experience teasing and harassment or pressure to dress in attire associated with their birth-assigned sex, especially when growing up in a nonsupportive and nonaccepting environment. Also in adolescents and adults, the distress resulting from gender incongruence often interferes with daily activities. Relationship difficulties, including sexual relationship problems, are common, and functioning at school or at work may be impaired. Gender dysphoria is associated with high levels of stigmatization, discrimination, and victimization, leading to negative self-concept, increased rates of depression, suicidality, and other mental disorder co-occurrence, school dropout, and economic marginalization, including unemployment, with attendant social and mental health risks, especially in individuals who lack family or social support. In addition, these individuals' access to health services and mental health services may be impeded by structural barriers, such as institutional discomfort about, inexperience with, or hostility toward working with this patient population (Bockting et al. 2013).

Differential Diagnosis

Nonconformity to gender roles

Gender dysphoria should be distinguished from simple nonconformity to stereotypical gender role behavior by the strong desire to be of another gender than the assigned one and by the extent and pervasiveness of gender-variant activities and interests. The diagnosis is not meant to merely describe nonconformity to stereotypical gender role behavior (e.g., "tomboyism" in girls, "girly-boy" behavior in boys, occasional cross-dressing in adult men). Given the increased openness of gender-diverse expressions by individuals across the entire range of the transgender spectrum, it is important that the clinical diagnosis be limited to those individuals whose distress and impairment meet the specified criteria.

Transvestic disorder

Transvestic disorder is diagnosed in heterosexual (or bisexual) adolescent and adult males (rarely in females) for whom women's clothing generates sexual excitement and causes distress and/or impairment without drawing their assigned gender into question. It is occasionally accompanied by gender dysphoria. An individual with transvestic disorder who also has clinically significant gender dysphoria can be given both diagnoses. In some cases of postpubertal-onset gender dysphoria in individuals assigned male at birth who are attracted to women, cross-dressing with sexual excitement is a precursor to the diagnosis of gender dysphoria (Blanchard 1985).

Body dysmorphic disorder

An individual with body dysmorphic disorder focuses on the alteration or removal of a specific body part because it is perceived as abnormally formed, not because it represents a repudiated assigned gender. When an individual's presentation meets criteria for both gender dysphoria and body dysmorphic disorder, both diagnoses can be given. Individuals wishing to have a healthy limb amputated (termed by some *body integrity identity disorder*) because it makes them feel more "complete" usually do not wish to change gender, but rather desire to live as an amputee or a disabled person.

Autism spectrum disorder

In individuals with autism spectrum disorder, diagnosing gender dysphoria can be challenging. It can be difficult to differentiate potential co-occurring gender dysphoria

from an autistic preoccupation because of the concrete and rigid thinking around gender roles and/or poor understanding of social relationships characteristic of autism spectrum disorder (de Vries et al. 2010).

Schizophrenia and other psychotic disorders

In schizophrenia, there may rarely be delusions of belonging to some other gender. In the absence of psychotic symptoms, insistence by an individual with gender dysphoria that he or she is another gender is not considered a delusion. Schizophrenia (or other psychotic disorders) and gender dysphoria may co-occur. Gender-themed delusions may occur in up to 20% of individuals with schizophrenia. They can usually be differentiated from gender dysphoria by their bizarre content and by waxing and waning with remissions and exacerbations of psychotic episodes (Campo et al. 2001; Meijer et al. 2017).

Other clinical presentations

Some individuals with an emasculinization desire who develop an alternative, nonmale/nonfemale gender identity do have a presentation that meets criteria for gender dysphoria. However, some males seek genital surgery for either aesthetic reasons or to remove psychological effects of androgens without changing male identity; in these cases, the criteria for gender dysphoria are not met.

Comorbidity

Clinically referred children with gender dysphoria show elevated levels of anxiety, disruptive, impulse-control, and depressive disorders. Autism spectrum disorder is more prevalent in clinically referred adolescents and adults with gender dysphoria than in the general population (Heylens et al. 2018; van der Miesen et al. 2016; van der Miesen et al. 2018). Clinically referred adolescents and adults with gender dysphoria often have high rates of associated mental disorders, with anxiety and depressive disorders being the most common. Individuals who have experienced harassment and violence may also develop posttraumatic stress disorder.

References: Gender Dysphoria

Bouman WP, Schwend AS, Motmans J, et al: Language and trans health (editorial). International Journal of Transgenderism 18(1):1–6, 2017

Bragg S, Renold E, Ringrose J, Jackson C: 'More than boy, girl, male, female': exploring young people's views on gender diversity within and beyond school contexts. Sex Education: Sexuality, Society and Learning 18(4):420–434, 2018

Hembree WC, Cohen-Kettenis PT, Gooren L, et al: Endocrine treatment of gender-dysphoric/gender-incongruent persons: an Endocrine Society Clinical Practice Guideline. J Clin Endocrinol Metab 102(11):3869–3903, 2017

Lee PA, Nordenström A, Houk CP, et al; Global DSD Update Consortium: Global disorders of sex development update since 2006: perceptions, approach and care. Horm Res Paediatr 85(3):158–180, 2016

White AE, Moeller J, Ivcevic Z, Brackett MA: Gender identity and sexual identity labels used by U.S. high school students: a co-occurrence network analysis. Psychology of Sexual Orientation and Gender Diversity 5(2):243–252, 2018

Åhs JW, Dhejne C, Magnusson C, et al: Proportion of adults in the general population of Stockholm County who want gender-affirming medical treatment. PLoS One 13(10):e0204606, 2018

Aitken M, Steensma TD, Blanchard R, et al: Evidence for an altered sex ratio in clinic-referred adolescents with gender dysphoria. J Sex Med 12(3):756–763, 2015

Aitken M, VanderLaan DP, Wasserman L, et al: Self-harm and suicidality in children referred for gender dysphoria. J Am Acad Child Adolesc Psychiatry 55(6):513–520, 2016

Arcelus J, Bouman WP, Van Den Noortgate W, et al: Systematic review and meta-analysis of prevalence studies in transsexualism. Eur Psychiatry 30:807–815, 2015

Blanchard R: Typology of male-to-female transsexualism. Arch Sex Behav 14:247–261, 1985

Bockting WO, Miner MH, Swinburne Romine RE, et al: Stigma, mental health, and resilience in an online sample of the US transgender population. Am J Public Health 103(5):943–951, 2013

Bränström R, Pachankis JE: Reduction in mental health treatment utilization among transgender individuals after gender-affirming surgeries: a total population study. Am J Psychiatry 177(8):727–734, 2020

Byne W, Bradley SJ, Coleman E, et al; American Psychiatric Association Task Force on Treatment of Gender Identity Disorder: Report of the American Psychiatric Association Task Force on Treatment of Gender Identity Disorder. Arch Sex Behav 41(4):759–796, 2012

Campbell MM, Fresán A, Addinall RM, et al: Experiences of gender incongruence and the relationship between social exclusion, psychological distress, and dysfunction among South African transgender adults: a field-study for ICD-11. Ann Clin Psychiatry 30(3):168–174, 2018

Campo JM, Nijman H, Evers C, et al: Gender identity disorders as a symptom of psychosis, schizophrenia in particular (in Dutch). Ned Tijdschr Geneeskd 145(39):1876–1880, 2001

Cohen-Kettenis PT, Owen A, Kaijser VG, et al: Demographic characteristics, social competence, and behavior problems in children with gender identity disorder: A cross-national, cross-clinic comparative analysis. J Abnorm Child Psychol 31(1):41–53, 2003

Conron KJ, Scott G, Stowell GS, Landers SJ: Transgender health in Massachusetts: results from a household probability sample of adults. Am J Public Health 102(1):118–122, 2012

Costa R, Dunsford M, Skagerberg E, et al: Psychological support, puberty suppression, and psychosocial functioning in adolescents with gender dysphoria. J Sex Med 12(11):2206–2214, 2015

Crissman HP, Berger MB, Graham LF, Dalton VK: Transgender demographics: a

household probability sample of US adults, 2014. Am J Public Health 107(2):213–215, 2017

de Graaf NM, Carmichael P, Steensma TD, Zucker KJ: Evidence for a change in the sex ratio of children referred for gender dysphoria: data from the Gender Identity Development Service in London (2000–2017). J Sex Med 15(10):1381–1383, 2018a

de Graaf NM, Cohen-Kettenis PT, Carmichael P, et al: Psychological functioning in adolescents referred to specialist gender identity clinics across Europe: a clinical comparison study between four clinics. Eur Child Adolesc Psychiatry 27(7):909–919, 2018b

de Graaf NM, Giovanardi G, Zitz C, Carmichael P: Sex ratio in children and adolescents referred to the Gender Identity Development Service in the UK (2009-2016). Arch Sex Behavior 47(5):1301–1304, 2018c

de Graaf NM, Steensma TD, Carmichael P, et al: Suicidality in clinic-referred transgender adolescents. Eur Child Adolesc Psychiatry Nov 9, 2020

de Vries ALC, Doreleijers TAH, Cohen-Kettenis PT: Disorders of sex development and gender identity outcome in adolescence and adulthood: understanding gender identity development and its clinical implications. Pediatr Endocrinol Rev 4(4):343–351, 2007

de Vries ALC, Noens ILJ, Cohen-Kettenis PT, et al: Autism spectrum disorders in gender dysphoric children and adolescents. J Autism Dev Disord 40(8):930–936, 2010

de Vries ALC, Steensma TD, Doreleijers TAH, Cohen-Kettenis PT: Puberty suppression in adolescents with gender identity disorder: a prospective follow-up study. J Sex Med 8(8):2276–2283, 2011

de Vries ALC, McGuire JK, Steensma TD, et al: Young adult psychological outcome after puberty suppression and gender reassignment. Pediatrics 134(4):696–704, 2014

Dhejne C, Lichtenstein P, Boman M, et al: Long-term follow-up of transsexual persons undergoing sex reassignment surgery: cohort study in Sweden. PLoS One 6(2):e16885,

2011

Dhejne C, Van Vlerken R, Heylens G, Arcelus J: Mental health and gender dysphoria: a review of the literature. Int Rev Psychiatry 28(1):44–57, 2016

Flores AR, Herman JL, Gates GJ, Brown TNT: How Many Adults Identify as Transgender in the United States? Los Angeles, CA, The Williams Institute, June 2016. Available at: http://williamsinstitute.law.ucla.edu/wp-content/uploads/How-Many-Adults-Identify-as-Transgender-in-the-United-States.pdf. Accessed July 12, 2019.

Foreman M, Hare L, York K, et al: Genetic link between gender dysphoria and sex hormone signaling. J Clin Endocrinol Metab 104(2):390–396, 2019 Erratum: J Clin Endocrinol Metab 105(1):dgz156, 2020

Glidden D, Bouman WP, Jones BA, Arcelus J, et al: Gender dysphoria and autism spectrum disorder: a systematic review of the literature. Sex Med Rev 4(1):3–14, 2016

Goodman M, Adams N, Cornell T, et al: Size and distribution of transgender and gender nonconforming populations: a narrative review. Endocrinol Metab Clin North Am 48(2):303–321, 2019

Heylens G, Aspeslagh L, Dierickx J, et al: The co-occurrence of gender dysphoria and autism spectrum disorder in adults: an analysis of cross-sectional and clinical chart data. J Autism Dev Disord 48(6):2217–2223, 2018

Jackman KB, Caceres BA, Kreuze EJ, Bockting WO: Suicidality among gender minority youth: analysis of 2017 Youth Risk Behavior Survey data. Arch Suicide Res Oct 23;1–16, 2019

Johansson A, Sundbom E, Höjerback T, Bodlund O: A five-year follow-up study of Swedish adults with gender identity disorder. Arch Sex Behavior 39(6):1429–1437, 2010

Jones RM, Wheelwright S, Farrell K, et al: Brief report: female-to-male transsexual people and autistic traits. J Autism Dev Disord 42(2):301–306, 2012

Khoury B, El Khoury J, Fresán Orellana A, et al: The ICD-11 classification of gender incongruence of adolescence and adulthood: adequacy among transgender people in Lebanon. Cult Health Sex Feb 7;1–12, 2020

Kreukels BPC, Köhler B, Nordenström A, et al; dsd-LIFE Group: Gender dysphoria and gender change in disorders of sex development/intersex conditions: results from the dsd-LIFE study. J Sex Med 15(5):777–785, 2018

Kuyper L, Wijsen C: Gender identities and gender dysphoria in the Netherlands. Arch Sex Behav 45(2):377–385, 2014

Lee PA, Nordenström A, Houk CP, et al; Global DSD Update Consortium: Global disorders of sex development update since 2006: perceptions, approach and care. Horm Res Paediatr 85(3):158–180, 2016

Leibowitz S, de Vries ALC: Gender dysphoria in adolescence. Int Rev Psychiatry 28(1):21–35, 2016

Littman L: Correction: Parent reports of adolescents and young adults perceived to show signs of a rapid onset of gender dysphoria. PLoS One 14(3):e0214157, 2019

Lobato MI, Soll BM, Costa AB, et al: Psychological distress among transgender people in Brazil: frequency, intensity and social causation—an ICD-11 field study. Braz J Psychiatry 41(4):310–315, 2019

Marshall E, Claes L, Bouman WP, et al: Non-suicidal self-injury and suicidality in trans people: a systematic review of the literature. Int Rev Psychiatry 28(1):58–69, 2016

Meijer JH, Eeckhout GM, van Vlerken RH, de Vries AL: Gender dysphoria and co-existing psychosis: review and four case examples of successful gender affirmative treatment. LGBT Health 4(2):106–114, 2017

Mueller SC, De Cuypere G, T'Sjoen G: Transgender research in the 21st century: a selective critical review from a neurocognitive perspective. Am J Psychiatry 174(12):1155–1162, 2017

Nieder TO, Herff M, Cerwenka S, et al: Age of onset and sexual orientation in transsexual males and females. J Sex Med 8(3):783–791, 2011

Olson KR, Durwood L, DeMeules M, McLaughlin KA: Mental health of transgender children who are supported in their identities. Pediatrics 137(3):e20153223, 2016 (erratum: Pediatrics 142(2):e20181436, 2018 30065001)

Ristori J, Steensma TD: Gender dysphoria in childhood. Int Rev Psychiatry 28(1):13–20, 2016

Schneider C, Cerwenka S, Nieder TO, 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. Arch Sex Behav 45(3):551–558, 2016

Singh D, Bradley SJ, Zucker KJ: A follow-up study on boys with gender identity disorder. Front Psychiatry 12:632784, 2021

Steensma TD: From Gender Variance to Gender Dysphoria: Psychosexual Development of Gender Atypical Children and Adolescents. Unpublished dssertation, VU University, Amsterdam, 2013

Steensma TD, Cohen-Kettenis PT: Gender transitioning before puberty? Arch Sex Behav 40(4):649–650, 2011

Steensma TD, Cohen-Kettenis PT: More than two developmental pathways in children with gender dysphoria? J Am Acad Child Adolescent Psychiatry 54(2):147–148, 2015

Steensma TD, Cohen-Kettenis PT: A critical commentary on follow-up studies and “desistence” theories about transgender and gender non-conforming children. International Journal of Transgenderism 19(1):225–230, 2018

Steensma TD, Biemond R, de Boer F, Cohen-Kettenis PT: Desisting and persisting gender dysphoria after childhood: a qualitative follow-up study. Clin Child Psychol Psychiatry 16(4):499–516, 2011

Steensma TD, McGuire JK, Kreukels BPC, et al: Factors associated with desistence and persistence of childhood gender dysphoria: a quantitative follow-up study. J Am Acad Child Adolesc Psychiatry 52(6):582–590, 2013

Steensma TD, Zucker KJ, Kreukels BPC, et al: Behavioral and emotional problems on the Teacher's Report Form: a cross-national, cross-clinic comparative analysis of gender dysphoric children and adolescents. J Abnorm Child Psychol 42(4):635–647, 2014

Steensma TD, Cohen-Kettenis PT, Zucker KJ: Evidence for a change in the sex ratio of children referred for gender dysphoria: data from the Center of Expertise on Gender Dysphoria in Amsterdam (1988–2016) (letter). J Sex Marital Ther 44(7):713–715, 2018

Temple Newhook J, Pyne J, Winters K, et al: A critical commentary on follow-up studies and "desistance" theories about transgender and gender-nonconforming children. International Journal of Transgenderism 19(1):212–224, 2018

Thrower E, Bretherton I, Pang KC, et al: Prevalence of autism spectrum disorder and attention-deficit hyperactivity disorder amongst individuals with gender dysphoria: a systematic review. J Autism Dev Disord 50(3):695–706, 2020

Toomey RB, Syvertsen AK, Shramko M: Transgender adolescent suicide behavior. Pediatrics 142(4):e20174218, 2018

Van Caenegem E, Wierckx K, Elaut E, et al: Prevalence of gender nonconformity in Flanders, Belgium. Arch Sex Behav 44(5):1281–1287, 2015

van der Miesen AI, Hurley H, De Vries AL: Gender dysphoria and autism spectrum disorder: a narrative review. Int Rev Psychiatry 28(1):70–80, 2016

van der Miesen AIR, Hurley H, Bal AM, de Vries ALC: Prevalence of the wish to be of the opposite gender in adolescents and adults with autism spectrum disorder. Arch Sex Behav 47(8):2307–2317, 2018

Wallien MSC, Cohen-Kettenis PT: Psychosexual outcome of gender-dysphoric children.

J Am Acad Child Adolesc Psychiatry 47(12):1413–1423, 2008

Wood H, Sasaki S, Bradley SJ, et al: Patterns of referral to a gender identity service for children and adolescents (1976–2011): age, sex ratio, and sexual orientation. J Sex Marital Ther 39(1):1–6, 2013

Zhang Q, Goodman M, Adams N, et al: Epidemiological considerations in transgender health: a systematic review with focus on higher quality data. International Journal of Transgender Health 21(2):125–137, 2020

Zucker KJ: The myth of persistence: response to "A Critical Commentary on Follow-up Studies and 'Desistance' Theories About Transgender and Gender Non-conforming Children" by Temple Newhook et al. International Journal of Transgenderism 19(1):231–245, 2018

Other Specified Gender Dysphoria

(F64.8)

This category applies to presentations in which symptoms characteristic of gender dysphoria that cause clinically significant distress or impairment in social, occupational, or other important areas of functioning predominate but do not meet the full criteria for gender dysphoria. The other specified gender dysphoria category is used in situations in which the clinician chooses to communicate the specific reason that the presentation does not meet the criteria for gender dysphoria. This is done by recording "other specified gender dysphoria" followed by the specific reason (e.g., "brief gender dysphoria," in which symptoms meet full criteria for gender dysphoria but the duration is less than the required 6 months).

Unspecified Gender Dysphoria

(F64.9)

This category applies to presentations in which symptoms characteristic of gender dysphoria that cause clinically significant distress or impairment in social,

occupational, or other important areas of functioning predominate but do not meet the full criteria for gender dysphoria. The unspecified gender dysphoria category is used in situations in which the clinician chooses *not* to specify the reason that the criteria are not met for gender dysphoria, and includes presentations in which there is insufficient information to make a more specific diagnosis.

Copyright

American Psychiatric Association

1.

DSM-5-TR

Officers 2021–2022

1. President Vivian B. Pender, M.D.President-Elect Rebecca W. Brendel, M.D., J.D.
2. Treasurer Richard F. Summers, M.D.
3. Secretary Sandra DeJong, M.D., M.Sc.

Assembly

1. Speaker Mary Jo Fitz-Gerald, M.D., M.B.A.
2. Speaker-Elect Adam P. Nelson, M.D.

Board of Trustees

1. Elie G. Aoun, M.D., M.R.O.
2. Jenny L. Boyer, M.D., Ph.D., J.D.
3. Kenneth Certa, M.D.
4. C. Freeman, M.D., M.B.A.
5. Mary Hasbah Roessel, M.D.

6. Glenn A. Martin, M.D.
7. Eric M. Plakun, M.D.
8. Michele Reid, M.D.
9. Felix Torres, M.D., M.B.A.
10. Sanya Virani, M.D., M.P.H.
11. Cheryl D. Wills, M.D.
12. Melinda Young, M.D.
13. Urooj Yazdani, M.D., Resident-Fellow Member Trustee-Elect

DSM-5

Officers 2012–2013

1. President Dilip V. Jeste, M.D.
2. President-Elect Jeffrey A. Lieberman, M.D.
3. Treasurer David Fassler, M.D.
4. Secretary Roger Peele, M.D.

Assembly

1. Speaker R. Scott Benson, M.D.
2. Speaker-Elect Melinda L. Young, M.D.

Board of Trustees

1. Jeffrey Akaka, M.D.
2. Carol A. Bernstein, M.D.
3. Brian Crowley, M.D.
4. Anita S. Everett, M.D.
5. Jeffrey Geller, M.D., M.P.H.
6. Marc David Graff, M.D.
7. James A. Greene, M.D.
8. Judith F. Kashtan, M.D.
9. Molly K. McVoy, M.D.
10. James E. Nininger, M.D.
11. John M. Oldham, M.D.

12. Alan F. Schatzberg, M.D.
13. Alik S. Widge, M.D., Ph.D.
14. Erik R. Vanderlip, M.D., Member-in-Training Trustee-Elect

Diagnostic and Statistical Manual of Mental Disorders

Fifth Edition DSM-5-TR

Manufactured in the United States of America on acid-free paper.

ISBN 978-0-89042-575-6 (Hardcover) 1st printing February 2022

ISBN 978-0-89042-576-3 (Paperback) 1st printing February 2022

American Psychiatric Association 800 Maine Avenue SW Suite 900 Washington, DC 20024-2812 www.psychiatry.org

The correct citation for this book is American Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, Text Revision. Washington, DC, American Psychiatric Association, 2022.

Library of Congress Cataloging-in-Publication Data Names: American Psychiatric Association, issuing body. Title: Diagnostic and statistical manual of mental

disorders : DSM-5-TR / American Psychiatric Association. Other titles: DSM-5-TR Description: Fifth edition, text revision. | Washington, DC : American Psychiatric Association Publishing, [2022] | Includes index. Identifiers: LCCN 2021051781 (print) | LCCN 2021051782 (ebook) | ISBN 9780890425756 (hardcover ; alk. paper) | ISBN 9780890425763 (paperback ; alk. paper) | ISBN 9780890425770 (ebook) Subjects: MESH: Diagnostic and statistical manual of mental disorders. 5th ed | Mental Disorders—classification | Mental Disorders—diagnosis Classification: LCC RC455.2.C4 (print) | LCC RC455.2.C4 (ebook) | NLM WM 15 | DDC 616.89/075—dc23/eng /20211209 LC record available at https://lccn.loc.gov/2021051781 LC ebook record available at https://lccn.loc.gov/2021051782.

British Library Cataloguing in Publication Data A CIP record is available from the British Library.

Text Design—Tammy J. Cordova

Manufacturing—Sheridan Books, Inc.

EXHIBIT 22

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

[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).

and Staphorsius et al. 2015) provided comparative evidence and the remaining 7 studies used within-person, before and after comparisons.

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

Study	Population	Intervention and comparison	Outcomes reported
	Participants were included in the study if they were diagnosed with gender dysphoria according to the DSM-5 criteria, registered at the clinic, were prepubertal and within the appropriate age range, and had started GnRH analogues. No concomitant treatments were reported.		
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
Joseph 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/cm^3 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/cm^3 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/cm^3 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/cm^3 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/m^2 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/cm^2 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/m^2 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?

<table>
<tr><th>Outcome</th><th>Evidence statement</th></tr>
<tr><td>Diagnostic criteria</td><td>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.</td></tr>
<tr><td>Age when GnRH analogues started</td><td>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):

<table>
<tr><th>Study</th><th>Mean age (±SD)</th></tr>
<tr><td>Costa et al. 2015</td><td>16.5 years (±1.3)</td></tr>
<tr><td>de Vries et al. 2011</td><td>13.6 years (±1.8)</td></tr>
<tr><td>Joseph et al. 2019</td><td>13.2 years (±1.4) in transfemales
12.6 years (±1.0) in transmales</td></tr>
<tr><td>Khatchadourian et al. 2014</td><td>14.7 years (±1.9)</td></tr>
</table>
</td></tr>
</table>

<table>
<tr><td></td><td>
<table>
<tr><td>Klink et al. 2015</td><td>14.9 years (±1.9) in transfemales
15.0 years (±2.0) in transmales</td></tr>
</table>
<table>
<tr><th>Study</th><th>Median age (range)</th></tr>
<tr><td>Brik et al. 2020</td><td>15.5 years (11.1–18.6) in transfemales
16.1 years (10.1–17.9) in transmales</td></tr>
<tr><td>Schagen et al. 2016</td><td>13.6 years (11.6–17.9) in transfemales
14.2 years (11.1–18.6) in transmales</td></tr>
<tr><td>Vlot et al. 2017</td><td>13.5 years (11.5–18.3) in transfemales
15.1 years (11.7–18.6) in transmales</td></tr>
</table>
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.
</td></tr>
<tr><td>Duration of treatment</td><td>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.</td></tr>
</table>

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: o Impact of the drug use such as its impact on bone density, arterial hypertension, cognitive development/functioning o 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/ (836)
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)

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)

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)

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)

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,ab 696
#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,ab 16

#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,ab 1
#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,ab 0
#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 (dysphori* 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)

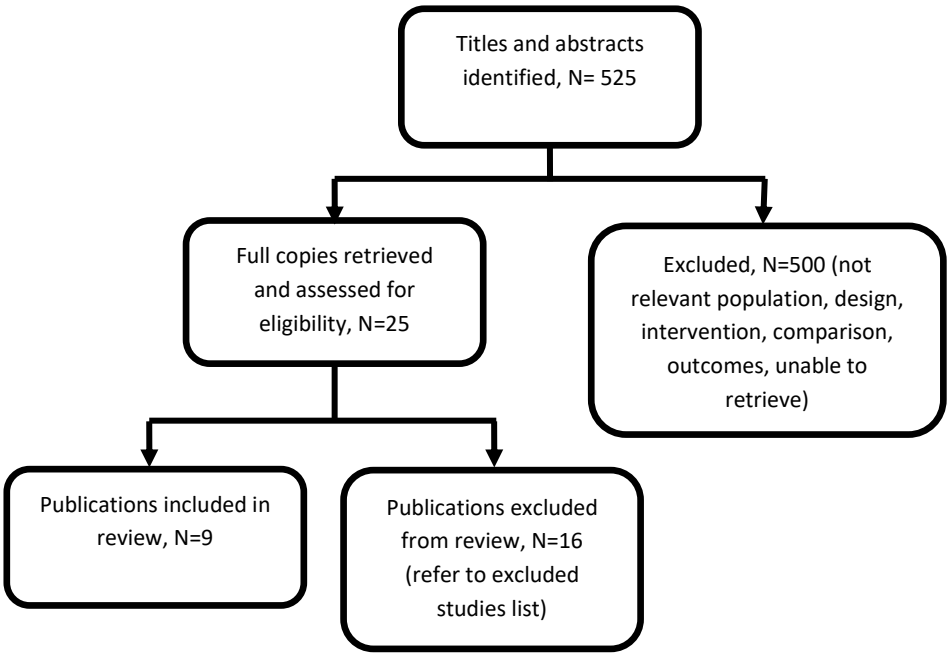
52 buserelin.ti,ab. (6)
53 bigonist.ti,ab. (0)
54 ("hoe 766" or hoe-766 or hoe766).ti,ab. (0)
55 profact.ti,ab. (0)
56 receptal.ti,ab. (0)
57 suprecur.ti,ab. (0)
58 suprefact.ti,ab. (0)
59 tiloryth.ti,ab. (0)
60 histrelin.ti,ab. (1)
61 "LHRH-hydrogel implant".ti,ab. (0)
62 ("RL 0903" or RL0903).ti,ab. (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 deslorelin.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)

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
Brik 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 2019).	**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. Six 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 analogues, 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]	

[1] The adolescent later indicated "I 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 16 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 16 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.	Adolescents with 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 160 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 56.1 [±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 [±11.38] versus 56.63 [±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.

			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<0.08 • 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	

			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	

			• 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, Doreleijers 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 196 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, errdf*), *P*: 15.98 (1,39), p<0.001. ***Impact on mental health*** Depressive symptoms were assessed using the Beck Depression Inventory (BDI-II). • 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, errdf*), *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

			birth males and sex assigned at birth females, *F* (*df, errdf*), *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, errdf*), *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, errdf*), *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, errdf*), *P*: 4.11 (1,55), p=0.047. • secondary sexual characteristics, *F* (*df, errdf*), *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, errdf*), *P*: 14.59 (1,55), p<0.001) and neutral characteristics, *F* (*df, errdf*), *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) 55.46 [±11.56] 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*: 7.26 (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 [±9.63], *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 [±1.79] 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.	0.235 (0.030) 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=0.865); 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=0.962); 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, −1.276 (1.410) at 1 year (p=0.000) **Lumbar spine BMD 0 to 2 years** Transfemales (mean [±SD]): 0.867 (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

Study details	Population	Interventions	Study outcomes	Appraisal and Funding
			0.695 (0.220) kg/m2 at baseline, 0.731 (0.209) kg/m2 at 2 years (p=0.058); 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 −0.863 (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, −0.600 (1.059) at 2 years (p=0.002) Transmales (mean [±SD]): 0.766 (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 (0.816) 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=31).
[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**

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 hormone)
			  ○ 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. in complete 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.

			• 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 analogues, 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, 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 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**.

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 (0.99) (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: 0.26 (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

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) 13.6 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**	This study was appraised using the Newcastle-Ottawa quality assessment checklist for cohort studies.

Study details	Population	Interventions	Study outcomes	Appraisal and Funding
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 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)

Study details	Population	Interventions	Study outcomes	Appraisal and Funding
Staphorsius A, Baudewijntje P, Kreukels P, et al. (2015) Puberty suppression and executive functioning: an fMRI-study	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 (triptorelin 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
in adolescents with gender dysphoria. 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 analogues 4.1 [±1.0]	subcutaneously or intramuscularly). **Comparison** The comparison was between adolescents with gender dysphoria receiving GnRH analogues and those without GnRH analogues.	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) 60.4 [±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 (15.6) • 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[3]** • 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 (0.18 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

Study details	Population	Interventions	Study outcomes	Appraisal and Funding
To investigate the course of 3 bone turnover markers in relation to bonemineral 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) (p≤0.1) 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 (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≤ 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 (p≤0.1); 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

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 (−1.30 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)	

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

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*=0.333	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: 17.88±5.24 *P*=0.503	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: 39.43±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 events/No of patients (n/N%)		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: 3.98±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).

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 events/No of patients (n/N%)		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 62.97 [±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=35 67.40 [±13.93]	n=36 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: 60.89±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).									

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[3] compared to baseline (higher scores indicate benefit).									

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[4] 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*<0.08	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).									

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[5]	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[5]	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[5]	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[5]	No serious indirectness	Not applicable	Not calculable	N=54	None	Baseline: 58.04±12.99 GnRH analogue: 53.81±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[5]	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).									

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[5]	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[5]	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[5]	No serious indirectness	Not applicable	Not calculable	N=54	None	Baseline: 53.30±11.87 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[5]	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[6]	No serious indirectness	Not applicable	Not calculable	N=8	N=10	GnRH analogue: 57.4 [±9.8] 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[6]	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

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		
Staphorsius 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 Brik 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 Brik 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 Brik 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[3]	No serious indirectness	Not applicable	Not calculable	11/27 (42%)	None	11/26 stopped GnRH analogues (42%)[4]	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 Brik 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 Brik 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 Brik 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 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), 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=31	None	Mean (SD), g/cm^3 Baseline: 0.235 (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/cm^3 Baseline: 0.196 (0.035) 1 year: 0.201 (0.033) p=0.074 z-score Baseline: −0.186 (1.230) 1 year: −0.541 (1.396) 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/cm^3 Baseline: 0.240 (0.027) 2 years: 0.240 (0.030) p=0.865 z-score Baseline: 0.486 (0.809) 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/cm^3 Baseline: 0.195 (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)							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±1.9) to starting gender-affirming hormones (mean age 16.6±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/cm^3 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±2.0) 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=18	None	Mean (SD), g/cm^3 GnRH analogue: 0.25 (0.03) Gender-affirming hormones: 0.24 (0.02) NS z-score GnRH analogue: 0.28 (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 Vlot et al. 2017	Serious limitations[3]	No serious indirectness	Not applicable	Not calculable	N=15	None	Median (range), g/cm^3 GnRH analogue: 0.21 (0.17 to 0.25) Gender-affirming hormones: 0.20 (0.18 to 0.24)	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		
							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 Vlot et al. 2017	Serious limitations[3]	No serious indirectness	Not applicable	Not calculable	N=5	None	Median (range), g/cm^3 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 Vlot et al. 2017	Serious limitations[3]	No serious indirectness	Not applicable	Not calculable	N=11	None	Median (range), g/cm^3 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 2.94) Gender-affirming hormones: −0.84 (−2.20 to 0.87) p-value: ≤0.01	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		
Change in lumbar BMAD from starting GnRH analogue to starting gender-affirming hormones in transmales (bone age of ≥14)									
1 observational study Vlot et al. 2017	Serious limitations[3]	No serious indirectness	Not applicable	Not calculable	N=23	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=39	None	Mean (SD), kg/m2 Baseline: 0.694 (0.149) 1 year: 0.718 (0.124) p=0.006 z-score Baseline: −0.395 (1.428) 1 year: −1.276 (1.410) p=0.000	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		
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 (1.384) p=0.000	IMPORTANT	VERY LOW
Change in lumbar BMD from starting GnRH analogue (mean age 14.9±1.9) to starting gender-affirming hormones (mean age 16.6±1.4) in transfemales									
1 observational study Klink et al. 2015	Serious limitations[2]	No serious indirectness	Not applicable	Not calculable	N=12 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 (0.98) NS	IMPORTANT	VERY LOW
Change in lumbar BMD from starting GnRH analogue (mean age 15.0±2.0) to starting gender-affirming hormones (mean age 16.4±2.3) in transmales									

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: 0.91 (0.10) p-value: 0.006 z-score GnRH analogue: 0.17 (1.18) Gender-affirming hormones: −0.72 (0.99) p-value: <0.001	IMPORTANT	VERY LOW
Bone density: change in femoral neck (hip) BMD									
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: 0.785 (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									

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=0.604 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)									

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 Vlot et al. 2017	Serious limitations[3]	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 Vlot et al. 2017	Serious limitations[3]	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 (−2.28 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 Vlot et al. 2017	Serious limitations[3]	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 (−1.39 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 14.9±1.9) to starting gender-affirming hormones (mean age 16.6±1.4) in transfemales									
1 observational study Klink et al. 2015	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: −0.66 (0.77) Gender-affirming hormones: −0.95 (0.63) NS	IMPORTANT	VERY LOW
Change in femoral BMD 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=18 N=13	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 14.9±1.9) to starting gender-affirming hormones (mean age 16.6±1.4) 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 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: 0.26 (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=18 N=18	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: −0.28 (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			IMPORTANCE	CERTAINTY
					No of events/No of patients% (n/N%)		Effect		
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 transfemales									
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 transfemales									
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 events/No of patients% (n/N%)		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 events/No of patients% (n/N%)		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 events/No of patients% (n/N%)		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 (12) 1 year: 66 (13) p-value: 0.20	IMPORTANT	VERY LOW
Change in serum creatinine (µmol/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
Other safety outcomes: adverse effects									
Proportion of patients reporting adverse effects									
1 cohort study Khatchadourian et al 2014	Serious limitations[2]	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 19 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 56.57 [±3.89] score at T1 56.62 [±4.0]	*F*-ratio 15.98 (*df, errdf*: 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, errdf*: 1,39), *P*=0.057	Critical	VERY LOW
Mean [±SD] Trait Anger (TPI), 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, errdf*: 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 [±2.68] 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, errdf*: 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
					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 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 [±0.16] 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, errdf*: 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 [±0.50] score at T1 2.39 [±0.69]	n-NR[2] score at T0 2.81 [±0.76] score at T1 3.18 [±0.42]	*F*-ratio 11.57 (*df, errdf*: 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
					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 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 0.081 (*df, errdf*: 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[4]	No serious indirectness	No serious inconsistency	Not calculable	n=not reported 55.4 [±12.7]	n=not reported 59.2 [±11.8]	*t*-test 2.15; *P*=0.03[5]	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[6] score at T0 73.10 [±8.84] score at T1 77.33 [±8.69]	n-NR[6] score at T0 67.25 [±11.06] score at T1 70.30 [±9.44]	*F*-ratio 5.77 (*df, errdf*: 1,39), *P*=0.021	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[7] score at T0 59.42 [±11.78] score at T1 50.38	n-NR[7] score at T0 61.73 [±13.60]	*F*-ratio 2.64 (*df, errdf*: 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		
					[±10.57]	score at T1 57.73 [±10.82]			
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[7] score at T0 60.00 [±9.51] score at T1 52.17 [±9.81]	n-NR[7] score at T0 61.80 [±14.12] score at T1 56.30 [±10.33]	*F*-ratio 1.16 (*df, errdf*: 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[7] score at T0 54.71 [±12.91] score at T1 48.75 [±10.22]	n-NR[7] score at T0 60.70 [±12.64] score at T1 57.87 [±11.66]	*F*-ratio 6.29 (*df, errdf*: 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[7] score at T0 53.56 [±12.26] score at T1 47.84 [±10.86]	n-NR[7] score at T0 57.10 [±10.87] score at T1 51.86 [±10.11]	*F*-ratio 1.99 (*df, errdf*: 1,52), *P*=0.164	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		
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[7] score at T0 55.88 [±11.81] score at T1 49.24 [±12.24]	n-NR[7] score at T0 56.17 [±13.25] score at T1 50.24 [±11.28]	*F*-ratio 0.049 (*df, errdf*: 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[7] score at T0 48.72 [±11.83] score at T1 46.52 [±9.23]	n-NR[7] score at T0 57.24 [±10.59] score at T1 52.97 [±8.51]	*F*-ratio 9.14 (*df, errdf*: 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, errdf), P: 14.59 (1,55), P<0.001) and neutral F (df, errdf), 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.

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

- Brik 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]

Copyright

EXHIBIT 23

Evidence review: Gender-affirming hormones 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 gender-affirming hormones 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 21 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 gender-affirming hormones 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).

Gender-affirming hormones are oestradiol for sex assigned at birth males (transfemales) and testosterone for sex assigned at birth females (transmales). The aim of gender-affirming hormones is to induce the development of the physical sex characteristics congruent with the individual's gender expression while aiming to improve mental health and quality of life outcomes.

No oestradiol-containing products are licensed for gender dysphoria and therefore any use for children and adolescents with gender dysphoria is off-label.

The only testosterone-containing product licensed for gender dysphoria is Sustanon 250 mg/ml solution for injection, which is indicated as supportive therapy for transmales, use of all other testosterone-containing products 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, to support young people and their families in managing the uncertainties inherent in gender identity development and to provide ongoing opportunities for exploration of gender identity. The plans may also include psychological support and exploration and, for some individuals, the use of gonadotrophin releasing hormone (GnRH) analogues in adolescence to suppress puberty; this may be followed later with gender-affirming hormones of the desired sex (NHS England 2013).

Currently NHS England, as part of the Gender Identity Development Service for Children and Adolescents, routinely commissions gender-affirming hormones for young people with continuing gender dysphoria from around their 16th birthday subject to individuals meeting the eligibility and readiness criteria (Clinical Commissioning Policy 2016).

[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).

2. Executive summary of the review

Ten observational studies were included in the evidence review. Seven studies were retrospective observational studies (Allen et al. 2019, Kaltiala et al. 2020, Khatchadourian et al. 2014, Klaver et Al. 2020, Klink et al. 2015, Stoffers et al. 2019, Vlot et al. 2017) and 3 studies were prospective longitudinal observational studies (Achille et al. 2020, Kuper et al. 2020, Lopez de Lara et al. 2020). No studies directly compared gender-affirming hormones to a control group (either placebo or active comparator). Follow-up was relatively short across all studies, with an average duration of treatment with gender-affirming hormones between around 1 year and 5.8 years.

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 saying natal or biological sex and 'cross sex hormones' are now referred to as 'gender-affirming hormones'. The research studies 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 gender-affirming hormones 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 impact on gender dysphoria, impact on mental health and quality of life. The quality of evidence for all these outcomes was assessed as very low certainty using modified GRADE.

Impact on gender dysphoria
The study by Lopez de Lara et al. 2020 in 23 adolescents with gender dysphoria found that during treatment with gender-affirming hormones, gender dysphoria (measured using the Utrecht Gender Dysphoria Scale [UGDS]) was statistically significantly reduced (improved) from a mean [±SD] score of 57.1 (±4.1) points at baseline to 14.7 (±3.2) points at 12 months, which is below the threshold (40 points) for gender dysphoria ($p<0.001$).

Impact on mental health
Depression
The study by Lopez de Lara et al. 2020 in 23 adolescents with gender dysphoria found that during treatment with gender-affirming hormones, depression (measured using the Beck Depression Inventory-II [BDI-II]) was statistically significantly reduced from a mean [±SD] score of 19.3 (±5.5) points at baseline to 9.7 (±3.9) points at 12 months ($p<0.001$).

The study by Achille et al. 2020 in 50 adolescents with gender dysphoria found that during treatment with gender-affirming hormones, depression was statistically significantly reduced from baseline to about 12 months follow-up:

- The Center for Epidemiologic Studies Depression (CESD-R) improved from a mean score of 21.4 points at baseline to 13.9 points ($p<0.001$).
- The Patient Health Questionnaire (PHQ 9) Modified for Teens improved, although absolute scores were not reported numerically ($p<0.001$).

The study by Kuper et al. 2020 in 148 adolescents with gender dysphoria (of whom 123 received gender-affirming hormones) found that during treatment with gender-affirming hormones for an average of 10.9 months, the impact on depression (measured using the Quick Inventory of Depressive Symptoms [QIDS]) was unclear as no statistical analysis was reported. The mean (±SD) self-reported score was 9.6 points (±5.0) at baseline and 7.4 (±4.5) at follow-up. The mean (±SD) clinician-reported score was 5.9 points (±4.1) at baseline and 6.0 (±3.8).

The study by Kaltiala et al. 2020 in 52 adolescents with gender dysphoria found that during treatment with gender-affirming hormones, statistically significantly fewer participants needed treatment for depression (54% at initial assessment compared with 15% at 12-month follow-up, p<0.001). No details of the treatments for depression are reported.

Anxiety

The study by Lopez de Lara et al. 2020 in 23 adolescents with gender dysphoria found that during treatment with gender-affirming hormones, state anxiety (measured using the State-Trait Anxiety Inventory [STAI] – State subscale) was statistically significantly reduced from a mean (±SD) score of 33.3 points (±9.1) at baseline to 16.8 points (±8.1) at 12 months (p<0.001). Trait anxiety (measured using STAI – Trait subscale) was also statistically significantly reduced from a mean (±SD) score of 33.0 (±7.2) points at baseline to 18.5 (±8.4) points at 12 months (p<0.001).

The study by Kuper et al. 2020 in 148 adolescents with gender dysphoria found that during treatment with gender-affirming hormones, small reductions were seen in anxiety, panic, generalised anxiety, social anxiety and separation anxiety symptoms and school avoidance (measured using the Screen for Child Anxiety Related Emotional Disorders [SCARED] questionnaire) from baseline to follow-up (mean duration of treatment 10.9 months). The statistical significance of these findings are unknown as no statistical analyses were reported.

The study by Kaltiala et al. 2020 in 52 adolescents with gender dysphoria found that during treatment with gender-affirming hormones, statistically significantly fewer participants needed treatment for anxiety (48% at initial assessment compared with 15% at 12-month follow-up, p<0.001). No details of treatments for anxiety are reported.

Suicidality and self-injury

The study by Allen et al. 2019 in 47 adolescents with gender dysphoria found that during treatment with gender-affirming hormones, suicide risk (measured using the Ask Suicide-Screening Questions [ASQ]) was statistically significantly reduced from an adjusted mean (±SE) score of 1.11 points (±0.22) at baseline to 0.27 points (±0.12) after about 12 months (p<0.001).

The study by Achille et al. 2020 in 50 adolescents with gender dysphoria (of whom 35 received gender-affirming hormones at follow-up) found that during treatment with gender-affirming hormones, the impact on suicidal ideation was unclear (measured using the PHQ 9_Modified for Teens with additional questions for suicidal ideation). At baseline 10%of participants had suicidal ideation and 6% had suicidal ideation after about 12 months, but it is unclear if these participants received gender-affirming hormones. No statistical analyses were reported.

The study by Kuper et al. 2020 in 148 adolescents with gender dysphoria reported the impact on suicidal ideation, suicide attempts and non-suicidal self-injury during treatment with gender-affirming hormones, after mean 10.9 months follow-up . The statistical significance of these findings are unknown as no statistical analyses were reported:

- Suicidal ideation was reported in 25% of participants 1 month before the initial assessment and in 38% of participants during follow-up.
- Suicide attempts were reported in 2% of participants at 3 months before the initial assessment and in 5% during follow-up.
- Self-injury was reported in 10% of participants at 3 months before the initial assessment and in 17% during follow-up.

The study by Kaltiala et al. 2020 in 52 adolescents with gender dysphoria reported that during treatment with gender-affirming hormones, statistically significantly fewer participants needed treatment for suicidal ideation or self-harm (35% at initial assessment compared with 4% at 12-month follow-up, p<0.001). No details of treatments for suicidal ideation or self-harm are reported.

Other related symptoms

The study by Kaltiala et al. 2020 in 52 adolescents with gender dysphoria found that during treatment with gender-affirming hormones, there was no statistically significant difference in the number of people needing treatment for either psychotic symptoms or psychosis, conduct problems or antisocial behaviour, substance abuse, autism, attention deficit hyperactivity disorder (ADHD) or eating disorders during the 12-month 'real life' phase compared with before or during the assessment. No details of the treatments received are reported.

Impact on quality of life

The study by Achille et al. 2020 in 50 adolescents with gender dysphoria (of whom 35 were receiving gender-affirming hormones at follow-up) found that during treatment with gender-affirming hormones, quality of life (measured using the Quality of Life Enjoyment and Satisfaction Questionnaire [QLES-Q-SF]) was statistically significantly improved from baseline to about 12 months, but absolute scores were not reported numerically (p<0.001).

The study by Allen et al. 2019 in 47 adolescents with gender dysphoria found that during treatment with gender-affirming hormones, quality of life (measured using the General Well-Being Scale [GWBS] of the Paediatric Quality of Life Inventory) was statistically significantly improved from an adjusted mean (±SE) score of 61.70 (±2.43) points at baseline to 70.23 (±2.15) points at about 12 months (p<0.002).

Important outcomes

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

Impact on body image

The study by Kuper et al. 2020 in 148 adolescents with gender dysphoria found that during treatment with gender-affirming hormones, the impact on body image is unclear (measured using the Body Image Scale [BIS]). The mean (±SD) BIS score was 70.7 points (±15.2) at baseline and 51.4 points (±18.3) at follow-up (mean duration of treatment 10.9 months; no statistical analysis was reported).

Psychosocial impact
The study by Lopez de Lara et al. 2020 in 23 adolescents with gender dysphoria found that during treatment with gender affirming hormones, family functioning is unchanged (measured using the Family Adaptability, Partnership, Growth, Affection and Resolve [APGAR] test). The mean score was 17.9 points at baseline and 18.0 points at 12-month follow-up (no statistical analysis was reported).

The study by Lopez de Lara et al. 2020 in 23 adolescents with gender dysphoria found that during treatment with gender affirming hormones, behavioural problems (measured using the Strengths and Difficulties Questionnaire [SDQ]) were statistically significantly improved from a mean (±SD) of 14.7 (±3.3) points at baseline to 10.3 points (±2.9) at 12-month follow-up ($p<0.001$).

The study by Kaltiala et al. 2020 in 52 adolescents with gender dysphoria found that about 12-months after starting treatment with gender-affirming hormones:

- Statistically significantly fewer participants were living with parents or guardians (73% versus 40%, $p=0.001$) and statistically significantly fewer participants had normal peer contacts (89% versus 81%, $p<0.001$).
- There were no statistically significant differences in:
    - progress in school or work (64% versus 60%, $p=0.69$),
    - the number of participants who had been dating or in steady relationships (62% versus 58%, $p=0.51$)
    - the ability to cope with matters outside of the home (for example, shopping and travelling alone on local public transport; 81% versus 81%, $p=1.0$)

Engagement with health care services
No evidence was identified.

Impact on extent of and satisfaction with surgery
No evidence was identified.

De-transition
No evidence was identified.

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

Important outcomes

The important outcomes for decision making are short- and long-term safety outcomes and adverse effects. The quality of evidence for all these outcomes was assessed as very low certainty using modified GRADE.

Bone density

The study by Klink et al. 2015 in 34 adolescents with gender dysphoria (who were previously treated with a GnRH analogue) found that gender-affirming hormones may increase lumbar spine and femoral neck bone density. However, not all results are statistically significant (particularly in transfemales). Z-scores suggest the average bone density at the end of follow-up was generally lower than in the equivalent cisgender population (transfemales compared with cis-males and transmales compared with cis-females). From starting gender-affirming hormones to age 22 years:

- There was no statistically significant difference in lumbar spine bone mineral apparent density (BMAD) z-score in transfemales, but this was statistically significantly higher in transmales (z-score [±SD]: start of hormones -0.50 [±0.81], age 22 years -0.033 [±0.95], p=0.002).
- There was no statistically significant difference in lumbar spine bone mineral density (BMD) z-score in transfemales or transmales.
- Actual lumbar spine BMAD and BMD values were statistically significantly higher in transfemales and transmales.
- There was no statistically significant difference in femoral neck BMD z-score in transfemales, but this was statistically significantly higher in transmales (z-score [SD]: start of hormones -0.35 [0.79], age 22 years -0.35 [0.74], p=0.006).
- There was no statistically significant difference in actual femoral neck BMAD values in transfemales, but this was statistically significantly higher in transmales.
- Actual femoral neck BMD values were statistically significantly higher in transfemales and transmales.

The study by Vlot et al. 2017 in 70 adolescents with gender dysphoria (who were previously treated with a GnRH analogue) found that gender-affirming hormones may increase lumbar spine and femoral neck bone density. However, not all results are statistically significant. Z-scores suggest the average bone density at the end of follow-up was generally lower than the equivalent cisgender population (transfemales compared with cis-males and transmales compared with cis-females). From starting gender-affirming hormones to 24-month follow-up:

- The z-score for lumbar spine BMAD was statistically significantly higher in transfemales with a bone age of less than 15 years (z-score [range]: start of hormones -1.52 [-2.36 to 0.42], 24-month follow-up -1.10 [-2.44 to 0.69], p≤ 0.05) and 15 years and older (z-score [range]: start of hormones -1.15 [-2.21 to 0.08], 24-month follow-up -0.66 [-1.66 to 0.54], p≤ 0.05).
- The z-score for lumbar spine BMAD was statistically significantly higher in transmales with a bone age of less than 14 years (z-score [range]: start of hormones -0.84 [-2.2 to 0.87], 24-month follow-up -0.15 [-1.38 to 0.94], p≤ 0.01) and 14 years and older (z-score [range]: start of hormones -0.29 [-2.28 to 0.90], 24-month follow-up -0.06 [-1.75 to 1.61], p≤ 0.01).
- Actual lumbar spine BMAD values were statistically significantly higher in transfemales and transmales of all bone ages.
- There was no statistically significant difference in femoral neck BMAD z-score in transfemales (all bone ages).

- The z-score for femoral neck BMAD was statistically significantly higher in transmales with a bone age of less than 14 years (z-score [range]: start of hormones -0.37 [-2.28 to 0.47], 24-month follow-up -0.37 [-2.03 to 0.85], p≤ 0.01) and 14 years and older (z-score [range]: start of hormones -0.27 [-1.91 to 1.29], 24-month follow-up 0.02 [-2.1 to 1.35], p≤0.05).
- There was no statistically significant difference in actual femoral neck BMAD values in transfemales (all bone ages), but this was statistically significantly higher in transmales (all bone ages).

The study by Stoffers et al. 2019 in 62 sex assigned at birth females (transmales) with gender dysphoria (who were previously treated with a GnRH analogue) found that during treatment with gender-affirming hormones there was no statistically significant difference in lumbar spine or femoral neck bone density (measured as BMD z-scores or actual values) from starting gender-affirming hormones to any timepoint (6, 12 and 24 months).

Change in clinical parameters

The study by Klaver et al. 2020 in 192 adolescents with gender dysphoria found that during treatment with gender-affirming hormones, from starting treatment to age 22 years:

- Glucose levels, insulin levels and insulin resistance were largely unchanged in transfemales and transmales.
- Total cholesterol, HDL cholesterol and LDL cholesterol levels were unchanged in transfemales, and there was a statistically significant improvement in triglyceride levels.
- Total cholesterol, HDL cholesterol, LDL cholesterol and triglyceride levels significantly worsened in transmales, but mean levels were within the UK reference range at the end of treatment.
- Diastolic blood pressure was statistically significantly increased in transfemales and transmales. Systolic blood pressure was also statistically significantly increased in transmales, but not in transfemales. The absolute increases in blood pressure were small.
- Body mass index was statistically significantly increased in transfemales and transmales, although most participants were within the healthy weight range (18.5 to 24.9 kg/m).

The study by Stoffers et al. 2019 in 62 sex assigned at birth females (transmales) with gender dysphoria found that during treatment with gender affirming hormones, from starting treatment to 24-month follow-up:

- There was no statistically significant change in glycosylated haemoglobin (HbA1c).
- There was no statistically significant change in aspartate aminotransferase (AST), alanine aminotransferase (ALT) and gamma-glutamyltransferase (GCT).
- There was a statistically significant increase in alkaline phosphatase (ALP) at some timepoints, but the difference was not statistically significant by 24-months.
- There was a statistically significant increase in serum creatinine levels at all timepoints up to 24 months, but these were within the UK reference range. Serum urea levels were unchanged (follow-up duration not reported).

Treatment discontinuation and adverse effects

The study by Khatchadourian et al. 2014 in 63 adolescents (24 transfemales and 39 transmales) with gender dysphoria found that during treatment with gender affirming hormones (duration of treatment not reported):

- No participants permanently discontinued treatment.
- No transfemales temporarily discontinued treatment, but 3 transmales temporarily discontinued treatment due to mental health comorbidities (n=2) and androgenic alopecia (n=1). All 3 participants eventually resumed treatment, although timescales were not reported
- No severe complications were reported.
- No transfemales reported minor complications, but 12 transmales developed minor complications which were: severe acne (n=7), androgenic alopecia (n=1), mild dyslipidaemia (n=3) and significant mood swings (n=1).

In children and adolescents with gender dysphoria, what is the cost-effectiveness of gender-affirming hormones 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 gender-affirming hormones for children and adolescents with gender dysphoria.

From the evidence selected, are there particular sub-groups of children and adolescents with gender dysphoria that derive comparatively more (or less) benefit from treatment with gender-affirming hormones than the wider population of children and adolescents with gender dysphoria?

Some studies reported data separately for the following subgroups of children and adolescents with gender dysphoria:

- Sex assigned at birth males (transfemales).
- Sex assigned at birth females (transmales).
- Tanner stage at which GnRH analogue or gender-affirming hormones started.
- Diagnosis of a mental health condition.

Some direct comparisons of transfemales and transmales were included. No evidence was found for other specified subgroups.

Sex assigned at birth males (transfemales)
Impact on mental health
In the study by Kuper et al. 2020 in 33 to 45 (number varies by outcome) sex assigned at birth males (transfemales) with gender dysphoria found that during treatment with gender-affirming hormones changes were seen in depression, anxiety and anxiety-related symptoms from baseline to follow-up (mean duration of treatment 10.9 months). The authors did not report any statistical analyses, so it is unclear if any changes were statistically significant.

The study by Allen et al. 2019 in 47 adolescents with gender dysphoria found that during treatment with gender-affirming hormones, suicide risk (measured using the ASQ) is not statistically significant different in transfemales compared with transmales, between baseline and the final assessment at about 12 months (p=0.79).

The study by Achille et al. 2020 in 17 transfemales with gender dysphoria found that during treatment with gender-affirming hormones, suicidal ideation (measured using the PHQ 9_Modified for Teens with additional questions for suicidal ideation) was reported in 11.8% (2/17) of transfemales at baseline compared with 5.9% (1/17) at about 12-months follow-up (no statistical analysis was reported).

Impact on quality of life

The study by Allen et al. 2019 in 47 adolescents with gender dysphoria found that during treatment with gender-affirming hormones, quality of life (measured using the GWBS of the Paediatric Quality of Life Inventory) was not statistically significant different in transfemales compared with transmales, between baseline and the final assessment at about 12 months (p=0.32).

Bone density

The studies by Klink et al. 2015 and Vlot et al. 2017 provided evidence on bone density in transfemales; see above for details.

Change in clinical parameters

The study by Klaver et al. 2020 provided evidence on the following clinical parameters in transfemales:

- Glucose levels, insulin levels and insulin resistance.
- Total cholesterol, HDL cholesterol and LDL cholesterol and triglycerides.
- Blood pressure.
- Body mass index.

See above for details.

Treatment discontinuation and adverse effects

The study by Khatchadourian et al. 2014 provided evidence on treatment discontinuation and adverse effects in transfemales; see above for details.

Sex assigned at birth females (transmales)

Impact on mental health

In the study by Kuper et al. 2020 in 65 to 78 (number varies by outcome) sex assigned at birth females (transmales) with gender dysphoria found that during treatment with gender-affirming hormones, changes were seen in depression, anxiety and anxiety-related symptoms from baseline to 10.9 month follow-up. The authors did not report any statistical analyses, so it is unclear if any changes were statistically significant.

The study by Allen et al. 2019 in 47 adolescents with gender dysphoria found that during treatment with gender-affirming hormones, suicide risk (measured using the ASQ) is not statistically significantly different in transmales compared with transfemales, between baseline and the final assessment (p=0.79).

The study by Achille et al. 2020 in 33 transmales with gender dysphoria found that during treatment with gender-affirming hormones, suicidal ideation (measured using the PHQ 9_Modified for Teens with additional questions for suicidal ideation) was reported in 9.1% (3/33) of transmales at baseline compared with 6.1% (2/33) at about 12-months follow-up (no statistical analysis reported).

Impact on quality of life
The study by Allen et al. 2019 in 47 adolescents with gender dysphoria found that during treatment with gender-affirming hormones, quality of life (measured using the GWBS of the Paediatric Quality of Life Inventory) was not statistically significantly different in transmales compared with transfemales, between baseline and the final assessment at about 12 months (p=0.32).

Bone density
The studies by Klink et al. 2015, Stoffers et al. 2019 and Vlot et al. 2017 provided evidence on bone density in transmales; see above for details.

Change in clinical parameters
The study by Klaver et al. 2020 provided evidence on the following clinical parameters in transmales:

- Glucose levels, insulin levels and insulin resistance.
- Total cholesterol, HDL cholesterol and LDL cholesterol and triglycerides.
- Blood pressure.
- Body mass index.

See above for details.

The study by Stoffers et al. 2019 provided evidence on HbA1c, liver enzymes and renal function in transmales; see above for details.

Treatment discontinuation and adverse effects
The study by Khatchadourian et al. 2014 provided evidence on treatment discontinuation and adverse effects in transmales; see above for details.

Tanner stage at which GnRH analogues or gender-affirming hormones started
The study by Kuper et al. 2020 stated that the impact of Tanner stage on outcomes was considered, but it is unclear if this refers to Tanner stage at the initial assessment, at the start of GnRH analogue treatment or another timepoint. No results were reported.

Diagnosis of a mental health condition
Impact on mental health
The study by Achille et al. 2020 in 50 adolescents with gender dysphoria found that during treatment with gender-affirming hormones, there was no statistically significant difference in depression (measured using the CESD-R and PHQ 9_Modified for Teens) when the results were adjusted for engagement in counselling and medicines for mental health problems, from baseline to about 12-months follow-up.

Impact on quality of life
The study by Achille et al. 2020 in 50 adolescents with gender dysphoria found that during treatment with gender-affirming hormones, there was no statistically significant difference in quality of life (measured using the QLES-Q-SF) when the results were adjusted for engagement in counselling and medicines for mental health problems, from baseline to about 12-months follow-up.

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 gender-affirming hormones?
(c) what was the duration of treatment with GnRH analogues?

The most commonly reported diagnostic criteria for gender dysphoria was the DSM criteria in use at the time (5/10 studies). In 3 studies (Klaver et al. 2020, Klink et al. 2015 and Vlot et al. 2017) DSM-IV-TR criteria was used. In 2 studies (Kuper et al. 2020 and Stoffers et al. 2019) DSM-V criteria was used. One study from Finland (Kaltiala et al. 2020) used the ICD-10 diagnosis of 'transexualism'. It was not reported how gender dysphoria was defined in the remaining 4 studies.

In the studies, treatment with gender-affirming hormones started at about 16 to 17 years, with a range of about 14 to 19 years. Most studies did not report the duration of treatment with GnRH analogues, but where this was reported there was a wide variation ranging from a few months up to about 5 years (Klaver et al. 2020, Klink et al. 2015 and Stoffers et al. 2019).

Discussion

The key limitation to identifying the effectiveness and safety of gender-affirming hormones for children and adolescents with gender dysphoria is the lack of reliable comparative studies.

All the studies included in the evidence review are uncontrolled observational studies, which are subject to bias and confounding and were of very low certainty using modified GRADE. A fundamental limitation of all the uncontrolled studies included in this review is that any changes in scores from baseline to follow-up could be attributed to a regression-to-the-mean.

The included studies have relatively short follow-up, with an average duration of treatment with gender-affirming hormones between around 1 year and 5.8 years. Further studies with a longer follow-up are needed to determine the long-term effect of gender-affirming hormones for children and adolescents with gender dysphoria.

Most studies included in this review did not report comorbidities (physical or mental health) and no study reported concomitant treatments in detail. Because of this it is not clear whether any changes seen were due to gender-affirming hormones or other treatments the participants may have received.

There is a degree of indirectness in some studies, with some participants included that fall outside of the population of this evidence review. Furthermore, participant numbers are poorly reported in some studies, with high numbers lost to follow-up or outcomes not reported for some participants. The authors provide no explanation for this incomplete reporting.

Details of the gender-affirming hormone treatment regimen are poorly reported in most of the included studies, with limited information provided about the medicines, doses and routes of administration used. It is not clear whether the interventions used in the studies are reflective of current UK practice for children and adolescents with gender dysphoria.

It is difficult to draw firm conclusions for many of the effectiveness and safety outcomes reported in the included studies because many different scoring tools and methods were used to assess the same outcome, often with conflicting results. In addition to this, most outcomes reported across the included studies do not have an accepted minimal clinically important difference (MCID), making it difficult the determine whether any statistically significant changes seen are clinically meaningful. However, the authors of some studies report thresholds to interpret the results of the scoring tools (for example, by linking scores to symptom severity), so some conclusions can be made.

Conclusion

Any potential benefits of gender-affirming hormones must be weighed against the largely unknown long-term safety profile of these treatments in children and adolescents with gender dysphoria.

Results from 5 uncontrolled, observational studies suggest that, in children and adolescents with gender dysphoria, gender-affirming hormones are likely to improve symptoms of gender dysphoria, and may also improve depression, anxiety, quality of life, suicidality, and psychosocial functioning. The impact of treatment on body image is unclear. All results were of very low certainty using modified GRADE.

Safety outcomes were reported in 5 observational studies. Statistically significant increases in some measures of bone density were seen following treatment with gender-affirming hormones, although results varied by bone region (lumber spine versus femoral neck) and by population (transfemales versus transmales). However, z-scores suggest that bone density remained lower in transfemales and transmales compared with an equivalent cisgender population. Results from 1 study of gender-affirming hormones started during adolescence reported statistically significant increases in blood pressure and body mass index, and worsening of the lipid profile (in transmales) at age 22 years, although longer term studies that report on cardiovascular event rates are required. Adverse events and discontinuation rates associated with gender-affirming hormones were only reported in 1 study, and no conclusions can be made on these outcomes.

This review did not identify sub-groups of patients who may benefit more from gender-affirming hormones.

No cost-effectiveness evidence was found to determine whether gender-affirming hormones are a cost-effective treatment for children and adolescents with gender dysphoria.

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 gender-affirming hormones 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 gender-affirming hormones 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 gender-affirming hormones 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 particular sub-groups of children and adolescents with gender dysphoria that derive comparatively more (or less) benefit from treatment with gender-affirming hormones 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 gender-affirming hormones?
   (c) what was the duration of GnRH analogues treatment?

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 and were conducted on 21 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 appendix E and appendix 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

Ten observational studies were included in the evidence review. Seven studies were retrospective observational studies (Allen et al. 2019, Kaltiala et al. 2020, Khatchadourian et al. 2014, Klaver et Al. 2020, Klink et al. 2015, Stoffers et al. 2019, Vlot et al. 2017) and three studies were prospective longitudinal observational studies (Achille et al. 2020, Kuper et al. 2020, Lopez de Lara et al. 2020).

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 saying natal or biological sex and 'cross sex hormones' are now referred to as 'gender-affirming hormones'. The research studies 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
Achille et al. 2020 Prospective longitudinal study Single centre, New York, United States	50 children, adolescents and young adults with gender dysphoria; 17 transfemales and 33 transmales Mean age at baseline was 16.2 years (SD 2.2)	**Intervention** Endocrine interventions (the collective term used for puberty suppression and gender-affirming hormones) were introduced as per Endocrine Society and the World Professional Association for Transgender Health (WPATH) guidelines Puberty suppression was: • GnRH analogue and/or anti-androgens (transfemales) • GnRH analogue or medroxyprogesterone (transmales) Once eligible, gender-affirming hormones were offered, these were: • Oestradiol (transfemales)	**Critical Outcomes** *Impact on mental health* • Depression- The Center for Epidemiologic Studies Depression Scale (CESD-R) • Depression- The Patient Health Questionnaire Modified for Teens (PHQ 9_Modified for Teens) *Impact on quality of life* • Quality of Life Enjoyment and Satisfaction Questionnaire (QLES-Q-SF) **Important Outcomes** *None reported*

Study	Population	Intervention and comparison	Outcomes reported
		• Testosterone (transmales) Doses and formulations not reported After about 12-months treatment ('wave 3'): • 24 people (48%) were on gender-affirming hormones alone • 12 people (24%) were on puberty suppression alone • 11 people (22%) were on both gender-affirming hormones and puberty suppression • 3 people (6%) were on no endocrine intervention **Comparison** No comparison group. Change over time reported	
Allen et al. 2019 Retrospective longitudinal study Single centre, Kansas City, USA	47 adolescents and young adults with gender dysphoria: 14 transfemales and 33 transmales Mean age at administration (start of treatment) 16.5 years	**Intervention** 39 participants received gender-affirming hormones only 8 participants received hormones and a GnRH analogue Mean duration of treatment with gender-affirming hormones was 349 days (range 113 to 1,016) **Comparison** No comparison group. Comparison over time reported	**Critical Outcomes** *Impact on mental health* • Suicidality- Ask Suicide-Screening Questions (ASQ) instrument *Impact on quality of life* • General Well-Being Scale (GWBS) of the Pediatric Quality of Life Inventory **Important Outcomes** *None reported*
Kaltiala et al. 2020	52 adolescents with gender dysphoria: 11 transfemales and 41 transmales.	**Intervention** Hormonal sex assignment treatment – details of	**Critical Outcomes** *Impact on mental health*

Study	Population	Intervention and comparison	Outcomes reported
Retrospective chart review Single centre, Tampere, Finland	Mean age at diagnosis 18.1 years (range 15.2 to 19.9)	intervention not reported, although all patients received gender-affirming hormones. **Comparison** No comparison group. Comparison over time reported	• Need for mental health treatment **Important Outcomes** *Psychosocial Impact* Measure of functioning in different domains of adolescent development, which were: • Living with parent(s)/ guardians • Normative peer contacts • Progresses normatively in school/ work • Has been dating or had steady relationships • Is age-appropriately able to deal with matters outside of the home
Khatchadourian et al. 2014 Retrospective chart review Single centre, Vancouver, Canada	84 young people with gender dysphoria, of whom 63 received gender-affirming hormones. Median age at start of gender-affirming hormones was: • 17.3 years (range 13.7-19.8) for testosterone • 17.9 years (range 13.3-22.3) for oestrogen	**Intervention** Transfemales: Oestrogen (oral micronized 17β-oestradiol) Transmales: Testosterone (injectable testosterone enanthate and/or cypionate) 19 participants (30%) had previously received a GnRH analogue **Comparison** No comparison group. Comparison over time reported.	**Critical Outcomes** *None reported* **Important Outcomes** *Safety:* • Adverse events • Discontinuation rates
Klaver et al. 2020 Retrospective chart review Single centre, Amsterdam, Netherlands	192 people with gender dysphoria who started GnRH analogues before the age of 18 years, and started gender-affirming hormones within 1.5 years of their 22nd birthday.	**Intervention** Oral oestrogen or intramuscular (IM) testosterone **Comparison**	**Critical Outcomes** *None reported* **Important Outcomes** *Safety* • Body mass index (BMI)

Study	Population	Intervention and comparison	Outcomes reported
	Mean age at start of gender-affirming hormones: • Transfemale – 16.4 years (SD 1.1) • Transmale – 16.9 years (SD 1.9)	No comparison group. Comparison over time reported	• Systolic blood pressure • Diastolic blood pressure • Glucose • Insulin • HOMA-IR • Total cholesterol • HDL cholesterol • LDL cholesterol • Triglycerides
Klink et al. 2015 Retrospective longitudinal study Single centre, Amsterdam, Netherlands	34 young people with gender dysphoria who had received GnRH analogues, gender-affirming hormones and gonadectomy. The study included 15 transfemales and 19 transmales; mean age at start of gender-affirming hormones was 16.6 years (SD 1.4) and 16.4 years (SD 2.3) respectively. At the start of gender-affirming hormone treatment, in the transfemale subgroup the median Tanner P was 4 (IQR 2) and the median Tanner G was 12 (IQR 11) In the transmale subgroup the median Tanner B was 5 (IQR 2) and the median Tanner P was 5 (IQR 0)	**Intervention** Transfemales – oral 17-β oestradiol (incremental dosing) Transmales – IM testosterone (Sustanon 250 mg/ml; incremental dosing) Median duration of treatment with gender-affirming hormones for transfemales was 5.8 years (range 3.0 to 8.0) and for transmales was 5.4 years (range 2.8 to 7.8) The GnRH analogue was subcutaneous (SC) triptorelin 3.75 mg every 4 weeks No details of gonadectomy reported **Comparison** No comparison group. Comparison over time reported.	**Critical Outcomes** None **Important Outcomes** *Safety* • Bone mineral apparent density (BMAD) • Bone mineral density (BMD) Measures reported at 3 timepoints: start of GnRH analogue treatment, start of gender-affirming hormone treatment and age 22 years.
Kuper et al. 2020 Prospective longitudinal study	Children and adolescents with gender dysphoria (9 to18 years), n=148, of whom: • 25 received puberty suppression only	**Intervention** Gender-affirming hormones, guided by Endocrine Society Clinical Practice Guidelines	**Critical Outcomes** *Impact on mental health* • Depression- Quick Inventory of Depressive

Study	Population	Intervention and comparison	Outcomes reported
Single centre, Texas, USA	• 93 received gender-affirming hormone therapy only • 30 received both Mean age 14.9 years	**Comparison** No comparison group. Comparison over time reported.	Symptoms (QIDS), self-reported • Depression- QIDS, clinician-reported • Anxiety- Screen for Child Anxiety Related Emotional Disorders (SCARED) • Panic- specific questions from SCARED • Generalised anxiety-specific questions from SCARED • Social anxiety - specific questions from SCARED • Separation anxiety-specific questions from SCARED • School avoidance-specific questions from SCARED **Important Outcomes** *Impact on body image* • Body Image Scale (BIS)
Lopez de Lara et al. 2020 Prospective analytical study Single centre, Madrid, Spain	23 adolescents with gender dysphoria: 7 transfemales and 16 transmales. Mean age at baseline was 16 years (range 14 to 18)	**Intervention** Gender-affirming hormones: • Oral oestradiol • Intramuscular testosterone Participants had previously received GnRH analogues in the intermediate pubertal stages (Tanner 2 to 3). Participants were assessed twice: • pre-treatment (T0), • after 12 months treatment with gender-affirming hormones (T1)	**Critical Outcomes** *Impact on gender dysphoria* • Utrecht Gender Dysphoria Scale (UGDS) *Impact on mental health* • Depression- Beck Depression Inventory II (BDI-II) • Anxiety- State-Trait Anxiety Inventory **Important Outcomes** *Psychosocial Impact* • Family functioning-Family APGAR test • Patient strengths and difficulties-Strengths and Difficulties Questionnaire,

Study	Population	Intervention and comparison	Outcomes reported
		Comparison No comparison group. Comparison over time reported.	Spanish Version (SDQ-Cas).
Stoffers et al. 2019 Retrospective chart review Single centre, Leiden, Netherlands	62 transmales with gender dysphoria. Patients had received a GnRH analogue and more than 6 months of testosterone treatment. Median age at start of testosterone was 17.23 years (range 14.9 to 18.4) Median treatment duration was 12 months (range 5 to 33) Change over time	**Intervention** Testosterone intramuscular injections (Sustanon 250 mg). Dose was titrated to a maintenance dose of 125 mg every 2 weeks. Participants who started GnRH analogues at 16 years or older had their dose increased more rapidly. Some participants chose to receive testosterone every 3-4 weeks, and participants could switch to transdermal preparations if needed. **Comparison** No comparison group. Comparison over time reported.	**Critical Outcomes** None **Important Outcomes** *Safety* • Body mass index (BMI) • Blood pressure • BMD • Acne • Liver enzymes • Creatinine • Urea • HbA1c
Vlot et al. 2017 Retrospective chart review Single centre, Amsterdam, Netherlands	70 children and adolescents with gender dysphoria Median age at baseline – • 13.5 years (11.5-18.3) for transfemales • 15.1 years (range 11.7-18.6) for transmales Comparison is change over time. 24 month follow-up.	**Intervention** Oestrogen or testosterone (had previously received triptorelin for puberty suppression) **Comparison** No comparison group. Comparison over time reported.	**Critical Outcomes** None **Important Outcomes** *Safety* • Bone mineral apparent density (BMAD)

5. Results

In children and adolescents with gender dysphoria, what is the clinical effectiveness of gender-affirming hormones 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 study (Lopez de Lara et al. 2020) provided evidence relating to the impact on gender dysphoria, measured using the Utrecht Gender Dysphoria Scale (UGDS) score during the first year of treatment with gender-affirming hormones. The UGDS is a validated, screening tool for both adolescents and adults, used 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 authors state that the cut-off point to identify gender dysphoria is 40 points. The higher the UGDS score the greater the gender dysphoria. In this study (n=23), the mean (±SD) UGDS score was statistically significantly reduced (improved) from 57.1 (±4.1) points at baseline to 14.7 points (±3.2) at 12 months (p<0.001). A UGDS score below 40 suggests an absence of gender dysphoria **(VERY LOW)**. **This study provides very low certainty evidence that gender-affirming hormones statistically significantly improve gender dysphoria from baseline to 12 months follow-up. The mean UGDS score was below the threshold for gender dysphoria at follow-up.**
Impact on mental health: depression **Certainty of evidence: very low**	This is a critical outcome because depression may impact on social, occupational, or other areas of functioning in children and adolescents. Four observational studies (Achille et al. 2020; Kaltiala et al. 2020; Kuper et al. 2020; Lopez de Lara et al. 2020) provided evidence relating to the impact on depression in children and adolescents with gender dysphoria, with follow-up of around 12 months. Five different outcome measures for depression were reported. **Beck Depression Inventory (BDI-II)** One uncontrolled, prospective, analytical study (Lopez de Lara et al. 2020) reported the change in 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. In Lopez de Lara et al. 2020 (n=23) the mean (±SD) BDI-II score was statistically significantly reduced (improved) from 19.3 (±5.5) points at baseline to 9.7 (±3.9) points at 12 months (p<0.001) **(VERY LOW)**. **Center for Epidemiologic Studies Depression (CESD-R)** One uncontrolled, prospective, longitudinal study (Achille et al. 2020) reported the change in CESD-R scale. The CESD-R is a valid, widely used tool to assess depressive symptoms. Total score ranges from 0 to 60, with higher scores indicating more depressive symptoms. There are no specific scores to categorise depression severity, although the authors of the study suggest that a total CESD-R score less than 16 suggests no clinical depression.

	In Achille et al. 2020 (n=50), the mean CESD-R score statistically significantly reduced (improved) from 21.4 points at baseline to 13.9 points at about 12 months follow-up (p<0.001; standard deviation not reported) **(VERY LOW)**. **Patient Health Questionnaire (PHQ 9) Modified for Teens** One uncontrolled, prospective, longitudinal study (Achille et al. 2020) reported the change in PHQ 9_Modified for Teens score. The PHQ 9_Modified for Teens is a validated tool to assess depression, dysthymia and suicide risk. The tool consists of 9 questions scored from 0 to 3 (total score 0 to 27), plus an additional 4 questions that are not scored. A score of 0 to 4 suggests no or minimal depressive symptoms, 5 to 9 mild, 10 to 14 moderate, 15 to 19 moderately severe, and 20-27 severe symptoms. In Achille et al. 2020 (n=50), the mean PHQ 9_Modified for Teens score statistically significantly reduced (improved) from baseline to around 12 months follow-up, although absolute scores were not reported numerically (p<0.001). From the visual representation of results, the PHQ-9_Modified for Teens score is about 9 at baseline and about 5 at final follow-up **(VERY LOW)**. **Quick Inventory of Depressive Symptoms (QIDS)** One uncontrolled, prospective, longitudinal study (Kuper et al. 2020) reported the change in QIDS, clinician-reported and self-reported. Both the clinician-reported and self-reported QIDS are validated tools to assess depressive symptoms. The tool consists of 16 items, with the highest score for 9 domains (sleep, weight, psychomotor changes, depressed mood, decreased interest, fatigue, guilt, concentration, and suicidal ideation) added to give a total score ranging from 0 to 27. A score of 0 to 5 suggests no depression, 6 to 10 mild symptoms, 11 to 15 moderate symptoms, 16 to 20 severe symptoms, and 21 to 27 very severe symptoms. In Kuper et al. 2020 (n=105), the mean (±SD) QIDS self-reported score was 9.6 points (±5.0) at baseline and 7.4 (±4.5) after 10.9 months of treatment with gender-affirming hormones (no statistical analysis reported). The mean (±SD) QIDS clinician-reported score was 5.9 points (±4.1) at baseline and 6.0 (±3.8) after 10.9 months of treatment with gender-affirming hormones (no statistical analysis was reported) **(VERY LOW)**. **Participants needing treatment for depression** One observational study (Kaltiala et al. 2020) reported the proportion of participants needing treatment for depression before or during the initial assessment and during the 12-month follow-up period after starting gender-affirming hormones. In Kaltiala et al. 2020 (n=52), statistically significantly fewer participants needed treatment for depression during the 12-month 'real life' phase (15%, 8/52) compared with before or during the assessment (54%, 28/52; p<0.001). No details of what treatments for depression the participants received are reported **(VERY LOW)**.

	These studies provide very low certainty evidence that during treatment with gender-affirming hormones depression is reduced from baseline to about 12 months follow-up. However, most participants had mild symptoms at the start of treatment.
Impact on mental health: anxiety **Certainty of evidence: very low**	This is a critical outcome because anxiety may impact on social, occupational, or other areas of functioning in children and adolescents. Three observational studies (Kaltiala et al. 2020; Kuper et al. 2020; Lopez de Lara et al. 2020) provided evidence relating to the impact on anxiety in children and adolescents with gender dysphoria. **State-Trait Anxiety Inventory (STAI)** One uncontrolled, prospective, analytical study (Lopez de Lara et al. 2020) reported the change in STAI scores. STAI 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. In Lopez de Lara et al. 2020 (n=23), the mean (±SD) STAI-State subscale was statistically significantly reduced (improved) with gender-affirming hormones from 33.3 points (±9.1) at baseline to 16.8 points (±8.1) at 12 months (p<0.001). The mean STAI-Trait subscale scores also statistically significantly reduced (improved) from 33.0 points (±7.2) at baseline to 18.5 points (±8.4) at 12 months (p<0.001) **(VERY LOW)**. **Screen for Child Anxiety Related Emotional Disorders (SCARED)** One uncontrolled, prospective, longitudinal study (Kuper et al. 2020) reported anxiety symptoms using the SCARED questionnaire. Other anxiety-related symptoms using specific questions from the SCARED questionnaire were also reported: panic, generalised anxiety, social anxiety, separation anxiety and school avoidance. SCARED is a validated, 41-point questionnaire, with each item scored 0 to 2. A total score of 25 or more is suggestive of anxiety disorder, with scores above 30 being more specific. Certain scores for specific questions may indicate the presence of other anxiety-related disorders: • A score of 7 or more in questions related to panic disorder or significant somatic symptoms may indicate the presence of these. • A score of 9 or more in questions related to generalised anxiety disorder may indicate the presence of this. • A score of 5 or more in questions related to separation anxiety may indicate the presence of this. • A score of 8 or more in questions related to social anxiety disorder may indicate the presence of this. • A score of 3 or more in questions related to significant school avoidance may indicate the presence of this. In Kuper et al. 2020 (n=80 to 82, varies by outcome), small reductions were seen in anxiety, panic, generalised anxiety, social anxiety and separation anxiety and school avoidance symptoms (measured using the SCARED questionnaire) from baseline to follow-up (mean duration of treatment 10.9 months). The statistical significance of these findings are unknown as no statistical analyses were reported **(VERY LOW)**.

	Participants needing treatment for anxiety One observational study (Kaltiala et al. 2020) reported the proportion of participants needing treatment for anxiety before or during initial assessment and during the 12-month follow-up period after starting gender-affirming hormones. In Kaltiala et al. 2020 (n=52), statistically significantly fewer participants needed treatment for anxiety during the 12-month 'real life' phase (15%, 8/52) compared with before or during the assessment (48%, 25/52; p<0.001). No details of what treatments for anxiety the participants received are reported **(VERY LOW)**. **These studies provide very low certainty evidence that during treatment with gender-affirming hormones anxiety symptoms may be reduced from baseline to around 12 months follow-up.**
Impact on mental health: suicidality and self-injury **Certainty of evidence: very low**	These are critical outcomes 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. Four observational studies (Achille et al. 2020; Allen et al. 2019; Kaltiala et al. 2020; Kuper et al. 2020) provided evidence relating to suicidal ideation in children and adolescents with gender dysphoria, with an average follow-up of around 12 months. **Ask Suicide-Screening Questions (ASQ)** One uncontrolled, retrospective, longitudinal study (Allen et al. 2019) reported the change in ASQ. This is a 4-item dichotomous (yes/no) response measure designed to identify risk of suicide. The authors of Allen et al. 2019 amended 1 question in the ASQ ("*Have you ever tried to kill yourself?*") by prefacing it with "*In the past few weeks . . .*" as they were not investigating lifetime incidence. A response of 'no' is scored as 0 and a response of 'yes' is scored as 1; each item is summed to give an overall score for suicidal ideation ranging from 0 to 4. A person is considered to have screened positive if they answer 'yes' to any item with higher scores indicating higher levels of suicidal ideation. In Allen et al. 2019 (n=39), the adjusted mean (±SE) ASQ score statistically significantly reduced from 1.11 points (±0.22) at baseline to 0.27 points (±0.12) after a mean duration of treatment of about 12 months (p<0.001) **(VERY LOW)**. **PHQ 9_Modified for Teens (additional questions for suicidal ideation)** One uncontrolled, prospective, longitudinal study (Achille et al. 2020) reported the change in suicidal ideation measured using additional questions from the PHQ 9_Modified for Teens. This is a validated tool to assess depression, dysthymia and suicide risk (see above for detailed description). In addition to the 9 scored questions, the PHQ 9_Modified Teens asked 4 additional questions relating to suicidal ideation and difficulty dealing with problems of life. Responses to the PHQ 9_Modified for Teens were used to determine if the participant had suicidal ideation or not, but specific details of how this was determined are not reported.

	In Achille et al. 2020 (n=50), 10% (5/50) of participants had suicidal ideation at baseline and 6% (3/50) had suicidal ideation after about 12 months treatment with gender-affirming hormones (no statistical analysis reported) **(VERY LOW)**. **Suicidality and non-suicidal self-injury** One uncontrolled, prospective, longitudinal study (Kuper et al. 2020) reported on suicidal ideation, suicide attempts and non-suicidal self-injury, although it was unclear how and when this outcome was measured. In Kuper et al. 2020 (n=130), 25% of participants reported suicidal ideation 1 month before the initial assessment and 38% reported this during the follow-up period (no statistical analysis reported). Suicide attempts were reported in 2% of participants at 3 months before the initial assessment and 5% during follow-up. Self-injury was reported in 10% of participants at 3 months before the initial assessment and 17% during follow-up. No statistical analysis was reported for any outcomes. Mean duration of gender-affirming hormone treatment was 10.9 months **(VERY LOW)**. **Participants needing treatment for suicidality or self-harm** One observational study (Kaltiala et al. 2020) reported the proportion of participants requiring treatment for suicidality or self-harm before or during initial assessment and during the 12-month follow-up period after starting gender-affirming hormones. In Kaltiala et al. 2020 (n=52) statistically significantly fewer participants needed treatment for suicidality or self-harm during the 12-month 'real life' phase (4%, 2/52) compared with before or during the assessment (35%, 18/52; p<0.001). No details of what treatments for suicidal ideation or self-harm the participants received are reported **(VERY LOW)**. **These studies provide very low certainty evidence that gender-affirming hormones may reduce suicidality from baseline to about 12 months follow-up. However, results are inconsistent and it is difficult to draw conclusions.**
Impact on mental health: other **Certainty of evidence: very low**	This is a critical outcome because mental health problems may impact on social, occupational, or other areas of functioning in children and adolescents. One observational study (Kaltiala et al. 2020) reported the proportion of participants needing treatment for either psychotic symptoms or psychosis, substance abuse, autism, attention deficit hyperactivity disorder (ADHD) or eating disorders before or during initial assessment and during the 12-month follow-up period after starting gender-affirming hormones. In Kaltiala et al. 2020 (n=52) there was no statistically significant difference in the number of people needing treatment for either psychotic symptoms / psychosis, substance abuse, autism, attention deficit hyperactivity disorder (ADHD) or eating disorders during the 12-month 'real life' phase compared with before or during the assessment.

	No details of which specific treatments the participants received are reported **(VERY LOW)**. **This study provides very low certainty evidence on the need for treatment for either psychotic symptoms or psychosis, conduct problems or antisocial behaviour, substance abuse, autism, attention deficit hyperactivity disorder (ADHD) or eating disorders during treatment with gender-affirming hormones. No conclusions could be drawn.**
Impact on quality of life score **Certainty of evidence: very low**	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. Two uncontrolled longitudinal studies Achille et al. 2020; Allen et al. 2019) provided evidence relating to quality of life in children and adolescents with gender dysphoria. **Quality of Life Enjoyment and Satisfaction Questionnaire (QLES-Q-SF)** One uncontrolled, prospective, longitudinal study (Achille et al. 2020) reported the change in QLES-Q-SF scores from baseline to about 12 months of treatment with gender-affirming hormones. QLES-Q-SF is a validated questionnaire, consisting of 15 questions that rate quality of life on a scale of 1 (poor) to 5 (very good). In Achille et al. 2020 (n=50), the mean QLES-Q-SF score was statistically significantly reduced from baseline to about 12 months ($p<0.001$). However, absolute scores are not reported numerically **(VERY LOW)**. **General Well-Being Scale (GWBS) of the Paediatric Quality of Life Inventory** One uncontrolled, retrospective, longitudinal study (Allen et al. 2019) reported the change in adjusted mean GWBS of the Paediatric Quality of Life Inventory score from baseline to about 12 months of treatment with gender-affirming hormones. The GWBS of the Paediatric Quality of Life Inventory contains 7 items that measure two dimensions: general wellbeing (6 items) and general health (1 item). Each item is scored from 0 to 4, and the total score is linearly transformed to a 0 to 100 scale. Higher scores reflect fewer perceived problems and greater well-being. In Allen et al. 2019 (n=47), the adjusted mean (±SE) GWBS of the Paediatric Quality of Life Inventory score was statistically significantly increased (improved) from 61.70 (±2.43) points at baseline to 70.23 (±2.15) points at about 12 months ($p<0.002$) **(VERY LOW)**. **This study provides very low certainty evidence that gender-affirming hormones statistically significantly improve quality of life and well-being from baseline to 12 months follow-up.**
Important outcomes	
Impact on body image	This is an important outcome because some children and adolescents with gender dysphoria may want to take steps to suppress features of

Certainty of evidence: very low	their physical appearance associated with their sex assigned at birth or accentuate physical features of their desired gender. One uncontrolled, prospective, longitudinal study (Kuper et al. 2020) provided evidence relating to the impact on body image in children and adolescents with gender dysphoria who started treatment with gender-affirming hormones (median duration 10.9 months; range 1 to 18), measured by the change in Body Image Scale (BIS) score. BIS is a validated 30-item scale covering 3 aspects: primary, secondary and neutral body characteristics. Higher scores represent a higher degree of body dissatisfaction. In Kuper et al. 2020 (n=86), the mean (±SD) BIS score was 70.7 points (±15.2) at baseline and 51.4 points (±18.3) at follow-up (no statistical analysis reported) **(VERY LOW)**. **This study provides very low certainty evidence on the effects of gender-affirming hormones on body image during treatment with gender-affirming hormones (mean duration of treatment 10.9 months). No conclusions could be drawn.**
Psychosocial impact **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 uncontrolled, observational studies (Kaltiala et al. 2020; Lopez de Lara et al. 2020) provided evidence related to psychosocial impact in children and adolescents with gender dysphoria. **Family APGAR (Adaptability, Partnership, Growth, Affection and Resolve) test** One uncontrolled, prospective, analytical study (Lopez de Lara et al. 2020) reported the Family APGAR test. The Family APGAR test is a 5-item questionnaire, with higher scores indicating better family functioning. The authors reported the following interpretation of the test: functional, 17 to 20 points; mildly dysfunctional, 16 to 13 points; moderately dysfunctional, 12 to 10 points; severely dysfunctional, <9 points. In Lopez de Lara et al. 2020 (n=23), the mean Family APGAR test score was unchanged from baseline (17.9 points) to 12-month follow-up (18.0 points; no statistical analysis or standard deviations reported) **(VERY LOW)**. **Strengths and Difficulties Questionnaire (SDQ)** One uncontrolled, prospective, analytical study (Lopez de Lara et al. 2020) reported on behaviour using the Strengths and Difficulties Questionnaire (SDQ, Spanish version). The SDQ includes 25-items covering emotional symptoms, conduct problems, hyperactivity/ inattention, peer relationship problems and prosocial behaviour. The authors state that a score of more than 20 suggests having a behavioural disorder (normal 0 to 15, borderline 16 to 19, abnormal 20 to 40).

	In Lopez de Lara et al. 2020 (n=23), the mean (±SD) SDQ score was statistically significantly reduced (improved) from 14.7 points (±3.3) at baseline to 10.3 points (±2.9) at 12-month follow-up (p<0.001) **(VERY LOW)**. **Psychosocial functioning** One uncontrolled, retrospective chart review (Kaltiala et al. 2020) reported various markers of functioning in adolescent development, covering living arrangements, peer contacts, school or work progress, relationships, and ability to cope with matters outside the home. These measures were reported during the gender identity assessment and at about 12 months after starting gender-affirming hormones (referred to as the 'real-life phase'). In Kaltiala et al. 2020 (n=52), from the gender identity assessment to the 12-month follow-up period: • statistically significantly fewer participants were living with parents or guardians (73% versus 40%, p=0.001) • statistically significantly fewer participants had normal peer contacts (89% versus 81%, p<0.001) • there was no statistically significant difference in progress in school or work (64% versus 60%, p=0.69) • there was no statistically significant difference in the number of participants who had been dating or in steady relationships (62% versus 58%, p=0.51) • there was no statistically significant difference in the participant's ability to cope with matters outside of the home (81% versus 81%, p=1.00) **(VERY LOW)**. **These studies provide very low certainty evidence that gender-affirming hormones statistically significantly improve behavioural problems (measured by SDQ score). However, the SDQ score was in the 'normal' range at baseline and at 12-month follow up. There was no significant impact on other measures of psychosocial functioning.**
Engagement with health care services	This is an important outcome because patient engagement with health care services will impact on their clinical outcomes. No evidence was identified.
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.
De-transition	This is an important outcome because there is uncertainty about the short- and long-term safety and adverse effects of gender-affirming hormones in children and adolescents with gender dysphoria No evidence was identified.

Abbreviations: APGAR: Adaptability, Partnership, Growth, Affection and Resolve; ASQ: Ask Suicide-Screening Questions; BDI-II: Beck Depression Inventory II; BIS: Body Image Scale; CESD-R: Center for Epidemiologic Studies Depression; GWBS: General Well-Being Scale; p: p-value; PHQ 9_Modified for Teens: Patient Health Questionnaire Modified for Teens; QIDS: Quick Inventory of Depressive Symptoms; QLES-Q-SF: Quality of Life Enjoyment and Satisfaction Questionnaire; SCARED: Screen for Child Anxiety Related Emotional Disorders;

SD: standard deviation; SE: standard error; SDQ: Strengths and Difficulties Questionnaire; STAI: State-Trait Anxiety Inventory; UGDS: Utrecht Gender Dysphoria Scale.

In children and adolescents with gender dysphoria, what is the short-term and long-term safety of gender-affirming hormones 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 spine **Certainty of evidence: very low**	This is an important outcome because childhood and adolescence is a key time for bone development and gender-affirming hormones may affect bone development, as shown by changes in lumbar spine bone density. Three uncontrolled, observational studies (2 retrospective and 1 prospective) provided evidence related to bone density: lumbar spine in children and adolescents with gender dysphoria. This was reported as either bone mineral density (BMD), bone mineral apparent density (BMAD), or both. One study reported change in bone density from start of treatment with gender-affirming hormones to age 22 years (Klink et al. 2015). Two studies reported change in bone density from start of gender-affirming hormones up to 24-month follow-up (Stoffers et al. 2019 and Vlot et al. 2017). All participants had previously been treated with a GnRH analogue. All outcomes were reported separately for transfemales and transmales; also see subgroups table below. **Bone mineral apparent density (BMAD)** Two uncontrolled, observational studies reported change in lumbar BMAD (Klink et al. 2015; Vlot et al. 2017). BMAD is a size adjusted value of BMD, incorporating bone size measurements using a UK reference population of growing cis-gender adolescents (up to age 17 years). BMAD is used to correct for height and height gain and may provide a more accurate estimate of bone density in growing adolescents. BMAD was reported as g/cm^3 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. A cis-gender population was used to calculate the bone density z-score, meaning transfemales were compared with cis-males and transmales were compared with cis-females. In Klink et al. 2015 (n=34): • There was no statistically significant difference in lumbar spine BMAD z-score from starting gender-affirming hormones to age 22 years in transfemales. • The z-score for lumbar spine BMAD was statistically significantly higher at age 22 years compared with the start of gender-affirming hormones in transmales (z-score [±SD]: start of hormones -0.50 [±0.81], age 22 years -0.033 [±0.95], p=0.002).

- Actual lumbar spine BMAD values in g/cm^3 were statistically significantly higher at age 22 years compared with the start of gender-affirming hormones in transfemales and transmales **(VERY LOW)**.

In Vlot et al. 2017 (n=70):

- The z-score for lumbar spine BMAD in transfemales with a bone age of <15 years was statistically significantly higher at 24-month follow-up compared with start of gender-affirming hormones (z-score [range]: start of hormones -1.52 [-2.36 to 0.42], 24-month follow-up -1.10 [-2.44 to 0.69], p≤ 0.05). Statistically significant improvements in z-score for lumbar spine BMAD in transfemales with a bone age of ≥15 years were also seen (z-score [range]: start of hormones -1.15 [-2.21 to 0.08], 24-month follow-up -0.66 [-1.66 to 0.54], p≤ 0.05).
- The z-score for lumbar spine BMAD in transmales with a bone age of <14 years was statistically significantly higher at 24-month follow-up compared with start of gender-affirming hormones (z-score [range]: start of hormones -0.84 [-2.2 to 0.87], 24-month follow-up -0.15 [-1.38 to 0.94], p≤ 0.01). Statistically significant improvements in z-score for lumbar spine BMAD in transmales with a bone age of ≥14 years were also seen (z-score [range]: start of hormones -0.29 [-2.28 to 0.90], 24-month follow-up -0.06 [-1.75 to 1.61], p≤ 0.01).
- Actual lumbar spine BMAD values in g/cm^3 were statistically significantly higher at 24-month follow-up compared with start of gender-affirming hormones in transfemales and transmales of all bone ages **(VERY LOW)**.

Bone mineral density (BMD)

Two uncontrolled, observational studies reported change in lumbar BMD (Klink et al. 2015; Stoffers et al. 2019). BMD was determined using dual energy x-ray absorptiometry (DXA-scan; HologicQDR4500, Hologic). BMD was reported as g/cm^2 and as z-scores – see BMAD above for more details).

In Klink et al. 2015 (n=34):

- There was no statistically significant difference in lumbar spine BMD z-score from starting gender-affirming hormones to age 22 years in transfemales or transmales.
- Actual lumbar spine BMD values in g/cm^2 were statistically significantly higher at age 22 years compared with the start of gender-affirming hormones in transfemales and transmales **(VERY LOW)**.

In Stoffers et al. 2019 (n=62 at 6-month follow-up; n=15 at 24-month follow-up):

- There was no statistically significant difference in lumbar spine BMD z-score in transmales from starting gender-affirming hormones to any timepoint (6, 12 and 24 months).
- There was also no statistically significant difference in actual lumbar spine BMD values in g/cm^2 from starting gender-affirming hormones to any timepoint (6, 12 and 24 months) **(VERY LOW)**.

	These studies provide very low certainty evidence that lumber spine bone density (measured by BMAD) increases during treatment with gender-affirming hormones (from baseline to follow-up of 2 to 5 years). Z-scores at the end of follow-up suggest the average lumbar spine bone density was generally lower than the equivalent cisgender population (transfemales compared with cis-males and transmales compared with cis-females). The results for bone density (measured by BMD) were inconsistent.
Change in bone density: femoral neck **Certainty of evidence: very low**	This is an important outcome because childhood and adolescence is a key time for bone development and gender-affirming hormones may affect bone development, as shown by changes in femoral neck bone density. Three uncontrolled, observational studies (2 retrospective and 1 prospective) provided evidence related to bone density: femoral neck in children and adolescents with gender dysphoria. This was reported as either bone mineral density (BMD), bone mineral apparent density (BMAD), or both. One study reported change in bone density from start of gender-affirming hormones to age 22 years (Klink et al. 2015). Two studies reported change in bone density from start of gender-affirming hormones up to 24-month follow-up (Stoffers et al. 2019 and Vlot et al. 2017). All participants had previously been treated with a GnRH analogue. All outcomes were reported separately for transfemales and transmales; also see subgroups table below. **Bone mineral apparent density (BMAD)** Two uncontrolled, observational studies reported change in femoral neck BMAD (Klink et al. 2015; Vlot et al. 2017). See above for more details on BMAD. In Klink et al. 2015 (n=34): • The z-score for femoral neck BMAD was reported for the start of gender-affirming hormones but not at age 22 years in transfemales or transmales. No statistical analysis reported. • In transfemales there was no statistically significant difference in actual femoral neck BMAD values in g/cm^3 at age 22 years compared with start of gender-affirming hormones. In transmales actual lumbar spine BMAD values in g/cm^3 were statistically significantly higher at age 22 years compared with start of gender-affirming hormones (mean [±SD]: start of hormones 0.31 [±0.04], age 22 years 0.33 [±0.05], p=0.010) **(VERY LOW)**. In Vlot et al. 2017 (n=70): • In transfemales (all bone ages), there was no statistically significant difference in femoral neck BMAD z-score from start of gender-affirming hormones to 24-month follow-up. • The z-score for femoral neck BMAD in transmales with a bone age of <14 years was statistically significantly higher at 24-month follow-up compared with start of gender-affirming hormones (z-score [range]: start of hormones -0.37 [-2.28 to 0.47], 24-month follow-up -0.37 [-2.03 to 0.85], p≤0.01). Statistically significant improvements in z-score for lumbar spine BMAD in transmales with a bone age of ≥14 years were also

	seen (z-score [range]: start of hormones -0.27 [-1.91 to 1.29], 24-month follow-up 0.02 [-2.1 to 1.35], p≤0.05). • In transfemales of all bone ages, there was no statistically significant change in actual femoral neck BMAD values in g/cm^3 from start of gender-affirming hormones to 24-month follow-up. In transmales of all bone ages, actual femoral neck BMAD values in g/cm^3 were statistically significantly higher at 24-month follow-up compared with start of gender-affirming hormones **(VERY LOW)**. **Bone mineral density (BMD)** Two uncontrolled, observational studies reported change in femoral neck BMD (Klink et al. 2015; Stoffers et al. 2019). See above for more details on BMD. In Klink et al. 2015 (n=34): • In transfemales, there was no statistically significant difference in femoral neck BMD z-score from start of gender-affirming hormones to age 22 years. In transmales, femoral neck BMD z-score was statistically significantly higher at age 22 years compared with start of gender-affirming hormones (z-score [SD]: start of hormones -0.35 [0.79], age 22 years -0.35 [0.74], p=0.006). • Actual femoral neck BMD values in g/cm^2 were statistically significantly higher at age 22 years compared with start of gender-affirming hormones in transfemales and transmales **(VERY LOW)**. In Stoffers et al. 2019 (n=62 at 6-month follow-up; n=15 at 24-month follow-up): • there was no statistically significant difference in right or left femoral neck BMD z-score in transmales, from the start of gender-affirming hormones to any timepoint (6, 12 and 24 months). • There was also no statistically significant difference in transmales in right or left actual femoral neck BMD values in g/cm^2 from start of gender-affirming hormones to any timepoint (6, 12 and 24 months) **(VERY LOW)**. **These studies provide very low certainty evidence that during treatment with gender-affirming hormones from baseline to follow-up of 2 to 5 years, femoral neck bone density (measured by BMAD) was unchanged in transfemales but was statistically significantly increased in transmales (although the absolute change was small). Z-scores at the end of follow-up suggest that average femoral neck bone density was lower in both transfemales and transmales than in the equivalent cisgender population (transfemales compared with cis-males and transmales compared with cis-females). The results for bone density (measured by BMD) were inconsistent.**
Change in clinical parameters: glucose, insulin and HbA1c	This is an important outcome because the effect of gender-affirming hormones on insulin sensitivity and cardiovascular risk in children and adolescents with gender dysphoria is unknown.

Certainty of evidence: very low	Two uncontrolled, retrospective chart reviews (Klaver et al. 2020; Stoffers et al. 2019) provided evidence on glucose, insulin and HbA1c. All outcomes were reported separately for transfemales and transmales; also see subgroups table below. **Glucose levels, insulin levels and insulin resistance** One retrospective chart review (Klaver et al. 2020) reported non-comparative evidence on the change in glucose levels, insulin levels and insulin resistance (measured using Homeostatic Model Assessment of Insulin Resistance [HOMA-IR]) between starting gender-affirming hormones and age 22 years. In Klaver et al. 2020 (n=192): • There was no statistically significant change in glucose levels, insulin levels and insulin resistance in transfemales. • There was no statistically significant change in glucose levels in transmales. • There was a statistically significant decrease in insulin levels in transmales (mean change [95% CI] -2.1 mU/L [-3.9 to -0.3], p<0.05; mean insulin level at 22 years [95% CI] 8.6 mU/L [6.9 to 10.2]). • There was a statistically significant decrease in insulin resistance in transmales (HOMA-IR; mean change [95% CI] -0.5 [-1.0 to -0.1], p<0.05; mean HOMA-IR at 22 years [95% CI] 1.8 [1.4 to 2.2]) **(VERY LOW)**. **HbA1c** One retrospective chart review (Stoffers et al. 2019; n=62) reported non-comparative evidence on the change in HbA1c in transmales between starting gender-affirming hormones and 24-month follow-up. There was no statistically significant change in HbA1c **(VERY LOW)**. **These studies provide very low certainty evidence that gender-affirming hormones do not affect HbA1c, glucose levels, insulin levels and insulin resistance.**
Change in clinical parameters: lipids **Certainty of evidence: very low**	This is an important outcome because the effect of gender-affirming hormones on lipid profiles and cardiovascular risk in children and adolescents with gender dysphoria is unknown. One retrospective chart review (Klaver et al. 2020) provided non-comparative evidence on the change in lipids (total cholesterol, HDL cholesterol, LDL cholesterol and triglycerides) between starting gender-affirming hormones and age 22 years. All outcomes were reported separately for transfemales and transmales; also see subgroups table below. In Klaver et al. 2020 (n=192): • There was no statistically significant change in total cholesterol, HDL cholesterol and LDL cholesterol in transfemales. • There was a statistically significant decrease (improvement) in triglycerides in transfemales (mean change [95% CI] +0.2 mmol/L [0.0 to 0.5], p<0.05; mean triglyceride level at 22 years [95% CI] 1.1 mmol/L [0.9 to 1.4]). • There was a statistically significant increase in total cholesterol in transmales (mean change [95% CI] +0.4 mmol/L [0.2 to 0.6],

	p<0.001; mean total cholesterol at 22 years [95% CI] 4.6 mmol/L [4.3 to 4.8]). • There was a statistically significant decrease (worsening) in HDL cholesterol (mean change in transmales [95% CI] -0.3 mmol/L [-0.4 to -0.1], p<0.001; mean HDL cholesterol at 22 years [95% CI] 1.3 mmol/L [1.2 to 1.3]). • There was a statistically significant increase (worsening) in LDL cholesterol in transmales (mean change [95% CI] +0.4 mmol/L [0.2 to 0.6], p<0.001; mean LDL cholesterol at 22 years [95% CI] 2.6 mmol/L [2.4 to 2.8]). • There was a statistically significant increase (worsening) in triglycerides in transmales (mean change [95% CI] +0.5 mmol/L [0.3 to 0.7], p<0.001; mean triglyceride level at 22 years [95% CI] 1.3 mmol/L [1.1 to 1.5]) **(VERY LOW)**. **This study provides very low certainty evidence that gender-affirming hormones do not affect lipid profiles in transfemales. In transmales, there was a small but statistically significant worsening in cholesterol levels from start of gender-affirming hormone treatment to age 22 years, but mean cholesterol and triglyceride levels were within the UK reference range at the end of treatment.**
Change in clinical parameters: blood pressure **Certainty of evidence: very low**	This is an important outcome because the effect of gender-affirming hormones on blood pressure and cardiovascular risk in children and adolescents with gender dysphoria is unknown. One retrospective chart review (Klaver et al. 2020) provided non-comparative evidence on the change in blood pressure between starting gender-affirming hormones and at age 22 years. All outcomes were reported separately for transfemales and transmales; also see subgroups table below. In Klaver et al. 2020 (n=192): • There was no statistically significant change in systolic blood pressure (SBP) in transfemales. However, there was a statistically significant increase in diastolic blood pressure (DBP) in transfemales (mean change [95% CI] +6 mmHg [3 to 10], p<0.001; mean DBP at 22 years [95% CI] 75 [72 to 78]). • In transmales, there was a statistically significant increase in SBP (mean change [95% CI] +5 mmHg [1 to 9], p<0.05; mean SBP at 22 years [95% CI] 126 [122 to 130]), and DBP (mean change [95% CI] +6 mmHg [4 to 9], p<0.001; mean DBP at 22 years [95% CI] 74 [72 to 77]) **(VERY LOW)**. **This study provides very low certainty evidence that gender-affirming hormones statistically significantly increase blood pressure from start of treatment to age 22 years, although the absolute increase was small.**
Change in clinical parameters: body mass index (BMI)	This is an important outcome because the effect of gender-affirming hormones on weight gain and cardiovascular risk in children and adolescents with gender dysphoria is unknown. One retrospective chart review (Klaver et al. 2020) provided non-comparative evidence on the change in body mass index (BMI) between starting gender-affirming hormones and age 22 years. All

Certainty of evidence: very low	outcomes were reported separately for transfemales and transmales; also see subgroups table below. In Klaver et al. 2020 (n=192): • There was a statistically significant increase in BMI in transfemales from the start of gender-affirming hormones to age 22 years (mean change [95% CI] +1.9 [0.6 to 3.2], p<0.005; mean BMI at 22 years [95% CI] 23.2 [21.6 to 24.8]. At age 22 years, 9.9% of transfemales were obese, compared with 3.0% in a reference population of cisgender men. • There was a statistically significant increase in BMI in transmales from the start of gender-affirming hormones to age 22 years (mean change [95% CI] +1.4 [0.8 to 2.0], p<0.005; mean BMI at 22 years [95% CI] 23.9 [23.0 to 24.7]). At age 22 years, 6.6% of transmales were obese, compared with 2.2% in a reference population of cisgender women **(VERY LOW)**. **This study provides very low certainty evidence that gender-affirming hormones statistically significantly increase BMI from start of treatment to age 22 years, although most participants were within the healthy weight range.**
Change in clinical parameters: 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, gender-affirming hormones may need to be stopped. One retrospective chart review (Stoffers et al. 2019) provided non-comparative evidence on the change in liver enzymes in transmales between starting gender-affirming hormones and up to 24-months follow-up. In Stoffers et al. 2019 (n=62): • There was no statistically significant change in aspartate aminotransferase (AST), alanine aminotransferase (ALT) and gamma-glutamyltransferase (GCT) in transmales. • There was a statistically significant increase in alkaline phosphatase (ALP) levels from starting gender-affirming hormones to 6- and 12-months follow-up, although by 24-months the difference was not statistically significant (median [IQR]: start of hormones 102 [78 to 136], 6-month follow-up 115 [102 to 147] p<0.001, 12-month follow-up 112 [88 to 143] p<0.001) **(VERY LOW)**. **This study provides very low certainty evidence that gender-affirming hormones do not affect liver function in transmales from baseline to 24 months follow-up.**
Change in clinical parameters: kidney function **Certainty of evidence: very low**	This is an important outcome because if renal damage (raised serum creatinine and urea are markers of this) is suspected, treatment with gender-affirming hormones may need to be stopped. One retrospective chart review (Stoffers et al. 2019) provided non-comparative evidence on the change in serum creatinine and serum urea levels in transmales between starting gender-affirming hormones and up to 24-months follow-up. In Stoffers et al. 2019 (n=62):

	• There was a statistically significant increase in creatinine levels in transmales at all timepoints up to 24 months (mean [SD]: start of hormones 62 umol/L [7], 6 months 70 umol/L [9] , 12 months 74 umol/L [10], 24 months 81 umol/L [10], p<0.001). • There was no statistically significant change in urea in transmales (follow-up duration not reported) **(VERY LOW)**. **This study provides very low certainty evidence on the effects of gender-affirming hormones on kidney function in transmales from baseline to 24 months follow-up. A statistically significant increase in creatinine levels was seen, but these were within the UK reference range. Urea levels were unchanged.**

Treatment discontinuation
 	
Certainty of evidence: very low
	This is an important outcome because there is uncertainty about the short- and long-term impact of stopping treatment with gender-affirming hormones in children and adolescents with gender dysphoria.
	
	One uncontrolled, retrospective chart review (Khatchadourian et al. 2014) provided evidence relating to permanent or temporary treatment discontinuation in children and adolescents with gender dysphoria.
	
	Khatchadourian et al. 2014 narratively reported treatment discontinuation in a cohort of 63 adolescents (24 transfemales and 39 transmales) who received gender-affirming hormones:
	• No participants permanently discontinued gender-affirming hormones.
	• No transfemales temporarily discontinued gender-affirming hormones.
	• Three transmales temporarily discontinued gender-affirming hormones due to:
	  ○ mental health comorbidities (n=2)
	  ○ androgenic alopecia (n=1).
	All 3 participants eventually resumed treatment, although timescales were not reported **(VERY LOW)**.
	
	This study provides very low certainty evidence that the rates of discontinuation during treatment with gender-affirming hormones are low (duration of treatment not reported).

Adverse effects
 	
Certainty of evidence: very low
	This is an important outcome because if there are adverse effects, gender-affirming hormones may need to be stopped.
	
	One uncontrolled, retrospective chart review (Khatchadourian et al. 2014) provided evidence relating to adverse effects from gender-affirming hormones in children and adolescents with gender dysphoria.
	
	Khatchadourian et al. 2014 narratively reported adverse effects in a cohort of 63 adolescents (24 transfemales and 39 transmales) receiving treatment with gender-affirming hormones:
	• No severe complications were reported.
	• No transfemales reported minor complications.
	• Twelve transmales developed minor complications, which were:
	  ○ severe acne, requiring isotretinoin treatment (n=7)
	  ○ androgenic alopecia (n=1)
	  ○ mild dyslipidaemia (further details not provided; n=3)
	  ○ significant mood swings (n=1) **(VERY LOW)**.

	This study provides very low certainty evidence about the potential adverse effects of gender-affirming hormones (duration of treatment not reported). No conclusions could be drawn.

Abbreviations: ALP: alkaline phosphatase; ALT: alanine aminotransferase; AST: aspartate aminotransferase; BMAD: bone mineral apparent density; BMD: bone mineral density; BMI: body mass index; DBP: diastolic blood pressure; GGT: gamma-glutamyl transferase; HbA1c: glycated haemoglobin; HDL: high-density lipoproteins; HOMA-IR: Homeostatic Model Assessment of Insulin Resistance; IQR: interquartile range; LDL: low-density lipoproteins; p: p-value; SBP: systolic blood pressure; SD: standard deviation.

In children and adolescents with gender dysphoria, what is the cost-effectiveness of gender-affirming hormones 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 gender-affirming hormones 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 gender-affirming hormones 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 mental health: depression and anxiety** One uncontrolled, prospective, longitudinal study (Kuper et al. 2020) reported the change in depression (measured using QIDS clinician-reported and self-reported), anxiety and anxiety-related symptoms (measured using SCARED) in transfemales. See the clinical effectiveness results above for full details. In Kuper et al. 2020 (n=33 to 45, varies by outcome), changes were seen in depression, anxiety and anxiety-related symptoms from baseline to follow-up but the authors did not report any statistical analyses, so it is unclear if was any changes were statistically significant **(VERY LOW)**. **This study provides very low certainty evidence on the effects of gender-affirming hormones on depression, anxiety and anxiety-related symptoms over time in sex assigned at birth males (transfemales; mean duration of treatment 10.9 months). No conclusions could be drawn.** **Impact on mental health: suicidality**

	One uncontrolled, retrospective, longitudinal study (Allen et al. 2019) reported the change in Ask Suicide-Screening Questions (ASQ) in transfemales compared with transmales. See the clinical effectiveness results above for full details. Between baseline and the final assessment, there was no statistically significant difference in change in ASQ score for transfemales compared with transmales (p=0.79; n=47) **(VERY LOW)**. One uncontrolled, prospective, longitudinal study (Achille et al. 2020) reported the change in suicidal ideation in transfemales measured using additional questions from the PHQ 9_Modified for Teens. See the clinical effectiveness results above for full details. At baseline, 11.8% (2/17) of transfemales had suicidal ideation, compared with 5.9% (1/17) at about 12-months follow-up (no statistical analysis reported) **(VERY LOW)**. **These studies provide very low certainty evidence that any change in suicidal ideation is not different between sex assigned at birth males (transfemales) and sex assigned at birth females (transmales) from baseline to follow-up of about 12 months.** **Impact on quality of life** One uncontrolled, retrospective, longitudinal study (Allen et al. 2019) reported the change in the GWBS of the Paediatric Quality of Life Inventory in transfemales compared with transmales. See the clinical effectiveness results above for full details. Between baseline and final assessment, there was no statistically significant difference in change in GWBS of the Paediatric Quality of Life Inventory for transfemales compared with transmales (p=0.32; n=47) **(VERY LOW)**. **This study provides very low certainty evidence that any change in general wellbeing is not different between sex assigned at birth males (transfemales) and sex assigned at birth females (transmales) from baseline to follow-up of about 12 months.** **Impact on body image** One uncontrolled, prospective, longitudinal study (Kuper et al. 2020) reported change in Body Image Scale (BIS) in transfemales. See the clinical effectiveness results above for full details. In Kuper et al. 2020 (n=30), the mean (±SD) BIS score was 67.5 points (±19.5) at baseline and 49.0 points (±21.6) at follow-up (no statistical analysis reported) **(VERY LOW).** **This study provides very low certainty evidence on the effects of gender-affirming hormones on body image over time in transfemales (mean duration of treatment 10.9 months). No conclusions could be drawn.** **Change in bone density: lumbar spine**

	Two uncontrolled, observational, retrospective studies provided evidence relating to the effect of gender-affirming hormones on lumber spine bone density in transfemales (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 lumbar spine bone density (measured by BMAD) increases during treatment with gender-affirming hormones in sex assigned at birth males (transfemales). Z-scores at the end of follow-up suggest average lumbar spine bone density was generally lower than in the equivalent cisgender population. The results for lumbar spine bone density (measured by BMD) were inconsistent.** **Change in bone density: femoral neck** Two uncontrolled, observational, retrospective studies provided evidence relating to the effect of gender-affirming hormones on femoral neck bone density in transfemales (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 femoral neck bone density (measured by BMAD) was unchanged in sex assigned at birth males (transfemales) during treatment with gender-affirming hormones (follow-up between 2 and 5 years). Z-scores at the end of follow-up suggest and the average femoral neck bone density was lower than in the equivalent cisgender population. The results for femoral neck bone density (measured by BMD) were inconsistent.** **Change in clinical parameters: glucose, insulin and HbA1c** One uncontrolled, retrospective chart review (Klaver et al. 2020) provided evidence on glucose, insulin and HbA1c in transfemales. See the safety results table above for a full description of the results. **This study provided very low certainty evidence that gender-affirming hormones do not affect HbA1c, glucose levels, insulin levels and insulin resistance in sex assigned at birth males (transfemales) from the start of treatment to age 22 years.** **Change in clinical parameters: lipids** One retrospective chart review (Klaver et al. 2020) provided evidence on the change in lipids (total cholesterol, HDL cholesterol, LDL cholesterol and triglycerides) in transfemales. See the safety results table above for a full description of the results. **This study provides very low certainty evidence that gender-affirming hormones do not affect lipid profiles in sex assigned at birth males (transfemales) from the start of treatment to age 22 years.** **Change in clinical parameters: blood pressure**

	One retrospective chart review (Klaver et al. 2020) provided evidence on the change in blood pressure in transfemales. See the safety results table above for a full description of the results. **This study provides very low certainty evidence that gender-affirming hormones statistically significantly increase blood pressure in sex assigned at birth males (transfemales), although the absolute increase was small from the start of treatment to age 22 years.** **Change in clinical parameters: body mass index (BMI)** One retrospective chart review (Klaver et al. 2020) provided evidence on the change in BMI in transfemales. See the safety results table above for a full description of the results. **This study provides very low certainty evidence that gender-affirming hormones statistically significantly increase BMI in sex assigned at birth males (transfemales), although most participants were within the healthy weight range from the start of treatment to age 22 years.** **Treatment discontinuation** One uncontrolled, retrospective chart review provided evidence relating to permanent or temporary discontinuation of gender-affirming hormones in transfemales (Khatchadourian et al. 2014). **This study provides very low certainty evidence that the rates of discontinuation during treatment with gender-affirming hormones in sex assigned at birth males (transfemales) are low. Duration of treatment with gender-affirming hormones was not reported.** **Adverse effects** One uncontrolled, retrospective chart review provided evidence relating to adverse effects from gender-affirming hormones in transfemales (Khatchadourian et al. 2014). **This study provides very low certainty evidence about the potential adverse effects of gender-affirming hormones in sex assigned at birth males (transfemales). No conclusions could be drawn. Duration of treatment with gender-affirming hormones was not reported.**
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 mental health: depression and anxiety** One uncontrolled, prospective, longitudinal study (Kuper et al. 2020) reported the change in depression (measured using QIDS clinician-reported and self-reported), anxiety and anxiety-related symptoms (measured using SCARED) in transmales. See the clinical effectiveness results above for full details. In Kuper et al. 2020 (n=65 to 78, varies by outcome), changes were seen in depression, anxiety and anxiety-related symptoms from

	baseline to follow-up but the authors did not report any statistical analysis, so it is unclear if any changes are statistically significant **(VERY LOW)**. **This study provides very low certainty evidence on the effects of gender-affirming hormones on depression, anxiety and anxiety-related symptoms over 10.9 months in transmales. No conclusions could be drawn.** **Impact on mental health: suicidality** One uncontrolled, retrospective, longitudinal study (Allen et al. 2019) reported the change in Ask Suicide-Screening Questions (ASQ) in transmales compared with transfemales. See the sex assigned at birth males (transfemales) row above for full details of the results. One uncontrolled, prospective, longitudinal study (Achille et al. 2020) reported the change in suicidal ideation in transmales measured using additional questions from the PHQ 9_Modified for Teens. See the clinical effectiveness results above for full details. At baseline, 9.1% (3/33) of transmales had suicidal ideation, compared with 6.1% (2/33) at about 12-months follow-up (no statistical analysis reported) **(VERY LOW)**. **These studies provide very low certainty evidence that any change in suicidal ideation is not different between sex assigned at birth females (transmales) and sex assigned at birth males (transfemales). Mean duration of treatment about 12 months.** **Impact on quality of life** One uncontrolled, retrospective, longitudinal study (Allen et al. 2019) reported the change in the GWBS of the Paediatric Quality of Life Inventory in transmales compared with transfemales. See the sex assigned at birth males (transfemales) row above for full details of the results. **This study provides very low certainty evidence that any change in general wellbeing is not different between sex assigned at birth females (transmales) and sex assigned at birth males (transfemales). Mean duration of treatment about 12 months.** **Impact on body image** One uncontrolled, prospective, longitudinal study (Kuper et al. 2020) reported change in Body Image Scale (BIS) in transmales. See the clinical effectiveness results above for full details. In Kuper et al. 2020 (n=66), the mean (±SD) BIS score was 71.1 points (±13.4) at baseline and 52.9 points (±16.8) at follow-up (no statistical analysis reported) **(VERY LOW)**. **This study provides very low certainty evidence on the effects of gender-affirming hormones on body image over 10.9 months in transmales. No conclusions could be drawn.** **Change in bone density: lumbar spine**

	Three uncontrolled, observational, retrospective studies provided evidence relating to the effect of gender-affirming hormones on lumber spine bone density in transmales (Klink et al. 2015, Stoffers et al. 2019 and Vlot et al. 2017). See the safety results table above for a full details of the results. **These studies provide very low certainty evidence that lumbar spine bone density (measured by BMAD) increases during 2 to 5 years treatment with gender-affirming hormones in sex assigned at birth females (transmales). Z-scores at the end of follow-up suggest the average lumbar spine bone density was generally lower than in the equivalent cisgender population. The results for lumbar spine bone density (measured by BMD) were inconsistent.** **Change in bone density: femoral neck** Three uncontrolled, observational, retrospective studies provided evidence relating to the effect of gender-affirming hormones on femoral neck bone density in transmales (Klink et al. 2015, Stoffers et al. 2019 and Vlot et al. 2017). See the safety results table above for a full details of the results. **These studies provide very low certainty evidence that femoral neck bone density (measured by BMAD) statistically significantly increased in sex assigned at birth females (transmales) during 2 to 5 years treatment with gender-affirming hormones. Z-scores at the end of follow-up suggest the average femoral neck bone density was generally lower than in the equivalent cisgender population. The results for femoral neck bone density (measured by BMD) were inconsistent.** **Change in clinical parameters: glucose, insulin and HbA1c** Two uncontrolled, retrospective chart reviews (Klaver et al. 2020; Stoffers et al. 2019) provided evidence on glucose, insulin and HbA1c in transmales. See the safety results table above for full details of the results. **This study provided very low certainty evidence that gender-affirming hormones do not affect HbA1c, glucose levels, insulin levels and insulin resistance in sex assigned at birth females (transmales). Reported from start of treatment to age 22 years.** **Change in clinical parameters: lipids** One retrospective chart review (Klaver et al. 2020) provided evidence on the change in lipids (total cholesterol, HDL cholesterol, LDL cholesterol and triglycerides) in transmales. See the safety results table above for full details of the results. **This study provides very low certainty evidence that treatment with gender-affirming hormones is associated with a small but statistically significant worsening of cholesterol levels in sex assigned at birth females (transmales), but mean cholesterol and triglyceride levels were within the UK reference range at end of treatment, from start of treatment to age 22 years.**

	Change in clinical parameters: blood pressure One retrospective chart review (Klaver et al. 2020) provided evidence on the change in blood pressure in transmales. See the safety results table above for full details of the results. **This study provides very low certainty evidence that gender-affirming hormones statistically significantly increase blood pressure in sex assigned at birth females (transmales), although the absolute increase was small, from start of treatment to age 22 years.** **Change in clinical parameters: body mass index (BMI)** One retrospective chart review (Klaver et al. 2020) provided evidence on the change in body mass index (BMI) in transmales. See the safety results table above for full details of the results. **This study provides very low certainty evidence that gender-affirming hormones statistically significantly increase BMI in sex assigned at birth females (transmales), although most participants were within the healthy weight range, from start of treatment to age 22 years.** **Change in clinical parameters: liver function** One retrospective chart review (Stoffers et al. 2019) provided non-comparative evidence on the change in liver enzymes in transmales between starting gender-affirming hormones and up to 24-months follow-up. See the safety results table above for full details of the results. **This study provides very low certainty evidence that gender-affirming hormones for about 12 months do not affect liver function in sex assigned at birth females (transmales).** **Change in clinical parameters: kidney function** One retrospective chart review (Stoffers et al. 2019) provided non-comparative evidence on the change in serum creatinine and serum urea levels in transmales between starting gender-affirming hormones and up to 24-months follow-up. See the safety results table above for full details of the results. **This study provides very low certainty evidence on the effects of gender-affirming hormones on kidney function in sex assigned at birth females (transmales). A statistically significant increase in creatinine levels was seen at about 12 months follow-up, but these were within the UK reference range. Urea levels were unchanged.** **Treatment discontinuation** One uncontrolled, retrospective chart review provided evidence relating to permanent or temporary discontinuation of gender-affirming hormones in transmales (Khatchadourian et al. 2014). See the safety results table above for full details of the results. **This study provides very low certainty evidence that the rates of treatment discontinuation with gender-affirming hormones in sex**

	assigned at birth females (transmales) is low. Duration of gender-affirming hormones not reported. **Adverse effects** One uncontrolled, retrospective chart review provided evidence for adverse effects from gender-affirming hormones in transmales (Khatchadourian et al. 2014). See the safety results table above for full details of the results. **This study provides very low certainty evidence about the potential adverse effects of gender-affirming hormones in sex assigned at birth females (transmales). No conclusions could be drawn. Duration of gender-affirming hormones not reported.**
Duration of gender dysphoria	No evidence was identified.
Age at onset of gender dysphoria	No evidence was identified.
Age at onset of puberty	No evidence was identified.
Tanner stage at which GnRH analogue or gender-affirming hormones started	One uncontrolled, prospective, longitudinal study (Kuper et al. 2020) reported the impact of Tanner stage on outcomes, although it is not clear whether this is referring to Tanner stage at initial assessment, at the start of GnRH analogues or at another timepoint.
Diagnosis of autistic spectrum disorder	No evidence was identified.
Diagnosis of a mental health condition	One uncontrolled, prospective, longitudinal study (Achille et al. 2020) reported outcomes that were adjusted for engagement in counselling and medicines for mental health problems. Information about diagnoses and treatment were not provided. Rates of mental health issues appear to be high in the cohort. **Impact on mental health** Achille et al. 2020 reported the change in depression scores, controlled for engagement in counselling and medicines for mental health problems (measured using the Center for Epidemiologic Studies Depression [CESD-R] scale and Patient Health Questionnaire Modified for Teens [PHQ 9_Modified for Teens] score: • There was no statistically significant change in CESD-R from baseline to about 12-months follow-up. • There was no statistically significant change in PHQ 9_Modified for Teens score from baseline to about 12-months follow-up **(VERY LOW)**. **Impact on quality of life** Achille et al. 2020 reported the change in quality of life scores, controlled for engagement in counselling and medicines for mental health problems (measured using the Quality of Life Enjoyment and Satisfaction Questionnaire [QLES-Q-SF] score: • There was no statistically significant change in QLES-Q-SF score from baseline to about 12-months follow-up **(VERY LOW)**.

	This study provides very low certainty evidence about outcomes that were adjusted for engagement in counselling and medicines for mental health problems. No conclusions could be drawn.

Abbreviations: ASQ: Ask Suicide-Screening Questions; CESD-R: Center for Epidemiologic Studies Depression; GnRH: Gonadotrophin releasing hormone; GWBS: General Well-Being Scale; HDL: high-density lipoproteins; LDL: low-density lipoproteins; p: p-value; PHQ 9_Modified for Teens: Patient Health Questionnaire Modified for Teens; QLES-Q-SF: Quality of Life Enjoyment and Satisfaction Questionnaire.

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 gender-affirming hormones?

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

<table>
<tr><th>Outcome</th><th>Evidence statement</th></tr>
<tr><td>Diagnostic criteria</td><td>The DSM-IV-TR criteria was used in 3 studies (Klaver et al. 2020, Klink et al. 2015 and Vlot et al. 2017).

The DSM-V criteria was used in 2 studies (Kuper et al. 2020 and Stoffers et al. 2019). 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 think of themselves which must have lasted at least 6 months.

The ICD-10 diagnosis of ‘transsexualism’ was used in 1 study (Kaltiala et al. 2020). The authors state that this is the corresponding diagnosis to ‘gender dysphoria’ in the DSM-V, and that diagnostic assessments in the study location (Finland) take place according to ICD-10.

It was not reported how gender dysphoria was defined in the remaining 4 studies (VERY LOW).

From the evidence selected, the most commonly reported diagnostic criteria for gender dysphoria (5/10 studies) was the DSM criteria in use at the time the study was conducted.</td></tr>
<tr><td>Age when gender-affirming hormones started</td><td>8/10 studies reported the age at which participants started treatment with gender-affirming hormones, either as the mean age (with SD) or median age (with the range):

<table>
<tr><th>Study</th><th>Mean age (± SD)</th></tr>
<tr><td>Allen et al. 2019</td><td>16.7 years (not reported)</td></tr>
<tr><td>Khatchadourian et al. 2014</td><td>17.4 years (1.9)</td></tr>
<tr><td>Klaver et al. 2020</td><td>16.4 years (1.1) in transfemales
16.9 years (0.9) in transmales</td></tr>
<tr><td>Kuper et al. 2020</td><td>16.2 (1.2)</td></tr>
<tr><td>Klink et al. 2015</td><td>16.6 years (1.4) in transfemales
16.4 years (2.3) in transmales</td></tr>
</table>
</td></tr>
</table>

<table>
<tr><td></td><td>
<table>
<tr><th>Study</th><th>Median age (range)</th></tr>
<tr><td>Stoffers et al. 2019</td><td>17.2 years (15 to 19.5)</td></tr>
<tr><td>Vlot et al. 2017</td><td>16.3 years (15.9 to 19.5) in transfemales
16.0 years (14.0 to 18.9) in transmales</td></tr>
</table>

Age at the start of treatment was not reported in 3 studies:
• In Achille et al. 2020 the mean age at initial assessment (baseline) was 16.2 years (SD ±2.2)
• In Kaltiala et al. 2020 the mean age at diagnosis was 18.1 years (range 15.2 to 19.9)
• In Lopez de Lara et al. 2020 the mean age of participants was 16 years (range 14 to 18), although it is not clear if this is at the initial assessment or at the start of gender-affirming hormones.

The evidence included showed that most children and adolescents started treatment with gender-affirming hormones at about 16 to 17 years, with a range of about 14 to 19 years.
</td></tr>
<tr><td>Duration of treatment with GnRH analogues</td><td>
The duration of treatment with GnRH analogues was reported in 3/10 studies:

<table>
<tr><th>Study</th><th>Median duration</th></tr>
<tr><td>Klaver et al. 2020</td><td>2.1 years (IQR 1.0 to 2.7) in transfemales
1.0 years (IQR 0.5 to 2.9) in transmales</td></tr>
<tr><td>Klink et al. 2015</td><td>1.3 years (range 0.5 to 3.8) in transfemales
1.5 years (range 0.25 to 5.2) in transmales
(GnRH analogue monotherapy)</td></tr>
<tr><td>Stoffers et al. 2019</td><td>8 months (range 3 to 39)</td></tr>
</table>

The evidence included showed wide variation in the duration of treatment with gender-affirming hormones, but most studies did not report this information. Treatment duration ranged from a few months up to about 5 years.
</td></tr>
</table>

Abbreviations: DSM, Diagnostic and Statistical Manual of Mental Disorders criteria; GnRH, Gonadotrophin-releasing hormone; ICD, International Statistical Classification of Diseases and Related Health Problems; IQR, interquartile range; SD, standard deviation.

6. Discussion

A key limitation to identifying the effectiveness and safety of gender-affirming hormones for children and adolescents with gender dysphoria is the lack of reliable comparative studies. All the studies included in this evidence review are uncontrolled observational studies, which are subject to bias and confounding and were of very low certainty using modified GRADE. 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 gender-affirming hormones, 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. A fundamental limitation of all the uncontrolled studies included in this review is that any changes in scores from baseline to follow-up could be attributed to a regression-to-the-mean.

The included studies have relatively short follow-up, with an average duration of treatment with gender-affirming hormones between around 1 year and 5.8 years. Further studies with a longer follow-up are needed to determine the long-term effect of gender-affirming hormones for children and adolescents with gender dysphoria.

Most studies included in this review did not report comorbidities (physical or mental health) and no study reported concomitant treatments in detail. Because of this it is not clear whether any changes observed were due to gender-affirming hormones or other treatments the participants may have received. For example, we do not know if any improvement in depression symptom score over time was the result of gender-affirming hormones or the mental health support the person may be receiving (including medicines or counselling). This may be of particular importance for the mental health outcomes discussed in this review, since depression, anxiety and other related symptoms are common in children and adolescents with gender dysphoria. In Achille et al. 2020, at baseline around one-third of participants were taking medicines for mental health problems and around two-thirds reported being depressed in the past year. In Kaltiala et al. 2020, half the participants needed mental health treatment during and before gender identity assessment, with the most common reasons for treatment being depression, anxiety and suicidality. Only 1 study reported outcomes adjusted for engagement in counselling and medicines for mental health problems (Achille et al. 2020). This study found that gender-affirming hormones had no significant impact on depression and quality of life when adjusted for mental health care, despite significant approvements reported for the unadjusted results. However, it is not possible to draw conclusions on the impact of concurrent mental health treatment on the effect of gender-affirming hormones based on this study alone. Details of the mental health care provided are not reported in the study and results are presented for transfemales and transmales separately, resulting in small patient numbers and possible underpowering.

In most of the included studies, details of the gender-affirming hormone treatment regimens are poorly reported, with limited information provided about the medicines, doses and routes of administration used. It is not clear whether the interventions used in the studies are reflective of current UK practice for children and adolescents with gender dysphoria. There is also the suggestion that the hormone dose used in 1 study may have been too low; the authors of Klink et al. 2015 suggest that the relatively low initial dose of oestrogen for transfemales may be the reason for the observed lack of effect on lumber spine bone density. Duration of treatment with a GnRH analogue is also poorly reported and is only stated in 3/10 studies.

There is a degree of indirectness in some studies, with some participants included that fall outside of the population of this evidence review. For example, in Kuper et al. 2020 17% of participants received puberty suppression alone, and in Achille et al. 2020, 30% of participants received no treatment or puberty suppression alone. Some results and statistical analyses are only reported for the whole cohort in these studies and not the subgroup of participants who received gender-affirming hormones.

Participant numbers are poorly reported in some of the included studies. In Achille et al. 2020, 47% (45/95) of the people who entered the study did not have follow-up data and were excluded from the analyses, with no explanation or description of those people lost to follow-up. In Kuper et al. 2020, the number of participants varied by outcome, with less than

two-thirds of participants providing data for some outcomes. The authors provide no explanation for this incomplete reporting.

It is not clear whether some outcome measures, specifically those related to psychosocial functioning, are relevant to the UK population. In Kaltiala et al. 2020, an observational study conducted in Finland, the proportion of participants living with parents or guardians is reported as marker of appropriate functioning. The authors state that in Finnish culture young people tend to leave the parental home early, with only around one-quarter of 20 to 24 year olds still living at home. This is lower than in the UK, where around half of 20 to 24 year olds live with their parents or guardians (ONS: Why are more young people living with their parents?).

It is difficult to draw firm conclusions for many of the effectiveness and safety outcomes reported in the included studies because many different scoring tools and methods were used to assess the same outcome, often with conflicting results. For example, bone density is reported as bone mineral density (BMD) and bone mineral apparent density (BMAD) in the same study, the latter being a size-adjusted measure often useful for people whose bones are still growing. For some populations (transfemale versus transmale) and bone regions (lumber spine versus femoral neck), statistically significant differences in BMD are reported but not for BMAD, and vice versa.

In addition to this, most outcomes reported across the included studies do not have an accepted minimal clinically important difference (MCID), making it difficult the determine whether any observed statistically significant changes are clinically meaningful. However, the authors of some studies report thresholds to interpret the results of the scoring tools, so some conclusions can be made. For example, the mean Utrecht Gender Dysphoria Scale (UGDS) score (a measure of gender dysphoria symptoms) reduced to about 15 points after treatment with gender-affirming hormones (Lopez de Lara et al. 2020). The authors state that scores of 40 points or above signify gender dysphoria, suggesting that after about 12 months of treatment with gender-affirming hormones, the majority of participants did not have symptoms of gender dysphoria.

The impact of gender-affirming hormones on bone density was reported in 3 studies (Klink et al. 2015, Stoffers et al. 2019 and Vlot et al. 2017). Although these studies did not include a control group, comparisons to a reference population are reported using z-scores. Comparisons were made to a cisgender population, meaning for example that bone density in transfemales was compared with bone density in cisgender males. The authors of Klink et al. 2015 note that this may not be the ideal comparison, because androgens and oestrogens affect bone differently, and that bone properties in a trans population differ from their age- and sex assigned at birth-matched controls. Beyond this, a major limitation when trying to determine the impact of gender-affirming hormones on the short- and long-term bone health of children and adolescents is the lack of data on fracture rates and other patient-orientated outcomes, including rates of osteoporosis. Studies of GnRH analogues in children and adolescents with gender dysphoria suggest that GnRH analogue treatment may reduce the expected increase in bone density (which is expected during puberty). Although improvements in bone density were reported following treatment with gender-affirming hormones, Z-scores suggest that bone density remained lower in transfemales and transmales compared with an equivalent cisgender population.

One study reported on cardiovascular risk factors at age 22 years in people who started gender-affirming hormones for gender dysphoria as adolescents. While glucose levels, insulin levels and insulin resistance were broadly unchanged at 22 years, statistically significant increases in blood pressure and body mass index were seen. A small but statistically significant worsening of the lipid profile in transmales who received testosterone was also seen at age 22 years. However, further studies with a considerably longer follow-up and a focus on patient-oriented outcomes, including cardiovascular events and mortality are needed to determine the long-term impact on cardiovascular health of starting gender-affirming hormones during childhood and adolescence.

Only 1 study reported adverse events and discontinuation rates with gender-affirming hormones in children and adolescents. Conclusions on these outcomes cannot be made based on this study alone.

This review did not identify sub-groups of people who may benefit more from gender-affirming hormones. Limited evidence from 2 studies suggests there was no difference in response to treatment between transfemales and transmales for mental health and quality of life (Achille et al. 2020 and Allen et al. 2019).

7. Conclusion

This evidence review found limited evidence for the effectiveness and safety of gender-affirming hormones in children and adolescents with gender dysphoria, with all studies being uncontrolled, observational studies, and all outcomes of very low certainty. Any potential benefits of treatment must be weighed against the largely unknown long-term safety profile of these treatments.

The results from 5 uncontrolled, observational studies (Achille et al. 2020, Allen et al. 2019, Kaltiala et al. 2020. Kuper et al. 2020, Lopez de Lara et al. 2020) suggest that, in children and adolescents with gender dysphoria, gender-affirming hormones are likely to improve symptoms of gender dysphoria, and may also improve depression, anxiety, quality of life, suicidality, and psychosocial functioning. The impact of treatment on body image is unclear. All results were of very low certainty. The clinical relevance of any improvements to the person is difficult to determine because most outcomes do not have a recognised minimal clinically important difference, and the authors do not present statistical analysis for some outcomes.

A further 5 uncontrolled, observational studies (Khatchadourian et al. 2014, Klaver et al. 2020, Klink et al. 2015, Stoffers et al. 2019 and Vlot et al. 2017) reported on safety outcomes, all of which provided very low certainty evidence. Statistically significant increases in some measures of bone density were seen following treatment with gender-affirming hormones, although results varied by bone region (lumber spine versus femoral neck) and by population (transfemales versus transmales). However, z-scores suggest that bone density remained lower in transfemales and transmales compared with an equivalent cisgender population. Results from 1 study of gender-affirming hormones started during adolescence reported statistically significant increases in blood pressure and body mass index, and worsening of the lipid profile (in transmales) at age 22 years, although longer term studies that report on cardiovascular event rates are needed. Adverse events and discontinuation rates associated with gender-affirming hormones were only reported in 1 study, and no conclusions can be made on these outcomes.

This review did not identify sub-groups of people who may benefit more from gender-affirming hormones. Limited evidence from 2 studies suggests there was no difference in response to treatment between transfemales and transmales for mental health and quality of life (Achille et al. 2020 and Allen et al. 2019).

No cost-effectiveness evidence was found to determine whether gender-affirming hormones are a cost-effective treatment for children and adolescents with gender dysphoria.

Appendix A PICO

The review questions for this evidence review are:

1. For children and adolescents with gender dysphoria, what is the clinical effectiveness of treatment with gender-affirming hormones 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 gender-affirming hormones 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 gender-affirming hormones 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 particular sub-groups of children and adolescents with gender dysphoria that derive comparatively more (or less) benefit from treatment with gender-affirming hormones 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 gender-affirming hormones?
   (c) what was the duration of GnRH analogues treatment?

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 the 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 at which treatment was initiated with GnRH analogues and with gender-affirming hormones. • The age of onset of gender dysphoria • The age of onset of puberty • Adolescents with gender dysphoria who have a pre-existing diagnosis of autistic spectrum disorder. • 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), psychosis, personality disorder, Attention Deficit Hyperactivity Disorder and eating disorders.
I – Intervention	Gender-affirming hormone treatments: • A testosterone preparation for sex assigned at birth female patients which may include testosterone in the form of Sustanon injections*; testosterone enantate injections; Tostran gel*; Testogel; Testim gel; oral testosterone capsules in the form of testosterone undecanoate (Restandol); Andriol testocaps; Nebido • An oestradiol preparation** for sex assigned at birth male patients which may include: oral estradiol valerate*; oestrogen patches (7β-oestradiol patches e.g. Evorel or Estradem); Estradot patches; ethinyloestradiol *** *These are the used by Leeds Hospital, England. ** Be aware that the American spelling is oestrogen without the ‘o’. ***Ethinyloestradiol is rarely used.
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 gender-affirming hormones 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, suicide, 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 measure. • **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 measures. *Important to decision making* • **Impact on body image** This outcome is important because some young people with gender dysphoria 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 experienced 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 and

	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 in adulthood proceed to transitioning surgery. Stated measures of the extent of surgery and satisfaction with surgery in studies may be reported. • **De-transition** The proportion of patients who de-transition following the commencement of gender-affirming hormone treatment 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 gender-affirming hormones in children and adolescents with gender dysphoria. **B: Safety** • Short and long -term safety and adverse effects of taking gender-affirming hormones is important to assess whether treatment causes acute side effects that may lead to withdrawing the treatment or long term effects that may impact on decisions for transitioning or de-transitioning. Aspects to be reported on should include Impact of the drug use such as clinically relevant derangement in renal and liver function tests, lipids, glucose, insulin and glycosylated haemoglobin, cognitive development and functioning. The clinical and physical impact of temporary and permanent withdrawal the drug such as when patients decide to de-transition – e.g. delay in the attainment of peak bone mass, attenuation of peak bone mass, permanent physical effects. **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 21 July 2020, limiting the search to papers published in English language in the last 20 years. Conference abstracts, non-systematic reviews, narrative reviews, commentaries, letters, editorials, guidelines, pre-publication prints, case reports and resource utilisation studies were excluded.

Database: Medline
Platform: Ovid
Version: Ovid MEDLINE(R) <1946 to July 17, 2020>
Search date: 21 Jul 2020
Number of results retrieved: 650
Search strategy:
Database: Ovid MEDLINE(R) <1946 to July 17, 2020>
Search Strategy:
---
1 Gender Dysphoria/ (485)
2 Gender Identity/ (18431)
3 "Sexual and Gender Disorders"/ (75)
4 Transsexualism/ (3758)
5 Transgender Persons/ (3134)
6 Health Services for Transgender Persons/ (136)
7 exp Sex Reassignment Procedures/ (835)
8 (gender* adj3 (dysphori* or incongru* or identi* or disorder* or confus* or minorit* or queer*)).tw. (7223)
9 (transgend* or transex* or transsex* or transfem* or transwom* or transma* or transmen* or transperson* or transpeopl*).tw. (12665)
10 (trans or crossgender* or cross-gender* or crossex* or cross-sex* or genderqueer*).tw. (102312)
11 ((sex or gender*) adj3 (reassign* or chang* or transform* or transition*)).tw. (6969)
12 (male-to-female or m2f or female-to-male or f2m).tw. (114785)
13 or/1-12 (252562)
14 exp Infant/ or Infant Health/ or Infant Welfare/ (1137237)
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. (852126)
16 exp Child/ or exp Child Behavior/ or Child Health/ or Child Welfare/ (1912796)
17 Minors/ (2572)
18 (child* or minor or minors or boy* or girl* or kid or kids or young*).ti,ab,in,jn. (2360626)
19 exp pediatrics/ (58102)
20 (pediatric* or paediatric* or peadiatric*).ti,ab,in,jn. (835833)
21 Adolescent/ or Adolescent Behavior/ or Adolescent Health/ (2023650)
22 Puberty/ (13277)

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. (424041)
24 Schools/ (38087)
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. (468784)
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. (89314)
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. (887443)
29 or/14-28 (5532185)
30 13 and 29 (79220)
31 (transchild* or transyouth* or transteen* or transadoles* or transgirl* or transboy*).tw. (7)
32 30 or 31 (79220)
33 Hormones/ad, tu, th (4514)
34 exp Progesterone/ad, tu, th (10899)
35 exp Estrogens/ad, tu, th (28936)
36 exp Gonadal Steroid Hormones/ad, tu, th (34137)
37 (progesteron* or oestrogen* or estrogen*).tw. (196074)
38 ((cross-sex or crosssex or gender-affirm*) and (hormon* or steroid* or therap* or treatment* or prescri* or pharm* or medici* or drug* or intervention* or care)).tw. (544)
39 exp Estradiol/ad, tu, th (10823)
40 exp Testosterone/ad, tu, th (8318)
41 (testosteron* or sustanon* or tostran or testogel or testim or restandol or andriol or testocaps* or nebido or testavan).tw. (74936)
42 (oestrad* or estrad* or evorel or ethinyloestrad* or ethinylestrad* or elleste or progynova or zumenon or bedol or femseven or nuvelle).tw. (90464)
43 or/33-42 (304239)
44 32 and 43 (3183)
45 limit 44 to yr="2000 -Current" (2019)
46 animals/ not humans/ (4685420)
47 45 not 46 (1194)
48 limit 47 to english language (1155)
49 (MEDLINE or pubmed).tw. (163678)
50 systematic review.tw. (121198)
51 systematic review.pt. (130231)
52 meta-analysis.pt. (117148)
53 intervention$.ti. (123904)
54 or/49-53 (380217)
55 randomized controlled trial.pt. (509468)
56 randomi?ed.mp. (796957)
57 placebo.mp. (194937)
58 or/55-57 (848627)
59 exp cohort studies/ or exp epidemiologic studies/ or exp clinical trial/ or exp evaluation studies as topic/ or exp statistics as topic/ (5562241)
60 ((control and (group* or study)) or (time and factors)).mp. (3274107)
61 (program or survey* or ci or cohort or comparative stud* or evaluation studies or follow-up*).mp. (4624419)
62 or/59-61 (9030680)
63 Observational Studies as Topic/ (5177)
64 Observational Study/ (81866)
65 Epidemiologic Studies/ (8358)

66 exp Case-Control Studies/ (1090891)
67 exp Cohort Studies/ (2011414)
68 Cross-Sectional Studies/ (332273)
69 Controlled Before-After Studies/ (526)
70 Historically Controlled Study/ (185)
71 Interrupted Time Series Analysis/ (913)
72 Comparative Study.pt. (1866044)
73 case control$.tw. (112152)
74 case series.tw. (59119)
75 (cohort adj (study or studies)).tw. (170281)
76 cohort analy$.tw. (6758)
77 (follow up adj (study or studies)).tw. (45131)
78 (observational adj (study or studies)).tw. (86247)
79 longitudinal.tw. (204239)
80 prospective.tw. (495367)
81 retrospective.tw. (442876)
82 cross sectional.tw. (284856)
83 or/63-82 (4368140)
84 54 or 58 or 62 or 83 (9402123)
85 48 and 84 (683)
86 limit 85 to (letter or historical article or comment or editorial or news or case reports) (33)
87 85 not 86 (650)

Database: Medline in-process
Platform: Ovid
Version: Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations <1946 to July 17, 2020>
Search date: 21 July 2020
Number of results retrieved: 122
Search strategy:
Database: Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations <1946 to July 17, 2020>
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 incongru* or identi* or disorder* or confus* or minorit* or queer*)).tw. (1473)
9 (transgend* or transex* or transsex* or transfem* or transwom* or transma* or transmen* or transperson* or transpeopl*).tw. (2315)
10 (trans or crossgender* or cross-gender* or crossex* or cross-sex* or genderqueer*).tw. (20821)
11 ((sex or gender*) adj3 (reassign* or chang* or transform* or transition*)).tw. (963)
12 (male-to-female or m2f or female-to-male or f2m).tw. (15453)
13 or/1-12 (39735)
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. (80295)

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. (320315)
19 exp pediatrics/ (0)
20 (pediatric* or paediatric* or peadiatric*).ti,ab,in,jn. (119124)
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. (59969)
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. (68979)
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. (10287)
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. (112220)
29 or/14-28 (523053)
30 13 and 29 (9143)
31 (transchild* or transyouth* or transteen* or transadoles* or transgirl* or transboy*).tw. (3)
32 30 or 31 (9144)
33 Hormones/ad, tu, th (0)
34 exp Progesterone/ad, tu, th (0)
35 exp Estrogens/ad, tu, th (0)
36 exp Gonadal Steroid Hormones/ad, tu, th (0)
37 (progesteron* or oestrogen* or estrogen*).tw. (13291)
38 ((cross-sex or crosssex or gender-affirm*) and (hormon* or steroid* or therap* or treatment* or prescri* or pharm* or medici* or drug* or intervention* or care)).tw. (241)
39 exp Estradiol/ad, tu, th (0)
40 exp Testosterone/ad, tu, th (0)
41 (testosteron* or sustanon* or tostran or testogel or testim or restandol or andriol or testocaps* or nebido or testavan).tw. (5458)
42 (oestrad* or estrad* or evorel or ethinyloestrad* or ethinylestrad* or elleste or progynova or zumenon or bedol or femseven or nuvelle).tw. (4772)
43 or/33-42 (19706)
44 32 and 43 (316)
45 limit 44 to yr="2000 -Current" (303)
46 animals/ not humans/ (1)
47 45 not 46 (303)
48 limit 47 to english language (303)
49 (MEDLINE or pubmed).tw. (36030)
50 systematic review.tw. (29830)
51 systematic review.pt. (1007)
52 meta-analysis.pt. (49)
53 intervention$.ti. (21354)
54 or/49-53 (68976)
55 randomized controlled trial.pt. (277)
56 randomi?ed.mp. (74978)
57 placebo.mp. (18290)
58 or/55-57 (81427)
59 exp cohort studies/ or exp epidemiologic studies/ or exp clinical trial/ or exp evaluation studies as topic/ or exp statistics as topic/ (455)

60 ((control and (group* or study)) or (time and factors)).mp. (214372)
61 (program or survey* or ci or cohort or comparative stud* or evaluation studies or follow-up*).mp. (339764)
62 or/59-61 (507046)
63 Observational Studies as Topic/ (0)
64 Observational Study/ (91)
65 Epidemiologic Studies/ (0)
66 exp Case-Control Studies/ (1)
67 exp Cohort Studies/ (1)
68 Cross-Sectional Studies/ (0)
69 Controlled Before-After Studies/ (0)
70 Historically Controlled Study/ (0)
71 Interrupted Time Series Analysis/ (0)
72 Comparative Study.pt. (46)
73 case control$.tw. (14451)
74 case series.tw. (13070)
75 (cohort adj (study or studies)).tw. (29119)
76 cohort analy$.tw. (1039)
77 (follow up adj (study or studies)).tw. (3540)
78 (observational adj (study or studies)).tw. (17421)
79 longitudinal.tw. (34485)
80 prospective.tw. (63689)
81 retrospective.tw. (73761)
82 cross sectional.tw. (60195)
83 or/63-82 (250805)
84 54 or 58 or 62 or 83 (687622)
85 48 and 84 (126)
86 limit 85 to (letter or historical article or comment or editorial or news or case reports) (4)
87 85 not 86 (122)

Database: Medline epubs ahead of print
Platform: Ovid
Version: Ovid MEDLINE(R) Epub Ahead of Print <July 17, 2020>
Search date: 21 July 2020
Number of results retrieved: 32
Search strategy:
Database: Ovid MEDLINE(R) Epub Ahead of Print <July 17, 2020>
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 incongru* or identi* or disorder* or confus* or minorit* or queer*)).tw. (430)
9 (transgend* or transex* or transsex* or transfem* or transwom* or transma* or transmen* or transperson* or transpeopl*).tw. (637)
10 (trans or crossgender* or cross-gender* or crossex* or cross-sex* or genderqueer*).tw. (1499)
11 ((sex or gender*) adj3 (reassign* or chang* or transform* or transition*)).tw. (179)
12 (male-to-female or m2f or female-to-male or f2m).tw. (2460)

13 or/1-12 (4883)
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. (15416)
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. (53285)
19 exp pediatrics/ (0)
20 (pediatric* or paediatric* or peadiatric*).ti,ab,in,jn. (22649)
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. (13005)
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. (12420)
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. (1407)
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. (20083)
29 or/14-28 (87968)
30 13 and 29 (1618)
31 (transchild* or transyouth* or transteen* or transadoles* or transgirl* or transboy*).tw. (1)
32 30 or 31 (1618)
33 Hormones/ad, tu, th (0)
34 exp Progesterone/ad, tu, th (0)
35 exp Estrogens/ad, tu, th (0)
36 exp Gonadal Steroid Hormones/ad, tu, th (0)
37 (progesteron* or oestrogen* or estrogen*).tw. (1876)
38 ((cross-sex or crosssex or gender-affirm*) and (hormon* or steroid* or therap* or treatment* or prescri* or pharm* or medici* or drug* or intervention* or care)).tw. (63)
39 exp Estradiol/ad, tu, th (0)
40 exp Testosterone/ad, tu, th (0)
41 (testosteron* or sustanon* or tostran or testogel or testim or restandol or andriol or testocaps* or nebido or testavan).tw. (846)
42 (oestrad* or estrad* or evorel or ethinyloestrad* or ethinylestrad* or elleste or progynova or zumenon or bedol or femseven or nuvelle).tw. (665)
43 or/33-42 (2850)
44 32 and 43 (64)
45 limit 44 to yr="2000 -Current" (61)
46 animals/ not humans/ (0)
47 45 not 46 (61)
48 limit 47 to english language (61)
49 (MEDLINE or pubmed).tw. (7948)
50 systematic review.tw. (7508)
51 systematic review.pt. (28)
52 meta-analysis.pt. (37)
53 intervention$.ti. (4267)
54 or/49-53 (15048)
55 randomized controlled trial.pt. (1)

56 randomi?ed.mp. (14113)
57 placebo.mp. (3097)
58 or/55-57 (15128)
59 exp cohort studies/ or exp epidemiologic studies/ or exp clinical trial/ or exp evaluation studies as topic/ or exp statistics as topic/ (34)
60 ((control and (group* or study)) or (time and factors)).mp. (31615)
61 (program or survey* or ci or cohort or comparative stud* or evaluation studies or follow-up*).mp. (65735)
62 or/59-61 (88222)
63 Observational Studies as Topic/ (0)
64 Observational Study/ (4)
65 Epidemiologic Studies/ (0)
66 exp Case-Control Studies/ (0)
67 exp Cohort Studies/ (0)
68 Cross-Sectional Studies/ (0)
69 Controlled Before-After Studies/ (0)
70 Historically Controlled Study/ (0)
71 Interrupted Time Series Analysis/ (0)
72 Comparative Study.pt. (0)
73 case control$.tw. (2577)
74 case series.tw. (2480)
75 (cohort adj (study or studies)).tw. (7959)
76 cohort analy$.tw. (287)
77 (follow up adj (study or studies)).tw. (632)
78 (observational adj (study or studies)).tw. (3763)
79 longitudinal.tw. (7079)
80 prospective.tw. (12148)
81 retrospective.tw. (16600)
82 cross sectional.tw. (9459)
83 or/63-82 (48534)
84 54 or 58 or 62 or 83 (119752)
85 48 and 84 (32)
86 limit 85 to (letter or historical article or comment or editorial or news or case reports) (0)
87 85 not 86 (32)

Database: Medline daily update
Platform: Ovid
Version: Ovid MEDLINE(R) Daily Update <July 21, 2020>
Search date: 22 July 2020
Number of results retrieved: 3
Search strategy

Database: Ovid MEDLINE(R) Daily Update <July 21, 2020>
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 incongru* or identi* or disorder* or confus* or minorit* or queer*)).tw. (22)

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 Hormones/ad, tu, th (3)
34 exp Progesterone/ad, tu, th (3)
35 exp Estrogens/ad, tu, th (8)
36 exp Gonadal Steroid Hormones/ad, tu, th (22)
37 (progesteron* or oestrogen* or estrogen*).tw. (161)
38 ((cross-sex or crosssex or gender-affirm*) and (hormon* or steroid* or therap* or treatment* or prescri* or pharm* or medici* or drug* or intervention* or care)).tw. (3)
39 exp Estradiol/ad, tu, th (8)
40 exp Testosterone/ad, tu, th (8)
41 (testosteron* or sustanon* or tostran or testogel or testim or restandol or andriol or testocaps* or nebido or testavan).tw. (79)
42 (oestrad* or estrad* or evorel or ethinyloestrad* or ethinylestrad* or elleste or progynova or zumenon or bedol or femseven or nuvelle).tw. (61)
43 or/33-42 (261)
44 32 and 43 (7)
45 limit 44 to yr="2000 -Current" (7)
46 animals/ not humans/ (3647)
47 45 not 46 (6)
48 limit 47 to english language (6)
49 (MEDLINE or pubmed).tw. (529)
50 systematic review.tw. (512)

51 systematic review.pt. (522)
52 meta-analysis.pt. (370)
53 intervention$.ti. (247)
54 or/49-53 (1065)
55 randomized controlled trial.pt. (595)
56 randomi?ed.mp. (1203)
57 placebo.mp. (219)
58 or/55-57 (1234)
59 exp cohort studies/ or exp epidemiologic studies/ or exp clinical trial/ or exp evaluation studies as topic/ or exp statistics as topic/ (7958)
60 ((control and (group* or study)) or (time and factors)).mp. (4307)
61 (program or survey* or ci or cohort or comparative stud* or evaluation studies or follow-up*).mp. (5828)
62 or/59-61 (11814)
63 Observational Studies as Topic/ (27)
64 Observational Study/ (449)
65 Epidemiologic Studies/ (7)
66 exp Case-Control Studies/ (2173)
67 exp Cohort Studies/ (3287)
68 Cross-Sectional Studies/ (837)
69 Controlled Before-After Studies/ (1)
70 Historically Controlled Study/ (0)
71 Interrupted Time Series Analysis/ (6)
72 Comparative Study.pt. (768)
73 case control$.tw. (182)
74 case series.tw. (139)
75 (cohort adj (study or studies)).tw. (561)
76 cohort analy$.tw. (22)
77 (follow up adj (study or studies)).tw. (40)
78 (observational adj (study or studies)).tw. (253)
79 longitudinal.tw. (429)
80 prospective.tw. (778)
81 retrospective.tw. (1032)
82 cross sectional.tw. (739)
83 or/63-82 (5471)
84 54 or 58 or 62 or 83 (12581)
85 48 and 84 (3)
86 limit 85 to (letter or historical article or comment or editorial or news or case reports) (0)
87 85 not 86 (3)

Database: Embase
Platform: Ovid
Version: Embase <1974 to 2020 July 22>
Search date: 23 July 2020
Number of results retrieved: 1207
Search strategy:

Database: Embase <1974 to 2020 July 22>
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/ (1108)
8 (gender* adj3 (dysphori* or incongru* or identi* or disorder* or confus* or minorit* or queer*)).tw. (12470)
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 (581748)
14 exp juvenile/ or Child Behavior/ or Child Welfare/ or Child Health/ or infant welfare/ or "minor (person)"/ or elementary student/ or adolescent health/ or middle school student/ or high school student/ (3440943)
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 (181778)
27 (transchild* or transyouth* or transteen* or transadoles* or transgirl* or transboy*).tw. (17)
28 26 or 27 (181778)
29 hormone/bd, ad, an, cr, do, it, dt, to, ei, ih, ia, ar, cv, dl, im, na, ip, ut, va, iv, ve, vi, po, pa, pr, sc, li, th, tp, td (5160)
30 exp progesterone derivative/bd, ad, an, cr, do, it, dt, to, ei, ih, ia, ar, cv, dl, im, na, ip, ut, va, iv, ve, vi, po, pa, pr, sc, li, th, tp, td (23479)
31 exp estrogen/bd, ad, an, cr, do, it, dt, to, ei, ih, ia, ar, cv, dl, im, na, ip, ut, va, iv, ve, vi, po, pa, pr, sc, li, th, tp, td (57641)
32 steroid hormone/bd, ad, an, cr, do, it, dt, to, ei, ih, ia, ar, cv, dl, im, na, ip, ut, va, iv, ve, vi, po, pa, pr, sc, li, th, tp, td (372)
33 sex hormone/bd, ad, an, cr, do, it, dt, to, ei, ih, ia, ar, cv, dl, im, na, ip, ut, va, iv, ve, vi, po, pa, pr, sc, li, th, tp, td (1984)
34 hormonal therapy/ (42222)
35 (progesteron* or oestrogen* or estrogen*).tw. (254142)
36 ((cross-sex or crosssex or gender-affirm*) and (hormon* or steroid* or therap* or treatment* or prescri* or pharm* or medici* or drug* or intervention* or care)).tw. (1224)
37 exp estradiol derivative/bd, ad, an, cr, do, it, dt, to, ei, ih, ia, ar, cv, dl, im, na, ip, ut, va, iv, ve, vi, po, pa, pr, sc, li, th, tp, td (30740)

38 exp testosterone derivative/bd, ad, an, cr, do, it, dt, to, ei, ih, ia, ar, cv, dl, im, na, ip, ut, va, iv, ve, vi, po, pa, pr, sc, li, th, tp, td (15868)
39 (testosteron* or sustanon* or tostran or testogel or testim or restandol or andriol or testocaps* or nebido or testavan).tw. (99596)
40 (oestrad* or estrad* or evorel or ethinyloestrad* or ethinylestrad* or elleste or progynova or zumenon or bedol or femseven or nuvelle).tw. (114290)
41 or/29-40 (438737)
42 28 and 41 (6053)
43 limit 42 to yr="2000 -Current" (4741)
44 nonhuman/ not human/ (4649157)
45 43 not 44 (3636)
46 limit 45 to english language (3513)
47 (MEDLINE or pubmed).tw. (261145)
48 exp systematic review/ or systematic review.tw. (302985)
49 meta-analysis/ (191173)
50 intervention$.ti. (200041)
51 or/47-50 (660206)
52 random:.tw. (1552336)
53 placebo:.mp. (455979)
54 double-blind:.tw. (210671)
55 or/52-54 (1807280)
56 cohort analysis/ (596360)
57 exp epidemiology/ (3434332)
58 exp clinical trial/ (1504711)
59 evaluation study/ (45870)
60 statistics/ (301181)
61 ((control and (group* or study)) or (time and factors)).mp. (3324555)
62 (program or survey* or ci or cohort or comparative stud* or evaluation studies or follow-up*).mp. (6067112)
63 or/56-62 (11048972)
64 Clinical study/ (155444)
65 Case control study/ (157943)
66 Family study/ (26047)
67 Longitudinal study/ (141660)
68 Retrospective study/ (937696)
69 comparative study/ (859061)
70 Prospective study/ (613138)
71 Randomized controlled trials/ (182542)
72 70 not 71 (606604)
73 Cohort analysis/ (596360)
74 cohort analy$.tw. (13020)
75 (Cohort adj (study or studies)).tw. (302159)
76 (Case control$ adj (study or studies)).tw. (137432)
77 (follow up adj (study or studies)).tw. (63423)
78 (observational adj (study or studies)).tw. (168428)
79 (epidemiologic$ adj (study or studies)).tw. (106448)
80 (cross sectional adj (study or studies)).tw. (220073)
81 case series.tw. (104089)
82 prospective.tw. (861922)
83 retrospective.tw. (886445)
84 or/64-69,72-83 (4047788)
85 51 or 55 or 63 or 84 (12494560)
86 46 and 85 (2151)
87 86 not (letter or editorial).pt. (2137)

88 87 not (conference abstract or conference paper or conference proceeding or "conference review").pt. (1207)

Database: APA PsycInfo

Platform: Ovid
Version: APA PsycInfo <1806 to July Week 2 2020>
Search date: 22 July 2020
Number of results retrieved: 581
Search strategy:

Database: APA PsycInfo <1806 to July Week 2 2020>
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 incongruen* or identi* or disorder* or confus* or minorit* or queer*)).tw. (15276)
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 (99498)
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/ (29181)
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 (1765408)
26 11 and 25 (49560)
27 (transchild* or transyouth* or transteen* or transadoles* or transgirl* or transboy*).tw. (14)

28 26 or 27 (49561)
29 hormones/ (8408)
30 sex hormones/ (1777)
31 exp progestational hormones/ (2409)
32 estrogens/ (3889)
33 steroids/ (3797)
34 (progesteron* or oestrogen* or estrogen*).tw. (11188)
35 ((cross-sex or crosssex or gender-affirm*) and (hormon* or steroid* or therap* or treatment* or prescri* or pharm* or medici* or drug* or intervention* or care)).tw. (457)
36 estradiol/ (3120)
37 testosterone/ (5606)
38 (testosteron* or sustanon* or tostran or testogel or testim or restandol or andriol or testocaps* or nebido or testavan).tw. (9625)
39 (oestrad* or estrad* or evorel or ethinyloestrad* or ethinylestrad* or elleste or progynova or zumenon or bedol or femseven or nuvelle).tw. (6741)
40 or/29-39 (30344)
41 28 and 40 (1005)
42 limit 41 to yr="2000 -Current" (749)
43 limit 42 to english language (692)
44 limit 43 to ("0200 book" or "0240 authored book" or "0280 edited book" or "0300 encyclopedia" or "0400 dissertation abstract") (111)
45 43 not 44 (581)

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: 22 July 2020
Number of results retrieved: CDSR 0 ; CENTRAL 67.

ID Search Hits
#1 MeSH descriptor: [Gender Dysphoria] this term only 3
#2 MeSH descriptor: [Gender Identity] this term only 227
#3 MeSH descriptor: [Sexual and Gender Disorders] this term only 2
#4 MeSH descriptor: [Transsexualism] this term only 27
#5 MeSH descriptor: [Transgender Persons] this term only 36
#6 MeSH descriptor: [Health Services for Transgender Persons] this term only 0
#7 MeSH descriptor: [Sex Reassignment Procedures] explode all trees 4
#8 (gender* near/3 (dysphori* or incongru* or identi* or disorder* or confus* or minorit* or queer*)):ti,ab,kw 702
#9 (transgend* or transex* or transsex* or transfem* or transwom* or transma* or transmen* or transperson* or transpeopl*):ti,ab,kw 959
#10 (trans or crossgender* or cross-gender* or crossex* or cross-sex* or genderqueer*):ti,ab,kw 3969
#11 ((sex or gender*) near/3 (reassign* or chang* or transform* or transition*)):ti,ab,kw 524
#12 (male-to-female or m2f or female-to-male or f2m):ti,ab,kw 516
#13 #1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10 or #11 or #12 6413
#14 MeSH descriptor: [Infant] explode all trees 28440
#15 MeSH descriptor: [Infant Health] this term only 49
#16 MeSH descriptor: [Infant Welfare] this term only 82

#17 (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,kw,so 89530
#18 MeSH descriptor: [Child] explode all trees 44089
#19 MeSH descriptor: [Child Behavior] explode all trees 2061
#20 MeSH descriptor: [Child Health] this term only 98
#21 MeSH descriptor: [Child Welfare] this term only 325
#22 MeSH descriptor: [Minors] this term only 8
#23 (child* or minor or minors or boy* or girl* or kid or kids or young*):ti,ab,kw,so 265417
#24 MeSH descriptor: [Pediatrics] explode all trees 661
#25 (pediatric* or paediatric* or peadiatric*):ti,ab,kw,so 57725
#26 MeSH descriptor: [Adolescent] this term only 102154
#27 MeSH descriptor: [Adolescent Behavior] this term only 1358
#28 MeSH descriptor: [Adolescent Health] this term only 29
#29 MeSH descriptor: [Puberty] this term only 295
#30 (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,kw,so 140927
#31 MeSH descriptor: [Schools] this term only 1914
#32 MeSH descriptor: [Child Day Care Centers] this term only 231
#33 MeSH descriptor: [Nurseries, Infant] explode all trees 17
#34 MeSH descriptor: [Schools, Nursery] this term only 37
#35 (pre-school* or preschool* or kindergar* or daycare or day-care or nurser* or school* or pupil* or student*):ti,ab,kw,so 97810
#36 (("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
#37 (("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
#38 #14 or #15 or #16 or #17 or #18 or #19 or #20 or #21 or #22 or #23 or #24 or #25 or #26 or #27 or #28 or #29 or #30 or #31 or #32 or #33 or #34 or #35 or #36 or #37 516067
#39 #13 and #38 2488
#40 (transchild* or transyouth* or transteen* or transadoles* or transgirl* or transboy*):ti,ab,kw 0
#41 #39 or #40 2488
#42 MeSH descriptor: [Hormones] this term only 2241
#43 MeSH descriptor: [Progesterone] explode all trees 3135
#44 MeSH descriptor: [Estrogens] explode all trees 1841
#45 MeSH descriptor: [Gonadal Steroid Hormones] explode all trees 10747
#46 (progesteron* or oestrogen* or estrogen*):ti,ab,kw 18387
#47 ((cross-sex or crosssex or gender-affirm*) and (hormon* or steroid* or therap* or treatment* or prescri* or pharm* or medici* or drug* or intervention* or care)):ti,ab,kw 24
#48 MeSH descriptor: [Estradiol] explode all trees 4434
#49 MeSH descriptor: [Testosterone] explode all trees 2945
#50 (testosteron* or sustanon* or tostran or testogel or testim or restandol or andriol or testocaps* or nebido or testavan):ti,ab,kw 7386
#51 (oestrad* or estrad* or evorel or ethinyloestrad* or ethinylestrad* or elleste or progynova or zumenon or bedol or femseven or nuvelle):ti,ab,kw 11410
#52 #42 or #43 or #44 or #45 or #46 or #47 or #48 or #49 or #50 or #51 31870
#53 #41 and #52 121
#54 "conference":pt or (clinicaltrials or trialsearch):so 492465
#55 #53 not #54 72

Database: HTA

Platform: Wiley
Version: up to 2018
Search date: 22nd July 2020
Number of results retrieved: 4
Search strategy:

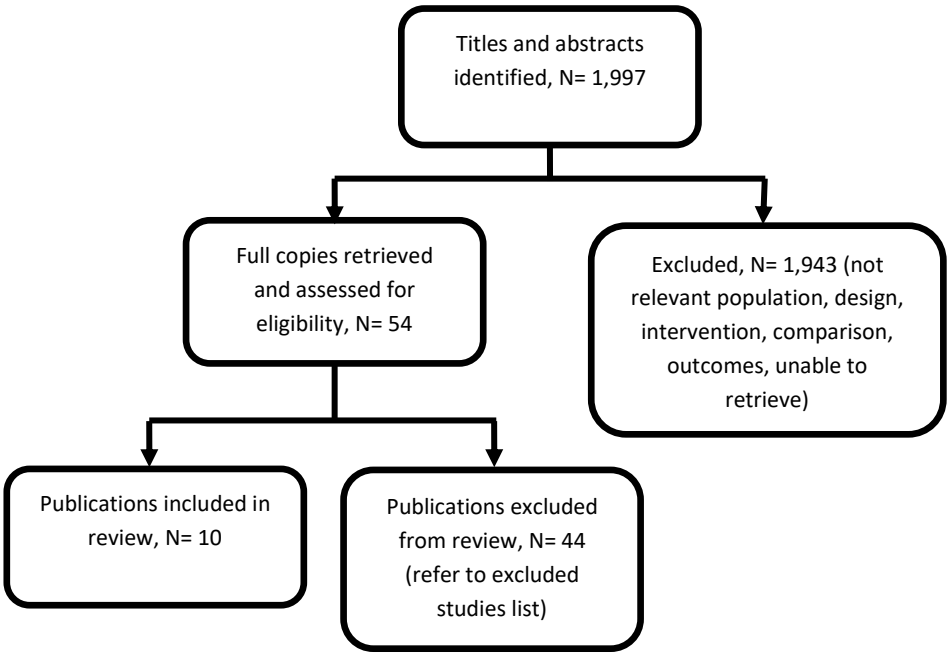
#1 MeSH DESCRIPTOR Gender Dysphoria 0
#2 MeSH DESCRIPTOR Gender Identity 12
#3 MeSH DESCRIPTOR Sexual and Gender Disorders 2
#4 MeSH DESCRIPTOR Transsexualism 12
#5 MeSH DESCRIPTOR Transgender Persons 3
#6 MeSH DESCRIPTOR Health Services for Transgender Persons 0
#7 MeSH DESCRIPTOR Sex Reassignment Procedures EXPLODE ALL TREES 1
#8 ((gender* near3 (dysphori* or incongru* 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*) near3 (reassign* or chang* or transform* or transition*))) 24
#12 ((male-to-female or m2f or female-to-male or f2m)) 86
#13 #1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 261
#14 MeSH DESCRIPTOR Infant EXPLODE ALL TREES 2964
#15 MeSH DESCRIPTOR Infant Health 0
#16 MeSH DESCRIPTOR Infant Welfare 22
#17 ((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*)) 5510
#18 MeSH DESCRIPTOR Child EXPLODE ALL TREES 4935
#19 MeSH DESCRIPTOR Child Behavior EXPLODE ALL TREES 64
#20 MeSH DESCRIPTOR Child Health 2
#21 MeSH DESCRIPTOR Child Welfare 80
#22 MeSH DESCRIPTOR Minors 2
#23 ((child* or minor or minors or boy* or girl* or kid or kids or young*)) 13575
#24 MeSH DESCRIPTOR Pediatrics EXPLODE ALL TREES 119
#25 ((pediatric* or paediatric* or peadiatric*)) 2842
#26 MeSH DESCRIPTOR Adolescent 4594
#27 MeSH DESCRIPTOR Adolescent Behavior 94
#28 MeSH DESCRIPTOR Adolescent Health 0
#29 MeSH DESCRIPTOR Puberty 3
#30 ((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*)) 5621
#31 MeSH DESCRIPTOR Schools 168
#32 MeSH DESCRIPTOR Child Day Care Centers 12
#33 MeSH DESCRIPTOR Schools, Nursery 3
#34 ((pre-school* or preschool* or kindergar* or daycare or day-care or nurser* or school* or pupil* or student*)) 4454
#35 ((("eight" or "nine" or "ten" or "eleven" or "twelve" or "thirteen" or "fourteen" or "fifteen" or "sixteen" or "seventeen" or "eighteen" or "nineteen") near2 (year or years or age or ages or aged))) 380
#36 ((("8" or "9" or "10" or "11" or "12" or "13" or "14" or "15" or "16" or "17" or "18" or "19") near2 (year or years or age or ages or aged))) 7996

#37 #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20 OR #21 OR #22 OR #23 OR #24 OR #25 OR #26 OR #27 OR #28 OR #29 OR #30 OR #31 OR #32 OR #33 OR #34 OR #35 OR #36 22640
#38 #13 AND #37 116
#39 (#13 AND #37) IN HTA 4

Appendix C Evidence selection

The literature searches identified 1,997 references. These were screened using their titles and abstracts and 54 references were obtained and assessed for relevance. Of these, 10 references are included in the evidence review. The remaining 44 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
Aranda G, Mora M, Hanzu FA et al. (2019) Effects of sex steroids on cardiovascular risk profile in transgender men under gender affirming hormone therapy. Endocrinologia, diabetes y nutricion 66(6): 385–392	Excluded on population – adult study, participants not 18 years or less (mean age 27.1 years).
Arnold, Justin D, Sarkodie, Eleanor P, Coleman, Megan E et al. (2016) Incidence of Venous Thromboembolism in Transgender Women	Excluded on population – adult study, participants not 18 years or less (mean age 33.2 years).

Study reference	Reason for exclusion
Receiving Oral Estradiol. The journal of sexual medicine 13(11): 1773–1777	
Asscheman, Henk, Giltay, Erik J, Megens, Jos A J et al. (2011) A long-term follow-up study of mortality in transsexuals receiving treatment with cross-sex hormones. European journal of endocrinology 164(4): 635–42	Excluded on population – although some participants started gender-affirming hormones when young, the study does not report the proportion who started treatment when 18 years or less. Mean ages at start of treatment were 31.4 years (transfemales) and 26.1 years (transmales), suggesting the majority of participants were older than 18 years at the start of treatment. Outcomes not reported separately for people aged 18 years or less.
Author not, found (2014) Hormone therapy for the treatment of gender dysphoria. Lansdale, PA: HAYES, Inc	Full text paper not available.
Baba, T., Endo, T., Honnma, H. et al. (2007) Association between polycystic ovary syndrome and female-to-male transsexuality. Human Reproduction 22(4): 1011–1016	Excluded on population – although study included some younger people (age range 17 to 47), most participants were adults (mean age around 25 years) and the proportion who started treatment when 18 years or less is not reported. Outcomes not reported separately for people aged 18 years or less.
Becerra-Fernandez A, Perez-Lopez G, Roman MM et al. (2014) Prevalence of hyperandrogenism and polycystic ovary syndrome in female to male transsexuals. Endocrinologia y Nutricion: Organo de la Sociedad Espanola de Endocrinologia y Nutricion 61(7): 351–8	Excluded on population – although study included some younger people (age range 18 to 45), most participants were adults (mean age around 25 years) and the proportion who started treatment when 18 years or less is not reported. Outcomes not reported separately for people aged 18 years or less.
Becker I, Auer M, Barkmann C et al. (2018) A Cross-Sectional Multicenter Study of Multidimensional Body Image in Adolescents and Adults with Gender Dysphoria Before and After Transition-Related Medical Interventions. Archives of Sexual Behavior 47(8): 2335–2347	Excluded on population – study included people aged 14 to 21 years. Outcomes not reported separately for people aged 18 years or less. Better evidence available – only 11 participants received gender-affirming hormones. The majority of the study cohort were either pre-treatment, received puberty suppression alone, or received hormones and underwent surgery.
Chew D, Anderson J, Williams K et al. (2018) Hormonal Treatment in Young People With Gender Dysphoria: A Systematic Review. Pediatrics 141(4): e20173742	Excluded on better available evidence - systematic review did not meta-analyse results from. Individual studies from this systematic review are either

Study reference	Reason for exclusion
	included, or excluded because they did not meet the PICO criteria.
Connolly MD, Zervos MJ, Barone CJ 2nd et al. (2016) The Mental Health of Transgender Youth: Advances in Understanding. The Journal of Adolescent Health: Official Publication of the Society for Adolescent Medicine 59(5): 489–495	Excluded on intervention - review did not investigate gender-affirming hormones
de Vries ALC, McGuire JK, Steensma TD et al. (2014) Young adult psychological outcome after puberty suppression and gender reassignment. Pediatrics 134(4): 696–704	Exclude on intervention – all participants had surgery after gender-affirming hormones. Unable to determine whether changes were due to hormones or surgery. Complete data only available for 40 patients. Details of gender-affirming hormones are poorly reported. Outcomes reported in other study (with a population that more closely matches PICO)
Elamin MB, Garcia MZ, Murad MH et al. (2010) Effect of sex steroid use on cardiovascular risk in transsexual individuals: a systematic review and meta-analyses. Clinical Endocrinology 72(1): 1–10	Exclude on population – all included studies conducted in adult population. Unclear whether hormones were started when participants were aged 18 years or less. Outcomes not reported by age at treatment initiation.
Fernandez JD and Tannock LR (2016) Metabolic effects of hormone therapy in transgender patients. Endocrine Practice: Official Journal of the American College of Endocrinology and the American Association of Clinical Endocrinologists 22(4): 383–8	Excluded on population – adult study, participants not 18 years or less (mean ages 31 and 27 years).
Fighera TM, Ziegelmann PK, Da Silva TR et al. (2019) Bone mass effects of cross-sex hormone therapy in transgender people: Updated systematic review and meta-analysis. Journal of the Endocrine Society 3(5): 943–964	Excluded on population – all included studies conducted in adult population. Unclear whether hormones were started when participants were aged 18 years or less. Outcomes not reported by age at treatment initiation.
Getahun D, Nash R, Flanders WD et al. (2018) Cross-sex Hormones and Acute Cardiovascular Events in Transgender Persons: A Cohort Study. Annals of Internal Medicine 169(4): 205–213	Excluded on population – adult study, participants not 18 years or less.
Gomez-Gil E, Zubiaurre-Elorza L, de Antonio IE et al. (2014) Determinants of quality of life in Spanish transsexuals attending a gender unit before genital sex reassignment surgery. Quality of Life Research: an International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation 23(2): 669–76	Excluded on population – although study included some younger people (age range 16 to 67), most participants were adults (mean age 31.2 years) and the proportion who started treatment when 18 years or less is not reported. Outcomes not reported separately for people aged 18 years or less.
Gomez-Gil E, Zubiaurre-Elorza L, Esteva I et al. (2012) Hormone-treated transsexuals report less	Excluded on population – adult study, participants not 18 years or less (mean age 24.6 years).

Study reference	Reason for exclusion
social distress, anxiety and depression. Psychoneuroendocrinology 37(5): 662–70	
Gooren LJ, van Trotsenburg MAA, Giltay EJ et al. (2013) Breast cancer development in transsexual subjects receiving cross-sex hormone treatment. The Journal of Sexual Medicine 10(12): 3129–34	Excluded on population – study reports on cancer rates in people aged 18-80 years. The 3 cases of cancer all started gender-affirming hormone treatment >18 years.
Grimstad FW, Boskey E, Grey M (2020) New-Onset Abdominopelvic Pain After Initiation of Testosterone Therapy Among TransMasculine Persons: A Community-Based Exploratory Survey. LGBT health 7(5): Published Online:13 Jul 2020https://doi.org/10.1089/lgbt.2019.0258	Excluded on population – adult study, participants not 18 years or less.
Hannema SE, Schagen SEE, Cohen-Kettenis PT et al. (2017) Efficacy and Safety of Pubertal Induction Using 17beta-Estradiol in Transgirls. The Journal of Clinical Endocrinology and Metabolism 102(7): 2356–2363	Excluded on better evidence available – small study (n=28) with high drop-out rate (n=16 at final follow-up). Same outcomes reported in larger studies.
Jarin J, Pine-Twaddell E, Trotman G et al. (2017) Cross-Sex Hormones and Metabolic Parameters in Adolescents With Gender Dysphoria. Pediatrics 139(5)	Excluded on population and better evidence available. Although the study included some younger people (age range 13 to 25; mean age 16 and 18), the proportion who started treatment when 18 years or less is not reported. Outcomes not reported separately for people aged 18 years or less. Outcomes were limited to physiological results (including haemoglobin, lipids and BMI). Follow-up only 6 months, other included studies report same outcomes with longer follow-up (12 to 31 months).
Keo-Meier CL, Herman LI, Reisner SL et al. (2015) Testosterone treatment and MMPI-2 improvement in transgender men: a prospective controlled study. Journal of consulting and clinical psychology 83(1): 143–56	Excluded on population – although study included some younger people (age range 18 to 54), most participants were adults (mean age 26.6 years) and the proportion who started treatment when 18 years or less is not reported. Outcomes not reported separately for people aged 18 years or less.
Klaver M, de Mutsert R, Wiepjes CM 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	Excluded on outcomes – reported outcomes not included in PICO document. The risk of obesity with gender-affirmed hormones was reported in an included study.
McFarlane T, Zajac JD, Cheung AS (2018) Gender-affirming hormone therapy and the risk of sex hormone-dependent tumours in transgender individuals-A systematic review. Clinical Endocrinology 89(6): 700-711	Exclude on population – all included studies conducted in adult population.

Study reference	Reason for exclusion
Meriggiola MC, Armillotta F, Costantino A et al. (2008) Effects of testosterone undecanoate administered alone or in combination with letrozole or dutasteride in female to male transsexuals. The Journal of Sexual Medicine 5(10): 2442–53	Excluded on population – adult study, participants not 18 years or less.
Nota NM, Wiepjes CM, de Blok, CJM et al. (2018) The occurrence of benign brain tumours in transgender individuals during cross-sex hormone treatment. Brain: A Journal of Neurology 141(7): 2047–2054	Excluded on population – adult study, participants not 18 years or less.
Oda H and Kinoshita T (2017) Efficacy of hormonal and mental treatments with MMPI in FtM individuals: Cross-sectional and longitudinal studies. BMC Psychiatry 17(1): 256	Excluded on population – although study included some younger people (age range 15 to 43), most participants were adults (mean age around 25.6 years) and the proportion who started treatment when 18 years or less is not reported. Outcomes not reported separately for people aged 18 years or less.
Olson-Kennedy J, Okonta V, Clark LF et al. (2018) Physiologic Response to Gender-Affirming Hormones Among Transgender Youth. The Journal of Adolescent Health: Official Publication of the Society for Adolescent Medicine 62(4): 397–401	Excluded on population – although study included some younger people (age range 12 to 23; mean age 18 years). Outcomes not reported separately for people aged 18 years or less. Outcomes limited to physiological results (including haemoglobin, lipids, liver enzymes and BMI). Same outcomes reported in included studies that had a less indirect population and a longer follow-up.
Ott J, Kaufmann U, Bentz K et al. (2010) Incidence of thrombophilia and venous thrombosis in transsexuals under cross-sex hormone therapy. Fertility and sterility 93(4): 1267–72	Excluded on population – adult study, participants not 18 years or less.
Pakpoor J, Wotton CJ, Schmierer K et al. (2016) Gender identity disorders and multiple sclerosis risk: A national record-linkage study. Multiple sclerosis. Multiple Sclerosis Journal. 22(13): 1759–1762	Excluded on population – although study included some younger people, outcomes not reported separately for people aged 18 years or less. Also exclude for intervention – unclear if people received gender-affirming hormones.
Pyra M, Casimiro I, Rusie L et al. (2020) An Observational Study of Hypertension and Thromboembolism among Transgender Patients Using Gender-Affirming Hormone Therapy. Transgender Health 5(1): 1–9	Excluded on population – adult study (age range 20-70). Age at which gender-affirming hormones started not reported.
Quiros C, Patrascioiu I, Mora M et al. (2015) Effect of cross-sex hormone treatment on cardiovascular risk factors in transsexual individuals. Experience in a specialized unit in Catalonia. Endocrinologia y nutricion : organo de la Sociedad Espanola de Endocrinologia y Nutricion 62(5): 210–6	Excluded on population – adult study, participants not 18 years or less.

Study reference	Reason for exclusion
Rowniak S, Bolt L, Sharifi C (2019) Effect of cross-sex hormones on the quality of life, depression and anxiety of transgender individuals: A quantitative systematic review. JBI Database of Systematic Reviews and Implementation Reports 17(9): 1826–1854	Exclude on population – all included studies conducted in adult population.
Sequeira GM, Kidd K, El Nokali NE et al. (2019) Early Effects of Testosterone Initiation on Body Mass Index in Transmasculine Adolescents. Journal of Adolescent Health 65(6): 818–820	Exclude on outcome - study only reports BMI z-score over 12 month testosterone treatment. BMI not listed as an outcome of interest in the PICO document. Other included studies have investigated the impact of gender-affirming hormone treatment on CV risk profile, including longer term obesity rates, with a longer follow-up and more participants.
Shim JY, Laufer MR, Grimstad FW (2020) Dysmenorrhea and Endometriosis in Transgender Adolescents. Journal of Pediatric and Adolescent Gynecology. Available online 11 June 2020. https://doi.org/10.1016/j.jpag.2020.06.001	Exclude on population – only 2 participants taking testosterone before diagnosis of dysmenorrhea.
Slabbekoorn D, Van Goozen SHM, Gooren, LJG et al. (2001) Effects of cross-sex hormone treatment on emotionality in transsexuals. International Journal of Transgenderism 5(3): http://www.symposion.com/ijt/ijtvo05no03_02.htm	Excluded on population – adult study (age range 21 to 28 years)
Smith YLS., Van Goozen SHM, Kuiper AJ et al. (2005) Sex reassignment: Outcomes and predictors of treatment for adolescent and adult transsexuals. Psychological Medicine 35(1): 89–99	Excluded on population – results on adults only used to assess hormone treatment.
Sutherland N, Espinel W, Grotzke M et al. (2020) Unanswered Questions: Hereditary breast and gynecological cancer risk assessment in transgender adolescents and young adults. Journal of Genetic Counseling 29(4): 625–633	Excluded on study type – narrative review of 3 case reports.
van Velzen DM, Paldino A, Klaver M et al. (2019) Cardiometabolic Effects of Testosterone in Transmen and Estrogen Plus Cyproterone Acetate in Transwomen. The Journal of Clinical Endocrinology and Metabolism 104(6): 1937–1947	Excluded on population – adult study, participants not 18 years or less.
White Hughto JM and Reisner SL (2016) A Systematic Review of the Effects of Hormone Therapy on Psychological Functioning and Quality of Life in Transgender Individuals. Transgender Health 1(1): 21–31	Exclude on population – all included studies conducted in adult population.
Wiepjes CM, de Blok CJM, Staphorsius AS et al. (2020) Fracture Risk in Trans Women and Trans Men Using Long-Term Gender-Affirming Hormonal Treatment: A Nationwide Cohort Study. Journal of Bone and Mineral Research 35(1): 64–70	Excluded on population – adult study, all participants started gender-affirming hormones after 18 years.
Wierckx K, Mueller S, Weyers S et al. (2012) Long-term evaluation of cross-sex hormone treatment in	Excluded on population – adult study, participants not 18 years or less.

Study reference	Reason for exclusion
transsexual persons. The Journal of Sexual Medicine 9(10): 2641–51	
Wierckx K, Van Caenegem E, Schreiner T et al. (2014) 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. The journal of sexual medicine 11(8): 1999–2011	Excluded on population – adult study, participants not 18 years or less.
Wilson R, Jenkins C, Miller H et al. (2006) The effect of oestrogen on cytokine and antioxidant levels in male to female transsexual patients. Maturitas 55(1): 14–8	Excluded on population – adult study, participants not 18 years or less.
Witcomb GL, Bouman WP, Claes L et al. (2018) Levels of depression in transgender people and its predictors: Results of a large matched control study with transgender people accessing clinical services. Journal of Affective Disorders 235: 308–315	Excluded on population – although study included some younger people (age range 15 to 79), most participants were adults (mean age around 30.4 years) and the proportion who started treatment when 18 years or less is not reported. Outcomes not reported separately for people aged 18 years or less.

Appendix E Evidence tables

Study details	Population	Interventions	Study outcomes	Appraisal and Funding
Full citation 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 **Study location** Single centre, New York, United States **Study type** Prospective longitudinal study **Study aim** To assess the psychological wellbeing and quality of life in children and adolescents who have sought endocrine	Inclusion and exclusion not reported- it appears from the description in the publication that all people referred for gender dysphoria were invited to participate, and the vast majority agreed. Of the 95 treatment naïve people who entered the study, 50 people completed all follow-up questionnaires and were included in the analysis. No description of the 45 people without follow-up data reported. The study included 50 children, adolescents and young adults with gender dysphoria.	**Intervention** Endocrine interventions (the collective term used by authors for puberty suppression and gender-affirming hormones) were introduced as per Endocrine Society and the World Professional Association for	**Critical Outcomes** ***Impact on mental health*** Depression symptoms were assessed using the Center for Epidemiologic Studies Depression Scale (CESD-R). Statistically significant improvements in CESD-R score were observed from baseline (initial assessment; 21.4 points) to about 12 months follow-up (13.9 points; p<0.001). Regression analysis, controlling for reported medicines for mental health problems and engagement in counselling, found no statistically significant change from baseline in transfemales (p=0.27) and transmales (p=0.43). The Patient Health Questionnaire Modified for Teens (PHQ 9_Modified for Teens) was also used to assess depression symptoms. Depression scores improved from baseline (p< 0.001; absolute scores not reported numerically). Regression analysis, controlling for reported medicines for mental health problems and engagement in counselling, found no statistically significant change from baseline in transfemales (p=0.07) and transmales (p=0.67). Suicidal ideation measured using the additional questions from the PHQ 9_Modified for Teens, was presented in 10% (5/50) of	This study was appraised using the Newcastle-Ottawa tool for cohort studies. **Domain 1: Selection domain** 1. b) somewhat representative 2. c) no-non exposed cohort 3. a) secure record 4. b) no **Domain 2: Comparability** 1. c) no comparator **Domain 3: Outcome** 1. c) self-report 2. a) yes – 6 monthly assessment up to 12 months (preliminary results from an ongoing study) 3. c) Follow up rate less than 80% and no description of those lost **Overall quality is assessed as poor** Other comments: Although regression analysis results for some outcomes were controlled for use of medicines for mental health problems,

intervention to help with gender dysphoria. **Study dates** Study recruitment ran from December 2013 to December 2018; study is ongoing	17 transfemales and 33 transmales. Diagnostic criteria for gender dysphoria not reported. Mean age at baseline was 16.2 years (SD 2.2). Mean age at the start of gender-affirming hormone treatment not reported.	Transgender Health (WPATH) guidelines. Puberty suppression was: • GnRH agonist and/or anti-androgens (transfemales) • GnRH agonist or medroxyprogesterone (transmales) Average duration of GnRH analogue treatment not reported. Once eligible, gender-affirming hormones were offered, these were: • Oestradiol (transfemales) • Testosterone (transmales) Doses and route of administration not reported. After about 12-months treatment ('wave 3' in the study): • 24 people (48%) were on gender-affirming hormones alone • 12 people (24%) were on puberty suppression alone	participants at baseline and 6% (3/50) at about 12-month follow-up, no statistical analysis reported. The study also reported results by gender: In transfemales, 11.8% (2/17) had suicidal ideation at baseline compared with 5.9% (1/17) at 12-month follow-up (no statistically analysis reported) In transmales, 9.1% (3/33) had suicidal ideation at baseline compared with 6.1% (2/33) at 12-month follow-up (no statistically analysis reported) ***Impact on quality of life*** Quality of Life Enjoyment and Satisfaction Questionnaire (QLES-Q-SF) scores: there was no statistically significant change in score from baseline to about 12-months (p=0.085; absolute scores not reported numerically). Regression analysis, controlling for reported medicines for mental health problems and engagement in counselling, found not statistically significant change from baseline in transfemales (p=0.06) and transmales (p=0.08). *No other critical or important outcomes reported*	details of these is not reported. Other co-morbidities not reported. Source of funding: None

Study details	Population	Interventions	Study outcomes	Appraisal and Funding
		• 11 people (22%) were on both gender-affirming hormones and puberty suppression • 3 people (6%) were on no endocrine intervention Results not represented separately for the sub-group of people who received gender-affirming hormones. Average duration of treatment with gender-affirming hormones not reported. **Comparison** No comparison group. Change overtime reported.		

Study details	Population	Interventions	Study outcomes	Appraisal and Funding
Full citation Allen, LR, Watson, LB, Egan, AM et al. (2019) Well-being and suicidality among transgender youth after gender-affirming hormones. Clinical Practice in Pediatric	The study included adolescents and young adults (age range 13-20 years) who received services for gender dysphoria in a clinic in the United States. Participants were required to have received gender-	39 participants received gender-affirming hormones only 8 participants received a GnRH analogue followed by gender-affirming hormones.	**Critical Outcomes** ***Impact on mental health*** The Ask Suicide-Screening Questions (ASQ) instrument was used to assess suicidality. Following an average of about 12 months treatment with gender-affirming hormones, adjusted mean ASQ score was statistically significantly lower (from 1.11 [standard error	This study was appraised using the Newcastle-Ottawa tool for cohort studies. **Domain 1: Selection domain** 1. b) somewhat representative 2. c) no-non exposed cohort

Study details	Population	Interventions	Study outcomes	Appraisal and Funding
Psychology 7(3): 302-311 **Study location** Single centre, Kansas City, United States **Study type** Retrospective longitudinal study **Study aim** To examine suicidality and general well-being following administration of gender-affirming hormones. **Study dates** Participants first presented to the clinic between 2015 and 2018.	affirming hormones for at least 3 months, and have pre-test and final assessment data points. No exclusion criteria reported. In total 47 adolescents and young adults with gender dysphoria were included: 14 transfemales (sex assigned at birth male) and 33 transmales (sex assigned at birth female). Diagnostic criteria for gender dysphoria not reported. Mean age at pre-test (before administration of gender-affirming hormones) was 16.59 years (range 13.73 to 19.04). Mean age at the start of treatment in the sub-group who received gender-affirming hormones-only was 16.72 years. Mean age at the start of treatment with gender-affirming hormones in people who previously	Mean duration of treatment in the gender-affirming hormones only subgroup was 366 days. Mean duration of gender-affirming hormone treatment in people who had previously received a GnRH analogue was not reported. Mean duration of treatment with a GnRH analogue was not reported. Participants were assessed at the start of treatment and at least 3 months after treatment.	(SE) 0.22] at baseline to 0.27 [SE 0.12] at final assessment; p<0.001). The authors also reported change in ASQ separately for transfemales (from 1.21 [SE 0.36] at baseline to 0.24 [SE 0.19] at final assessment) and transmales (from 1.01 [SE 0.36] at baseline to 0.29 [0.13] at final assessment). There was no statistically significant difference in change from baseline between transfemales and transmales (p=0.79) ***Impact on quality of life*** Assessed using the General Well-Being Scale (GWBS) of the Pediatric Quality of Life Inventory. Following an average of about 12 months treatment with gender-affirming hormones, adjusted mean GWBS score was statistically significantly higher (from 61.7 [SE 2.43] at baseline to 70.23 [2.15] at final assessment; p<0.002). The authors also reported change in GWBS of the Pediatric Quality of Life Inventory for transfemales (from 58.44 [SE 4.09] at baseline to 69.52 [SE 3.62] at final assessment) and transmales (from 64.95 [SE 2.66] at baseline to 70.94 [2.35] at final assessment). There was no statistically significant difference in change from baseline between transfemales and transmales (p=0.32) *No other critical or important outcomes reported*	3. a) secure record 4. b) no **Domain 2: Comparability** 2. c) no comparator **Domain 3: Outcome** 1. b) record linkage 2. a) yes – mean duration of treatment was 366 days 3. a) complete follow up - all subjects accounted for **Overall quality is assessed as poor** Other comments: None Source of funding: Not reported

Study details	Population	Interventions	Study outcomes	Appraisal and Funding
	received a GnRH analogue was not reported.			

Study details	Population	Interventions	Study outcomes	Appraisal and Funding
Full citation Kaltiala, R., Heino, E., Tyolajarvi, M. et al. (2020) Adolescent development and psychosocial functioning after starting cross-sex hormones for gender dysphoria. Nordic Journal of Psychiatry 74(3): 213-219 **Study location** Single centre, Tampere, Finland **Study type** Retrospective chart review **Study aim** To evaluate the psychosocial functioning and need for mental health treatment during the gender identity diagnostic phase and after about	The study included adolescents who were referred to the gender identity service before they 18 years old, were diagnosed with gender dysphoria, received gender-affirming hormones and completed a follow-up of approximately 12 months after starting hormones. In total 52 adolescents were included, comprising of 11 transfemales and 41 transmales. Gender dysphoria was diagnosed according to International Classification of Disease 10 (ICD-10). The authors state that the corresponding diagnosis to 'gender dysphoria' in	Intervention referred to as 'hormonal sex reassignment treatment' – details of intervention not reported, although gender-affirming hormones were prescribed to all participants. It is not clear from the study whether additional interventions were prescribed. Medical records reviewed for the 'real-life phase' – the approximately 12 months follow-up period for this population in Finland.	**Critical Outcomes** ***Impact on mental health*** Of the 52 people who received gender-affirming hormones, 50% (26/52) needed mental health treatment before or during the assessment and 46% (24/51) needed mental health treatment during the 12-month 'real life' phase (no statistically significant difference). For specific symptoms / conditions: • depression: 54% (28/52) needed treatment before or during the assessment and 15% (8/52) needed treatment during the 12-month 'real life' phase (statistically significant reduction, p<0.001) • anxiety: 48% (25/52) needed treatment before or during the assessment and 15% (8/52) needed treatment during the 12-month 'real life' phase (statistically significant reduction, p<0.001) • suicidality/self-harm: 35% (18/52) needed treatment before or during the assessment and 4% (2/52) needed treatment during the 12-month 'real life' phase (statistically significant reduction, p<0.001) • conduct problems/antisocial: 14% (7/52) needed treatment before or during the	This study was appraised using the Newcastle-Ottawa tool for cohort studies. **Domain 1: Selection domain** 1. b) somewhat representative 2. c) no-non exposed cohort 3. a) secure record 4. b) no **Domain 2: Comparability** 1. c) cohorts are not comparable on the basis of the design or analysis controlled for confounders **Domain 3: Outcome** 1. b) record linkage 2. a) yes – 12 month follow-up 3. a) complete follow up - all subjects accounted for **Overall quality is assessed as poor**

Study details	Population	Interventions	Study outcomes	Appraisal and Funding
a year on gender-affirming hormones. **Study dates** 2011 to 2017	the ICD-10 is 'transsexualism'. Mean age at diagnosis 18.1 years (range 15.2 to 19.9)		assessment and 6% (3/52) needed treatment during the 12-month 'real life' phase (no statistically significant difference, p= 0.18) • psychotic symptoms/psychosis: 2% (1/52) needed treatment before or during the assessment and 4% (2/52) needed treatment during the 12-month 'real life' phase (no statistically significant difference, p= 0.56) • substance abuse: 4% (2/52) needed treatment before or during the assessment and 2% (1/52) needed treatment during the 12-month 'real life' phase (no statistically significant difference, p= 0.56) • autism: 12% (6/52) needed treatment before or during the assessment and 6% (3/52) needed treatment during the 12-month 'real life' phase (no statistically significant difference, p= 0.30) • ADHD: 10% (5/52) needed treatment before or during the assessment and 2% (1/52) needed treatment during the 12-month 'real life' phase (no statistically significant difference, p= 0.09) • eating disorder: 2% (1/52) needed treatment before or during the assessment and 2% (1/52) needed treatment during the 12-month 'real life' phase (no statistically significant difference, p= 1.0). No details of actual treatment reported. **Important Outcomes** ***Psychosocial Impact*** Study reported on measures of functioning in different domains of adolescent development,	Other comments: None Source of funding: No source of funding reported

			reported over the approximately 12-month period after starting gender-affirming hormones (referred to as the 'real-life phase' in Finland) Significantly fewer participants were living with parent(s)/ guardians during the real-life phase (40%; 21/50) compared with during gender identity assessment (73%; 38/52; p=0.001)) There was a statistically significant reduction in the number of participants with normative peer contacts, from gender identity assessment (89%; 46/52) to the real-life phase (81%; 42/52; p<0.001). There was no significant difference in the number of participants who were progressing normally in school or work during gender identity assessment (64%; 33/52) compared with the real-life phase (60%; 31/52). There was no significant difference in the number of participants who have been dating or were in steady relationships during gender identity assessment (62%; 32/50) compared with the real-life phase (58%; 30/52). There was no significant difference in the number of participants who were able to deal with matters outside of the home in an age-appropriate manner during gender identity assessment (81% (42/52) compared with the real-life phase (81%; 42/52)	

Study details	Population	Interventions	Study outcomes	Appraisal and Funding
			No other critical or important outcomes reported	

Study details	Population	Interventions	Study outcomes	Appraisal and Funding

Full citation
Khatchadourian K, Amed S, Metzger DL (2014) Clinical management of youth with gender dysphoria in Vancouver. The Journal of pediatrics 164(4): 906-11
 	
Study location
Single centre study, Vancouver, Canada
 	
Study type
Retrospective chart review
 	
Study aim
To describe the patient characteristics, clinical management, and response to treatment in a cohort of people seen in a single clinic.
 	
Study dates
1998 to 2011
	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.
	
	63 children, adolescents and young people with gender dysphoria who started gender-affirming hormones, out of 84 young people seen in the unit between 1998 and 2011.
	39 transfemales and 24 transmales.
	
	Diagnostic criteria for gender dysphoria not reported.
	
	Mean age at the start of gender-affirming hormone treatment was 17.4 years (SD 1.9).
		Intervention
		Transfemales: Oestrogen (oral micronized 17β-oestradiol)
		Transmales: Testosterone (injectable testosterone enanthate and/or cypionate)
		
		19 participants (30%) had previously received a GnRH analogue. The median time from start of GnRH analogue to start of gender-affirming hormones was 11.3 months (range 2.2 to 42.0). 11 participants continued GnRH analogues after starting gender-affirming hormones.
		
		Average duration of treatment with a GnRH analogue not reported
		
		Comparison
		No comparator
			Critical Outcomes
			
			No critical outcomes assessed.
			
			Important outcomes
			
			Safety
			Of the 63 participants who received gender-affirming hormones:
			• No participants permanently discontinued gender-affirming hormones
			• 3 participants (5%) temporarily discontinued treatment:
			  ○ 2 transmales due to concomitant mental health comorbidities
			  ○ 1 transmale due to androgenic alopecia.
			  ○ No transfemale stopped treatment.
			The authors report that all patients eventually restarted gender-affirming hormones, although they do not report how long treatment was
				This study was appraised using the Newcastle-Ottawa tool for cohort studies.
				
				Domain 1: Selection domain
				1. b) somewhat representative
				2. c) no-non exposed cohort
				3. a) secure record*
				4. b) no
				Domain 2: Comparability
				1. c) cohorts are not comparable on the basis of the design or analysis controlled for confounders
				Domain 3: Outcome
				1. b) record linkage
				2. b) no – although follow-up time is reported for patients with more than 1 clinic visit, duration of treatment with gender-affirming hormones is not reported
				3. c) incomplete - missing data
				
				Overall quality is assessed as poor

Study details	Population	Interventions	Study outcomes	Appraisal and Funding
			stopped for, or what the effect of stopped treatment was. • No participants reported major complications • 12 participants (19%) had minor complications: ○ 7 transmales had severe acne (requiring isotretinoin) ○ 1 transmale had andogenic alopecia ○ 3 transmales had mild dyslipidaemia (levels not reported) ○ 1 transmale had significant mood swings ○ No transfemales had minor complications	Other comments: Mental health comorbidity was reported for all participants but not for the gender-affirming hormone cohort separately. Concomitant use of other medicines was not reported. Source of funding: No source of funding identified.

Study details	Population	Interventions	Study outcomes	Appraisal and Funding
Full citation 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) **Study location** Single centre, Amsterdam, Netherlands	Participants were included if i) they had started GnRH analogue treatment before 18 years, ii) if whole body dual-energy radiograph absorptiometry was performed at least once during treatment (4 months before or after the start of GnRH analogues or gender-affirming hormones, or	Transfemales: Oestrogen (17-β oestradiol [E2]) orally, starting with 5 mcg/kg body weight per day, which was increased every 6 months until the maintenance dose of 2 mg per day was reached. Transmales: mixed testosterone esters (Sustanon), 25 mg/m^2 body surface area every 2 weeks intramuscularly,	**Critical Outcomes** No critical outcomes assessed. **Important outcomes** ***Safety*** Safety outcomes reported separately for transfemales and transmales. **For transfemales**, from the start of gender-affirming hormone treatment to age 22 years: • Mean BMI statistically significantly increased (mean change +1.9, 95% CI 0.6 to 3.2, p<0.005; mean BMI at	This study was appraised using the Newcastle-Ottawa tool for cohort studies. **Domain 1: Selection domain** 1. b) somewhat representative 2. c) no-non exposed cohort 3. a) secure record* 4. b) no **Domain 2: Comparability** 1. c) cohorts are not comparable on the basis

Study details	Population	Interventions	Study outcomes	Appraisal and Funding
Study type Retrospective chart review **Study aim** To examine the effects of treatment on changes in cardiovascular risk factors, including BMI, blood pressure, insulin sensitivity, and lipid levels. **Study dates** 1998-2015	within 1.5 years before or after the 22nd birthday), iii) if they were likely to have had at least 1 medical consultation in young adulthood. The study included 192 young people with dysphoria who met the above inclusion criteria: 71 transfemales and 121 transmales. Gender dysphoria was diagnosed according to the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition criteria. Mean age at the start of gender-affirming hormones was 16.4 years (SD 1.1) for transfemales and 16.9 years (SD 0.9) for transmales.	increased every 6 months to maintenance dose of 250 mg every 3 to 4 weeks. When GnRH analogues were started after the age of 16 years a different hormone starter dose was used (1 mg oestrogen daily and 75 mg testosterone weekly). Median (IQR) duration of GnRH analogue (monotherapy) was 2.1 years (1.0 to 2.7) in transfemales and 1.0 (0.5 to 2.9) for transmales.	22 years= 23.2, 95% CI 21.6 to 24.8). At age 22 years, 9.9% of the cohort were obese, compared with 3.0% in reference cisgender population[1]. • Mean systolic blood pressure (SBP) did not significantly change (mean change -3 mmHg, 95% CI -8 to 2; mean SBP at 22 years= 117 mmHg, 95% CI 113 to 122) • Mean diastolic blood pressure (DBP) statistically significantly increased (mean change +6 mmHg, 95% CI 3 to 10, p<0.001; mean DBP at 22 years= 75 mmHg, 95% CI 72 to 78) • Mean glucose level did not significantly change (mean change +0.1 mmol/L, 95% CI -0.1 to 0.2; mean glucose level at 22 years= 5.0 mmol/L, 95% CI 4.8 to 5.1) • Mean insulin level did not significantly change (mean change +2.7 mU/L, 95% CI -1.7 to 7.1; mean insulin level at 22 years= 5.0 mU/L (4.8 to 5.1) • Insulin resistance (mean Homeostatic Model Assessment of Insulin Resistance [HOMA-IR]) did not significantly change (mean change +0.7, 95% CI -0.2 to 1.5; mean HOMA-IR at 22 years 2.9, 95% CI 1.9 to 3.9) • Mean total cholesterol did not significantly change (mean change +0.1 mmol/L, 95% CI -0.2 to 0.4; mean total cholesterol at 22 years 4.1 mmol/L, 95% CI 3.8 to 4.4) • Mean HDL cholesterol did not significantly change (mean change +0.0 mmol/L, 95% CI -0.1 to 0.2; mean HDL cholesterol at 22 years 1.6 mmol/L, 95% CI 1.4 to 1.7) • Mean LDL cholesterol did not significantly change (mean change +0.0 mmol/L, 95%	of the design or analysis controlled for confounders **Domain 3: Outcome** 1. b) record linkage 2. a) yes- follow-up from start of gender-affirming hormones to age 22 years, around 5 years 3. a) complete follow up - all subjects accounted for **Overall quality is assessed as poor** Other comments: None Source of funding: No external funding

			CI -0.3 to 0.2; mean LDL cholesterol at 22 years 2.0 mmol/L, 95% CI 1.8 to 2.3) • Mean triglycerides statistically significantly increased (mean change +0.2 mmol/L, 95% CI 0.0 to 0.5, p<0.05; triglyceride level at 22 years 1.1 mmol/L, 95% CI 0.9 to 1.4) **For transmales**, from the start of gender-affirming hormone treatment to age 22 years: • Mean BMI statistically significantly increased (mean change +1.4, 95% CI 0.8 to 2.0, p<0.005; mean BMI at 22 years= 23.9, 95% CI 23.0 to 24.7). At age 22 years, 6.6% of the cohort were obese, compared with 2.2% in reference cisgender population[1]. • Mean systolic blood pressure (SBP) statistically significantly increased (mean change +5 mmHg, 95% CI 1 to 9; mean SBP at 22 years= 126 mmHg, 95% CI 122 to 130) • Mean diastolic blood pressure (DBP) statistically significantly increased (mean change +6 mmHg, 95% CI 4 to 9, p<0.001; mean DBP at 22 years= 74 mmHg, 95% CI 72 to 77) • Mean glucose level did not significantly change (mean change 0.0 mmol/L, 95% CI -0.2 to 0.2; mean glucose level at 22 years= 4.8 mmol/L, 95% CI 4.7 to 5.0) • Mean insulin level statistically significantly decreased (mean change -2.1 mU/L, 95% CI -3.9 to -0.3, p<0.05; mean insulin level at 22 years= 8.6 mU/L (6.9 to 10.2) • Insulin resistance (mean Homeostatic Model Assessment of Insulin Resistance [HOMA-IR]) statistically significantly	

Study details	Population	Interventions	Study outcomes	Appraisal and Funding
			decreased (mean change -0.5, 95% CI -1.0 to -0.1, p<0.05; mean HOMA-IR at 22 years 1.8, 95% CI 1.4 to 2.2) • Mean total cholesterol statistically significantly increased (mean change +0.4 mmol/L, 95% CI 0.2 to 0.6, p<0.001; mean total cholesterol at 22 years 4.6 mmol/L, 95% CI 4.3 to 4.8) • Mean HDL cholesterol statistically significantly decreased (mean change -0.3 mmol/L, 95% CI -0.4 to -0.2, p<0.001; mean HDL cholesterol at 22 years 1.3 mmol/L, 95% CI 1.2 to 1.3) • Mean LDL cholesterol statistically significantly increased (mean change +0.4 mmol/L, 95% CI 0.2 to 0.6, p<0.001; mean LDL cholesterol at 22 years 2.6 mmol/L, 95% CI 2.4 to 2.8) • Mean triglycerides statistically significantly increased (mean change +0.5 mmol/L, 95% CI 0.3 to 0.7, p<0.001; triglyceride level at 22 years 1.3 mmol/L, 95% CI 1.1 to 1.5)	

[1] Reference population taken from Fredriks et al. (2000)

Study details	Population	Interventions	Study outcomes	Appraisal and Funding
Full citation 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 **Study location** Single centre, Amsterdam, Netherlands **Study type** Retrospective longitudinal study **Study aim** To assess peak bone mass in young adults with gender dysphoria who had received GnRH analogues and gender-affirming hormones during their pubertal years. **Study dates**	34 young people with gender dysphoria who received GnRH analogues, gender-affirming hormones and gonadectomy. The study included 15 transfemales and 19 transmales; mean age at start of gender-affirming hormones was 16.6 years (SD 1.4) and 16.4 years (SD 2.3) respectively. Participants were required to meet the DSM-IV-TR criteria for gender identity disorder of adolescence. Participants were included if they had undergone gonadectomy between June 1998 and August 2012, and they were at least 21 years old when they had the surgery. Bone mineral density data were also required at the start of GnRH analogue, gender-affirming hormones and at the age of 22 years. No concomitant treatments were reported.	**Intervention** Transfemales - oral 17-β oestradiol (incremental dosing) Transmales – IM testosterone (Sustanon 250 mg/ml; incremental dosing) Median duration of treatment with gender-affirming hormones for transfemales was 5.8 years (range 3.0 to 8.0) and for transmales was 5.4 years (range 2.8 to 7.8). The GnRH analogue was SC triptorelin 3.75 mg every 4 weeks. No details of gonadectomy reported. **Comparison** No comparison group. Comparison over time reported.	**Critical outcomes** No critical outcomes reported **Important outcomes** **Safety** ***Bone density: lumbar spine*** **Lumbar spine bone mineral apparent density (BMAD)** Change from starting gender-affirming hormones to age 22 years in transfemales- Mean (SD); g/m^3 • Start of gender-affirming hormones: 0.22 (0.02) • Age 22 years: 0.23 (0.03) • p=0.003 z-score (range) • Start of gender-affirming hormones: -0.90 (0.80) • Age 22 years: -0.78 (1.03) • No statistically significant difference Change from starting gender-affirming hormones to age 22 years in transmales- Mean (SD); g/m^3 • Start of gender-affirming hormones: 0.24 (0.02) • Age 22 years: 0.25 (0.28 • p=0.001 z-score (SD) • Start of gender-affirming hormones: -0.50 (0.81) • Age 22 years: -0.033 (0.95) • p=0.002	This study was appraised using the Newcastle-Ottawa tool for cohort studies. **Domain 1: Selection domain** 1. b) somewhat representative 2. c) no-non exposed cohort 3. a) secure record* 4. b) no **Domain 2: Comparability** 1. c) cohorts are not comparable on the basis of the design or analysis controlled for confounders **Domain 3: Outcome** 1. b) record linkage 2. a) yes – mean duration of gender-affirming hormone treatment was 5.8 and 5.4 years. 3. c) follow-up rate variable across timepoints and no description of those lost **Overall quality is assessed as poor** Other comments: Within person comparison. Small numbers of participants in each subgroup. No

Study details	Population	Interventions	Study outcomes	Appraisal and Funding
Gonadectomy took place between June 1998 and August 2012	At the start of gender-affirming hormone treatment, in the transfemale subgroup the median Tanner P was 4 (IQR 2) and the median Tanner G was 12 (IQR 11). In the transmale subgroup the median Tanner B was 5 (IQR 2) and the median Tanner P was 5 (IQR 0).		**Lumbar spine bone mineral density (BMD)** Change from starting gender-affirming hormones to age 22 years in transfemales-Mean (SD); g/m^2 • Start of gender-affirming hormones: 0.84 (0.11) • Age 22 years: 0.93 (0.10) • p<0.001 z-score (range) • Start of gender-affirming hormones: -1.01 (0.98) • Age 22 years: -1.36 (0.83) • No statistically significant difference Change from starting gender-affirming hormones to age 22 years in transmales-Mean (SD); g/m^2 • Start of gender-affirming hormones: 0.91 (0.10) • Age 22 years: 0.99 (0.13) • P<0.001 z-score (range) • Start of gender-affirming hormones: -0.72 (0.99) • Age 22 years: -0.33 (1.12) • No statistically significant difference ***Bone density: femoral region, nondominant side*** **Femoral region, nondominant side BMAD** Change from starting gender-affirming hormones to age 22 years in transfemales-Mean (SD); g/m^3 • Start of gender-affirming hormones: 0.26 (0.04) • Age 22 years: 0.28 (0.05)	concomitant treatments or comorbidities were reported. Source of funding: None disclosed

Study details	Population	Interventions	Study outcomes	Appraisal and Funding
			• No statistically significant difference z-score (SD) • Start of gender-affirming hormones: -1.57 (1.74) • Age 22 years: Not reported • No statistical analysis reported Change from starting gender-affirming hormones to age 22 years in transmales- Mean (SD); g/m^3 • Start of gender-affirming hormones: 0.31 (0.04) • Age 22 years: 0.33 (0.05) • p=0.010 z-score (SD) • Start of gender-affirming hormones: -0.28 (0.74) • Age 22 years: Not reported • No statistical analysis reported **Femoral region, nondominant side BMD** Change from starting gender-affirming hormones to age 22 years in transfemales- Mean (SD); g/m^2 • Start of gender-affirming hormones: 0.87 (0.08) • Age 22 years: 0.94 (0.11) • P=0.009 z-score (SD) • Start of gender-affirming hormones: -0.95 (0.63) • Age 22 years: -0.69 (0.74) • No statistically significant difference Change from starting gender-affirming hormones to age 22 years in transmales- Mean (SD); g/m^2 • Start of gender-affirming hormones: 0.88 (0.09) • Age 22 years: 0.95 (0.10)	

Study details	Population	Interventions	Study outcomes	Appraisal and Funding
			• P<0.001 z-score (SD) • Start of gender-affirming hormones: -0.35 (0.79) • Age 22 years: -0.35 (0.74) • p=0.006	

Study details	Population	Interventions	Study outcomes	Appraisal and Funding
Full citation Kuper, Laura E, Stewart, Sunita, Preston, Stephanie et al. (2020) Body Dissatisfaction and Mental Health Outcomes of Youth on Gender-Affirming Hormone Therapy. Pediatrics 145(4) **Study location** Single centre, Texas, USA **Study type** Prospective longitudinal study **Study aim** To: • explore how baseline body dissatisfaction, depression, and anxiety symptoms vary by gender,	148 children and adolescents with gender dysphoria, n=148, of whom: • 25 received puberty suppression only • 93 received gender-affirming hormone therapy only • 30 received both Results for treatments reported separately. Mean age at initial assessment was 15.4 years (range 9 to 18). Mean age at start of gender-affirming hormone therapy was 16.2 years (range 13.2 to 18.6). All participants met the Diagnostic and Statistical	Hormone therapy, guided by Endocrine Society Clinical Practice Guidelines Follow-up at least 18 months from initial assessment at the clinic. Mean duration of gender-affirming hormone therapy before follow-up was 10.9 months (range 1 to 18; SD 3.3)	**Critical Outcomes** ***Impact on mental health*** Mean depression score, assessed using the Quick Inventory of Depressive Symptoms (QIDS), self-reported was 9.6 (SD 5.0) at baseline and 7.4 (SD 4.5) at follow-up. The authors did not present statistical analysis for the sub-group of participants receiving gender-affirming hormones and it is unclear whether the change in score was statistically significant. Mean depression score, assessed using the QIDS, clinician-reported was 5.9 (SD 4.1) at baseline and 6.0 (SD 3.8) at follow-up. The authors did not present statistical analysis for the sub-group of participants receiving gender-affirming hormones and it is unclear whether the change in score was statistically significant. Mean anxiety score, assessed using the Screen for Child Anxiety Related Emotional Disorders (SCARED) questionnaire was 32.6 (SD 16.3) at baseline and 28.4 (SD 15.9) at	This study was appraised using the Newcastle-Ottawa tool for cohort studies. **Domain 1: Selection domain** 1. b) somewhat representative 2. c) no-non exposed cohort 3. a) secure record 4. b) no **Domain 2: Comparability** 1. c) cohorts are not comparable on the basis of the design or analysis controlled for confounders **Domain 3: Outcome** 1. d) assessors not blinded to treatment 2. a) yes – follow-up at least 18 months from initial assessment. Mean duration of gender-affirming hormone

Study details	Population	Interventions	Study outcomes	Appraisal and Funding
age at initial assessment, and Tanner stage at first medical visit • examine how body dissatisfaction, depression, and anxiety symptoms change over the first year of gender-affirming hormone treatment • explore how any changes vary by affirmed gender, Tanner stage, age, type of treatment, months on gender-affirming hormone therapy, mental health treatment received, and whether chest surgery was also obtained (among transmales). **Study dates** Initial participant assessments took place between August 2014 and March 2018.	Manual of Mental Disorders, Fifth Edition criteria for gender dysphoria. Specific inclusion and exclusion criteria for the study are not reported. It would appear that all children and adolescents eligible for gender-affirming hormones were considered eligible for the study. The authors state that before initial assessment with a psychologist, psychiatrist, and/or clinical therapist, parents completed a phone intake survey. Around one-third of families did not follow-up after the phone intake.		follow-up. The authors did not present statistical analysis for the sub-group of participants receiving gender-affirming hormones and it is unclear whether the change in score was statistically significant. Mean panic score, assessed using specific questions from the SCARED questionnaire was 8.1 (SD 6.3) at baseline and 7.1 (SD 6.5) at follow-up. The authors did not present statistical analysis for the sub-group of participants receiving gender-affirming hormones and it is unclear whether the change in score was statistically significant. Mean generalised anxiety score, assessed using specific questions from the SCARED questionnaire was 10.0 (SD 5.1) at baseline and 8.8 (SD 6.5) at follow-up. The authors did not present statistical analysis for the sub-group of participants receiving gender-affirming hormones and it is unclear whether the change in score was statistically significant. Mean social anxiety score, assessed using specific questions from the SCARED questionnaire was 8.5 (SD 4.1) at baseline and 7.7 (SD 4.2) at follow-up. The authors did not present statistical analysis for the sub-group of participants receiving gender-affirming hormones and it is unclear whether the change in score was statistically significant. Mean separation anxiety score, assessed using specific questions from the SCARED	treatment was 10.9 months. 3. c) patient numbers vary by outcome with no explanation **Overall quality is assessed as poor** Other comments: None Source of funding: Supported by Children's Health. The Research Electronic Data Capture database was funded by the Clinical and Translational Science Awards program

			questionnaire was 3.5 (SD 3.0) at baseline and 3.1 (SD 2.5) at follow-up. The authors did not present statistical analysis for the sub-group of participants receiving gender-affirming hormones and it is unclear whether the change in score was statistically significant. Mean school avoidance score, assessed using specific questions from the SCARED questionnaire was 2.6 (SD 2.1) at baseline and 2.0 (SD 2.0) at follow-up. The authors did not present statistical analysis for the sub-group of participants receiving gender-affirming hormones and it is unclear whether the change in score was statistically significant. The authors also reported results separately for transfemales and transmales: **Transfemales** No statistical analyses were reported for this sub-group and it is unclear whether any changes in score were statistically significant. • Mean depression symptoms, assessed using the QIDS, self-reported was 7.5 (SD 4.9) at baseline and 6.6 (SD 4.4) at follow-up. • Mean depression symptoms, assessed using the QIDS, clinician-reported was 4.2 (SD 3.2) at baseline and 5.4 (SD 3.4) at follow-up. • Mean anxiety symptoms, assessed using the SCARED questionnaire was 26.4 (SD 14.2) at baseline and 24.3 (SD 15.4) at follow-up.	

			• Mean panic symptoms, assessed using specific questions from the SCARED questionnaire was 5.7 (SD 4.9) at baseline and 5.1 (SD 4.9) at follow-up. • Mean generalised anxiety symptoms, assessed using specific questions from the SCARED questionnaire was 8.6 (SD 5.1) at baseline and 8.0 (SD 5.1) at follow-up. • Mean social anxiety symptoms, assessed using specific questions from the SCARED questionnaire was 7.1 (SD 3.9) at baseline and 6.8 (SD 4.4) at follow-up. • Mean separation anxiety symptoms, assessed using specific questions from the SCARED questionnaire was 3.4 (SD 3.3) at baseline and 2.7 (SD 2.3) at follow-up. • Mean school avoidance symptoms, assessed using specific questions from the SCARED questionnaire was 1.8 (SD 1.7) at baseline and 1.9 (SD 2.1) at follow-up. **Transmales** No statistical analyses were reported for this sub-group and it is unclear whether any changes in score were statistically significant. • Mean depression symptoms, assessed using the QIDS, self-reported was 10.4 (SD 5.0) at baseline and 7.5 (SD 4.5) at follow-up. • Mean depression symptoms, assessed using the QIDS, clinician-reported was 6.7 (SD 4.4) at baseline and 6.2 (SD 4.1) at follow-up. • Mean anxiety symptoms, assessed using the SCARED questionnaire was 35.4 (SD	

			16.5) at baseline and 29.8 (SD 15.5) at follow-up. • Mean panic symptoms, assessed using specific questions from the SCARED questionnaire was 9.3 (SD 6.5) at baseline and 7.9 (SD 6.5) at follow-up. • Mean generalised anxiety symptoms, assessed using specific questions from the SCARED questionnaire was 10.4 (SD 5.0) at baseline and 9.0 (SD 5.1) at follow-up. • Mean social anxiety symptoms, assessed using specific questions from the SCARED questionnaire was 8.5 (SD 4.0) at baseline and 7.8 (SD 4.1) at follow-up. • Mean separation anxiety symptoms, assessed using specific questions from the SCARED questionnaire was 4.2 (SD 3.4) at baseline and 3.4 (SD 2.6) at follow-up. • Mean school avoidance symptoms, assessed using specific questions from the SCARED questionnaire was 2.6 (SD 2.1) at baseline and 2.0 (SD 2.0) at follow-up. No difference in impact on mental health found by Tanner age. Numerical results, statistical analysis and information on specific outcomes not reported. It is unclear from the paper whether Tanner age is at initial assessment, start of GnRH analogues, start of gender-affirming hormones, or another timepoint. **Important Outcomes** ***Impact on body image***	

Study details	Population	Interventions	Study outcomes	Appraisal and Funding
			Mean Body Image Scale (BIS) score was 70.7 (SD 15.2) at baseline and 51.4 (SD 18.3) at follow-up. The authors do not present statistical analysis for this population and it is unclear whether the change in score was statistically significant. The authors also reported body image results separately for transfemales and transmales. No statistical analyses were reported for this sub-groups and it is unclear whether changes in score were statistically significant. • In transfemales, BIS score was 67.5 (SD 19.5) at baseline and 49.0 (SD 21.6) at follow-up. • In transmales, BIS score was 71.1 (SD 13.4) at baseline and 52.9 (SD 16.8) at follow-up. No difference in body image score found by Tanner age. Numerical results, statistical analysis and information on specific outcomes not reported. It is unclear from the paper whether Tanner age is at initial assessment, start of GnRH analogues, start of gender-affirming hormones, or another timepoint. *No other critical or important outcomes reported*	

Study details	Population	Interventions	Study outcomes	Appraisal and Funding
Study dates Lopez de Lara, D., Perez Rodriguez, O., Cuellar Flores, I. et al. (2020) Psychosocial assessment in transgender adolescents. Anales de Pediatria **Study location** Single centre in Madrid, Spain **Study type** Prospective analytical study **Study aim** To assess the psychosocial status of patients seeking care in the paediatric endocrinology clinic for gender dysphoria, and the impact on psychosocial status of gender-affirming hormone therapy at 12 months of treatment **Study dates** Not reported	23 adolescents with gender dysphoria; 16 transmale and 7 transfemale. Participants were required to be at a stage of pubertal development of Tanner 2 or higher. People with mental health comorbidity that could affect the experience of gender dysphoria were excluded. Mean age at baseline was 16 years (range 14 to 18). 30 cisgender controls, matched for age, ethnicity, and socioeconomic status	Gender-affirming hormones- • Oral oestradiol • Intramuscular testosterone Participants had previously received gonadotropin-releasing hormone (GnRH) analogues in the intermediate pubertal stages (Tanner 2---3).	**Critical Outcomes** ***Impact on gender dysphoria*** Following gender-affirming hormones for 12 months, mean (±SD) Utrecht Gender Dysphoria Scale (UGDS) score statistically significantly improved, from 57.1 (±4.1) at baseline to 14.7 (±3.2; p<0.001) ***Impact on mental health*** Mean depression score statistically significantly improved following treatment with gender-affirming hormones. Mean Beck Depression Inventory II (BDI-II) score (±SD) reduced from 19.3 points (±5.5) at baseline to 9.7 points (±3.9) at 12 months (p<0.001). Mean anxiety scores statistically significantly improved following treatment with gender-affirming hormones. Mean (±SD) State-Trait Anxiety Inventory (STAI) State subscale score improved from 33.3 points (±9.1) at baseline to 16.8 points (±8.1) at 12 months (p<0.001). Mean (±SD) State-Trait Anxiety Inventory (STAI) Trait subscale score improved from 33.0 points (±7.2) at baseline to 18.5 points (±8.4) at 12 months (p<0.001). **Important Outcomes** ***Psychosocial Impact*** There was not change in family functioning, measured using the Family APGAR test, from baseline (17.9 points) to 1 year after starting	This study was appraised using the Newcastle-Ottawa tool for cohort studies. **Domain 1: Selection domain** 1. b) somewhat representative 2. Not applicable – although a control group is reported on, people in this group did not have gender dysphoria. 3. a) secure record* 4. b) no **Domain 2: Comparability** 1. Not applicable – although a control group is reported on, people in this group did not have gender dysphoria. **Domain 3: Outcome** 1. d) assessors not blinded to treatment 2. a) yes – 12 months treatment with gender-affirming hormones 3. a) complete follow up - all subjects accounted for **Overall quality is assessed as poor**

Study details	Population	Interventions	Study outcomes	Appraisal and Funding
			gender-affirming hormones (18.0 points; no statistical analysis reported). Results from the Strengths and Difficulties Questionnaire, Spanish Version (SDQ-Cas) showed statistically significant improvements from baseline (14.7 points; SD±3.3) to 12 months after gender-affirming hormones (10.3 points; SD±2.9; p<0.001) *No other critical or important outcomes reported*	Other comments: None Source of funding: Not reported

Study details	Population	Interventions	Study outcomes	Appraisal and Funding
Full citation Stoffers, Iris E; de Vries, Martine C; Hannema, Sabine E (2019) Physical changes, laboratory parameters, and bone mineral density during testosterone treatment in adolescents with gender dysphoria. The journal of sexual medicine 16(9): 1459-1468 **Study location** Single centre, Leiden, Netherlands **Study type** Retrospective chart review **Study aim** To report changes in height, BMI, blood pressure, laboratory parameters and bone density. **Study dates** November 2010 to August 2018	62 transmales with gender dysphoria. participants were required to have been receiving testosterone therapy for at least 6 months. Further inclusion or exclusion criteria not reported. Gender dysphoria was diagnosed according to the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition criteria. .	Testosterone intramuscular injection (Sustanon 250 mg). Dose escalated every 6 months up to the standard adult dose of 125 mg every 2 weeks or 250 mg every 3-4 weeks. A more rapid dose escalation was using in patients who started GnRH analogue treatment at 16 years or older. Median age at start of testosterone treatment was 17.2 years (range 14.9 to 18.4) Median duration of testosterone treatment was 12 months (range 5 to 33) Median duration of GnRH analogue treatment was 8 months (range 3 to 39)	**Critical Outcomes** No critical outcomes assessed. **Important outcomes** ***Safety*** ***Bone mineral density (BMD): lumbar spine*** There was no statistically significant difference in lumber spine bone mineral density (BMD) from start of testosterone treatment to any timepoint, up to 24 months follow-up. Mean (±SD), g/cm^2: • Start of testosterone: 0.90 (±0.11) • 6 months: 0.94 (±0.10) • 12 months: 0.95 (±0.09) • 24 months: 0.95 (±0.11) z-score (±SD): • Start of testosterone: -0.81 (±1.02) • 6 months: -0.67 (±0.95) • 12 months: -0.66 (±0.81) • 24 months: -0.74 (±1.17) ***Bone mineral density (BMD): femoral neck (hip)*** There was no statistically significant difference in right or left femoral neck (hip) bone mineral density (BMD) from start of	This study was appraised using the Newcastle-Ottawa tool for cohort studies. **Domain 1: Selection domain** 1. b) somewhat representative 2. c) no-non exposed cohort 3. a) secure record* 4. b) no **Domain 2: Comparability** 1. c) cohorts are not comparable on the basis of the design or analysis controlled for confounders **Domain 3: Outcome** 1. b) record linkage 2. a) yes – mean duration of gender-affirming hormone treatment was 5.8 and 5.4 years. 3. a) complete follow up - all subjects accounted for **Overall quality is assessed as poor** Other comments: None Source of funding: None

			testosterone treatment to any timepoint, up to 24 months follow-up.
			Right
			Mean (±SD), g/cm^2:
			• Start of testosterone: 0.77 (±0.08)
			• 6 months: 0.84 (±0.11)
			• 12 months: 0.82 (±0.08)
			• 24 months: 0.85 (±0.11)
			z-score (±SD):
			• Start of testosterone: -0.97 (0.79)
			• 6 months: -0.54 (±0.96)
			• 12 months: -0.80 (±0.69)
			• 24 months: -0.31 (±0.84)
			Left
			Mean (±SD), g/cm^2:
			• Start of testosterone: 0.76 (±0.09)
			• 6 months: 0.83 (±0.12)
			• 12 months: 0.81 (±0.08)
			• 24 months: 0.86 (±0.09)
			z-score (±SD):
			• Start of testosterone: -1.07 (0.85)
			• 6 months: -0.62 (±1.12)
			• 12 months: -0.93 (±0.63)
			• 24 months: -0.20 (±0.70)
			
			Other safety-related outcomes
			• Alkaline phosphatase: statistically significant increases observed from start of testosterone treatment to 6 months and 12 months ($p<0.001$), although difference at 24 months was not statistically significant. Median (IQR), U/L
			  ○ Start of testosterone: 102 (78 to 136)
			  ○ 6 months: 115 (102 to 147)
			  ○ 12 months: 112 (88 to 143)
			  ○ 24 months: 81 (range 69 to 98)
			• Creatinine: statistically significant increases observed from start of	

Study details	Population	Interventions	Study outcomes	Appraisal and Funding
			testosterone treatment to 6, 12 and 24 months (p<0.001). Mean (±SD), umol/L ○ Start of testosterone: 62 (±7) ○ 6 months: 70 (±9) ○ 12 months: 74 (±10) ○ 24 months: 81 (±10) There was no statistically significant change from start of testosterone treatment in: • HbA1c • Aspartate aminotransferase (AST) • Alanine aminotransferase (ALT) • Gamma-glutamyl transferase • Urea Numerical results, follow-up duration and further details of statistical analysis not reported.	

Study details	Population	Interventions	Study outcomes	Appraisal and Funding
Full citation Vlot MC, Klink DT, den Heijer M 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 **Study location** Single centre, Amsterdam, Netherlands **Study type** Retrospective chart review **Study aim** To investigate the impact of GnRH analogues and gender-affirming hormones on bone mineral apparent density (BMAD) in transgender adolescents. The study also report on levels of bone turnover markers, although the authors concluded that the	70 adolescents with gender dysphoria (42 transmales and 28 transfemales). Median age (range) at the start of gender-affirming hormones was 16.3 years (15.9 to 19.5) for transmales and 16.0 years (14.0 to 18.9) for transfemales. Participants were included if they had a diagnosis of gender dysphoria according to DSM-IV-TR criteria who received GnRH analogues and then gender-affirming hormones. No concomitant treatments were reported. The study categorised 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	Transfemales: Oestradiol oral Dose escalated every 6 months until standard adult dose of 2 mg daily was reached Transmales: Testosterone intramuscular injection (Sustanon 250 mg). Dose escalated every 6 months up to the standard adult dose of 250 mg every 4 weeks or 250 mg every 3-4 weeks. All participants previously received a GnRH analogue (triptorelin 3.75 mg subcutaneously every 4 weeks) Median duration of GnRH analogue therapy not reported.	**Critical outcomes** No critical outcomes reported **Important outcomes** ***Bone density: lumbar spine*** **Lumbar spine bone mineral apparent density (BMAD)** Transfemales (bone age <15 years), change from starting gender-affirming hormones to 24 months follow-up. Median (range), g/m^3 • Start of gender-affirming hormones (C0): 0.20 (0.18 to 0.24) • 24-month follow-up (C24): 0.22 (0.19 to 0.27) • Statistically significant increase (p≤0.01) z-score (range) • Start of gender-affirming hormones (C0): -1.52 (-2.36 to 0.42) • 24-month follow-up (C24): • Statistically significant increase (p≤0.05) Transfemales (bone age ≥15 years), change from starting gender-affirming hormones to 24 months follow-up. Median (range), g/m^3 • Start of gender-affirming hormones: 0.22 (0.19 to 0.24) • 24-months: 0.23 (0.21 to 0.26) • Statistically significant increase (p≤0.05) z-score (range) • Start of gender-affirming hormones: -1.15 (-2.21 to 0.08) • 24-months: -0.66 (-1.66 to 0.54)	This study was appraised using the Newcastle-Ottawa tool for cohort studies. **Domain 1: Selection domain** 1. b) somewhat representative 2. c) no-non exposed cohort 3. a) secure record* 4. b) no **Domain 2: Comparability** 1. c) cohorts are not comparable on the basis of the design or analysis controlled for confounders **Domain 3: Outcome** 1. b) record linkage 2. a) yes- 24 month follow-up 3. a) complete follow up - all subjects accounted for **Overall quality is assessed as poor.** Other comments: None Source of funding: grant from Abbott diagnostics

Study details	Population	Interventions	Study outcomes	Appraisal and Funding
added value of these seems to be limited. **Study dates** Participants started gender-affirming therapy between 2001 and 2011	group had a bone age of <15 years and the old transfemales group ≥15 years.		Statistically significant increase (p≤0.05) Transmales (bone age <14 years), change from starting gender-affirming hormones to 24 months follow-up. Median (range), g/m^3 • Start of gender-affirming hormones: 0.23 (0.19 to 0.28) • 24-months: 0.25 (0.22 to 0.28) • Statistically significant increase (p≤0.01) z-score (range) • Start of gender-affirming hormones: -0.84 (-2.2 to 0.87) • 24-months: -0.15 (-1.38 to 0.94) Statistically significant increase (p≤0.01) Transmales (bone age ≥14 years), change from starting gender-affirming hormones to 24 months follow-up. Median (range), g/m^3 • Start of gender-affirming hormones: 0.24 (0.20 to 0.28) • 24-months: 0.25 (0.21 to 0.30) • Statistically significant increase (p≤0.01) z-score (range) • Start of gender-affirming hormones: -0.29 (-2.28 to 0.90) • 24-months: -0.06 (-1.75 to 1.61) Statistically significant increase (p≤0.01) ***Bone density: femoral neck*** **Femoral neck BMAD** Transfemales (bone age <15 years), change from starting gender-affirming hormones to 24 months follow-up.	

Study details	Population	Interventions	Study outcomes	Appraisal and Funding
			Median (range), g/m^3 • Start of gender-affirming hormones: 0.27 (0.20 to 0.33) • 24-months: 0.27 (0.20 to 0.36) • No statistically significant change z-score (range) • Start of gender-affirming hormones: -1.32 (-3.39 to 0.21) • 24-months: -1.30 (-3.51 to 0.92) • No statistically significant change Transfemales (bone age ≥15 years), change from starting gender-affirming hormones to 24 months follow-up. Median (range), g/m^3 • Start of gender-affirming hormones: 0.30 (0.26 to 0.34) • 24-months: 0.29 (0.24 to 0.38) • No statistically significant change z-score (range) • Start of gender-affirming hormones: -0.36 (-1.50 to 0.46) • 24-months: -0.56 (-2.17 to 1.29) • No statistically significant change Transmales (bone age <14 years), change from starting gender-affirming hormones to 24 months follow-up. Median (range), g/m^3 • Start of gender-affirming hormones: 0.30 (0.22 to 0.35) • 24-months: 0.33 (0.23 to 0.37) • Statistically significant increase (p≤0.01) z-score (range) • Start of gender-affirming hormones: -0.37 (-2.28 to 0.47) • 24-months: -0.37 (-2.03 to 0.85)	

Study details	Population	Interventions	Study outcomes	Appraisal and Funding
			• Statistically significant increase (p≤0.01) Transmales (bone age ≥14 years), change from starting gender-affirming hormones to 24 months follow-up. • Start of gender-affirming hormones: 0.30 (0.23 to 0.41) • 24-months: 0.32 (0.23 to 0.41) • Statistically significant increase (p≤0.01) z-score (range) • Start of gender-affirming hormones: -0.27 ((-1.91 to 1.29) • 24-months: 0.02 (-2.1 to 1.35) • Statistically significant increase (p≤0.05)	

Appendix F Quality appraisal checklists

Newcastle-Ottawa Quality Assessment Form for Cohort Studies

Note: A study can be given a maximum of one star for each numbered item within the Selection and Outcome categories. A maximum of two stars can be given for Comparability.

Selection

1) Representativeness of the exposed cohort
    a) Truly representative (one star)
    b) Somewhat representative (one star)
    c) Selected group
    d) No description of the derivation of the cohort
2) Selection of the non-exposed cohort
    a) Drawn from the same community as the exposed cohort (one star)
    b) Drawn from a different source
    c) No description of the derivation of the non exposed cohort
3) Ascertainment of exposure
    a) Secure record (e.g., surgical record) (one star)
    b) Structured interview (one star)
    c) Written self report
    d) No description
    e) Other
4) Demonstration that outcome of interest was not present at start of study
    a) Yes (one star)
    b) No

Comparability

1) Comparability of cohorts on the basis of the design or analysis controlled for confounders
    a) The study controls for age, sex and marital status (one star)
    b) Study controls for other factors (list) _____ (one star)
    c) Cohorts are not comparable on the basis of the design or analysis controlled for confounders

Outcome

1) Assessment of outcome
    a) Independent blind assessment (one star)
    b) Record linkage (one star)
    c) Self report
    d) No description
    e) Other
2) Was follow-up long enough for outcomes to occur
    a) Yes (one star)
    b) No
    Indicate the median duration of follow-up and a brief rationale for the assessment above:_____
3) Adequacy of follow-up of cohorts
    a) Complete follow up- all subject accounted for (one star)

b) Subjects lost to follow up unlikely to introduce bias- number lost less than or equal to 20% or description of those lost suggested no different from those followed. (one star)
c) Follow up rate less than 80% and no description of those lost
d) No statement

Thresholds for converting the Newcastle-Ottawa scales to AHRQ standards (good, fair, and poor):

Good quality: 3 or 4 stars in selection domain AND 1 or 2 stars in comparability domain AND 2 or 3 stars in outcome/exposure domain

Fair quality: 2 stars in selection domain AND 1 or 2 stars in comparability domain AND 2 or 3 stars in outcome/exposure domain

Poor quality: 0 or 1 star in selection domain OR 0 stars in comparability domain OR 0 or 1 stars in outcome/exposure domain

Appendix G Grade profiles

Table 2: Question 1: For children and adolescents with gender dysphoria, what is the clinical effectiveness of treatment with gender-affirming hormones 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 patients		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result		
Impact on gender dysphoria (1 uncontrolled, prospective observational study)									
Change from baseline in mean gender dysphoria score, measured using the UGDS (duration of treatment 12 months). Higher scores indicate greater gender dysphoria.									
1 cohort study Lopez de Lara et al. 2020	Serious limitations[1]	No serious indirectness	No serious inconsistency	Not calculable	N=23	None	T0 (baseline) = 57.1 (SD 4.1) T1 (12 months) = 14.7 (SD 3.2) Statistically significant improvement, p<0.001	Critical	VERY LOW

Abbreviations: p: p-value; SD: standard deviation; UGDS: Utrecht Gender Dysphoria Scale

1 Downgraded 1 level - the cohort study by Lopez de Lara et al. 2020 was assessed 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 gender-affirming hormones 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		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result		
Impact on mental health (3 uncontrolled, prospective observational studies and 2 uncontrolled, retrospective observational studies)									
Change from baseline in mean depression score, measured using the BDI-II (duration of treatment 12 months). Higher scores indicate more severe depression.									

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of events		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result		
1 cohort study Lopez de Lara et al. 2020	Serious limitations[1]	No serious indirectness	No serious inconsistency	Not calculable	N=23	None	T0 (baseline) = 19.3 (SD 5.5) T1 (12 months) = 9.7 (SD 3.9) Statistically significant improvement, p<0.001	Critical	VERY LOW
Change from baseline in mean depression score, measured using the CESD-R (approximately 12-month follow-up). Higher scores indicate more severe depression.									
1 cohort study Achille et al. 2020	Serious limitations[2]	Serious indirectness[3]	No serious inconsistency	Not calculable	N=50	None	Wave 1 (baseline) = 21.4 Wave 3 (approx. 12 months) = 13.9 Statistically significant improvement (p<0.001)	Critical	VERY LOW
Change from baseline in depression score, measured using the Patient Health Questionnaire Modified for Teens (PHQ 9_Modified for Teens) (approximately 12-month follow-up). Higher scores indicate more severe depression.									
1 cohort study Achille et al. 2020	Serious limitations[2]	Serious indirectness[3]	No serious inconsistency	Not calculable	N=50	None	Statistically significant reductions in mean score, p<0.001 Results presented diagrammatically, numerical results for mean score not reported	Critical	VERY LOW
Change from baseline in depression symptoms, measured using the Quick Inventory of Depressive Symptoms (QIDS), self-reported (mean duration of gender-affirming hormone treatment 10.9 months). Higher scores indicate more severe depression.									
1 cohort study Kuper et al. 2020	Serious limitations[4]	No serious indirectness	No serious inconsistency	Not calculable	N=105	None	Baseline = 9.6 (SD 5.0) Follow-up = 7.4 (SD 4.5) No statistical analysis reported for the sub-group of participants receiving gender-affirming hormones	Critical	VERY LOW
Change from baseline in depression symptoms, measured using the Quick Inventory of Depressive Symptoms (QIDS), clinician-reported (mean duration of gender-affirming hormone treatment 10.9 months). Higher scores indicate more severe depression.									
1 cohort study	Serious limitations[4]	No serious indirectness	No serious inconsistency	Not calculable	N=106	None	Baseline = 5.9 (SD 4.1) Follow-up = 6.0 (SD 3.8)	Critical	VERY LOW

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of events		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result		
Kuper et al. 2020							No statistical analysis reported for the sub-group of participants who received gender-affirming hormones		
Need for treatment due to depression, during and before gender identity assessment, and during real life phase (approximately 12 months follow-up)									
1 cohort study Kaltiala et al. 2020	Serious limitations[7]	No serious indirectness	No serious inconsistency	Not calculable	N=52	None	During and before gender identity assessment 54% (28/52) During real life phase 15% (8/52) Statistically significant reduction (p<0.001)	Critical	VERY LOW
Change from baseline in anxiety score, measured using the STAI-State subscale (duration of treatment 12 months). Higher scores indicate more severe anxiety.									
1 cohort study Lopez de Lara et al. 2020	Serious limitations[1]	No serious indirectness	No serious inconsistency	Not calculable	N=23	None	T0 (baseline) = 33.3 (SD 9.1) T1 (12 months) = 16.8 (SD 8.1) Statistically significant improvement, p<0.001	Critical	VERY LOW
Change from baseline in anxiety score, measured using the STAI-Trait subscale (duration of treatment 12 months). Higher scores indicate more severe anxiety.									
1 cohort study Lopez de Lara et al. 2020	Serious limitations[1]	No serious indirectness	No serious inconsistency	Not calculable	N=23	None	T0 (baseline) = 33.0 (SD 7.2) T1 (12 months) = 18.5 (SD 8.4) Statistically significant improvement, p<0.001	Critical	VERY LOW
Change from baseline in anxiety symptoms, measured using the SCARED questionnaire (mean duration of gender-affirming hormone treatment 10.9 months). Higher scores indicate more severe anxiety.									
1 cohort study Kuper et al. 2020	Serious limitations[4]	No serious indirectness	No serious inconsistency	Not calculable	N=80	None	Baseline = 32.6 (SD 16.3) Follow-up = 28.4 (SD 15.9) No statistical analysis reported for the sub-group of participants	Critical	VERY LOW

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of events		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result		
							who received gender-affirming hormones		
Change from baseline in panic symptoms, measured using specific questions from the SCARED questionnaire (mean duration of gender-affirming hormone treatment 10.9 months). Higher scores indicate more severe symptoms.									
1 cohort study Kuper et al. 2020	Serious limitations[4]	No serious indirectness	No serious inconsistency	Not calculable	N=82	None	Baseline = 8.1 (SD 6.3) Follow-up = 7.1 (SD 6.5) No statistical analysis reported for the sub-group of participants who received gender-affirming hormones	Critical	VERY LOW
Change from baseline in generalised anxiety symptoms, measured using specific questions from the SCARED questionnaire (mean duration of gender-affirming hormone treatment was 10.9 months). Higher scores indicate more severe symptoms.									
1 cohort study Kuper et al. 2020	Serious limitations[4]	No serious indirectness	No serious inconsistency	Not calculable	N=82	None	Baseline = 10.0 (SD 5.1) Follow-up = 8.8 (SD 5.0) No statistical analysis reported for the sub-group of participants who received gender-affirming hormones	Critical	VERY LOW
Change from baseline in social anxiety symptoms, measured using specific questions from the SCARED questionnaire (mean duration of gender-affirming hormone treatment was 10.9 months). Higher scores indicate more severe symptoms.									
1 cohort study Kuper et al. 2020	Serious limitations[4]	No serious indirectness	No serious inconsistency	Not calculable	N=82	None	Baseline = 8.5 (SD 4.1) Follow-up = 7.7 (SD 4.2) No statistical analysis reported for the sub-group of participants who received gender-affirming hormones	Critical	VERY LOW
Change from baseline in separation anxiety symptoms, measured using specific questions from the SCARED questionnaire (mean duration of gender-affirming hormone treatment was 10.9 months). Higher scores indicate more severe symptoms.									
1 cohort study Kuper et al. 2020	Serious limitations[4]	No serious indirectness	No serious inconsistency	Not calculable	N=81	None	Baseline = 3.5 (SD 3.0) Follow-up = 3.1 (SD 2.5) No statistical analysis reported for the sub-group of participants	Critical	VERY LOW

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of events		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result		
							who received gender-affirming hormones		
Change from baseline in school avoidance, measured using specific questions from the SCARED questionnaire (mean duration of gender-affirming hormone treatment was 10.9 months). Higher scores indicate more severe symptoms.									
1 cohort study Kuper et al. 2020	Serious limitations[4]	No serious indirectness	No serious inconsistency	Not calculable	N=80	None	Baseline = 2.6 (SD 2.1) Follow-up = 2.0 (SD 2.0) No statistical analysis reported for the sub-group of participants who received gender-affirming hormones	Critical	VERY LOW
Need for treatment due to anxiety, during and before gender identity assessment, and during real life phase (approximately 12 months follow-up)									
1 cohort study Kaltiala et al. 2020	Serious limitations[7]	No serious indirectness	No serious inconsistency	Not calculable	N=52	None	During and before gender identity assessment 48% (25/52) During real life phase 15% (8/52) Statistically significant reduction (p<0.001)	Critical	VERY LOW
Change from baseline in adjusted mean suicidality score, measured using the ASQ instrument (mean treatment duration 349 days). Higher scores indicate a greater degree of suicidality.									
1 cohort study Allen et al. 2019	Serious limitations[5]	No serious indirectness	No serious inconsistency	Not calculable	N=39	None	T0 (baseline) = 1.11 (SE 0.22) T1 (final assessment) = 0.27 (SE 0.12) Statistically significant improvement in score from T0 to T1, p<0.001	Critical	VERY LOW
Change from baseline in percentage of participants with suicidal ideation, measured using the additional questions from the PHQ 9_Modified for Teens (approximately 12-month follow-up)									
1 cohort study Achille et al. 2020	Serious limitations[2]	Serious indirectness[3]	No serious inconsistency	Not calculable	N=50	None	Wave 1 (baseline) = 10% (5/50) Wave 3 (approx. 12 months) = 6% (3/50)	Critical	VERY LOW

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of events		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result		
							No statistical analysis reported		
Change from baseline in suicidal ideation (passive), information on which was collected by clinician, exact methods / tools not reported (mean duration of gender-affirming hormone treatment was 10.9 months)									
1 cohort study Kuper et al. 2020	Serious limitations[4]	Serious indirectness [6]	No serious inconsistency	Not calculable	N=130	None	Lifetime = 81% (105 people) 1 month before initial assessment = 25% (33 people) Follow-up period = 38% (51 people) No statistical analysis reported	Critical	VERY LOW
Change from baseline in suicide attempts, information on which was collected by clinician, exact methods / tools not reported (mean duration of gender-affirming hormone treatment was 10.9 months)									
1 cohort study Kuper et al. 2020	Serious limitations[4]	Serious indirectness[6]	No serious inconsistency	Not calculable	N=130	None	Lifetime = 15% (20 people) 3 months before initial assessment = 2% (3 people) Follow-up period = 5% (6 people) No statistical analysis reported	Critical	VERY LOW
Change from baseline in non-suicidal self-injury, information on which was collected by clinician, exact methods / tools not reported (mean duration of gender-affirming hormone treatment was 10.9 months)									
1 cohort study Kuper et al. 2020	Serious limitations[4]	Serious indirectness[6]	No serious inconsistency	Not calculable	N=130	None	Lifetime = 52% (68 people) 3 months before initial assessment = 10% (13 people) Follow-up period = 17% (23 people) No statistical analysis reported	Critical	VERY LOW
Need for treatment due to suicidality / self-harm, during and before gender identity assessment, and during real life phase (approximately 12 months follow-up)									
1 cohort study Kaltiala et al. 2020	Serious limitations[7]	No serious indirectness	No serious inconsistency	Not calculable	N=52	None	During and before gender identity assessment 35% (18/52) During real life phase	Critical	VERY LOW

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of events		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result		
							4% (2/52) Statistically significant reduction (p<0.001)		
Need for mental health treatment, during and before gender identity assessment, and during real life phase (approximately 12 months follow-up)									
1 cohort study Kaltiala et al. 2020	Serious limitations[7]	No serious indirectness	No serious inconsistency	Not calculable	N=52	None	During and before gender identity assessment 50% (26/52) During real life phase 46% (24/51) No statistically significant difference (p= 0.77)	Critical	VERY LOW
Need for treatment due to conduct problems / antisocial, during and before gender identity assessment, and during real life phase (approximately 12 months follow-up)									
1 cohort study Kaltiala et al. 2020	Serious limitations[7]	No serious indirectness	No serious inconsistency	Not calculable	N=52	None	During and before gender identity assessment 14% (7/52) During real life phase 6% (3/52) No statistically significant difference (p= 0.18)	Critical	VERY LOW
Need for treatment due to psychotic symptoms or psychosis, during and before gender identity assessment, and during real life phase (approximately 12 months follow-up)									
1 cohort study Kaltiala et al. 2020	Serious limitations[7]	No serious indirectness	No serious inconsistency	Not calculable	N=52	None	During and before gender identity assessment 2% (1/52) During real life phase 4% (2/52) No statistically significant difference (p= 0.56)	Critical	VERY LOW
Need for treatment due to substance abuse, during and before gender identity assessment, and during real life phase (approximately 12 months follow-up)									

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of events		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result		
1 cohort study Kaltiala et al. 2020	Serious limitations[7]	No serious indirectness	No serious inconsistency	Not calculable	N=52	None	During and before gender identity assessment 4% (2/52) During real life phase 2% (1/52) No statistically significant difference (p= 0.56)	Critical	VERY LOW
Need for treatment due to autism, during and before gender identity assessment, and during real life phase (approximately 12 months follow-up)									
1 cohort study Kaltiala et al. 2020	Serious limitations[7]	No serious indirectness	No serious inconsistency	Not calculable	N=52	None	During and before gender identity assessment 12% (6/52) During real life phase 6% (3/52) No statistically significant difference (p= 0.30)	Critical	VERY LOW
Need for treatment due to ADHD, during and before gender identity assessment, and during real life phase (approximately 12 months follow-up)									
1 cohort study Kaltiala et al. 2020	Serious limitations[7]	No serious indirectness	No serious inconsistency	Not calculable	N=52	None	During and before gender identity assessment 10% (5/52) During real life phase 2% (1/52) No statistically significant difference (p= 0.09)	Critical	VERY LOW
Need for treatment due to eating disorder, during and before gender identity assessment, and during real life phase (approximately 12 months follow-up)									
1 cohort study Kaltiala et al. 2020	Serious limitations[7]	No serious indirectness	No serious inconsistency	Not calculable	N=52	None	During and before gender identity assessment 2% (1/52)	Critical	VERY LOW

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of events		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result		
							During real life phase 2% (1/52) No statistically significant difference (p=1.0)		

Abbreviations: ADHD: attention deficit hyperactivity disorder; ASQ: Ask Suicide-Screening Questions; CESD-R: Center for Epidemiologic Studies Depression Scale; BDI-II: Beck Depression Inventory II (BDI-II); p: p-value; PHQ 9_Modified for Teens: Patient Health Questionnaire Modified for Teens; SCARED: Screen for Child Anxiety Related Emotional Disorders; SD: standard deviation; STAI: State-Trait Anxiety Inventory

1 Downgraded 1 level - the cohort study by Lopez de Lara et al. (2020) was assessed at high risk of bias (poor quality; lack of blinding and no control group).
2 Downgraded 1 level - the cohort study by Achille et al (2020) was assessed at high risk of bias (poor quality; lack of blinding, no control group and high number of participants lost to follow-up).
3 Serious indirectness in Achille 2020- Outcome reported for full study cohort, of whom 30% were taking no treatment or puberty suppression alone at follow-up. Results for people taking gender-affirming hormones not reported separately.[4] Downgraded 1 level - the cohort study by Kuper et al. (2020) was assessed at high risk of bias (poor quality).
5 Downgraded 1 level - the cohort study by Allen et al. (2019) was assessed at high risk of bias (poor quality; lack of blinding and no control group).
6 Serious indirectness in Kuper et al. 2020- Outcome reported for full study cohort, of whom approximately 17% received puberty suppression alone and did not receive gender-affirming hormones
7 Downgraded 1 level - the cohort study by Kaltiala et al. (2020) was assessed at high risk of bias (poor quality; 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 gender-affirming hormones compared with one or a combination of psychological support, social transitioning to the desired gender or no intervention? – Quality of life

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of patients		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result		
Impact on quality of life (1 uncontrolled, prospective observational study and 1 uncontrolled, retrospective observational study)									

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of patients		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result		
Change from baseline in mean quality of life score, measured using the QLES-Q-SF) (approximately 12-month follow-up). Higher scores indicated better quality of life.									
1 cohort study Achille et al. 2020	Serious limitations[1]	Serious indirectness[2]	No serious inconsistency	Not calculable	N=50	None	Numerical improvements in mean score reported from wave 1 (baseline) to wave 3 (approx. 12 months), but difference not statistically significant (p = 0.085) Results presented diagrammatically, numerical results for mean score not reported	Critical	VERY LOW
Change from baseline in adjusted mean well-being score, measured using the GWBS of the Pediatric Quality of Life Inventory (mean treatment duration 349 days). Higher scores indicated better well-being.									
1 cohort study Allen et al. 2019	Serious limitations[3]	No serious indirectness	No serious inconsistency	Not calculable	N=39	None	T0 (baseline) = 61.70 (SE 2.43) T1 (final assessment) = 70.23 (SE 2.15) Statistically significant improvement in well-being score, p<0.002	Critical	VERY LOW

Abbreviations: GWBS: General Well-Being Scale; p: p-value; QLES-Q-SF: Quality of Life Enjoyment and Satisfaction Questionnaire; SE: standard error

1 Downgraded 1 level - the cohort study by Achille et al (2020) was assessed at high risk of bias (poor quality; lack of blinding, no control group and high number of participants lost to follow-up).
2 Serious indirectness in Achille et al. 2020 - Outcome reported for full study cohort, of whom 30% were taking no treatment or puberty suppression alone at follow-up. Results for people taking gender-affirming hormones not reported separately.
3 Downgraded 1 level - the cohort study by Allen et al. (2019) was assessed at high risk of bias (poor quality; lack of blinding and no control group).

Table 5: Question 1: For children and adolescents with gender dysphoria, what is the clinical effectiveness of treatment with gender-affirming hormones 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 patients		Effect		
Study	**Risk of bias**	**Indirectness**	**Inconsistency**	**Imprecision**	**Intervention**	**Comparator**	**Result**		
Impact on body image *(1 uncontrolled, prospective observational study)*									
Change from baseline in mean body image, measured using the BIS (mean duration of gender-affirming hormone treatment was 10.9 months). Higher scores represent a higher degree of body dissatisfaction.									
1 cohort study Kuper et al. 2020	Serious limitations[1]	No serious indirectness	No serious inconsistency	Not calculable	N=86	None	Baseline = 70.7 (SD 15.2) Follow-up = 51.4 (SD 18.3) No statistical analysis reported for the sub-group of participants who received gender-affirming hormones	Important	VERY LOW

Abbreviations: BIS: Body Image Scale; p: p-value; SD: standard deviation

1 Downgraded 1 level - the cohort study by Kuper et al. (2020) was assessed at high risk of bias (poor quality; lack of blinding, no control group and high number of participants lost to follow-up).

Table 6: Question 1: For children and adolescents with gender dysphoria, what is the clinical effectiveness of treatment with gender-affirming hormones compared with one or a combination of psychological support, social transitioning to the desired gender or no intervention? – Psychological impact

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of patients		Effect		
Study	**Risk of bias**	**Indirectness**	**Inconsistency**	**Imprecision**	**Intervention**	**Comparator**	**Result (95% CI)**		
Psychosocial Impact (1 uncontrolled, prospective observational study and 1 uncontrolled, retrospective observational study)									
Change from baseline in family functioning, measured using the Family APGAR test. Higher scores suggest more family dysfunction.									
1 cohort study Lopez de Lara et al. 2020	Serious limitations[1]	No serious indirectness	No serious inconsistency	Not calculable	N=23	None	T0 (baseline) = 17.9 T1 (12 months) = 18.0 No statistical analysis reported	Important	VERY LOW
Change from baseline in mean patient strengths and difficulties score, measured using the SDQ, Spanish Version (total difficulties score) (duration of treatment 12 months). Higher scores suggest the presence of a behavioural disorder.									
1 cohort study	Serious limitations[1]	No serious indirectness	No serious inconsistency	Not calculable	N=23	None	T0 (baseline) = 14.7 (SD 3.3) T1 (12 months) = 10.3 (SD 2.9)	Important	VERY LOW

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of patients		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result (95% CI)		
Lopez de Lara et al. 2020							Statistically significant improvement p<0.001		
Functioning in adolescent development: Living with parent(s)/ guardians[2] (outcome reported for the approximately 12-month period after starting gender-affirming hormones; referred to as the 'real-life phase' in Finland). Not living with parent(s) or guardian in your early 20s is a marker of age-appropriate functioning in Finnish culture.									
1 cohort study Kaltiala et al. 2020	Serious limitations[3]	No serious indirectness	No serious inconsistency	Not calculable	N=52	None	During gender identity assessment = 73% (38/52) During real life phase = 40% (21/50) Statistically significant reduction (p=0.001)	Important	VERY LOW
Functioning in adolescent development: Normative peer contacts[4] (outcome reported for the approximately 12-month period after starting gender-affirming hormones; referred to as the 'real-life phase' in Finland)									
1 cohort study Kaltiala et al. 2020	Serious limitations[3]	No serious indirectness	No serious inconsistency	Not calculable	N=52	None	During gender identity assessment = 89% (46/52) During real life phase = 81% (42/52) Statistically significant reduction (p<0.001)	Important	VERY LOW
Functioning in adolescent development: Progresses normatively in school/ work[5] (outcome reported for the approximately 12-month period after starting gender-affirming hormones; referred to as the 'real-life phase' in Finland)									
1 cohort study Kaltiala et al. 2020	Serious limitations[3]	No serious indirectness	No serious inconsistency	Not calculable	N=52	None	During gender identity assessment = 64% (33/52) During real life phase = 60% (31/52) No statistically significant difference (p=0.69)	Important	VERY LOW
Functioning in adolescent development: Has been dating or had steady relationships[6] (outcome reported for the approximately 12-month period after starting gender-affirming hormones; referred to as the 'real-life phase' in Finland)									
1 cohort study	Serious limitations[3]	No serious indirectness	No serious inconsistency	Not calculable	N=52	None	During gender identity assessment = 62% (32/50)	Important	VERY LOW

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of patients		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result (95% CI)		
Kaltiala et al. 2020							During real life phase = 58% (30/52) No statistically significant difference (p=0.51)		
Functioning in adolescent development: Is age-appropriately able to deal with matters outside of the home[7] (outcome reported for the approximately 12-month period after starting gender-affirming hormones; referred to as the 'real-life phase' in Finland)									
1 cohort study Kaltiala et al. 2020	Serious limitations[2]	No serious indirectness	No serious inconsistency	Not calculable	N=52	None	During gender identity assessment = 81% (42/52) During real life phase = 81% (42/52) No statistically significant difference (p=1.00)	Important	VERY LOW

Abbreviations: APGAR: Adaptability, Partnership, Growth, Affection and Resolve; p: p-value; SD: standard deviation; SDQ: Strengths and Difficulties Questionnaire

1 Downgraded 1 level - the cohort study by Lopez de Lara et al. (2020) was assessed at high risk of bias (poor quality; lack of blinding and no control group).
2 Living arrangements were classified as (1) living with at least one parent/guardian, (2) living in a boarding school, with an adult relative, in some form of supported accommodation or the like, where supervision and guidance by a responsible adult is provided, (3) independently alone or in a shared household with a peer, (4) with a romantic partner. In the analyses dichotomised living arrangements as (a) parent(s)/guardian(s) vs. in other arrangements.
3 Downgraded 1 level - the cohort study by Kaltiala et al. (2020) was assessed at high risk of bias (poor quality; lack of blinding and no control group).
4 Peer relationships were classified as: (1) socialises with friends in leisure time, outside of activities supervised by adults, (2) socialises with peers only at school or in the context of rehabilitative activity, (3) spends time close to peers, for example in school or rehabilitative activity, but does not connect with them, (4) does not meet peers at all. In the analyses, peer relationships during (a) gender identity assessment and (b) the real-life phase were dichotomized to age-appropriate (normative) (1) vs. restricted or lacking (2–4).
5 School/work participation was classified as (1) age appropriate participation in mainstream curriculum, progresses without difficulties, (2) participates in mainstream curriculum with difficulty, (3) participates in rehabilitative educational or work activity, (4) not involved in education and working life. Age-appropriate participation during (1) was recorded if the adolescent attended mainstream secondary education or upper secondary education at a regular rate (a class per year in comprehensive school; has not changed more than once between tracks in upper secondary education) or had proceeded to work life after completing vocational education. Participation with difficulty (2) was recorded if the adolescent was enrolled in mainstream education but had to repeat a class, studied with special arrangements (for example, in a special small group), or followed some form of adjusted curriculum. In the analyses, school/work life during (a) gender identity assessment and (b) real-life phase was dichotomised to normative (1) vs. any other (2, 3 or 4).
6 Romantic involvement was recorded (1) has or has had a dating or steady relationship, not only online, (2) has had a romantic relationship only online, (3) has not had dating or steady relationships. In the analyses we compared has or has had (1) vs. has not had (2,3) a dating or steady relationship during (a) gender identity assessment and (b) real-life phase. Sexual history was recorded in more detail in case histories during gender identity assessment, and for this period we also collected the experiences of (French) kissing (yes/no), intercourse (yes/no) and experience of any genitally intimate contact with a partner (petting under clothes or naked, intercourse, oral sex) (yes/no).

7 In recording age-appropriate competence in managing everyday matters it was expected that early adolescents (up to 14 years) would be able, for example, to do shopping and travel alone on local public transport, and to help with household duties assigned by their parents. Middle adolescents (15–17 years) were further assumed, for example, to be able make telephone calls in matters important to them (for example, when seeking a summer job), to deal with school-related issues with school personnel without parental participation, to select and start new hobbies independently and to fulfil their role in summer jobs and in similar responsibilities of young people. Late adolescents (18 years and over), legally adults, were expected to have, in addition to the above, competence to talk to authorities such as professionals in health and social services, employment or educational institutions, to deal with banks or health insurance, to manage their financial issues and to manage their housekeeping if they chose to move to live independently of parents/guardians. Competence in managing everyday matters was recorded as follows: (1) the adolescent is able to cope age appropriately outside home, (2) the adolescent needs support in age-appropriate matters outside home but functions age-appropriately in the home (manages her/his own hygiene, clothing and nutrition, participates in (younger subjects) or takes responsibility for (older subjects) housekeeping) and (3) the adolescent's functioning is inadequate both at home and outside home. For the analyses, participants were determined to be able to age-appropriately able cope with matters outside of the home (1) vs. not (2,3).

Table 7: Question 2: For children and adolescents with gender dysphoria, what is the short-term and long-term safety of gender-affirming hormones 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 patients		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result (95% CI)		
Lumbar spine bone mineral apparent density (BMAD) (2 uncontrolled, retrospective observational studies)									
Change from start of gender-affirming hormones to age 22 years in lumber spine BMAD in transfemales									
1 cohort study Klink et al. 2015	Serious limitations[1]	Serious indirectness[2]	Not applicable	Not calculable	N=13 (Mean) N=14 (z-score)	None	Mean (SD), g/m^3 Start of gender-affirming hormones: 0.22 (0.02) Age 22 years: 0.23 (0.03) P=0.003 z-score (SD) Start of gender-affirming hormones: -0.90 (0.80) Age 22 years: -0.78 (1.03) No statistically significant difference	Important	VERY LOW
Change from baseline in lumbar spine BMAD in transfemales with a bone age less than 15 years ('young'; 24 months follow-up)									
1 cohort study Vlot et al. 2017	Serious limitations[3]	No serious indirectness	Not applicable	Not calculable	N=15	None	Median (range), g/m^3 Start of gender-affirming hormones (C0): 0.20 (0.18 to 0.24)	Important	VERY LOW

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of patients		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result (95% CI)		
							24-month follow-up (C24): 0.22 (0.19 to 0.27) Statistically significant increase (p≤0.01) z-score (range) Start of gender-affirming hormones (C0): -1.52 (-2.36 to 0.42) 24-month follow-up (C24): -1.10 (-2.44 to 0.69) Statistically significant increase (p≤0.05)		
Change from baseline in lumbar spine BMAD in transfemales with a bone age of 15 years or more ('old'; 24 months follow-up)									
1 cohort study Vlot et al. 2017	Serious limitations[3]	No serious indirectness	Not applicable	Not calculable	N=5	None	Median (range), g/m^3 Start of gender-affirming hormones (C0): 0.22 (0.19 to 0.24) 24-month follow-up (C24): 0.23 (0.21 to 0.26) Statistically significant increase (p≤0.05) z-score (range) Start of gender-affirming hormones (C0): -1.15 (-2.21 to 0.08) 24-month follow-up (C24): -0.66 (-1.66 to 0.54) Statistically significant increase (p≤0.05)	Important	VERY LOW
Change from start of gender-affirming hormones to age 22 years in lumber spine BMAD in transmales									

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of patients		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result (95% CI)		
1 cohort study Klink et al. 2015	Serious limitations[1]	Serious indirectness[2]	Not applicable	Not calculable	N=19 (Mean and z-score)	None	Mean (SD), g/m^3 Start of gender-affirming hormones: 0.24 (0.02) Age 22 years: 0.25 (0.28) P=0.001 z-score Start of gender-affirming hormones: -0.50 (0.81) Age 22 years: -0.033 (0.95) P=0.002	Important	VERY LOW
Change from baseline in lumbar spine BMAD in transmales with a bone age of less than 14 years ('young'; 24 months follow-up)									
1 cohort study Vlot et al. 2017	Serious limitations[3]	No serious indirectness	Not applicable	Not calculable	N=11	None	Median (range), g/m^3 Start of gender-affirming hormones (C0): 0.23 (0.19 to 0.28) 24-month follow-up (C24): 0.25 (0.22 to 0.28) Statistically significant increase (p≤0.01) z-score (range) Start of gender-affirming hormones (C0): -0.84 (-2.2 to 0.87) 24-month follow-up (C24): -0.15 (-1.38 to 0.94) Statistically significant increase (p≤0.01)	Important	VERY LOW
Change from baseline in lumbar spine BMAD in transmales with a bone age of 14 years or more ('old'; 24 months follow-up)									
1 cohort study	Serious limitations[3]	No serious indirectness	Not applicable	Not calculable	N=23	None	Median (range), g/m^3	Important	VERY LOW

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of patients		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result (95% CI)		
Vlot et al. 2017							Start of gender-affirming hormones (C0): 0.24 (0.20 to 0.28) 24-month follow-up (C24): 0.25 (0.21 to 0.30) Statistically significant increase (p≤0.01) z-score (range) Start of gender-affirming hormones (C0): -0.29 (-2.28 to 0.90) 24-month follow-up (C24): -0.06 (-1.75 to 1.61) Statistically significant increase (p≤0.01)		
Change in femoral neck BMAD (2 uncontrolled, retrospective observational studies)									
Change from start of gender-affirming hormones to age 22 years in femoral neck BMAD in transfemales									
1 cohort study Klink et al. 2015	Serious limitations[1]	Serious indirectness[2]	Not applicable	Not calculable	N=14 (Mean) N=10 (z-score)	None	Mean (SD), g/m^3 Start of gender-affirming hormones: 0.26 (0.04) Age 22 years: 0.28 (0.05) No statistically significant difference z-score (SD) Start of gender-affirming hormones: -1.57 (1.74) Age 22 years: Not reported	Important	VERY LOW
Change from baseline in femoral neck BMAD in transfemales with a bone age less than 15 years ('young'; 24 months follow-up)									
1 cohort study	Serious limitations[3]	No serious indirectness	Not applicable	Not calculable	N=16	None	Median (range), g/m^3 C0: 0.27 (0.20 to 0.33) C24: 0.27 (0.20 to 0.36)	Important	VERY LOW

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of patients		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result (95% CI)		
Vlot et al. 2017							No statistically significant change z-score (range) C0: -1.32 (-3.39 to 0.21) C24: -1.30 (-3.51 to 0.92) No statistically significant change		
Change from baseline in femoral neck BMAD in transfemales with a bone age of 15 years or more ('old'; 24 months follow-up)									
1 cohort study Vlot et al. 2017	Serious limitations[3]	No serious indirectness	Not applicable	Not calculable	N=6	None	Median (range), g/m^3 C0: 0.30 (0.26 to 0.34) C24: 0.29 (0.24 to 0.38) No statistically significant change z-score (range) C0: -0.36 (-1.50 to 0.46) C24: -0.56 (-2.17 to 1.29) No statistically significant change	Important	VERY LOW
Change from start of gender-affirming hormones to age 22 years in femoral neck BMAD in transmales									
1 cohort study Klink et al. 2015	Serious limitations[1]	Serious indirectness[2]	Not applicable	Not calculable	N=19 (Mean) N=18 (z-score)	None	Mean (SD), g/m^3 Start of gender-affirming hormones: 0.31 (0.04) Age 22 years: 0.33 (0.05) P=0.010 z-score (SD) Start of gender-affirming hormones: -0.28 (0.74) Age 22 years: Not reported	Important	VERY LOW
Change from baseline in femoral neck BMAD in transmales with a bone age of less than 14 years ('young'; 24 months follow-up)									

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of patients		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result (95% CI)		
1 cohort study Vlot et al. 2017	Serious limitations[3]	No serious indirectness	Not applicable	Not calculable	N=10	None	Median (range), g/m^3 C0: 0.30 (0.22 to 0.35) C24: 0.33 (0.23 to 0.37) Statistically significant increase (p≤0.01) z-score (range) C0: -0.37 (-2.28 to 0.47) C24: -0.37 (-2.03 to 0.85) Statistically significant increase (p≤0.01)	Important	VERY LOW
Change from baseline in femoral neck BMAD in transmales with a bone age of 14 years or more ('old'; 24 months follow-up)									
1 cohort study Vlot et al. 2017	Serious limitations[3]	No serious indirectness	Not applicable	Not calculable	N=23	None	Median (range), g/m^3 C0: 0.30 (0.23 to 0.41) C24: 0.32 (0.23 to 0.41) Statistically significant increase (p≤0.01) z-score (range) C0: -0.27 ((-1.91 to 1.29) C24: 0.02 (-2.1 to 1.35) Statistically significant increase (p≤0.05)	Important	VERY LOW
Change in lumbar spine BMD (2 uncontrolled, retrospective observational studies)									
Change from start of gender-affirming hormones to age 22 years in lumbar spine BMD in transfemales									
1 cohort study Klink et al. 2015	Serious limitations[1]	Serious indirectness[2]	Not applicable	Not calculable	N=15 (Mean) N=13 (z-score)	None	Mean (SD), g/m^2 Start of gender-affirming hormones: 0.84 (0.11) Age 22 years: 0.93 (0.10) P<0.001 z-score (SD)	Important	VERY LOW

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of patients		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result (95% CI)		
							Start of gender-affirming hormones: -1.01 (0.98) Age 22 years: -1.36 (0.83) No statistically significant difference		
Change from start of gender-affirming hormones to age 22 years in lumbar spine BMD in transmales									
1 cohort study Klink et al. 2015	Serious limitations[1]	Serious indirectness[2]	Not applicable	Not calculable	N=19 (Mean and z-score)	None	Mean (SD), g/m^2 Start of gender-affirming hormones: 0.91 (0.10) Age 22 years: 0.99 (0.13) P<0.001 z-score (SD) Start of gender-affirming hormones: -0.72 (0.99) Age 22 years: -0.33 (1.12) No statistically significant difference	Important	VERY LOW
Change from start of testosterone treatment in lumbar spine BMD in transmen (follow-up 6 to 24 months)									
1 cohort study Stoffers et al. 2019	Serious limitations[4]	No serious indirectness	Not applicable	Not calculable	N=62 (T0 and T6) N=37 (T12) N=15 (T24)	None	Mean (SD), g/cm^2 T0: 0.90 (0.11) T6: 0.94 (0.10) T12: 0.95 (0.09) T24: 0.95 (0.11) No statistically significant difference from T0 to any timepoint z-score (SD) T0: -0.81 (1.02) T6: -0.67 (0.95) T12: -0.66 (0.81) T24: -0.74 (1.17)	Important	VERY LOW

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of patients		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result (95% CI)		
							No statistically significant difference from T0 to any timepoint		
Change in femoral neck BMD (2 uncontrolled, retrospective observational studies)									
Change from start of gender-affirming hormones to age 22 years in femoral neck BMD in transfemales									
1 cohort study Klink et al. 2015	Serious limitations[1]	Serious indirectness[2]	Not applicable	Not calculable	N=15 (Mean) N=11 (z-score)	None	Mean (SD), g/m^2 Start of gender-affirming hormones: 0.87 (0.08) Age 22 years: 0.94 (0.11) P=0.009 z-score (SD) Start of gender-affirming hormones: -0.95 (0.63) Age 22 years: -0.69 (0.74) No statistically significant difference	Important	VERY LOW
Change from start of gender-affirming hormones to age 22 years in femoral neck BMD in transmales									
1 cohort study Klink et al. 2015	Serious limitations[1]	Serious indirectness[2]	Not applicable	Not calculable	N=19 (Mean) N=16 (z-score)	None	Mean (SD), g/m^2 Start of gender-affirming hormones: 0.88 (0.09) Age 22 years: 0.95 (0.10) P<0.001 z-score (SD) Start of gender-affirming hormones: -0.35 (0.79) Age 22 years: -0.35 (0.74) P=0.006	Important	VERY LOW
Change from start of testosterone treatment in right femoral neck (hip) BMD in transmales (follow-up 6 to 24 months)									
1 cohort study	Serious limitations[4]	No serious indirectness	Not applicable	Not calculable	N=62 (T0 and T6)	None	Mean (SD), g/cm^2 T0: 0.77 (0.08)	Important	VERY LOW

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of patients		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result (95% CI)		
Stoffers et al. 2019					N=37 (T12) N=15 (T24)		T6: 0.84 (0.11) T12: 0.82 (0.08) T24: 0.85 (0.11) No statistically significant difference from T0 to any timepoint z-score (SD) T0: -0.97 (0.79) T6: -0.54 (0.96) T12: -0.80 (0.69) T24: -0.31 (0.84) No statistically significant difference from T0 to any timepoint		
Change from start of testosterone treatment in left femoral neck (hip) BMD in transmales (follow-up 6 to 24 months)									
1 cohort study Stoffers et al. 2019	Serious limitations[4]	No serious indirectness	Not applicable	Not calculable	N=62 (T0 and T6) N=37 (T12) N=15 (T24)	None	Mean (SD), g/cm^2 T0: 0.76 (0.09) T6: 0.83 (0.12) T12: 0.81 (0.08) T24: 0.86 (0.09) No statistically significant difference from T0 to any timepoint z-score (SD) T0: -1.07 (0.85) T6: -0.62 (1.12) T12: -0.93 (0.63) T24: -0.20 (0.70) No statistically significant difference from T0 to any timepoint	Important	VERY LOW

Abbreviations: BMAD: bone mineral apparent density; BMD: bone mineral density; g: grams; m: metre; SD: standard deviation

1 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 control group and high number of participants lost to follow-up)
2 Outcomes reported after gender reassignment surgery and not after gender-affirming hormones alone. Unclear whether observed changes are due to hormones or surgery
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)
4 Downgraded 1 level - the cohort study by Stoffers et al. (2019) was assessed as at high risk of bias (poor quality overall; lack of blinding and no control group)

Table 8: Question 2: For children and adolescents with gender dysphoria, what is the short-term and long-term safety of gender-affirming hormones compared with one or a combination of psychological support, social transitioning to the desired gender or no intervention? – Cardiovascular risk factors

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of patients		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result (95% CI)		
Change in body mass index (1 uncontrolled, retrospective observational study)									
Change from start of gender-affirming hormones to age 22 years in BMI in transfemales									
1 cohort study Klaver et al. 2020	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=71	None	Mean change (95% CI) +1.9 (0.6 to 3.2) Statistically significant increase (p<0.005) Mean BMI at 22 years (95% CI): 23.2 (21.6 to 24.8)	Important	VERY LOW
Change from start of gender-affirming hormones to age 22 years in BMI in transmales									
1 cohort study Klaver et al. 2020	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=121	None	Mean change (95% CI) +1.4 (0.8 to 2.0) Statistically significant increase (p<0.005) Mean BMI at 22 years (95% CI): 23.9 (23.0 to 24.7)	Important	VERY LOW

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of patients		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result (95% CI)		
Obesity rates at age 22 years (1 uncontrolled, retrospective observational study)									
Obesity rates at age 22 years in transfemales who started gender-affirming hormones as adolescents (1 uncontrolled, retrospective observational study)									
1 cohort study Klaver et al. 2020	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=71	None	At 22 years, 9.9% of transfemales were obese, compared with 3.0% in reference cisgender population No statistically analysis reported	Important	VERY LOW
Obesity rates at age 22 years in transfemales who started gender-affirming hormones as adolescents (1 uncontrolled, retrospective observational study)									
1 cohort study Klaver et al. 2020	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=121	None	At 22 years, 6.6% of transmales were obese, compared with 2.2% in reference cisgender population No statistically analysis reported	Important	VERY LOW
Change in blood pressure (1 uncontrolled, retrospective observational study)									
Change from start of gender-affirming hormones to age 22 years in systolic blood pressure (SBP) in transfemales									
1 cohort study Klaver et al. 2020	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=71	None	Mean change (95% CI) -3 (-8 to 2) No statistically significant difference Mean SBP at 22 years (95% CI): 117 (113 to 122)	Important	VERY LOW
Change from start of gender-affirming hormones to age 22 years in diastolic blood pressure (DBP) in transfemales									

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of patients		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result (95% CI)		
1 cohort study Klaver et al. 2020	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=71	None	Mean change (95% CI) +6 (3 to 10) Statistically significant increase (p<0.001) Mean DBP at 22 years (95% CI): 75 (72 to 78)	Important	VERY LOW
Change from start of gender-affirming hormones to age 22 years in systolic blood pressure (SBP) in transmales									
1 cohort study Klaver et al. 2020	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=121	None	Mean change (95% CI): +5 (1 to 9) Statistically significant increase (p<0.05) Mean SBP at 22 years (95% CI): 126 (122 to 130)	Important	VERY LOW
Change from start of gender-affirming hormones to age 22 years in diastolic blood pressure (DBP) in transmales									
1 cohort study Klaver et al. 2020	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=121	None	Mean change (95% CI): +6 (4 to 9) Statistically significant increase (p<0.001) Mean DBP at 22 years (95% CI): 74 (72 to 77)	Important	VERY LOW
Change in glucose levels, insulin levels, insulin resistance and HbA1c (2 uncontrolled, retrospective observational studies)									
Change from start of gender-affirming hormones to age 22 years in glucose level (mmol/L) in transfemales									
1 cohort study Klaver et al. 2020	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=71	None	Mean change (95% CI): +0.1 (-0.1 to 0.2)	Important	VERY LOW

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of patients		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result (95% CI)		
							No statistically significant difference Mean glucose level at 22 years (95% CI): 5.0 (4.8 to 5.1)		
Change from start of gender-affirming hormones to age 22 years in insulin level (mU/L) in transfemales									
1 cohort study Klaver et al. 2020	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=71	None	Mean change (95% CI) +2.7 (-1.7 to 7.1) No statistically significant difference Mean insulin level at 22 years (95% CI): 13.0 (8.4 to 17.6)	Important	VERY LOW
Change from start of gender-affirming hormones to age 22 years in insulin resistance (HOMA-IR) in transfemales. Higher scores indicate more insulin resistance.									
1 cohort study Klaver et al. 2020	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=71	None	Mean change (95% CI) +0.7 (-0.2 to 1.5) No statistically significant difference Mean HOMA-IR at 22 years (95% CI): 2.9 (1.9 to 3.9)	Important	VERY LOW
Change from start of gender-affirming hormones to age 22 years in glucose level (mmol/L) in transmales									
1 cohort study Klaver et al. 2020	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=121	None	Mean change (95% CI) 0.0 (-0.2 to 0.2) No statistically significant difference	Important	VERY LOW

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of patients		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result (95% CI)		
							Mean glucose level at 22 years (95% CI): 4.8 (4.7 to 5.0)		
Change from start of gender-affirming hormones to age 22 years in insulin level (mU/L) in transmales									
1 cohort study Klaver et al. 2020	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=121	None	Mean change (95% CI) -2.1 (-3.9 to -0.3) Statistically significant decrease (p<0.05) Mean insulin level at 22 years (95% CI): 8.6 (6.9 to 10.2)	Important	VERY LOW
Change from start of gender-affirming hormones to age 22 years in insulin resistance (HOMA-IR) in transmales. Higher scores indicate more insulin resistance.									
1 cohort study Klaver et al. 2020	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=121	None	Mean change (95% CI): -0.5 (-1.0 to -0.1) Statistically significant decrease (p<0.05) Mean HOMA-IR at 22 years (95% CI): 1.8 (1.4 to 2.2)	Important	VERY LOW
Change from start of testosterone in HbA1c in transmales (up to 24 months follow-up)									
1 cohort study Stoffers et al. 2019	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N= Not reported	None	No statistically significant change from start of testosterone treatment Numerical results, follow-up duration and further details of	Important	VERY LOW

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of patients		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result (95% CI)		
							statistical analysis not reported.		
Change in lipid profile (1 uncontrolled, retrospective observational study)									
Change from start of gender-affirming hormones to age 22 years in total cholesterol (mmol/L) in transfemales									
1 cohort study Klaver et al. 2020	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=71	None	Mean change (95% CI): +0.1 (-0.2 to 0.4) No statistically significant difference Mean total cholesterol at 22 years (95% CI): 4.1 (3.8 to 4.4)	Important	VERY LOW
Change from start of gender-affirming hormones to age 22 years in HDL cholesterol (mmol/L) in transfemales									
1 cohort study Klaver et al. 2020	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=71	None	Mean change (95% CI): 0.0 (-0.1 to 0.2) No statistically significant difference Mean HDL cholesterol at 22 years (95% CI): 1.6 (1.4 to 1.7)	Important	VERY LOW
Change from start of gender-affirming hormones to age 22 years in LDL cholesterol (mmol/L) in transfemales									
1 cohort study Klaver et al. 2020	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=71	None	Mean change (95% CI): 0.0 (-0.3 to 0.2) No statistically significant difference Mean LDL cholesterol at 22 years (95% CI): 2.0 (1.8 to 2.3)	Important	VERY LOW

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of patients		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result (95% CI)		
Change from start of gender-affirming hormones to age 22 years in triglycerides (mmol/L) in transfemales									
1 cohort study Klaver et al. 2020	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=71	None	Mean change (95% CI): +0.2 (0.0 to 0.5) Statistically significant increase (p<0.05) Mean triglycerides at 22 years (95% CI): 1.1 (0.9 to 1.4)	Important	VERY LOW
Change from start of gender-affirming hormones to age 22 years in total cholesterol (mmol/L) in transmales									
1 cohort study Klaver et al. 2020	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=121	None	Mean change (95% CI): +0.4 (0.2 to 0.6) Statistically significant increase (p<0.001) Mean total cholesterol at 22 years (95% CI): 4.6 (4.3 to 4.8)	Important	VERY LOW
Change from start of gender-affirming hormones to age 22 years in HDL cholesterol (mmol/L) in transmales									
1 cohort study Klaver et al. 2020	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=121	None	Mean change (95% CI) -0.3 (-0.4 to -0.2) Statistically significant decrease (p<0.001) Mean HDL cholesterol at 22 years (95% CI): 1.3 (1.2 to 1.3)	Important	VERY LOW
Change from start of gender-affirming hormones to age 22 years in LDL cholesterol (mmol/L) in transmales									
1 cohort study Klaver et al. 2020	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=121	None	Mean change (95% CI): +0.4 (0.2 to 0.6)	Important	VERY LOW

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of patients		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result (95% CI)		
							Statistically significant increase (p<0.001) Mean LDL cholesterol at 22 years (95% CI): 2.6 (2.4 to 2.8)		
Change from start of gender-affirming hormones to age 22 years in triglycerides (mmol/L) in transmales									
1 cohort study Klaver et al. 2020	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=121	None	Mean change (95% CI) +0.5 (0.3 to 0.7) Statistically significant increase (p<0.001) Mean triglycerides at 22 years (95% CI): 1.3 (1.1 to 1.5)	Important	VERY LOW

Abbreviations: BMI: boss mass index; CI: confidence interval; DBP: diastolic blood pressure; HbA1c: glycated haemoglobin; HDL: high-density lipoproteins; HOMA-IR: Homeostatic Model Assessment of Insulin Resistance; LDL: low-density lipoproteins; mmol/L: millimoles per litre; mU/L: milliunits per litre; SBP: systolic blood pressure; SD: standard deviation

1 Downgraded 1 level - the cohort study by Klaver et al. (2020) 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 Stoffers et al. (2019) was assessed as at high risk of bias (poor quality overall; lack of blinding and no control group)

Table 9: Question 2: For children and adolescents with gender dysphoria, what is the short-term and long-term safety of gender-affirming hormones 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 patients		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result (95% CI)		
Liver enzymes (1 uncontrolled, retrospective observational study)									

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of patients		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result (95% CI)		
Change from start of testosterone in aspartate aminotransferase (AST) level in transmales (up to 24 months follow-up)									
1 cohort study Stoffers et al. 2019	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N= Not reported	None	No statistically significant change from start of testosterone treatment Numerical results, follow-up duration and further details of statistical analysis not reported.	Important	VERY LOW
Change from start of testosterone in alanine aminotransferase (ALT) level in transmales (up to 24 months follow-up)									
1 cohort study Stoffers et al. 2019	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N= Not reported	None	No statistically significant change from start of testosterone treatment Numerical results, follow-up duration and further details of statistical analysis not reported.	Important	VERY LOW
Change from start of testosterone in gamma-glutamyl transferase (GGT) level in transmales (up to 24 months follow-up)									
1 cohort study Stoffers et al. 2019	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N= Not reported	None	No statistically significant change from start of testosterone treatment Numerical results, follow-up duration and further details of statistical analysis not reported.	Important	VERY LOW
Change from start of testosterone in alkaline phosphatase (ALP) level in transmales (up to 24 months follow-up)									
1 cohort study Stoffers et al. 2019	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=62 (T0 and T1) N=37 (T12)	None	Median (IQR), U/L T0: 102 (78 to 136) T6: 115 (102 to 147) T12: 112 (88 to 143) T24: 81 (range 69 to 98)	Important	VERY LOW

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of patients		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result (95% CI)		
					N-15 (T24)		Statistically significant increase from T0 at T6 and T12 (p<0.001)		
Kidney markers (1 uncontrolled, retrospective observational study)									
Change from start of testosterone in serum creatinine level in transmales (up to 24 months follow-up)									
1 cohort study Stoffers et al. 2019	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N=62 (T0 and T1) N=37 (T12) N=15 (T24)	None	Mean (SD), umol/L T0: 62 (7) T6: 70 (9) T12: 74 (10) T24: 81 (10) Statistically significant increase from T0 at all timepoints (p<0.001)	Important	VERY LOW
Change from start of testosterone in serum urea[2] level in transmales (up to 24 months follow-up)									
1 cohort study Stoffers et al. 2019	Serious limitations[1]	No serious indirectness	Not applicable	Not calculable	N= Not reported	None	No statistically significant change from start of testosterone treatment Numerical results, follow-up duration and further details of statistical analysis not reported.	Important	VERY LOW
Adverse effects (1 uncontrolled, retrospective observational study)									
Permanent discontinuation of gender-affirming hormones (median follow-up 2.0 years (range 0.0 to 11.3)									
1 cohort study Khatchadourian et al. 2014	Serious limitations[3]	No serious indirectness	Not applicable	Not calculable	N=63	None	No participants permanently discontinued gender-affirming hormones.	Important	VERY LOW
Temporary discontinuation of gender-affirming hormones (median follow-up 2.0 years (range 0.0 to 11.3)									
1 cohort study	Serious limitations[3]	No serious indirectness	Not applicable	Not calculable	N=63	None	3/37 transmales receiving testosterone temporarily	Important	VERY LOW

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of patients		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result (95% CI)		
Khatchadourian et al. 2014							discontinued treatment, 2 due to concomitant mental health comorbidities and 1 due to androgenic alopecia. All eventually resumed treatment. No transfemales receiving oestrogen temporarily discontinued treatment		
Minor complications during treatment with gender-affirming hormones (median follow-up 2.0 years (range 0.0 to 11.3)									
1 cohort study Khatchadourian et al. 2014	Serious limitations[3]	No serious indirectness	Not applicable	Not calculable	N=63	None	12/63 participants had minor complications during treatment with gender-affirming hormones All 12 were transmales receiving testosterone. Complications were severe acne (n=7), androgenic alopecia (n=1) mild dyslipidaemia (n=3) and significant mood swings (n=1) No transfemales receiving oestrogen had minor complications	Important	VERY LOW
Severe complications during treatment with gender-affirming hormones (median follow-up 2.0 years (range 0.0 to 11.3)									
1 cohort study Khatchadourian et al. 2014	Serious limitations[3]	No serious indirectness	Not applicable	Not calculable	N=63	None	No severe complications reported during gender-affirming treatment	Important	VERY LOW

Abbreviations: ALP: alkaline phosphatase; ALT: alanine aminotransferase; AST: aspartate aminotransferase; GGT: gamma-glutamyl transferase; IQR: interquartile range; SD: standard deviation; U/L: units per litre; umol/L: micromole per litre

1 Downgraded 1 level - the cohort study by Stoffers et al. (2019) was assessed as at high risk of bias (poor quality overall; lack of blinding and no control group)
2 Referred to as 'ureum' in original publication
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)

Table 10: From the evidence selected, are there particular sub-groups of children and adolescents with gender dysphoria that derive comparatively more (or less) benefit from treatment with gender-affirming hormones than the wider population of children and adolescents with gender dysphoria? – Transfemales compared with transmales

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of patients		Effect		
Study	**Risk of bias**	**Indirectness**	**Inconsistency**	**Imprecision**	**Transfemales**	**Transmales**	**Result (95% CI)**		
Impact on mental health (1 uncontrolled, retrospective observational study)									
Change from baseline in adjusted mean suicidality score, measured using the ASQ tool (mean treatment duration 349 days). Higher scores indicate a greater degree of suicidality.									
1 cohort study Allen et al. 2019	Serious limitations[4]	No serious indirectness	No serious inconsistency	Not calculable	N=14	N=33	**Transfemales** T0 (baseline) = 1.21 (SE 0.36) T1 (final assessment) = 0.24 (SE 0.19) **Transmales** T0 (baseline) = 1.01 (SE 0.23) T1 (final assessment) = 0.29 (SE 0.13) No statistically significant difference in change from baseline between transfemales and transmales (p=0.79)	Critical	VERY LOW
Impact on quality of life (1 uncontrolled, retrospective observational study)									
Change from baseline in adjusted mean well-being score, measured using the GWBS of the Pediatric Quality of Life Inventory (mean treatment duration 349 days). Higher scores indicate better well-being.									
1 cohort study Allen et al. 2019	Serious limitations[4]	No serious indirectness	No serious inconsistency	Not calculable	N=14	N=33	**Transfemales** T0 (baseline) = 58.44 (SE 4.09) T1 (final assessment) = 69.52 (SE 3.62)	Critical	VERY LOW

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of patients		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Transfemales	Transmales	Result (95% CI)		
							Transmales T0 (baseline) = 64.95 (SE 2.66) T1 (final assessment) = 70.94 (SE 2.35) No statistically significant difference in change from baseline between transfemales and transmales (p=0.32)		

Abbreviations: ASQ: Ask Suicide-Screening Questions; GWBS: General Well-Being Scale; SE: standard error

1 The cohort study by Allen et al. 2019 was assessed at high risk of bias (poor quality; lack of blinding and no control group).

Table 11: From the evidence selected, are there particular sub-groups of children and adolescents with gender dysphoria that derive comparatively more (or less) benefit from treatment with gender-affirming hormones than the wider population of children and adolescents with gender dysphoria? – Sex assigned at birth males (transfemales)

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of events/No of patients% (n/N%)		Effect		
Study type and number of studies Author year	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result (95% CI)		
Change from baseline in mean depression symptoms in transfemales, measured using the Quick Inventory of Depressive Symptoms (QIDS), self-reported (mean duration of gender-affirming hormone treatment 10.9 months). Higher scores indicate more depression.									
1 cohort study Kuper et al. 2020	Serious limitations[1]	No serious indirectness	No serious inconsistency	Not calculable	N=40	None	Baseline = 7.5 (SD 4.9) Follow-up = 6.6 (SD 4.4) No statistical analysis reported for this sub-group	Critical	VERY LOW
Change from baseline in mean depression symptoms in transfemales, measured using the Quick Inventory of Depressive Symptoms (QIDS), clinician-reported (mean duration of gender-affirming hormone treatment 10.9 months). Higher scores indicate more severe depression.									
1 cohort study	Serious limitations[1]	No serious indirectness	No serious inconsistency	Not calculable	N=45	None	Baseline = 4.2 (SD 3.2) Follow-up = 5.4 (SD 3.4)	Critical	VERY LOW

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of events/No of patients% (n/N%)		Effect		
Study type and number of studies Author year	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result (95% CI)		
Kuper et al. 2020							No statistical analysis reported for this sub-group		
Change from baseline in mean anxiety symptoms in transfemales, measured using the SCARED questionnaire (mean duration of gender-affirming hormone treatment 10.9 months). Higher scores indicate more severe anxiety.									
1 cohort study Kuper et al. 2020	Serious limitations[1]	No serious indirectness	No serious inconsistency	Not calculable	N=33	None	Baseline = 26.4 (SD 14.2) Follow-up = 24.3 (SD 15.4) No statistical analysis reported for this sub-group	Critical	VERY LOW
Change from baseline in mean panic symptoms in transfemales, measured using specific questions from the SCARED questionnaire (mean duration of gender-affirming hormone treatment 10.9 months). Higher scores indicate more severe symptoms.									
1 cohort study Kuper et al. 2020	Serious limitations[1]	No serious indirectness	No serious inconsistency	Not calculable	N=34	None	Baseline = 5.7 (SD 4.9) Follow-up = 5.1 (SD 4.9) No statistical analysis reported for this sub-group	Critical	VERY LOW
Change from baseline in mean generalised anxiety symptoms in transfemales, measured using specific questions from the SCARED questionnaire (mean duration of gender-affirming hormone treatment was 10.9 months). Higher scores indicate more severe symptoms.									
1 cohort study Kuper et al. 2020	Serious limitations[1]	No serious indirectness	No serious inconsistency	Not calculable	N=34	None	Baseline = 8.6 (SD 5.1) Follow-up = 8.0 (SD 5.1) No statistical analysis reported for this sub-group	Critical	VERY LOW
Change from baseline in mean social anxiety symptoms in transfemales, measured using specific questions from the SCARED questionnaire (mean duration of gender-affirming hormone treatment was 10.9 months). Higher scores indicate more severe symptoms.									
1 cohort study Kuper et al. 2020	Serious limitations[1]	No serious indirectness	No serious inconsistency	Not calculable	N=34	None	Baseline = 7.1 (SD 3.9) Follow-up = 6.8 (SD 4.4) No statistical analysis reported for this sub-group	Critical	VERY LOW
Change from baseline in mean separation anxiety symptoms in transfemales, measured using specific questions from the SCARED questionnaire (mean duration of gender-affirming hormone treatment was 10.9 months). Higher scores indicate more severe symptoms.									
1 cohort study Kuper et al. 2020	Serious limitations[1]	No serious indirectness	No serious inconsistency	Not calculable	N=34	None	Baseline = 3.4 (SD 3.3) Follow-up = 2.7 (SD 2.3) No statistical analysis reported for this sub-group	Critical	VERY LOW

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of events/No of patients% (n/N%)		Effect		
Study type and number of studies Author year	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result (95% CI)		
Change from baseline in mean school avoidance symptoms in transfemales, measured using specific questions from the SCARED questionnaire (mean duration of gender-affirming hormone treatment was 10.9 months). Higher scores indicate more severe symptoms.									
1 cohort study Kuper et al. 2020	Serious limitations[1]	No serious indirectness	No serious inconsistency	Not calculable	N=33	None	Baseline = 1.8 (SD 1.7) Follow-up = 1.9 (SD 2.1) No statistical analysis reported for this sub-group	Critical	VERY LOW
Change from baseline in percentage of participants with suicidal ideation in transfemales, measured using the additional questions from the PHQ 9_Modified for Teens (approximately 12-month follow-up)									
1 cohort study Achille et al. 2020	Serious limitations[2]	Serious indirectness[2]	No serious inconsistency	Not calculable	N=17	None	Wave 1 (baseline) = 11.8% (2/17) Wave 2 (approx. 12 months) = 5.9% (1/17) No statistical analysis reported	Critical	VERY LOW
Impact on body image (1 uncontrolled, prospective observational study)									
Change from baseline in mean body image in transfemales, measured using the BIS (mean duration of gender-affirming hormone treatment was 10.9 months). Higher scores represent a higher degree of body dissatisfaction.									
1 cohort study Kuper et al. 2020	Serious limitations[1]	No serious indirectness	No serious inconsistency	Not calculable	N=30	None	Baseline = 67.5 (SD 19.5) Follow-up = 49.0 (SD 21.6) No statistical analysis reported for this sub-group	Important	VERY LOW

Abbreviations: BIS: Body Image Scale; PHQ 9: Patient Health Questionnaire 9; SCARED: Screen for Child Anxiety Related Emotional Disorders; SD: standard deviation

1 Downgraded 1 level - the cohort study by Kuper et al. (2020) was assessed at high risk of bias (poor quality; lack of blinding, no control group and high number of participants lost to follow-up).
2 Downgraded 1 level - the cohort study by Achille et al. 2020 was assessed at high risk of bias (poor quality; lack of blinding, no control group and high number of participants lost to follow-up).
3 Serious indirectness in Achille 2020- Approximately 30% of the full sample received puberty suppression alone or were receiving no treatment at final follow-up.

Table 12: From the evidence selected, are there particular sub-groups of children and adolescents with gender dysphoria that derive comparatively more (or less) benefit from treatment with gender-affirming hormones than the wider population of children and adolescents with gender dysphoria? – Sex assigned at birth females (transmales)

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of patients		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result (95% CI)		
Change from baseline in mean depression symptoms in transmales, measured using the Quick Inventory of Depressive Symptoms (QIDS), self-reported (mean duration of gender-affirming hormone treatment 10.9 months). Higher scores indicate more severe depression.									
1 cohort study Kuper et al. 2020	Serious limitations[1]	No serious indirectness	No serious inconsistency	Not calculable	N=76	None	Baseline = 10.4 (SD 5.0) Follow-up = 7.5 (SD 4.5) No statistical analysis reported for this sub-group	Critical	VERY LOW
Change from baseline in mean depression symptoms in transmales, measured using the Quick Inventory of Depressive Symptoms (QIDS), clinician-reported (mean duration of gender-affirming hormone treatment 10.9 months). Higher scores indicate more severe depression.									
1 cohort study Kuper et al. 2020	Serious limitations[1]	No serious indirectness	No serious inconsistency	Not calculable	N=78	None	Baseline = 6.7 (SD 4.4) Follow-up = 6.2 (SD 4.1) No statistical analysis reported for this sub-group	Critical	VERY LOW
Change from baseline in mean anxiety symptoms in transmales, measured using the SCARED questionnaire (mean duration of gender-affirming hormone treatment 10.9 months). Higher scores indicate more severe anxiety.									
1 cohort study Kuper et al. 2020	Serious limitations[1]	No serious indirectness	No serious inconsistency	Not calculable	N=65	None	Baseline = 35.4 (SD 16.5) Follow-up = 29.8 (SD 15.5) No statistical analysis reported for this sub-group	Critical	VERY LOW
Change from baseline in mean panic symptoms in transmales, measured using specific questions from the SCARED questionnaire (mean duration of gender-affirming hormone treatment 10.9 months). Higher scores indicate more severe symptoms.									
1 cohort study Kuper et al. 2020	Serious limitations[1]	No serious indirectness	No serious inconsistency	Not calculable	N=66	None	Baseline = 9.3 (SD 6.5) Follow-up = 7.9 (SD 6.5) No statistical analysis reported for this sub-group	Critical	VERY LOW
Change from baseline in mean generalised anxiety symptoms in transmales, measured using specific questions from the SCARED questionnaire (mean duration of gender-affirming hormone treatment was 10.9 months). Higher scores indicate more severe symptoms.									
1 cohort study Kuper et al. 2020	Serious limitations[1]	No serious indirectness	No serious inconsistency	Not calculable	N=66	None	Baseline = 10.4 (SD 5.0) Follow-up = 9.0 (SD 5.1) No statistical analysis reported for this sub-group	Critical	VERY LOW

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of patients		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result (95% CI)		
Change from baseline in mean social anxiety symptoms in transmales, measured using specific questions from the SCARED questionnaire (mean duration of gender-affirming hormone treatment was 10.9 months). Higher scores indicate more severe symptoms.									
1 cohort study Kuper et al. 2020	Serious limitations[1]	No serious indirectness	No serious inconsistency	Not calculable	N=66	None	Baseline = 8.5 (SD 4.0) Follow-up = 7.8 (SD 4.1) No statistical analysis reported for this sub-group	Critical	VERY LOW
Change from baseline in mean separation anxiety symptoms in transmales, measured using specific questions from the SCARED questionnaire (mean duration of gender-affirming hormone treatment was 10.9 months). Higher scores indicate more severe symptoms.									
1 cohort study Kuper et al. 2020	Serious limitations[1]	No serious indirectness	No serious inconsistency	Not calculable	N=65	None	Baseline = 4.2 (SD 3.4) Follow-up = 3.4 (SD 2.6) No statistical analysis reported for this sub-group	Critical	VERY LOW
Change from baseline in mean school avoidance symptoms in transmales, measured using specific questions from the SCARED questionnaire (mean duration of gender-affirming hormone treatment was 10.9 months). Higher scores indicate more severe symptoms.									
1 cohort study Kuper et al. 2020	Serious limitations[1]	No serious indirectness	No serious inconsistency	Not calculable	N=65	None	Baseline = 2.9 (SD 2.3) Follow-up = 2.0 (SD 2.3) No statistical analysis reported for this sub-group	Critical	VERY LOW
Change from baseline in percentage of participants with suicidal ideation in transmales, measured using the additional questions from the PHQ 9_Modified for Teens (approximately 12-month follow-up)									
1 cohort study Achille et al. 2020	Serious limitations[2]	Serious indirectness[3]	No serious inconsistency	Not calculable	N=33	None	Wave 1 (baseline) = 9.1% (3/33) Wave 2 (approx. 12 months) = 6.1% (2/33) No statistical analysis reported	Critical	VERY LOW
Impact on body image (1 uncontrolled, prospective observational study)									
Change from baseline in mean body image in transmales, measured using the BIS (mean duration of gender-affirming hormone treatment was 10.9 months). Higher scores represent a higher degree of body dissatisfaction.									
1 cohort study Kuper et al. 2020	Serious limitations[1]	No serious indirectness	No serious inconsistency	Not calculable	N=66	None	Baseline = 71.1 (SD 13.4) Follow-up = 52.9 (SD 16.8) No statistical analysis reported for this sub-group	Important	VERY LOW

Abbreviations: BIS: Body Image Scale; PHQ 9: Patient Health Questionnaire 9; SCARED: Screen for Child Anxiety Related Emotional Disorders; SD: standard deviation

1 Downgraded 1 level - the cohort study by Kuper et al. (2020) was assessed at high risk of bias (poor quality; lack of blinding, no control group and high number of participants lost to follow-up).
2 Downgraded 1 level - the cohort study by Achille et al. 2020 was assessed at high risk of bias (poor quality; lack of blinding, no control group and high number of participants lost to follow-up).
3 Serious indirectness in Achille 2020- Approximately 30% of the full sample received puberty suppression alone or were receiving no treatment at final follow-up.

Table 14: From the evidence selected, are there particular sub-groups of children and adolescents with gender dysphoria that derive comparatively more (or less) benefit from treatment with gender-affirming hormones than the wider population of children and adolescents with gender dysphoria? – Outcomes controlled for concurrent counselling and medicines for mental health problems

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of patients		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result (95% CI)		
Impact on mental health (1 uncontrolled, retrospective observational study)									
Change from baseline in mean depression score in transfemales, measured using the CESD-R (approximately 12-month follow-up; controlled for engagement in counselling and medicines for mental health problems). Higher scores indicate more depression.									
1 cohort study Achille et al. 2020	Serious limitations[1]	Serious indirectness[2]	No serious inconsistency	Not calculable	N=17	None	No statistically significant change from baseline (p=0.27) Numerical scores not reported	Critical	VERY LOW
Change from baseline in mean depression score in transmales, measured using the CESD-R (approximately 12-month follow-up; controlled for engagement in counselling and medicines for mental health problems). Higher scores indicate more severe depression.									
1 cohort study Achille et al. 2020	Serious limitations[1]	Serious indirectness[2]	No serious inconsistency	Not calculable	N=33	None	No statistically significant change from baseline (p=0.43) Numerical scores not reported	Critical	VERY LOW
Change from baseline in depression score in transfemales, measured using the Patient Health Questionnaire Modified for Teens (PHQ 9_Modified for Teens) (approximately 12-month follow-up; controlled for engagement in counselling and medicines for mental health problems). Higher scores indicate more severe depression.									
1 cohort study Achille et al. 2020	Serious limitations[1]	Serious indirectness[2]	No serious inconsistency	Not calculable	N=17	None	No statistically significant change from baseline (p=0.07) Numerical scores not reported	Critical	VERY LOW

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of patients		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result (95% CI)		
Change from baseline in depression score in transmales, measured using the Patient Health Questionnaire Modified for Teens (PHQ 9_Modified for Teens) (approximately 12-month follow-up; controlled for engagement in counselling and medicines for mental health problems). Higher scores indicate more severe depression.									
1 cohort study Achille et al. 2020	Serious limitations[1]	Serious indirectness[2]	No serious inconsistency	Not calculable	N=33	None	No statistically significant change from baseline (p=0.67) Numerical scores not reported	Critical	VERY LOW
Impact on quality of life (1 uncontrolled, retrospective observational study)									
Change from baseline in mean quality of life score in transfemales, measured using the QLES-Q-SF (approximately 12-month follow-up; controlled for engagement in counselling and medicines for mental health problems). Higher scores indicated better quality of life.									
1 cohort study Achille et al. 2020	Serious limitations[1]	Serious indirectness[2]	No serious inconsistency	Not calculable	N=17	None	No statistically significant change from baseline (p=0.06)	Critical	VERY LOW
Change from baseline in mean quality of life score in transmales, measured using the QLES-Q-SF (approximately 12-month follow-up; controlled for engagement in counselling and medicines for mental health problems). Higher scores indicated better quality of life.									
1 cohort study Achille et al. 2020	Serious limitations[1]	Serious indirectness[2]	No serious inconsistency	Not calculable	N=33	None	No statistically significant change from baseline (p=0.08)	Critical	VERY LOW
Psychosocial Impact (1 uncontrolled, retrospective observational study)									
Functioning in adolescent development: Progresses normatively in school/ work during the real-life phase – impact on need for mental health treatment before or during gender identity assessment									
1 cohort study Kaltiala et al. 2020	Serious limitations[3]	No serious indirectness	No serious inconsistency	Not calculable	N=49	None	Needed mental health treatment: 47% (15/32) functioning well Did not need mental health treatment: 82% (14/17) functioning well Statistically significant difference p=0.02	Important	VERY LOW
Functioning in adolescent development: Is age-appropriately able to deal with matters outside of the home during the real-life phase – impact on need for mental health treatment before or during gender identity assessment									

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of patients		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result (95% CI)		
1 cohort study Kaltiala et al. 2020	Serious limitations[3]	No serious indirectness	No serious inconsistency	Not calculable	N=49	None	Needed mental health treatment: 72% (23/32) managing well Did not need mental health treatment: 94% (16/17) managing well No statistically significant difference p=0.06	Important	VERY LOW
Functioning in adolescent development: Progresses normatively in school/ work during the real-life phase – impact on need for mental health treatment during the real-life phase									
1 cohort study Kaltiala et al. 2020	Serious limitations[3]	No serious indirectness	No serious inconsistency	Not calculable	N=51	None	Needed mental health treatment: 42% (10/24) functioning well Did not need mental health treatment: 74% (20/27) functioning well Statistically significant difference p=0.02	Important	VERY LOW
Functioning in adolescent development: Is age-appropriately able to deal with matters outside of the home during the real-life phase – impact on need for mental health treatment during the real-life phase									
1 cohort study Kaltiala et al. 2020	Serious limitations[3]	No serious indirectness	No serious inconsistency	Not calculable	N=51	None	Needed mental health treatment: 67% (16/24) managing well Did not need mental health treatment: 93% (25/27) managing well Statistically significant difference p=0.02	Important	VERY LOW

Abbreviations: CESD-R: Center for Epidemiologic Studies Depression; p: p-value; PHQ 9: Patient Health Questionnaire 9; QLES-Q-SF: Quality of Life Enjoyment and Satisfaction Questionnaire

1 Downgraded 1 level - the cohort study by Achille et al 2020 was assessed at high risk of bias (poor quality; lack of blinding, no control group and high number of participants lost to follow-up).
2 Serious indirectness in Achille 2020- Approximately 30% of the full sample received puberty suppression alone or were receiving no treatment at final follow-up.
3 Downgraded 1 level - the cohort study by Kaltiala et al. 2020 was assessed at high risk of bias (poor quality; lack of blinding and no control).

Table 15: From the evidence selected, are there particular sub-groups of children and adolescents with gender dysphoria that derive comparatively more (or less) benefit from treatment with gender-affirming hormones than the wider population of children and adolescents with gender dysphoria? – Tanner age

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of patients		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator	Result (95% CI)		
Impact on mental health (1 uncontrolled, retrospective observational study)									
Change from baseline in mental health problems – depression, anxiety and anxiety-related symptoms (mean duration of gender-affirming hormone treatment was 10.9 months)									
1 cohort study Kuper et al. 2020	Serious limitations[1]	No serious indirectness	No serious inconsistency	Not calculable	N=105	None	No difference in outcomes found by Tanner age. Numerical results, statistical analysis and information on specific outcomes not reported. It is unclear from the paper whether Tanner age is at initial assessment, start of GnRH analogues, start of gender-affirming hormones, or another timepoint	Critical	VERY LOW
Impact on body image (1 uncontrolled, prospective observational study)									
Change from baseline in mean body image, measured using the BIS (mean duration of gender-affirming hormone treatment was 10.9 months). Higher scores represent a higher degree of body dissatisfaction.									
1 cohort study Kuper et al. 2020	Serious limitations[1]	No serious indirectness	No serious inconsistency	Not calculable	N=105	None	No difference in body image score found by Tanner age. Numerical results, statistical analysis and information on specific outcomes not reported.	Important	VERY LOW

							It is unclear from the paper whether Tanner age is at initial assessment, start of GnRH analogues, start of gender-affirming hormones, or another timepoint		

Abbreviations: BIS: Body Image Scale

1 Downgraded 1 level - the cohort study by Kuper et al. 2020 was assessed at high risk of bias (poor quality; lack of blinding, no control group and high number of participants lost to follow-up).

Glossary

Ask Suicide-Screening Questions (ASQ)	ASQ is a four-item dichotomous (yes, no) response measure with high sensitivity, designed to identify risk of suicide. A patient is considered to have screened positive if they answered yes to any item. The authors of Allen et al. 2019 altered the fourth item of the ASQ ("Have you ever tried to kill yourself?") and prefaced it with "In the past few weeks . . ." as they were not investigating lifetime suicidality. A response of 'no' was scored as 0 and a response of 'yes' was scored as 1; each item was summed, generating an overall score for suicidality on a scale ranging from 0 to 4, with higher scores indicating greater levels of suicidal ideation.
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 bone size measurements using UK norms in growing adolescents.
Center for Epidemiologic Studies Depression scale (CESD-R)	The CESD-R is a valid, widely used tool to access depressive symptoms. The CESD-R asks about how frequently a person has felt or behaved in a certain way; with 20 questions scored from 0 score is calculated as a sum of 20 questions, ranging from 0 ("not at all or less than one day") to 3 ("5–7 days" and/or "nearly every day for 2 weeks"). Total score ranges from 0 to 60, with higher scores indicating more depressive symptoms.
Cisgender	Cisgender is a term for someone whose gender identity matches their birth-registered sex.
Family APGAR (Adaptability, Partnership, Growth, Affection and Resolve) test	The Family APGAR test is a 5-item questionnaire, with higher scores indicating better family functioning. The authors reported the following interpretation of the score: functional, 17-20 points; mildly dysfunctional, 16-13 points; moderately dysfunctional, 12-10 point; severely dysfunctional, <9 points.
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).

General Well-Being Scale (GWBS) of the Pediatric Quality of Life Inventory score	The GWBS of the Pediatric Quality of Life Inventory uses uses a 5-point response scale, contains seven items, and measures two dimensions: general wellbeing (6 items) and general health (1 item). Each item is scored from 0 to 4, and the total score is linearly transformed to a 0 to 100 scale. High scores reflect fewer perceived problems and greater well-being.
GnRH analogue	GnRH analogues competitively block GnRH receptors to prevent the spontaneous release of two gonadotropin hormones, Follicular Stimulating Hormone (FSH) and Luteinising Hormone (LH) from the pituitary gland. The reduction in LH and FSH 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.
Patient Health Questionnaire Modified for Teens score (PHQ 9_Modified for Teens)	The PHQ 9_Modified for Teens is a validated tool to assess depression, dysthymia and suicide risk. The tool consists of 9 questions scored from 0 to 3 (total score 0 to 27), plus an additional 4 questions that are not scored. A score of 0 to 4 suggests no or minimal depressive symptoms, 5 to 9 mild, 10-14 moderate, 15-19 moderate and 20-27 severe symptoms.
Quick Inventory of Depressive Symptoms (QIDS)	Both the clinician- and self-reported QIDS are validated tools to assess depressive symptoms. The tool consists of 16 items, with the highest score for 9 items (sleep, weight, psychomotor changes, depressed mood, decreased interest, fatigue, guilt, concentration, and suicidal ideation) are added to give a total score ranging from 0 to 27. A score of 0 to 5 is suggestive of no depressive symptoms, 6 to 10 mild symptoms, 11 to 15 moderate symptoms, 16-20 severe symptoms and 21 to 27 very severe symptoms.
Quality of Life Enjoyment and Satisfaction Questionnaire (QLES-Q-SF)	QLES-Q-SF is a validated questionnaire, consisting of 15 questions that rate quality of life on a scale of 1 (poor) to 5 (very good).
Screen for Child Anxiety Related Emotional Disorders (SCARED) questionnaire	SCARED is a validated, 41-point questionnaire, with each item scored 0 to 2. A total score of 25 or more is suggestive of anxiety disorder, with scores above 30 being more specific. Certain scores for specific questions may indicate the presence of other anxiety-related disorders: A score of 7 or more in questions related to panic disorder or significant somatic symptoms may indicate the presence of these. A score of 9 or more in questions related to generalised anxiety disorder may indicate the presence of this. A score of 5 or more in questions related to separation anxiety may indicate the presence of this. A score of 8 or more in questions related to social anxiety disorder may indicate the presence of this. A score of 3 or more in questions related to significant school avoidance may indicate the presence of this.
State-Trait Anxiety Inventory (STAI) score	STAI is a validated and commonly used measure of state anxiety (current state of anxiety) and trait anxiety (general state of calmness, confidence and security). It has 40 items, the first 20 covering state anxiety, the second 20 covering trait anxiety. STAI

	can be used in clinical settings to diagnose anxiety and to distinguish it from depressive illness. Each subtest (state and trait) is scored between 20 and 80, with higher scores indicating greater anxiety. There is no published minimal clinically meaningful difference (MCID) for STAI or thresholds for anxiety severity.
Strengths and Difficulties Questionnaire (SDQ, Spanish version	The SDQ, Spanish version includes 25-items covering emotional symptoms, conduct problems, hyperactivity/ inattention, peer relationship problems and prosocial behaviour. The authors state that a score of more than 20 is considered indicative of risk of having a disorder (normal: 0-15; borderline: 16-19, abnormal: 20-40).
Tanner stage	Tanner staging is a scale of physical development.
Transgender (including transmale and transfemale)	Transgender is a term for someone whose gender identity is not congruent with their birth-registered sex. A transfemale is a person who identifies as female and a transmale is a person who identifies as male.
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. Higher scores indicate higher levels of gender dysphoria.

References

Included studies

- 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
- Allen, LR, Watson, LB, Egan, AM et al. (2019) Well-being and suicidality among transgender youth after gender-affirming hormones. Clinical Practice in Pediatric Psychology 7(3): 302-311
- Kaltiala, R., Heino, E., Tyolajarvi, M. et al. (2020) Adolescent development and psychosocial functioning after starting cross-sex hormones for gender dysphoria. Nordic Journal of Psychiatry 74(3): 213-219
- Khatchadourian K, Amed S, Metzger DL (2014) Clinical management of youth with gender dysphoria in Vancouver. The Journal of pediatrics 164(4): 906-11
- 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)
- 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

- Kuper, Laura E, Stewart, Sunita, Preston, Stephanie et al. (2020) Body Dissatisfaction and Mental Health Outcomes of Youth on Gender-Affirming Hormone Therapy. Pediatrics 145(4)
- Lopez de Lara, D., Perez Rodriguez, O., Cuellar Flores, I. et al. (2020) Psychosocial assessment in transgender adolescents. Anales de Pediatria
- Stoffers, Iris E; de Vries, Martine C; Hannema, Sabine E (2019) Physical changes, laboratory parameters, and bone mineral density during testosterone treatment in adolescents with gender dysphoria. The journal of sexual medicine 16(9): 1459-1468
- Vlot MC, Klink DT, den Heijer M 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/ [accessed 27 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. Available from: https://www.psychiatry.org/patients-families/gender-dysphoria/what-is-gender-dysphoria [accessed 27 August 2020]
- NHS England (2013). NHS Standard contract for gender identity development service for children and adolescents https://www.england.nhs.uk/wp-content/uploads/2017/04/gender-development-service-children-adolescents.pdf [accessed 27 August 2020]
- NHS England (2016). Clinical Commissioning Policy: Prescribing of Cross-Sex Hormones as part of the Gender Identity Development Service for Children and Adolescents https://www.england.nhs.uk/wp-content/uploads/2018/07/Prescribing-of-cross-sex-hormones-as-part-of-the-gender-identity-development-service-for-children-and-adolesce.pdf [accessed 27 August 2020]

Copyright

EXHIBIT 24

The Cass Review

Independent review of gender identity services for children and young people: Interim report

February 2022

Independent review of gender identity services for children and young people: Interim report

February 2022

Contents

About this report

This interim report represents the work of the independent review of gender identity services for children and young people to date. It reflects a point in time. It does not set out final recommendations; these will be developed over the coming months, informed by our formal research programme.

This Review is forward looking. Its role is to consider how to improve and develop the future clinical approach and service model. However, in order to do this, it is first necessary to understand the current landscape and the reasons why change is needed, so that any future model addresses existing challenges, whilst retaining those features that service users and the professionals supporting them most value.

This report is primarily for the commissioners and providers of services for children and young people needing support around their gender. However, because of the wide interest in this topic, we have included some explanations about how clinical service development routinely takes place in the NHS, which sets the context for some of our interim advice.

The care of this group of children and young people is everyone's business. We therefore encourage the wider clinical community to take note of our work and consider their own roles in providing the best holistic support to this population.

Since the Review began, it has focused on hearing a wide range of perspectives to better understand the challenges within the current system and aspirations for how these could be addressed. This report does not contain all that we have heard during our listening sessions but summarises consistent themes. These conversations will continue throughout the course of the Review and there will be further opportunities for stakeholders to engage and contribute.

It is important to note that the references cited in this report do not constitute a comprehensive literature review and are included only to clarify why specific lines of enquiry are being pursued, and where there are unanswered questions that will be addressed more fully during the life of the Review. A formal literature review is one strand of the Review's commissioned work, and this will be reported in full when complete.

A note about language

There is sometimes no consensus on the best language to use relating to this subject. The language surrounding this area has also changed rapidly and young people have developed varied ways of describing their experiences using different terms and constructs that are relevant to them.

The Review tries as far as possible to use language and terms that are respectful and acknowledge diversity, but that also accurately illustrate the complexity of what we are trying to describe and articulate.

The terms we have used may not always feel right to some; nevertheless, it is important to emphasise that the language used is not an indication of a position being taken by the Review. A glossary of terms is included.

The Review is cognisant of the broader cultural and societal debates relating to the rights of transgender adults. It is not the role of the Review to take any position on the beliefs that underpin these debates. Rather, this Review is strictly focused on the clinical services provided to children and young people who seek help from the NHS to resolve their gender-related distress.

A letter to children and young people

Children and young people accessing the NHS deserve safe, timely and supportive services, and clinical staff with the training and expertise to meet their healthcare needs.

Dr Hilary Cass

I understand that as you read this letter some of you may be anxious because you are waiting to access support from the NHS around your gender identity. Maybe you have tried to get help from your local services, or from the Gender Identity Development Service (GIDS), and because of the long waiting lists they have not yet been able to see you. I hope that some of you have had help – maybe from a supportive GP, a local Child and Adolescent Mental Health Service (CAMHS), or from GIDS.

I have heard that young service users are particularly worried that I will suggest that services should be reduced or stopped. I want to assure you that this is absolutely not the case – the reverse is true. I think that more services are needed for you, closer to where you live. The GIDS staff are working incredibly hard and doing their very best to see you as quickly as possible but providing supportive care is not something that can be rushed – each young person needs enough time and space for their personal needs to be met. So, with the best will in the world, one service is not going to be able to respond to the growing demand in a timely way.

I am advising that more services are made available to support you. But I must be honest; this is not something that can happen overnight, and I can't come up with a solution that will fix the problems immediately. However, we do need to start now.

The other topic that I know is worrying some of you is whether I will suggest that hormone treatments should be stopped. On this issue, I have to share my thoughts as a doctor. We know quite a bit about hormone treatments, but there is still a lot we don't know about the long-term effects.

Whenever doctors prescribe a treatment, they want to be as certain as possible that the benefits will outweigh any adverse effects so that when you are older you don't end up saying 'Why did no-one tell me that that might happen?' This includes understanding both the risks and benefits of having treatment and not having treatment.

Therefore, what we will be doing over the next few months is trying to make sense of all the information that is available, as well as seeing if we can plug any of the gaps in the research. I am currently emphasising the importance of making decisions about prescribing as safe as possible. This means making sure you have all the information you need – about what we do know and what we don't know.

Finally, some of you may want the chance to talk to me and share your thoughts about how services should look in the future. Over the coming months we will need your help and there will be opportunities to get involved with the Review, so please keep an eye on our website (www.cass.independent-review.uk), where we will provide updates on our work.

Dr Hilary Cass, OBE

Introduction from the Chair

Anyone with an interest in the care of gender-questioning children and young people, as well as those with lived experience, may have wondered what qualifies me to take on this Review, and whether I have a pre-existing position on this subject.

I am a paediatrician who was in clinical practice until 2018, my area of specialism being children and young people with disability. I have also held many management and policy roles throughout my career, most notably as President of the Royal College of Paediatrics and Child Health (RCPCH) from 2012-15.

Children's services are often at a disadvantage in healthcare because health services are usually designed around the needs of adults. As President of RCPCH, a key part of my role was to advocate for services to be planned with children and families at their heart.

I have not worked in gender services during my career, but my strong focus on hearing the voice of service users, supporting vulnerable young people, equity of access, and strong clinical standards applies in this area as much as in my other work.

With this in mind, the aim of the Review is to ensure that children and young people who are experiencing gender incongruence or gender-related distress receive a high standard of NHS care that meets their needs and is safe, holistic and effective.

I have previously set out the principles governing this Review process, namely that:

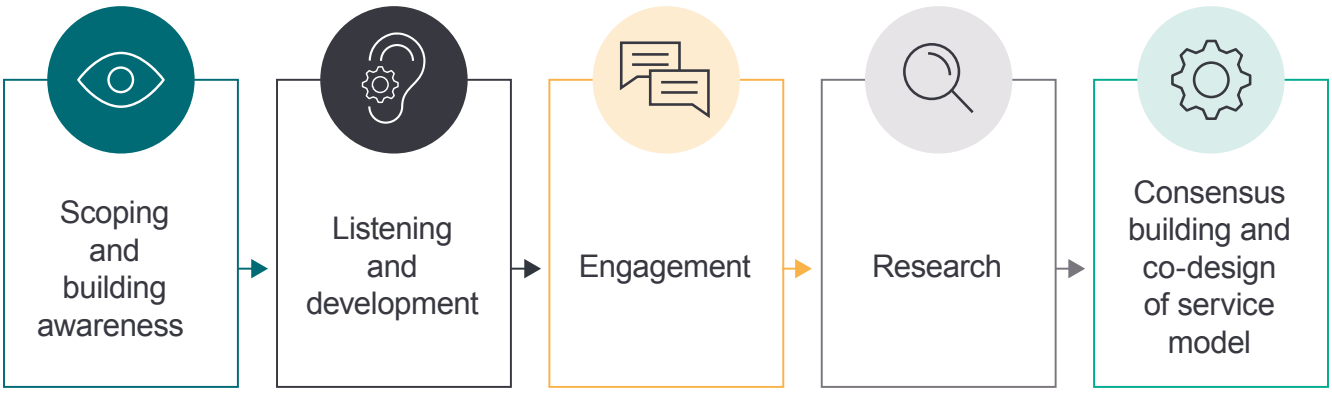
- The welfare of the child and young person will be paramount in all considerations.
- Children and young people must receive a high standard of care that meets their needs.
- There will be extensive and purposeful stakeholder engagement, including ensuring that children and young people can express their own views through a supportive process.
- The Review will be underpinned by research and evidence, including international models of good practice where available.
- There will be transparency in how the Review is conducted and how recommendations are made.
- There are no pre-determined outcomes with regards to the recommendations the Review will make.

The Review's terms of reference (**Appendix 1**) are wide ranging in scope, looking at different aspects of gender identity services across the whole pathway through primary, secondary and specialist services, up to the point of transition to adult services. This includes consideration of referral pathways, assessment, appropriate clinical management and workforce recommendations.

I have also been asked to explore the reasons for the considerable increase in the number of referrals, which have had a significant impact on waiting times, as well as the changing case-mix of gender-questioning children and young people presenting to clinical services.

The Review is taking an investigative approach to understanding what the future service model should look like for children and young people. This means that its outcomes are not being developed in isolation or by committee but rather through an ongoing dialogue aimed at building a shared understanding of the current situation and how it can and should be improved.

The key aspects of the approach to the Review are:



My starting point has been to hear from a variety of experts with relevant expertise and those with lived experience to understand as many perspectives as possible. To date, this has included hearing directly from those with lived experience, from professionals and support and advocacy groups. This listening process will continue.

We have been very fortunate in the generosity of all those who have been prepared to talk to the Review and share their experiences. In addition to some divergent opinions, there are also some themes and views which seem to be widely shared. The commitment of professionals at all levels is striking and I genuinely believe that with collective effort we can improve services for the children and young people who are at the heart of this Review.

These discussions have been valuable to get an in-depth sense of the current situation and different viewpoints on how it may be improved. However, it is essential that this initial understanding is underpinned by more detailed data and an enhanced evidence base, which is being delivered through the Review's academic research programme.

Providing this evidence base for the Review is going to take some time. I recognise there is a pressing need to enhance the services currently available for children, young people, their

parents and carers, some of whom are experiencing considerable distress. Clinicians providing their treatment and care are also under pressure and cannot sustain the current workload. As such, I know the time I am taking to complete this Review and make recommendations will be difficult for some, but it is necessary.

I wrote to NHS England in May 2021 (**Appendix 2**) setting out some more immediate considerations whilst awaiting my full recommendations. This report builds on that letter and looks to provide some further interim advice.

Through our research programme, the Review team will continue to examine the literature and, where possible, will fill gaps in the existing evidence base. However, there will be persisting evidence gaps and areas of uncertainty. We need the engagement of service users, support and advocacy groups, and professionals across the wider workforce to work with us in the coming months in a collaborative and open-minded manner in order to reach a shared understanding of the problems and an agreed way forward that is in the best interests of children and young people.

My measure of success for this Review will be that this group of children and young people receive timely, appropriate and excellent care, not just from specialists but from every healthcare professional they encounter as they take the difficult journey from childhood to adulthood.

1. Summary and interim advice

Summary

1.1. In recent years, there has been a significant increase in the number of referrals to the Gender Identity Development Service (GIDS) at the Tavistock and Portman NHS Foundation Trust. This has contributed to long waiting lists and growing concern about how the NHS should most appropriately assess, diagnose and care for this population of children and young people.

1.2. Within the UK, the single specialist service has developed organically, and the clinical approach has not been subjected to some of the usual control measures that are typically applied when new or innovative treatments are introduced. Many of the challenges and knowledge gaps that we face in the UK are echoed internationally,[1] and there are significant gaps in the research and evidence base.

1.3. This Review was commissioned by NHS England to make recommendations on how to improve services provided by the NHS to children and young people who are questioning their gender identity or experiencing gender incongruence and ensure that the best model for safe and effective services is commissioned (**Appendix 1**).

1.4. This interim report represents the Review's work to date. It sets out what we have heard so far and the approach we are taking moving forward. There is still much evidence to be gathered, questions to be answered, and voices to be heard, and our perspective will evolve as more evidence comes to light. However, there is sufficient clarity on several areas for the Review to be able to offer advice at this stage so that action can be taken more quickly.

1.5. The Review is not able to provide definitive advice on the use of puberty blockers and feminising/masculinising hormones at this stage, due to gaps in the evidence base; however, recommendations will be developed as our research programme progresses.

Every gender-questioning child or young person who seeks help from the NHS must receive the support they need to get on the appropriate pathway for them as an individual.

Children and young people with gender incongruence or dysphoria must receive the same standards of clinical care, assessment and treatment as every other child or young person accessing health services.

[1] Vrouenraets LJ, Fredriks AM, Hannema SE, Cohen-Kettenis PT, de Vries MC (2015). Early medical treatment of children and adolescents with gender dysphoria: an empirical ethical study. J Adolesc Health 57(4): 367-73. DOI: 10.1016/.2015.04.004.

Conceptual understanding and consensus about the meaning of gender dysphoria

1.6. In clinical practice, a diagnosis of gender dysphoria is currently based on an operational definition, using the criteria set out in DSM-5 (**Appendix 3**). Some of these criteria are seen by some as outdated in the context of current understanding about the flexibility of gender expression.

1.7. At primary, secondary and specialist level, there is a lack of agreement, and in many instances a lack of open discussion, about the extent to which gender incongruence in childhood and adolescence can be an inherent and immutable phenomenon for which transition is the best option for the individual, or a more fluid and temporal response to a range of developmental, social, and psychological factors. Professionals' experience and position on this spectrum may determine their clinical approach.

1.8. Children and young people can experience this as a 'clinician lottery', and failure to have an open discussion about this issue is impeding the development of clear guidelines about their care.

Service capacity and delivery

1.9. A rapid change in epidemiology and an increase in referrals means that the number of children seeking help from the NHS is now outstripping the capacity of the single national specialist service, the Gender Identity Development Service (GIDS) at The Tavistock and Portman NHS Foundation Trust.

1.10. The mix of young people presenting to the service is more complex than seen previously, with many being neurodiverse and/or having a wide range of psychosocial and mental health needs. The largest group currently comprises birth-registered females first presenting in adolescence with gender-related distress.

1.11. Until very recently, any local professional, including non-health professionals, could refer to GIDS, which has meant that the quality and appropriateness of referrals lacks consistency, and local service provision has remained patchy and scarce.

1.12. The staff working within the specialist service demonstrate a high level of commitment to the population they serve. However, the waiting list pressure and lack of consensus development on the clinical approach, combined with criticism of the service, have all resulted in rapid turnover of staff and inadequate capacity to deal with the increasing workload. Capacity constraints cannot be addressed through financial investment alone; there are some complex workforce (recruitment; retention; and training) and cultural issues to address.

1.13. Our initial work has indicated that many professionals working at primary and secondary level feel that they have the transferable skills and the commitment to offer more robust support to this group of children and young people, but are nervous about doing so, partly because of the lack of formal clinical guidance, and partly due to the broader societal context.

1.14. Primary and secondary care staff have told us that they feel under pressure to adopt an unquestioning affirmative approach and that this is at odds with the standard process of clinical assessment and diagnosis that they have been trained to undertake in all other clinical encounters.

1.15. Children and young people are waiting lengthy periods to access GIDS, during which time some may be at considerable risk. By the time they are seen, their distress may have worsened, and their mental health may have deteriorated.

1.16. Another significant issue raised with us is one of diagnostic overshadowing – many of the children and young people presenting have complex needs, but once they are identified as having gender-related distress, other important healthcare issues that would normally be managed by local services can sometimes be overlooked.

1.17. The current move to adult services at age 17-18 may fall at a critical time in the young person's gender management. In contrast, young people with neurodiversity often remain under children's services until age 19 and some other clinical services continue to mid-20s. Further consideration will be needed regarding the age of transfer to adult services.

Service standards

1.18. The Multi-Professional Review Group (MPRG), set up by NHS England to ensure that procedures for assessment and for informed consent have been properly followed, has stated that the following areas require consideration:

- From the point of entry to GIDS there appears to be predominantly an affirmative, non-exploratory approach, often driven by child and parent expectations and the extent of social transition that has developed due to the delay in service provision.
- From documentation provided to the MPRG, there does not appear to be a standardised approach to assessment or progression through the process, which leads to potential gaps in necessary evidence and a lack of clarity.
- There is limited evidence of mental health or neurodevelopmental assessments being routinely documented, or of a discipline of formal diagnostic or psychological formulation.
- Of 44 submissions received by the MPRG, 31% were not initially assured due to lack of safeguarding information. And in a number of cases there were specific safeguarding concerns. There do not appear to be consistent processes in place to work with other agencies to identify children and young people and families who may be vulnerable, at risk and require safeguarding.

- Appropriate clinical experts need to be involved in informing decision making.

1.19. Many of these issues were also highlighted by the Care Quality Commission (CQC) in 2020.[2]

International comparisons

1.20. The Netherlands was the first country to provide early endocrine interventions (now known internationally as the Dutch Approach). Although GIDS initially reported its approach to early endocrine intervention as being based on the Dutch Approach,[3] there are significant differences in the NHS approach. Within the Dutch Approach, children and young people with neurodiversity and/or complex mental health problems are routinely given therapeutic support in advance of, or when considered appropriate, instead of early hormone intervention. Whereas criteria to have accessed therapeutic support prior to starting hormone blocking treatment do not appear to be integral to the current NHS process.

1.21. NHS endocrinologists do not systematically attend the multi-disciplinary meetings where the complex cases that may be referred to them are discussed, and until very recently did not routinely have direct contact with the clinical staff member who had assessed the child or young person. This is not consistent with some international approaches for this group of children and young people, or in other multi-disciplinary models of care across paediatrics and adult medicine where challenging decisions about life-changing interventions are made.[4,5]

1.22. In the NHS, once young people are started on hormone treatment, the frequency of appointments drops off rather than intensifies, and review usually takes place quarterly. Again, this is different to the Dutch Approach.[6] GIDS staff would recommend more frequent contact during this period, but the fall-off in appointments reflects a lack of service capacity, with the aspiration being for more staff time to remedy this situation.

Existing evidence base

1.23. Evidence on the appropriate management of children and young people with gender incongruence and dysphoria is inconclusive both nationally and internationally.

---

[2] Care Quality Commission (2021). The Tavistock and Portman NHS Foundation Trust Gender Identity Service Inspection Report. London: CQC.
[3] de Vries ALC, Cohen-Kettenis PT (2012). Clinical management of gender dysphoria in children and adolescents: the Dutch approach. J Homosex 59: 301–320. DOI: 10.1080/00918369.2012.653300.
[4] Ibid.
[5] Kyriakou A, Nicolaides NC, Skordis N (2020). Current approach to the clinical care of adolescents with gender dysphoria. Acta Biomed 91(1): 165–75. DOI: 10.23750/abm.v91i1.9244.
[6] de Vries ALC, Cohen-Kettenis PT (2012). Clinical management of gender dysphoria in children and adolescents: the Dutch approach. J Homosex 59: 301–320. DOI: 10.1080/00918369.2012.653300.

1.24. A lack of a conceptual agreement about the meaning of gender dysphoria hampers research, as well as NHS clinical service provision.

1.25. There has not been routine and consistent data collection within GIDS, which means it is not possible to accurately track the outcomes and pathways that children and young people take through the service.

1.26. Internationally as well as nationally, longer-term follow-up data on children and young people who have been seen by gender identity services is limited, including for those who have received physical interventions; who were transferred to adult services and/or accessed private services; or who desisted, experienced regret or detransitioned.

1.27. There has been research on the short-term mental health outcomes and physical side effects of puberty blockers for this cohort, but very limited research on the sexual, cognitive or broader developmental outcomes.[7]

1.28. Much of the existing literature about natural history and treatment outcomes for gender dysphoria in childhood is based on a case-mix of predominantly birth-registered males presenting in early childhood. There is much less data on the more recent case-mix of predominantly birth-registered females presenting in early teens, particularly in relation to treatment and outcomes.

1.29. Aspects of the literature are open to interpretation in multiple ways, and there is a risk that some authors interpret their data from a particular ideological and/or theoretical standpoint.

The mismatch between service user expectations and clinical standards

1.30. By the time children and young people reach GIDS, they have usually had to experience increasingly long, challenging waits to be seen.[8] Consequently, some feel they want rapid access to physical interventions and find having a detailed assessment distressing.

1.31. Clinical staff are governed by professional, legal and ethical guidance which demands that certain standards are met before a treatment can be provided. Clinicians carry responsibility for their assessment and recommendations, and any harm that might be caused to a patient under their care. This can create a tension between the aspirations of the young person and the responsibilities of the clinician.

---

[7] National Institute for Health and Care Excellence (2020). Evidence Review: Gonadotrophin Releasing Hormone Analogues for Children and Adolescents with Gender Dysphoria.
[8] Care Quality Commission (2021). The Tavistock and Portman NHS Foundation Trust Gender Identity Service Inspection Report. London: CQC.

Interim advice

1.32. The Review considers that there are some areas where there is sufficient clarity about the way forward and we are therefore offering some specific observations and interim advice. The Review will work with NHS England, providers and the broader stakeholder community to progress action in these areas.

Service model

1.33. It has become increasingly clear that a single specialist provider model is not a safe or viable long-term option in view of concerns about lack of peer review and the ability to respond to the increasing demand.

1.34. Additionally, children and young people with gender-related distress have been inadvertently disadvantaged because local services have not felt adequately equipped to see them. It is essential that they can access the same level of psychological and social support as any other child or young person in distress, from their first encounter with the NHS and at every level within the service.

1.35. A fundamentally different service model is needed which is more in line with other paediatric provision, to provide timely and appropriate care for children and young people needing support around their gender identity. This must include support for any other clinical presentations that they may have.

1.36. The Review supports NHS England's plan to establish regional services, and welcomes the move from a single highly specialist service to regional hubs.

1.37. Expanding the number of providers will have the advantages of:

- creating networks within each area to improve early access and support;
- reducing waiting times for specialist care;
- building capacity and training opportunities within the workforce;
- developing a specialist network to ensure peer review and shared standards of care; and
- providing opportunities to establish a more formalised service improvement strategy.

Service provision

1.38. The primary remit of NHS England's proposed model is for the regional hubs to provide support and advice to referrers and professionals. However, it includes limited provision for direct contact with children and young people and their families.

1: The Review advises that the regional centres should be developed, as soon as feasibly possible, to become direct service providers, assessing and treating children and young people who may need specialist care, as part of a wider pathway. The Review team will work with NHS England and stakeholders to further define the proposed model and workforce implications.

2: Each regional centre will need to develop links and work collaboratively with a range of local services within their geography to ensure that appropriate clinical, psychological and social support is made available to children and young people who are in early stages of experiencing gender distress.

3: Clear criteria will be needed for referral to services along the pathway from primary to tertiary care so that gender-questioning children and young people who seek help from the NHS have equitable access to services.

4: Regional training programmes should be run for clinical practitioners at all levels, alongside the online training modules developed by Health Education England (HEE). In the longer-term, clearer mapping of the required workforce, and a series of competency frameworks will need to be developed in collaboration with relevant professional organisations.

Data, audit and research

1.39. A lack of routine and consistent data collection means that it is not possible to accurately track the outcomes and pathways children and young people take through the service. Standardised data collection is required in order to audit service standards and inform understanding of the epidemiology, assessment and treatment of this group. This, alongside a national network which brings providers together, will help build knowledge and improve outcomes through shared clinical standards and systematic data collection. In the longer-term, formalisation of such a network into a learning health system[9] with an academic host would mean that there was systematised use of data to produce a continuing research programme with rapid translation into clinical practice and a focus on training.

[9] Scobie S, Castle-Clarke S (2019). Implementing learning health systems in the UK NHS: Policy actions to improve collaboration and transparency and support innovation and better use of analytics. Learning Health Systems 4(1): e10209. DOI:10.1002/lrh2.10209.

5: The regional services should have regular co-ordinated national provider meetings and operate to shared standards and operating procedures with a view to establishing a formal learning health system.

6: Existing and future services should have standardised data collection in order to audit standards and inform understanding of the epidemiology, assessment and treatment of this group of children and young people.

7: Prospective consent of children and young people should be sought for their data to be used for continuous service development, to track outcomes, and for research purposes. Within this model, children and young people put on hormone treatment should be formally followed up into adult services, ideally as part of an agreed research protocol, to improve outcome data.

Clinical approach

Assessment processes

1.40. We have heard that there are inconsistencies and gaps in the assessment process. Our work to date has also demonstrated that clinical staff have different views about the purpose of assessment and where responsibility lies for different components of the process within the pathway of care. The Review team has commenced discussions with clinical staff across primary, secondary and tertiary care to develop a framework for these processes.

8: There needs to be agreement and guidance about the appropriate clinical assessment processes that should take place at primary, secondary and tertiary level.

9: Assessments should be respectful of the experience of the child or young person and be developmentally informed. Clinicians should remain open and explore the patient's experience and the range of support and treatment options that may best address their needs, including any specific needs of neurodiverse children and young people.

Hormone treatment

1.41. The issues raised by the Multi-Professional Review Group echo several of the problems highlighted by the CQC. It is essential that principles of the General Medical Council's Good Practice in Prescribing and Managing Medicine's and Devices[10] are closely followed, particularly given the gaps in the evidence base regarding hormone treatment. Standards for decision making regarding endocrine treatment should also be consistent with international best practice.[11,12,13]

10: Any child or young person being considered for hormone treatment should have a formal diagnosis and formulation, which addresses the full range of factors affecting their physical, mental, developmental and psychosocial wellbeing. This formulation should then inform what options for support and intervention might be helpful for that child or young person.

11: Currently paediatric endocrinologists have sole responsibility for treatment, but where a life-changing intervention is given there should also be additional medical responsibility for the differential diagnosis leading up to the treatment decision.

1.42. Paediatric endocrinologists develop a wide range of knowledge within their paediatric training, including safeguarding, child mental health, and adolescent development. Being party to the discussions and deliberations that have led up to the decision for medical intervention supports them in carrying out their legal responsibility for consent to treatment and the prescription of hormones.

12: Paediatric endocrinologists should become active partners in the decision making process leading up to referral for hormone treatment by participating in the multidisciplinary team meeting where children being considered for hormone treatment are discussed.

---

[10] General Medical Council (2021). Good practice in prescribing and managing medicines and devices (76-78).
[11] Hembree WC, Cohen-Kettenis PT, Gooren L, Hannema SE, Meyer WJ, Murad MH, et al (2017). Endocrine treatment of gender-dysphoric/gender-incongruent persons: an Endocrine Society clinical practice guideline. J Clin Endocrinol Metab 102(11): 3869–903. DOI: 10.1210/jc.2017-01658.
[12] Cohen-Kettenis PT, Steensma TD, de Vries ALC (2001). Treatment of adolescents with gender dysphoria in the Netherlands. Child Adolesc Psychiatr Clin N Am 20: 689–700. DOI: 10.1016/j.chc.2011.08.001.
[13] Kyriakou A, Nicolaides NC, Skordis N (2020). Current approach to the clinical care of adolescents with gender dysphoria. Acta Biomed 91(1): 165–75. DOI: 10.23750/abm.v91i1.9244.

1.43. Given the uncertainties regarding puberty blockers, it is particularly important to demonstrate that consent under this circumstance has been fully informed and to follow GMC guidance[14] by keeping an accurate record of the exchange of information leading to a decision in order to inform their future care and to help explain and justify the clinician's decisions and actions.

13: Within clinical notes, the stated purpose of puberty blockers as explained to the child or young person and parent should be made clear. There should be clear documentation of what information has been provided to each child or young person on likely outcomes and side effects of all hormone treatment, as well as uncertainties about longer-term outcomes.

14: In the immediate term the Multi-Professional Review Group (MPRG) established by NHS England should continue to review cases being referred by GIDS to endocrine services.

[14] General Medical Council (2020). Decision making and consent.

2. Context

Transgender, non-binary and gender fluid adults

2.1. NHS clinical services to support transgender adults with hormone treatment and subsequent surgery began in 1966.

2.2. Services were initially established within a mental health model, in conjunction with endocrinology and surgical services.

2.3. Currently, NHS services for transgender adults do not have adequate capacity to cope with demand.[15] In addition, the broader healthcare needs of this group are not well met. This is important in the context of the current generation of gender-questioning children and young people in that there are now two inflows into adult services – individuals transitioning in adulthood, and those moving through from children's services.

2.4. Legal rights and protections for transgender people lagged behind the provision of medical services, with the Gender Recognition Act 2004 coming into force in April 2005. Over the last few years, broader discussions about transgender issues have been played out in public, with discussions becoming increasingly polarised and adversarial. This polarisation is such that it undermines safe debate and creates difficulties in building consensus.

2.5. It is not the role of this Review to take any position on the cultural and societal debates relating to transgender adults. However, in achieving its objectives there is a need to consider the information and support that children and young people access from whatever source, as well as any pressures that they are subject to, before they access clinical services.

Terminology and diagnostic frameworks

2.6. The Office for National Statistics defines sex as "referring to the biological aspects of an individual as determined by their anatomy, which is produced by their chromosomes, hormones and their interactions; generally male or female; something that is assigned at birth".[16]

2.7. The Office for National Statistics defines gender as "a social construction relating to behaviours and attributes based on labels of masculinity and femininity; gender identity is a personal, internal perception of oneself and so the gender category someone identifies with may not match the sex they were assigned at birth".[17]

2.8. Societal attitudes towards gender roles and gender expression are changing. Children, teenagers and younger adults may more commonly see gender as a fluid, multi-faceted phenomenon which

[15] Gender Identity Clinic, The Tavistock and Portman NHS Foundation Trust. Waiting times.
[16] Office for National Statistics (2019). What is the difference between sex and gender?
[17] Ibid.

does not have to be binary, whereas older generations have tended to see gender as binary and fixed. It is not unusual for young people to explore both their sexuality and gender as they go through adolescence and early adulthood before developing a more settled identity. Many achieve this without experiencing significant distress or requiring support from the NHS, but this is not the case for all.

2.9. For those who require support from the NHS, there are two widely used frameworks which provide diagnostic criteria. The International Classification of Diseases (ICD), which is the World Health Organization (WHO) mandated health data standard, and the Diagnostic and Statistical Manual of Mental Disorders (DSM), which is the classification system for mental health disorders produced by the American Psychiatric Association. The current editions of these manuals – ICD-11 and DSM-5 – came into effect in January 2022 and 2013 respectively.

2.10. ICD-11[18] has attempted to depathologise gender diversity, removing the term 'gender identity disorders' from its mental health section and creating a new section for gender incongruence and transgender identities in a chapter on sexual health. These changes are part of a much broader societal drive to remove the stigma previously associated with transgender healthcare. ICD-11 defines gender incongruence as being "characterised by a marked incongruence between an individual's experienced/ expressed gender and the assigned sex." Gender variant behaviour and preferences alone are not a basis for assigning the diagnosis. The full criteria for gender incongruence of childhood and gender incongruence of adolescence or adulthood are listed in **Appendix 3**.

2.11. DSM-5[19] is currently the framework used to diagnose gender dysphoria. This diagnostic category describes gender dysphoria as "the distress that may accompany the incongruence between one's experienced or expressed gender and one's assigned gender". A diagnosis of gender dysphoria is usually deemed necessary before a young person can access hormone treatment, and criteria are listed in **Appendix 3**.

Conceptual understanding of gender incongruence in children and young people

2.12. Children and young people presenting to gender identity services are not a homogeneous group. They vary in their age at presentation, their cultural background, whether they identify as binary, non-binary, or gender fluid, whether they are neurodiverse and in a host of other ways.

[18] World Health Organization (2022). International Classification of Diseases Eleventh Revision.
[19] American Psychiatric Association (2013). Diagnostic and Statistical Manual of Mental Health Disorders: DSM-5™, 5th ed.

2.13. Some children and young people may thrive during a period of gender-questioning whilst for others it can be accompanied with a level of distress that can have a significant impact on their functioning and development.

2.14. Alongside these very varied presentations, it is highly unlikely that a single cause for gender incongruence will be found. Many authors view gender expression as a result of a complex interaction between biological, cultural, social and psychological factors.

2.15. Despite a high level of agreement about these points, there are widely divergent and, in some instances, quite polarised views among service users, parents, clinical staff and the wider public about how gender incongruence and gender-related distress in children and young people should be interpreted, and this has a bearing on expectations about clinical management.

2.16. These views will be influenced by how each individual weighs the balance of factors that may lead to gender incongruence, and the distress that may accompany it. Beliefs about whether it might be inherent and/or immutable, whether it might be a transient response to adverse experiences, whether it might be highly fluid and/or likely to change in later adolescence/early adulthood, etc will have a profound influence on expectations about treatment options.[20]

2.17. All of these views may be overlaid with strongly held concerns about children's and young people's rights, autonomy, and/or protection.

2.18. The disagreement and polarisation is heightened when potentially irreversible treatments are given to children and young people, when the evidence base underlying the treatments is inconclusive, and when there is uncertainty about whether, for any particular child or young person, medical intervention is the best way of resolving gender-related distress.

2.19. As with many other contemporary polarised disagreements, the situation is exacerbated when there is no space to have open, non-judgemental discussions about these differing perspectives. A key aim of this review process will be to encourage such discussions in a safe and respectful manner so that progress can be made in finding solutions.

[20] Wren B (2019). Notes on a crisis of meaning in the care of gender-diverse children. In: Hertzmann L, Newbigin J (eds) Sexuality and Gender Now: Moving Beyond Heteronormativity. Routledge.

3. Current services

Current service model for gender-questioning children and young people

3.1. Currently there are no locally or regionally commissioned services for children and young people who seek help from the NHS in managing their gender-related distress. Within primary and secondary care, some clinical staff have more interest and expertise in initial management of this group of young people, but such individuals are few and far between.

3.2. The pathway for NHS support around gender identity for children and young people is designated as a highly specialised service.[21] The Gender Identity Development Service (GIDS) at the Tavistock and Portman NHS Foundation Trust is commissioned by NHS England to provide specialist assessment, support and, where appropriate, hormone intervention for children and young people with gender dysphoria. It is the only NHS provider of specialist gender services for children and young people in England. The Trust runs satellite bases in Leeds and Bristol. Until recently GIDS accepted referrals from multiple sources, for example, GPs, secondary care, social care, schools, and support and advocacy groups, which is unusual for a specialist service.

3.3. Children and young people are assessed by two members of the GIDS team who may be any combination of psychologists, psychotherapists, family therapists, or social workers. If there is uncertainty about the right approach, individual cases may be discussed in a complex case meeting. Those deemed appropriate for physical interventions are referred on to the endocrine team; under the current Standard Operating Procedure (SOP), this decision requires a multi-disciplinary team (MDT) discussion within GIDS. A member of the GIDS team attends new appointments in the endocrine clinic, but they will not routinely be the member of staff who saw the young person for assessment. However, very recently a triage meeting has been piloted to enable endocrinologists to discuss upcoming appointments with the clinician who saw the young person for assessment. The young person then attends an education session prior to their endocrine appointment. The endocrinologist will assess any medical contraindications prior to seeking consent from the patient for any hormone treatments.

3.4. For many years, the GIDS approach was to offer assessment and support, and to only start puberty blockers when children reached sexual maturity at about age 15 (Tanner Stage 5) as the first step in the treatment process to feminise or masculinise the young person, with

[21] National Health Service Commissioning Board and Clinical Commissioning Groups (Responsibilities and Standing Rules) Regulations 2012.

oestrogen or testosterone given from age 16. Feminising/masculinising hormones are not given at an earlier stage because of the irreversibility of some of their actions in developing secondary sex characteristics of the acquired gender.[22,23]

3.5. In 1998, a new protocol was published by the Amsterdam gender identity clinic.[24] It was subsequently named the Dutch Approach.[25] This involved giving puberty blockers much earlier, from the time that children showed the early signs of puberty (Tanner Stage 2), to pause further pubertal changes of the sex at birth. This stage of pubertal development was chosen because it was felt that although many younger children experienced gender incongruence as a transient developmental phenomenon, those who expressed early gender incongruence which continued into puberty were unlikely to desist at that stage.

3.6. It was felt that blocking puberty would buy time for children and young people to fully explore their gender identity and help with the distress caused by the development of their secondary sexual characteristics. The Dutch criteria for treating children with early puberty blockers were: (i) a presence of gender dysphoria from early childhood; (ii) an increase of the gender dysphoria after the first pubertal changes; (iii) an absence of psychiatric comorbidity that interferes with the diagnostic work-up or treatment; (iv) adequate psychological and social support during treatment; and (v) a demonstration of knowledge and understanding of the effects of gonadotropin-releasing hormones (puberty blockers), feminising/masculinising hormones, surgery, and the social consequences of sex reassignment.[26]

3.7. Under the Dutch Approach, feminising/masculinising hormones were started at age 16 and surgery was permitted to be undertaken from age 18, as in England.

3.8. From 2011, early administration of puberty blockers was started in England under a research protocol, which partially paralleled the Dutch Approach (the Early Intervention Study). From 2014, this protocol was adopted by GIDS as routine clinical practice. Results of the Early Intervention Study were published in December 2021.[27]

---

[22] Delemarre-van de Wall HA, Cohen-Kettinis PT (2006). Clinical management of gender identity disorder in adolescents: a protocol on psychological and paediatric endocrinology aspects. Eur J Endocrinol 155 (Suppl 1): S131–7. DOI: 10.1530/eje.1.02231.
[23] de Vries ALC, Cohen-Kettenis PT (2012). Clinical management of gender dysphoria in children and adolescents: the Dutch approach. J Homosex 59: 301–320. DOI: 10.1080/00918369.2012.653300.
[24] Cohen-Kettenis PT, Van Goozen S (1998). Pubertal delay as an aid in diagnosis and treatment of a transsexual adolescent. Eur Child Adolesc Psychiatry 7: 246–8. DOI: 10.1007/s007870050073.
[25] de Vries ALC, Cohen-Kettenis PT (2012). Clinical management of gender dysphoria in children and adolescents: the Dutch approach. J Homosex 59: 301–320. DOI: 10.1080/00918369.2012.653300.
[26] Ibid.
[27] Carmichael P, Butler G, Masic U, Cole TJ, De Stavola BL, Davidson S, et al (2021). Short-term outcomes of pubertal suppression in a selected cohort of 12 to 15 year old young people with persistent gender dysphoria in the UK. PLoS One. 16(2):e0243894. DOI:10.1371/journal.pone.0243894.

3.9. However, the Dutch Approach differs from the GIDS approach in having stricter requirements about provision of psychological interventions. For example, under the Dutch Approach, if young people have gender confusion, aversion towards their sexed body parts, psychiatric comorbidities or Autism Spectrum Disorder (ASD) related diagnostic difficulties, they may receive psychological interventions only, or before, or in combination with medical intervention. Of note, in 2011, the Amsterdam team were reporting that up to 10% of their referral base were young people with ASD.[28]

Changing epidemiology

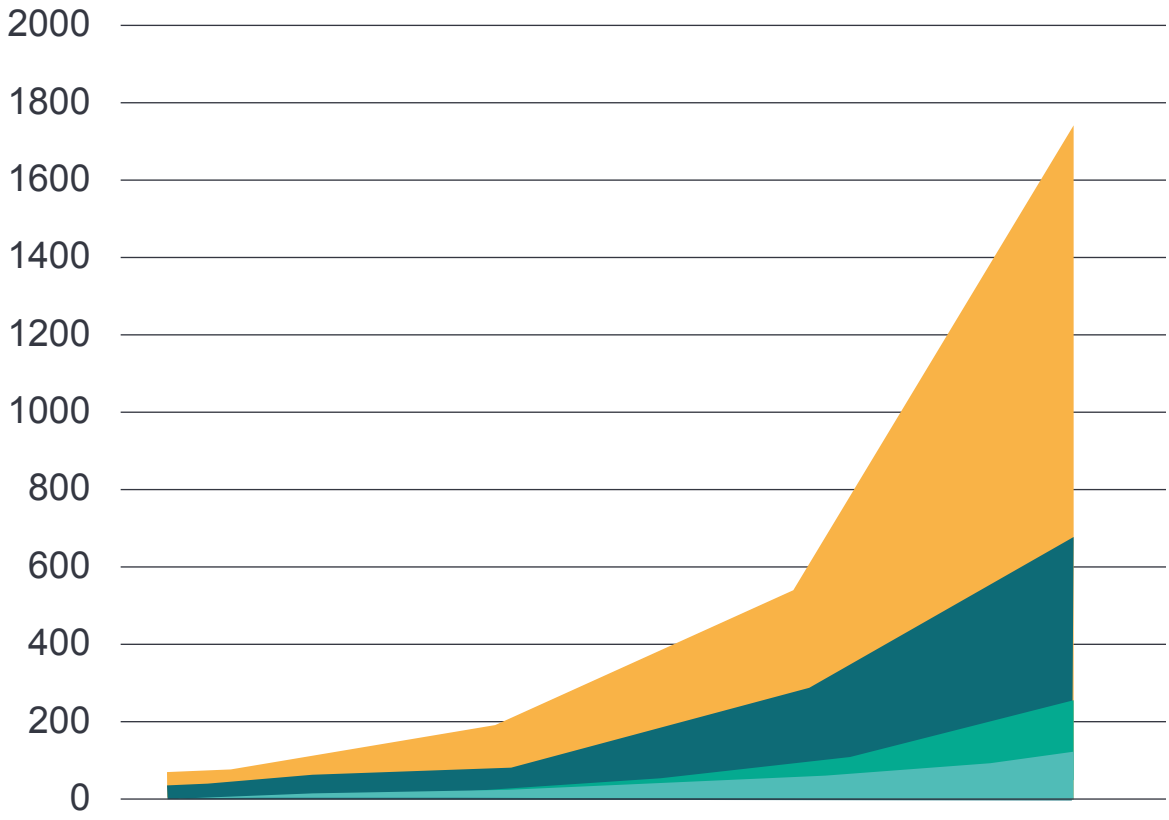
3.10. In the last few years, there has been a significant change in the numbers and case-mix of children and young people being referred to GIDS.[29] From a baseline of approximately 50 referrals per annum in 2009, there was a steep increase from 2014-15, and at the time of the CQC inspection of the Tavistock and Portman NHS Foundation Trust in October 2020 there were 2,500 children and young people being referred per annum, 4,600 children and young people on the waiting list, and a waiting time of over two years to first appointment.[30] This has severely impacted on the capacity of the existing service to manage referrals in the safe and responsive way that they aspire to and has led to considerable distress for those on the waiting list.

3.11. This increase in referrals has been accompanied by a change in the case-mix from predominantly birth-registered males presenting with gender incongruence from an early age, to predominantly birth-registered females presenting with later onset of reported gender incongruence in early teen years. In addition, approximately one third of children and young people referred to GIDS have autism or other types of neurodiversity. There is also an over-representation percentage wise (compared to the national percentage) of looked after children.[31]

---

[28] Cohen-Kettenis PT, Steensma TD, de Vries ALC (2001). Treatment of adolescents with gender dysphoria in the Netherlands. Child Adolesc Psychiatr Clin N Am 20: 689–700. DOI: 10.1016/j.chc.2011.08.001.
[29] de Graaf NM, Giovanardi G, Zitz C, Carmichael P (2018). Sex ratio in children and adolescents referred to the gender identity development service in the UK (2009-2016). Arch Sex Behav 47(5): 1301–4.
[30] Care Quality Commission (2021). The Tavistock and Portman NHS Foundation Trust Gender Identity Service Inspection Report. London: CQC.
[31] Matthews T, Holt V, Sahin S, Taylor A, Griksaitis (2019). Gender Dysphoria in looked-after and adopted young people in a gender identity development service. Clinical Child Psychol Psychiatry 24: 112-128. DOI: 10.1177/1359104518791657.

Figure 1: Sex ratio in children and adolescents referred to GIDS in the UK (2009-16)



	2009	2010	2011	2012	2013	2014	2015	2016
Adolescents F	15	48*	78*	141*	221*	314*	689*	1071*
Adolescents M	24	44*	41	77*	120*	185*	293*	426*
Children F	2	7	12	17	22	36	77*	138*
Children M	10	19	29	30	31	55*	103*	131

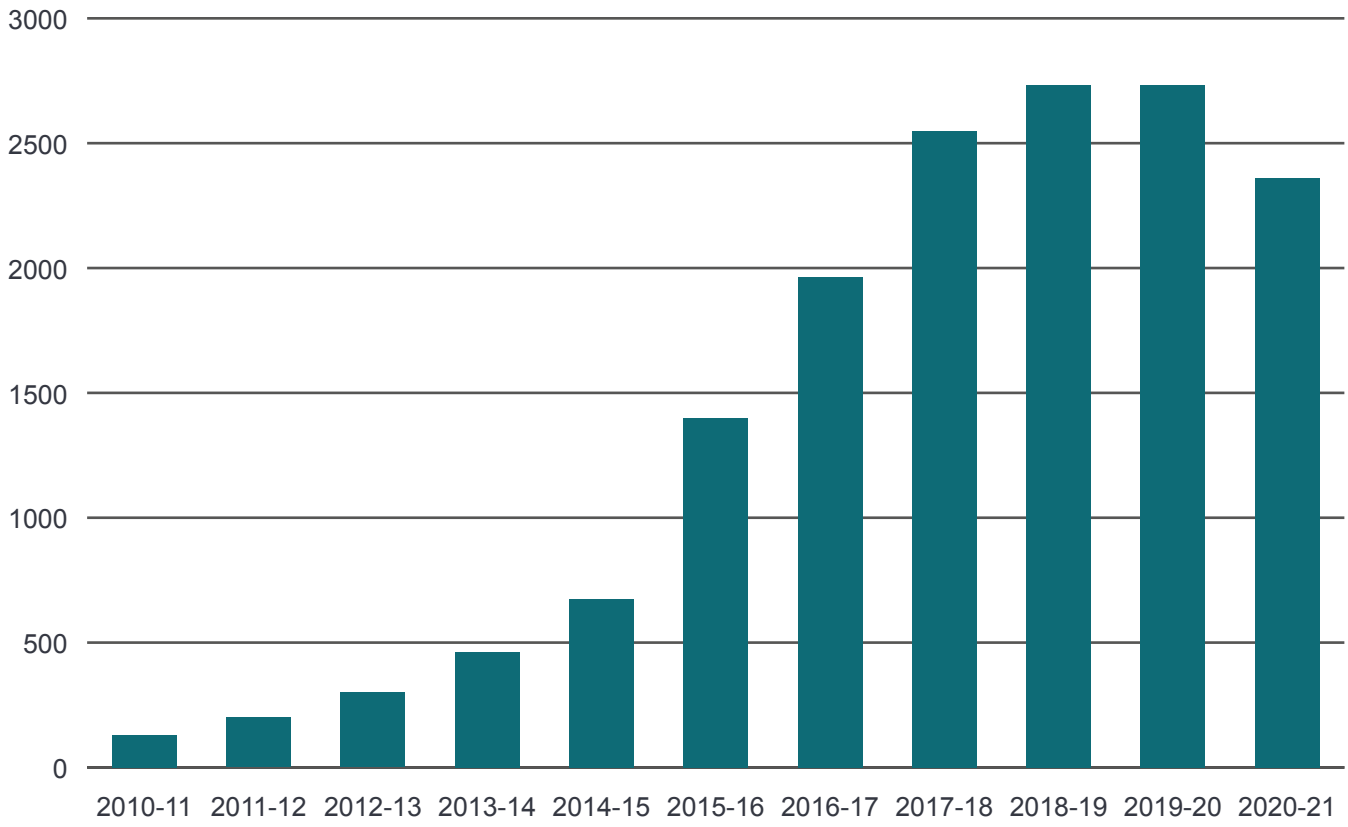
AFAB = assigned female at birth; AMAB = assigned male at birth

***Indicates** p<.05 which shows a significant increase of referrals compared to the previous year

Source: de Graaf NM, Giovanardi G, Zitz C, Carmichael P (2018).[32]

[32] de Graaf NM, Giovanardi G, Zitz C, Carmichael P (2018). Sex ratio in children and adolescents referred to the gender identity development service in the UK (2009-2016). Arch Sex Behav 47(5): 1301–4.

Figure 2: Referrals to GIDS, 2010-11 to 2020-21



Source: Gender Identity Development Service.[33]

3.12. In 2019, GIDS reported that about 200 children and young people from a referral base of 2,500 were referred on to the endocrine pathway. There is no published data on how the other children and young people from this referral baseline were managed, for example if: their gender dysphoria was resolved; they were still being assessed or receiving ongoing psychological support and input; they were not eligible for puberty blockers due to age; they were referred to endocrine services at a later stage; they were transferred to adult services; or they accessed private services.

Challenges to the service model and clinical approach

3.13. Over a number of years, in parallel with the increasing numbers of referrals, GIDS faced increasing challenges, both internally and externally. There were different views held within the staff group about the appropriate clinical approach, with some more strongly affirmative and some more cautious and concerned about the use of physical intervention. The complexity of the cases had also increased, so clinical decision making had become more difficult. There was also a high staff

[33] Gender Identity Development Service. Referrals to GIDS, financial years 2010-11 to 2020-21.

turnover, and accounts from staff concerned about the clinical care, which were picked up in both mainstream and social media. This culminated in 2018 with an internal report by a staff governor.

3.14. Following that report, a review was carried out in 2019 by the Trust's medical director. This set out the need for clearer processes for the service's referral management, safeguarding, consent, and clinical approach, and an examination of staff workload and support, and a new Standard Operating Procedure (SOP) was put in place.

NHS England Policy Working Group

3.15. In January 2020, a Policy Working Group (PWG) was established by NHS England to undertake a review of the published evidence on the use of puberty blockers and feminising/masculinising hormones in children and young people with gender dysphoria to inform a policy position on their future use. Given the increasingly evident polarisation among clinical professionals, Dr Cass was asked to chair the group as a senior clinician with no prior involvement or fixed views in this area. The PWG comprised an expert group including endocrinologists, child and adolescent psychiatrists and paediatricians representing their respective Royal Colleges, an ethicist, a GP, senior clinicians from the NHS GIDS, a transgender adult and parents of gender-questioning young people. The process was supported by a public health consultant and policy, pharmacy and safeguarding staff from NHS England.

3.16. NHS England uses a standardised protocol for developing clinical policies. The first step of this involves defining the PICO (the Population being treated, the Intervention, a Comparator treatment, and the intended Outcomes). This of itself was challenging, with a particular difficulty being definition of the intended outcomes of puberty blockers, and suitable comparators for both hormone interventions. However, agreement was reached on what should be included in the PICO and subsequently the National Institute for Health and Care Excellence (NICE) was commissioned to review the published evidence,[34,35] again following a standardised protocol which has strict criteria about the quality of studies that can be included.

3.17. Unfortunately, the available evidence was not strong enough to form the basis of a policy position. Some of the challenges and outstanding uncertainties are summarised as follows.

---

[34] National Institute for Health and Care Excellence (2020). Evidence Review: Gonadotrophin Releasing Hormone Analogues for Children and Adolescents with Gender Dysphoria.
[35] National Institute for Health and Care Excellence (2020). Evidence review: gender-affirming hormones for children and adolescents with gender dysphoria.

Feminising/masculinising hormones

3.18. Sex hormones have been prescribed for transgender adults for several decades, and the long-term risks and side effects are well understood. These include increased cardiovascular risk, osteoporosis, and hormone-dependent cancers.

3.19. In young people, consideration also needs to be given to the impact on fertility, with the need for fertility counselling and preservation.

3.20. The additional physical risk of starting these treatments at age 16+ rather than age 18+ is unlikely to add significantly to the total lifetime risk, although data on this will not be available for many years. However, as evidenced by take-up of treatment with feminising/masculinising hormones, where there is a high level of certainty that physical transition is the right option, the child or young person may be more accepting of these risks, which can seem remote from the immediate gender distress.

3.21. The most difficult question in relation to feminising/masculinising hormones therefore is not about long-term physical risk which is tangible and easier to understand. Rather, given the irreversible nature of many of the changes, the greatest difficulty centres on the decision to proceed to physical transition; this relies on the effectiveness of the assessment, support and counselling processes, and ultimately the shared decision making between clinicians and patients. Decisions need to be informed by long-term data on the range of outcomes, from satisfaction with transition, through a range of positive and negative mental health outcomes, through to regret and/or a decision to detransition. The NICE evidence review demonstrates the poor quality of these data, both nationally and internationally.

3.22. Regardless of the nature of the assessment process, some children and young people will remain fluid in their gender identity up to early to mid-20s, so there is a limit as to how much certainty one can achieve in late teens. This is a risk that needs to be understood during the shared decision making process with the young person.

3.23. It is also important to note that any data that are available do not relate to the current predominant cohort of later-presenting birth-registered female teenagers. This is because the rapid increase in this subgroup only began from around 2014-15. Since young people may not reach a settled gender expression until their mid-20s, it is too early to assess the longer-term outcomes of this group.

Puberty blockers

3.24. The administration of puberty blockers is arguably more controversial than administration of the feminising/ masculinising hormones, because there are more uncertainties associated with their use.

3.25. There has been considerable discussion about whether the treatment is 'experimental'; strictly speaking an experimental treatment is one that is being given as part of a research protocol, and this is not the case with puberty blockers, because the GIDS research protocol was stopped in 2014. At that time, the treatment was experimental and innovative, because the drug was licensed for use in children, but specifically for children with precocious puberty. This was therefore the first time it was used 'off-label' in the UK for children with gender dysphoria. If a drug is used 'off-label' it means it is being used for a condition that is different from the one for which it was licensed. The many uncertainties around the 'off-label' use were recognised, but given that this was not a new drug, it did not need Medicines and Healthcare products Regulatory Agency (MHRA) approval at that time.

3.26. The important question now, as with any treatment, is whether the evidence for the use and safety of the medication is strong enough as judged by reasonable clinical standards.

3.27. One of the challenges that NHS England's PWG faced in considering this question was the lack of clarity about intended outcomes, several of which have been proposed including:

- providing time/space for the young person to make a decision about continuing with transition;
- reducing or preventing worsening of distress;
- improving mental health; and
- stopping potentially irreversible pubertal changes which might later make it difficult for the young person to 'pass' in their intended gender role.

3.28. Proponents for the use of puberty blockers highlight the distress that young people experience through puberty and the risk of self-harm or suicide.[36] However, some clinicians do not feel that distress is actually alleviated until children and young people are able to start feminising/ masculinising hormones. The Review will seek to gain a better understanding of suicide data and the impact of puberty blockers through its research programme.

3.29. On the other hand, it has been asserted that starting puberty blockers at an older age provides children and young people with more time to achieve fertility preservation. In the case of birth-registered males, there is an argument that it also

[36] Turban JL, King D, Carswell JM, et al (2020). Pubertal suppression for transgender youth and risk of suicidal ideation. Pediatrics 145 (2): e20191725. DOI: 10.1542/peds.2019-1725.

allows more time to achieve adequate penile growth for successful vaginoplasty.

3.30. In the short-term, puberty blockers may have a range of side effects such as headaches, hot flushes, weight gain, tiredness, low mood and anxiety, all of which may make day-to-day functioning more difficult for a child or young person who is already experiencing distress. Short-term reduction in bone density is a well-recognised side effect, but data is weak and inconclusive regarding the long-term musculoskeletal impact.[37]

3.31. The most difficult question is whether puberty blockers do indeed provide valuable time for children and young people to consider their options, or whether they effectively 'lock in' children and young people to a treatment pathway which culminates in progression to feminising/masculinising hormones by impeding the usual process of sexual orientation and gender identity development. Data from both the Netherlands[38] and the study conducted by GIDS[39] demonstrated that almost all children and young people who are put on puberty blockers go on to sex hormone treatment (96.5% and 98% respectively). The reasons for this need to be better understood.

3.32. A closely linked concern is the unknown impacts on development, maturation and cognition if a child or young person is not exposed to the physical, psychological, physiological, neurochemical and sexual changes that accompany adolescent hormone surges. It is known that adolescence is a period of significant changes in brain structure, function and connectivity.[40] During this period, the brain strengthens some connections (myelination) and cuts back on others (synaptic pruning). There is maturation and development of frontal lobe functions which control decision making, emotional regulation, judgement and planning ability. Animal research suggests that this development is partially driven by the pubertal sex hormones, but it is unclear whether the same is true in humans.[41] If pubertal sex hormones are essential to these brain maturation processes, this raises a secondary question of whether there is a critical time window for the processes to take place, or whether catch up is possible when oestrogen or testosterone is introduced later.

---

[37] National Institute for Health and Care Excellence (2020). Evidence Review: Gonadotrophin Releasing Hormone Analogues for Children and Adolescents with Gender Dysphoria.
[38] Brik T, Vrouenraets LJJJ, de Vries MC, Hannema SE (2020). Trajectories of adolescents treated with gonadotropin-releasing hormone analogues for gender dysphoria. Arch Sex Behav 49: 2611–8. DOI: 10.1007/s10508-020-01660-8.
[39] Carmichael P, Butler G, Masic U, Cole TJ, De Stavola BL, Davidson S, et al (2021). Short-term outcomes of pubertal suppression in a selected cohort of 12 to 15 year old young people with persistent gender dysphoria in the UK. PLoS One. 16(2):e0243894. DOI:10.1371/journal.pone.0243894.
[40] Delevichab K, Klinger M, Nana OJ, Wilbrecht L (2021). Coming of age in the frontal cortex: The role of puberty in cortical maturation. Semin Cell Dev Biol 118: 64–72. DOI: 10.1016/j.semcdb.2021.04.021.
[41] Goddings A-L, Beltz A, Jiska S, Crone EA, Braams BR (2019). Understanding the role of puberty in structural and functional development of the adolescent brain. J Res Adolesc 29(1): 32–53. DOI: 10.1111/jora.12408.

3.33. An international interdisciplinary panel[42] has highlighted the importance of understanding the neurodevelopmental outcomes of pubertal suppression and defined an appropriate approach for investigating this further. However, this work has not yet been undertaken.

Initiation of Cass Review

3.34. Dr Cass' own reflections on the PWG process, the available literature, and the issues it highlighted were as follows:

- Firstly, that hormone treatment is just one possible outcome for gender-questioning children and young people. A much better understanding is needed about: the increasing numbers of children and young people with gender-related distress presenting for help; the appropriate clinical pathway for each individual; their support needs; and the full range of potential treatment options.
- Secondly, there is very limited follow-up of the subset of children and young people who receive hormone treatment, which limits our understanding about the long-term outcomes of these treatments and this lack of follow up data should be corrected.
- Thirdly, the assessment process is inconsistent across the published literature. The outcome of hormone treatment is highly influenced by whether the assessment process accurately selects those children and young people most likely to benefit from medical treatment. This makes it difficult to draw conclusions from published studies.

3.35. In light of the above, NHS England commissioned this independent review to make recommendations on how the clinical management and service provision for children and young people who are experiencing gender incongruence or gender-related distress can be improved.

CQC inspection

3.36. In October and November 2020, the Care Quality Commission (CQC) inspectors carried out an announced, focused inspection of GIDS due to concerns reported to them by healthcare professionals and the Children's Commissioner for England. Concerns related to clinical practice, safeguarding procedures, and assessments of capacity and consent to treatment.

[42] Chen D, Strang JF, Kolbuck VD, Rosenthal SM, Wallen K, Waber DP, et al (2020). Consensus parameter: research methodologies to evaluate neurodevelopmental effects of pubertal suppression in transgender youth. Transgender Health 5(4). DOI: 10.1089/trgh.2020.0006.

3.37. The CQC report, published in January 2021,[43] gave the service an overall rating of inadequate. The report noted the high level of commitment and caring approach of the staff but identified a series of issues that needed improvement. In addition to the growing waiting list pressures, the CQC identified problems in several other areas including: the assessment and management of risk; the variations in clinical approach; the lack of clarity and consistency of care plans; the lack of any clear written rationale for decision making in individual cases; and shortfalls in the multidisciplinary mix required for some patient groups. Recording of capacity, competency and consent had improved since the new SOP in January 2020; however, there remained a culture in which staff reported feeling unable to raise concerns.

3.38. The CQC reported that when it inspected GIDS, there did not appear to be a formalised assessment process, or standard questions to explore at each session, and it was not possible to tell from the notes why an individual child might have been referred to endocrinology whilst another had not. Current GIDS data demonstrate that a majority of children and young people seen by the service do not get referred for endocrine treatment, but there is no clear information about what other diagnoses they receive, and what help or support they might need.

3.39. Since the CQC report, NHS England and The Tavistock and Portman NHS Foundation Trust management team have been working to address the issues raised. However, whilst some problems require a focused Trust response, the waiting list requires a system-wide response. This was noted in the letter from the Review to NHS England in May 2021 (**Appendix 2**).

Legal background

3.40. This section sets out the chronology of recent case law. In October 2019, a claim for Judicial Review was brought against The Tavistock and Portman NHS Foundation Trust. The claimants' case was summarised by the High Court as follows: "The claimants' case is that children and young persons under 18 are not competent to give consent to the administration of puberty blocking drugs. Further, they contend that the information given to those under 18 by the defendant [GIDS] is misleading and insufficient to ensure such children or young persons are able to give informed consent. They further contend that the absence of procedural safeguards, and the inadequacy of the information provided, results in an infringement of the rights of such children and young persons under Article 8 of the European Convention

[43] Care Quality Commission (2021). The Tavistock and Portman NHS Foundation Trust Gender Identity Service Inspection Report. London: CQC.

for the Protection of Human Rights and Fundamental Freedoms."[44]

3.41. In December 2020, three judges in the High Court of England and Wales handed down judgment in *Bell v Tavistock.*[45] (Most cases in the High Court are heard by a single judge sitting alone, and when a case is heard by more than one judge in the High Court, it is described as the Divisional Court.) The Divisional Court recognised that the Tavistock's policies and practices as set out in the service specification were not unlawful. However, the Court made a declaration that set out in detail a series of implications of treatment that a child would need to understand to be *Gillick* competent[46] to consent to puberty blockers. Specifically, because most children put on puberty blockers go on to have feminising/masculinising hormones, the judgment said a child would need to understand not only the full implications of puberty blocking drugs, but also the implications of the full pathway of medical and surgical transition. The judges concluded that it will be "very doubtful" that 14-15 year-olds have such competence, and "highly unlikely" that children aged 13 or under have competence for that decision. Under the Mental Capacity Act 2005, 16-17 year-olds are presumed to have capacity, and they are effectively treated as adults for consent to medical treatment under the Family Law Reform Act 1969 section 8, but the judges suggested that it would be appropriate for clinicians to involve the court in any case where there were doubts as to whether the proposed treatment would be in the long term best interests of a 16-17 year-old.

3.42. Following the Divisional Court judgment in *Bell v Tavistock*, a claim was brought against the Tavistock in the High Court Family Division by the mother of a child for a declaration that she and the child's father had the ability in law to consent on behalf of their child to the administration of puberty blockers *(AB v CD)*.[47] The Court concluded that "the parents' right to consent to treatment on behalf of the child continues even when the child is *Gillick* competent to make the decision, save where the parents are seeking to override the decision of the child" [para 114] and that there is no "general rule that puberty blockers should be placed in a special category by which parents are unable in law to give consent" [para 128].

---

[44] Bell v Tavistock. [2020] EWHC 3274 (Admin).
[45] Ibid.
[46] Gillick v West Norfolk and Wisbech AHA [1986] AC 112.
[47] AB v CD & Ors [2021] EWHC 741.

3.43. Subsequently, the Tavistock appealed the Divisional Court's earlier decision in *Bell v Tavistock* and was successful.[48] The Court of Appeal held that it was not appropriate for the Divisional Court to provide the guidance about the likelihood of having *Gillick* competence at particular ages, or about the need for court approval [para 91]. The Court of Appeal went on to say "The Divisional Court concluded that Tavistock's policies and practices (as expressed in the service specification and the SOP) were not unlawful and rejected the legal criticism of its materials. In those circumstances, the claim for judicial review is dismissed." [para 91]. However, clinicians should "take great care before recommending treatment to a child and be astute to ensure that the consent obtained from both child and parents is properly informed" [para 92].

3.44. The Court of Appeal in *Bell v Tavistock* recognised the lawfulness of treating children for gender dysphoria in this jurisdiction. Recognising the divergences in medical opinion, morality and ethics, it indicated that the question of whether treatment should be made available is a matter of policy "for the National Health Service, the medical profession and its regulators and Government and Parliament" [para 3].

3.45. Following the Divisional Court decision in *Bell v Tavistock*, new referrals for puberty blockers were suspended and a requirement was put in place that children currently on puberty blockers were reviewed with a view to court proceedings for a judge to determine the best interests for children in whom these medications were considered essential. This requirement was changed following *AB v CD*, with the reinstatement of the hormone pathway in March 2021. However, an external panel, the Multi Professional Review Group (MPRG), was established to ensure that procedures for assessment and for informed consent had been properly followed. The outcome of the *Bell* appeal has not changed this requirement, which is contingent not just on the legal processes but on the concerns raised by CQC regarding consent, documentation and clarity about decision making within the service.[49]

[48] EWCA [2021] Civ 1363.
[49] Care Quality Commission (2021). The Tavistock and Portman NHS Foundation Trust Gender Identity Service Inspection Report. London: CQC.

The Multi-Professional Review Group

3.46. NHS England has established a Multi-Professional Review Group (MPRG) to review whether the agreed process has been followed for a child to be referred into the endocrinology clinic and to be prescribed treatment. The Review has spoken directly to the MPRG, which has reported its observations of current practice.

3.47. The MPRG has stated that its work has been impeded by delays in the provision of clinical information, the lack of structure in the documentation received, and gaps in the necessary evidence. This means that when reviewing the documents provided it is not always easy to determine if the process for referral for endocrine treatment has been fully or safely followed for a particular child or young person.

3.48. The MPRG indicates that there does not appear to be a standardised approach to assessment. They are particularly concerned about safeguarding shortfalls within the assessment process. There is also limited evidence of systematic, formal mental health or neurodevelopmental assessments being routinely documented, or of a discipline of formal diagnostic formulation in relation to co-occurring mental health difficulties. This issue was also highlighted by the Care Quality Commission (CQC).[50]

3.49. Additionally, there is concern that communications to GPs and parents regarding prescribed treatment with puberty blockers sometimes come from non-medical staff.

[50] Care Quality Commission (2021). The Tavistock and Portman NHS Foundation Trust Gender Identity Service Inspection Report. London: CQC.

4. What the review has heard so far

Listening sessions

4.1. Since its establishment, the Review has met with an extensive range of stakeholders, including professionals, their respective governing organisations and those with lived experience, both directly and through support and advocacy groups, to understand the broad range of views and experiences surrounding the delivery of gender identity services.

What we have heard from service users, their families and support and advocacy groups

Issues for children and young people

4.2. What we understand most clearly from all we have heard is that at the centre of a difficult and complex debate are children, young people and families in great distress. We have heard concerns about children and young people facing the stress of being on a prolonged waiting list with limited support available from statutory services, lack of certainty about when and if they might reach the top of that list and subsequent impacts on mental health. Also, the particular issues that have followed the *Bell v Tavistock* litigation.

4.3. We have heard about the anxiety that birth-registered males face as they come closer to the point where they will grow facial hair and their voice drops, and the fear that it will make it harder for them to pass as a transgender woman in later life. We have also heard about the distress experienced by birth-registered females as they reach puberty, including the use of painful, and potentially harmful, binding processes to conceal their breasts.

4.4. When children and young people are able to access the service, there is often a sense of frustration with what several describe as the “gatekeeping” medical model and a “clinician lottery”. This can feel like a series of barriers and hurdles designed to add to, rather than alleviate, distress. Most children and young people seeking help do not see themselves as having a medical condition; yet to achieve their desired intervention they need to engage with clinical services and receive a medical diagnosis of gender dysphoria. By the time they are seen in the GIDS clinic, they may feel very certain of their gender identity and be anxious to start hormone treatment as quickly as possible. However, they can then face a period of what can seem like intrusive, repetitive and unnecessary questioning. Some feel that this undermines their autonomy and right to self-determination.

4.5. We have heard that some young people learn through peers and social media what they should and should not say to therapy staff in order to access hormone treatment; for example, that they are advised not to admit to previous abuse or trauma, or uncertainty about their sexual orientation. We have also heard that many of those seeking NHS support identify as non-binary, gender non-conforming, or gender fluid. We understand that some

young people who identify as non-binary feel their needs are not met by clinical services unless they give a binary narrative about their gender preferences.

Issues for parents

4.6. We have also heard about the distress parents may feel as they try to work out how best to support their children and how tensions and conflict may arise where parents and their children have different views. For example, some parents have highlighted the importance of ensuring that children and young people are able to keep their options fluid until such time as it becomes essential to commit to a hormonal course of action, whilst their children may want more rapid hormone intervention.

4.7. We have heard about families trying to balance the risks of obtaining unregulated and potentially dangerous hormone supplies over the internet or from private providers versus the ongoing trauma of prolonged waits for assessment.

4.8. Parents have also raised concerns about the vulnerability of neurodiverse children and young people and expressed that the communication needs of these children and young people are not adequately reflected during assessment processes or treatment planning.

4.9. GIDS has always required consent/ assent from both the child and parents/ carers and has sought ways to resolve family conflict, which in the worst-case scenario can lead to family breakdown. It has been highlighted to us that the future service model should provide more targeted support for parents and carers.

Service issues

4.10. Another significant issue raised with us is one of diagnostic overshadowing – many of the children and young people presenting have complex needs, but once they are identified as having gender-related distress, other important healthcare issues that would normally be managed by local services can sometimes be subsumed by the label of gender dysphoria. This issue is compounded by the waiting list, which means that there can be a significant period of time without appropriate assessment, treatment or care.

4.11. Stakeholders have spoken of the need for appropriate assessment when first accessing NHS services to aid both the exploration of the child or young person's wellbeing and gender distress and any other challenges they may be facing.

Information

4.12. We have also heard about the lack of access to accurate, balanced information upon which children, young people and their families/carers can inform their decisions.

4.13. We have heard that distress may be exacerbated by pressure to identify with societal stereotyping and concerns over the influence of social media, which can be seen to perpetuate unrealistic images of gender and set unhealthy expectations, especially given how long

children and young people are waiting to access services.

Other issues

4.14. Several issues that were raised with us are not explored further in this interim report, but we have taken note of them. These will be considered further during the lifetime of the Review and include:

- The important role of schools and the challenges they face in responding appropriately to gender-questioning children and young people.
- The complex interaction between sexuality and gender identity, and societal responses to both; for example, we have heard from young lesbians who felt pressured to identify as transgender male, and conversely transgender males who felt pressured to come out as lesbian rather than transgender. We have also heard from adults who identified as transgender through childhood, and then reverted to their birth-registered gender in teen years.
- The issues faced by detransitioners highlight the need for better services and pathways for this group, many of whom are living with irreversible effects of transition but for whom there is no clear access to services as they fall outside the responsibility of NHS gender identity services.
- The age at which adult gender identity clinics can receive referrals, with concerns about the inclusion of 17-year-olds. The service offer in adult services is perceived to be quite different from that of GIDS, and young people presenting later may therefore not be afforded the same level of therapeutic input under the adult service model. There is also concern about the impact on the young person of changing clinicians at a crucial point in their care. The movement of young people with special educational needs between children's and adult services raises particular concerns.

What we have heard from healthcare professionals

Lack of professional consensus

4.15. Clinicians and associated professionals we have spoken to have highlighted the lack of an agreed consensus on the different possible implications of gender-related distress – whether it may be an indication that the child or young person is likely to grow up to be a transgender adult and would benefit from physical intervention, or whether it may be a manifestation of other causes of distress. Following directly from this is a spectrum of opinion about the correct clinical approach, ranging broadly between those who take a more gender-affirmative approach to those who take a more cautious, developmentally-informed approach.

4.16. Speaking to current and ex-GIDS staff, we have heard about the pressure on GIDS clinicians, many of whom feel overwhelmed by the numbers of children and young people being referred and who are demoralised by the media coverage of their service. Although the clinical team attempt to manage risk on the waiting list by engaging with local services, there is limited capacity and/or capability to respond appropriately to the needs of this group in primary and secondary care. The Review has already referred to this issue as the most pressing priority in its letter to NHS England (**Appendix 2**), alongside potential risks relating to safeguarding and/or mental health issues, and diagnostic overshadowing.

4.17. With respect to GIDS, we have been told that although there are forums for staff to discuss difficult cases with senior colleagues, it is still difficult for staff to raise concerns about the clinical approach. Also that many individuals who are more cautious and advocate the need for an exploratory approach have left the service.

Consistency and standards

4.18. GIDS staff have confirmed that judgements are very individual, with some clinicians taking a more gender-affirmative approach and others emphasising the need for caution and for careful exploration of broader issues. The Review has been told that there is considerable variation in the approach taken between the London, Leeds and Bristol teams.

4.19. Speaking to professionals outside GIDS, we have heard widespread concern about the lack of guidance and evidence on how to manage this group of young people.

4.20. Some secondary care providers told us that their training and professional standards dictate that when working with a child or young person they should be taking a mental health approach to formulating a differential diagnosis of the child or young person's problems. However, they are afraid of the consequences of doing so in relation to gender distress because of the pressure to take a purely affirmative approach. Some clinicians feel that they are not supported by their professional body on this matter. Hence the practice of passing referrals straight through to GIDS is not just a reflection of local service capacity problems, but also of professionals' practical concerns about the appropriate clinical management of this group of children and young people.

4.21. GPs have expressed concern about being pressurised to prescribe puberty blockers or feminising/masculinising hormones after these have been initiated by private providers.

4.22. This also links to professional concerns about parents being anxious for hormone treatment to be initiated when the child or young person does not seem ready.

Other issues

4.23. We have also heard that parents and carers play a huge role and are instrumental in helping young people

to keep open their developmental opportunities. In discussion with social workers, we heard concerns about how looked after children are supported in getting the help and support they need.

4.24. Therapists who work with detransitioners and people with regret have highlighted a lack of services and pathways and a need for services to support this population. There is also the need for more research to understand what factors contribute to the decision to detransition.

4.25. The importance of broad holistic interventions to help reduce distress has been emphasised to the Review, with therapists and other clinicians advocating the importance of careful developmentally informed assessment and of showing children and young people a range of different narratives, experiences and outcomes.

4.26. Clinicians have raised concerns about children and young people's NHS numbers being changed inconsistently, as there is no specific guidance for GPs and others as to when this should be done for this population and under what consent. This has implications for safeguarding and clinical management of these children and young people and it also makes it difficult to do research exploring long-term outcomes.

4.27. As with the comments made by service users, their families and support and advocacy groups, we have heard similar views from professionals about the transition from children's to adult services, and the role of schools.

Structured engagement with primary, secondary and specialist clinicians

4.28. The Review's letter to NHS England (**Appendix 2**) set out some of the immediate issues with the current provision of gender identity services for children and young people and suggested how its work might help with the challenging problem of establishing an infrastructure outside GIDS. This included looking at the capacity, capability and confidence of the wider workforce and how this could be built and sustained, and the establishment of potential assessment frameworks for use in primary and/or secondary care.

Professional panel – primary and secondary care

4.29. In order to understand the challenges and establish a picture of current competency, capacity and confidence among the workforce outside the specialist gender development service, an online professional panel was established to explore issues around gender identity services for children and young people. The role of the panel was aimed at better comprehending how it looks and feels for clinicians and other professionals working with these young people, as well as any broader thoughts about the work, and to start exploring how the care of these

children and young people can be better managed in the future.

4.30. The project was designed to capture a broad mix of professional views and experiences, recruiting from the professional groups that are most likely to have a role in the care pathway – GPs, paediatricians, child psychiatrists, child psychologists and child psychotherapists, nurses and social workers.

4.31. A total of 102 clinicians and other professionals were involved in the panel. The panel represented a balanced professional mix, and participant ages and gender were broadly representative of the overall sector workforce. Participants were self-selecting and were recruited via healthcare professional networks and Royal Colleges.

4.32. Each week the panel was set an independent activity comprised of two or more tasks. Additionally, a sub-set of the panel was invited to participate in focus groups at the midway and endpoint of the project. Activities were designed to capture an understanding of:

- experiences of working with gender-questioning children and young people and panel members' confidence and competence to manage their care;
- changes they may have experienced in the presentation of children and young people with gender-related distress;
- areas where professionals feel they require more information in order to support gender-questioning children and young people;
- where professionals currently go to find that information;
- the role of different professions in the care pathway;
- the role of professionals in the assessment framework; and
- what participants felt should be included in an assessment framework across the whole service pathway.

Gender specialist questionnaire

4.33. Having concluded the professional panel exercise, we wanted to triangulate what we had heard with the thoughts and views of professionals working predominantly or exclusively with gender-questioning children and young people.

4.34. To do this in a systematic way, we conducted an online survey which contained some service-specific questions, but also reflected and sought to test some of what we had heard from primary and secondary care professionals.

Findings

4.35. This structured engagement has yielded valuable insights from clinicians and professionals with experience working with gender-questioning children and young people both within and outside the specialist gender service. It has contributed to the thinking of the Review and informed some of the interim advice set out in this report.

4.36. There are a number of consistent messages arising from these activities:

- The current long waiting lists that gender-questioning children and young people and their families/carers face are unacceptable for all parties involved, including professionals.
- Many professionals in our sample said that not only are gender-questioning children and young people having to wait a long time before receiving treatment, but they also do not receive appropriate support during this waiting period.
- Another impact of the long wait that clinicians reported is that when a child or young person is seen at GIDS, they may have a more fixed view of what they need and are looking for action to be taken quickly. This reportedly can lead to frustration with the assessment process.
- When considering the more holistic support that children and young people may need, gender specialists further highlighted the difficulties that children and young people face accessing local support, for example, from CAMHS, whilst being seen at GIDS.
- It is clear from the professionals who took part in these activities that there is a strong professional commitment to provide quality care to gender-questioning children and young people and their families/carers. However, this research indicates that levels of confidence and competence do vary among primary and secondary care professionals in our sample.
- Concerns were expressed by professionals who took part in this research about the lack of consensus among the clinical community on the right clinical approach to take when working with a gender-questioning child or young person and their families/carers.
- In order to support clinicians and professionals more widely, participants felt there is a need for a robust evidence base, consistent legal framework and clinical guidelines, a stronger assessment process and different pathway options that holistically meet the needs of each gender-questioning child or young person and their families/carers.

4.37. There are also several areas where further discussion and consensus is needed:

- There is not a consistent view among the professionals participating in the panel and questionnaire about the nature of gender dysphoria and therefore the role of assessment for children and young people experiencing gender dysphoria.

- Some clinicians felt that assessment should be focused on whether medical interventions are an appropriate course of action for the individual. Other clinicians believe that assessment should seek to make a differential diagnosis, ruling out other potential causes of the child or young person's distress.
- There are different perspectives on the roles of primary, secondary and specialist services in the care pathway(s) and what support or action might best be provided at different levels.
- While there was general consensus that diagnostic or psychological formulation needs to form part of the assessment process, there were differing views as to whether a mental state assessment is needed, and should it be, where in the pathway and by whom this should be done.

4.38. It is important to note that the information gathered represents the views and insights of the panel participants and survey respondents at a moment in time and findings should be read in the context of a developing narrative on the subject, where perspectives may evolve. This relates to both the experiences of professionals, but also the extent to which this subject matter is discussed in the public sphere.

4.39. The Review is grateful to all the participants for their time and high level of engagement. The Review will build on the work we have undertaken and, alongside our academic research, will continue with a programme of engagement with professionals, service users and their families, which will help to further develop the evidence base.

The full reports from the professional panel and gender specialist questionnaire are on the Review's website (https://cass.independent-review.uk/).

5. Principles of evidence based service development

Evidence based service development

5.1. This chapter integrates the information regarding the development of the current service (see Chapter 3) with the views we have heard to date (see Chapter 4) and sets this in the context of how evidence is routinely used to develop and improve services in the NHS.

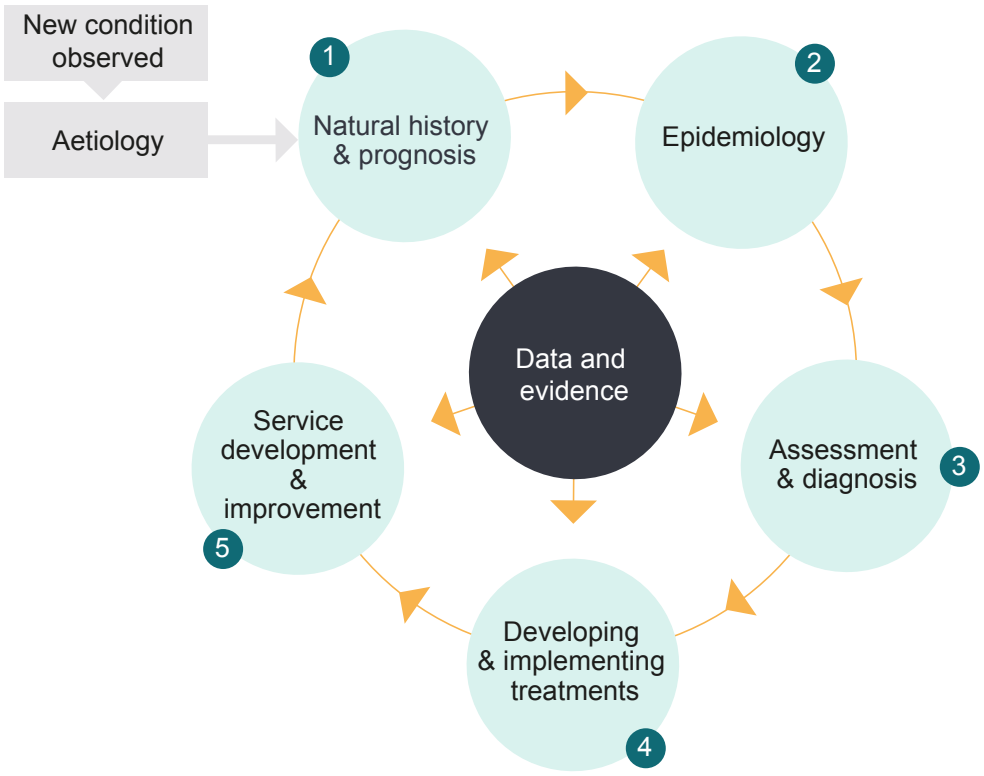
5.2. Some earlier information is necessarily repeated here, but this is with the intention of providing a more accessible explanation of the standards and processes which govern clinical service development. This is essential to an understanding of the rationale for the Review's recommendations.

5.3. Because the specialist service has evolved rapidly and organically in response to demand, the clinical approach and overall service design has not been subjected to some of the normal quality controls that are typically applied when new or innovative treatments are introduced. This Review now affords everyone concerned the opportunity to step back and consider from first principles what this cohort of children and young people now need from NHS services, based on the evidence that exists, or additional evidence that the Review hopes to collect.

5.4. In **Appendix 4** we have described the service development process for three different conditions which may help to illustrate what would be expected to happen at each different stage of developing a clinical service. The steps may proceed in a different sequence for different conditions, but each step is important in the development of evidence based care.

5.5. We recognise that for some of those reading this report it may feel wrong to compare gender incongruence or dysphoria to clinical conditions, and indeed this approach would not be justified if individuals presenting with these conditions did not require clinician intervention. However, where a clinical intervention is given, the same ethical, professional and scientific standards have to be applied as to any other clinical condition.

Key stages of service development



New condition observed: This often begins with a few case reports and then clinicians begin to recognise a recurring pattern and key clinical features, and to develop fuller descriptions of the condition.

Aetiology: Clinicians and scientists try to work out the cause of the condition or the underlying physical or biological basis. Sometimes the answers to this are never found.

Natural history and prognosis: It is important to understand how a condition usually evolves over time, with or without treatment. The latter is important if treatment has limited efficacy and the condition is 'self-limiting' (that is, it resolves without treatment), because otherwise there is a risk that treatments create more difficulties than the condition itself.

5.6. The first UK service for gender-questioning children and young people was established in 1989. At that time there were very few children and young people being seen by medical services internationally. The most common presentation in the early years of the service was of birth-registered

boys who had demonstrated gender incongruence from an early age.[51,52,53]

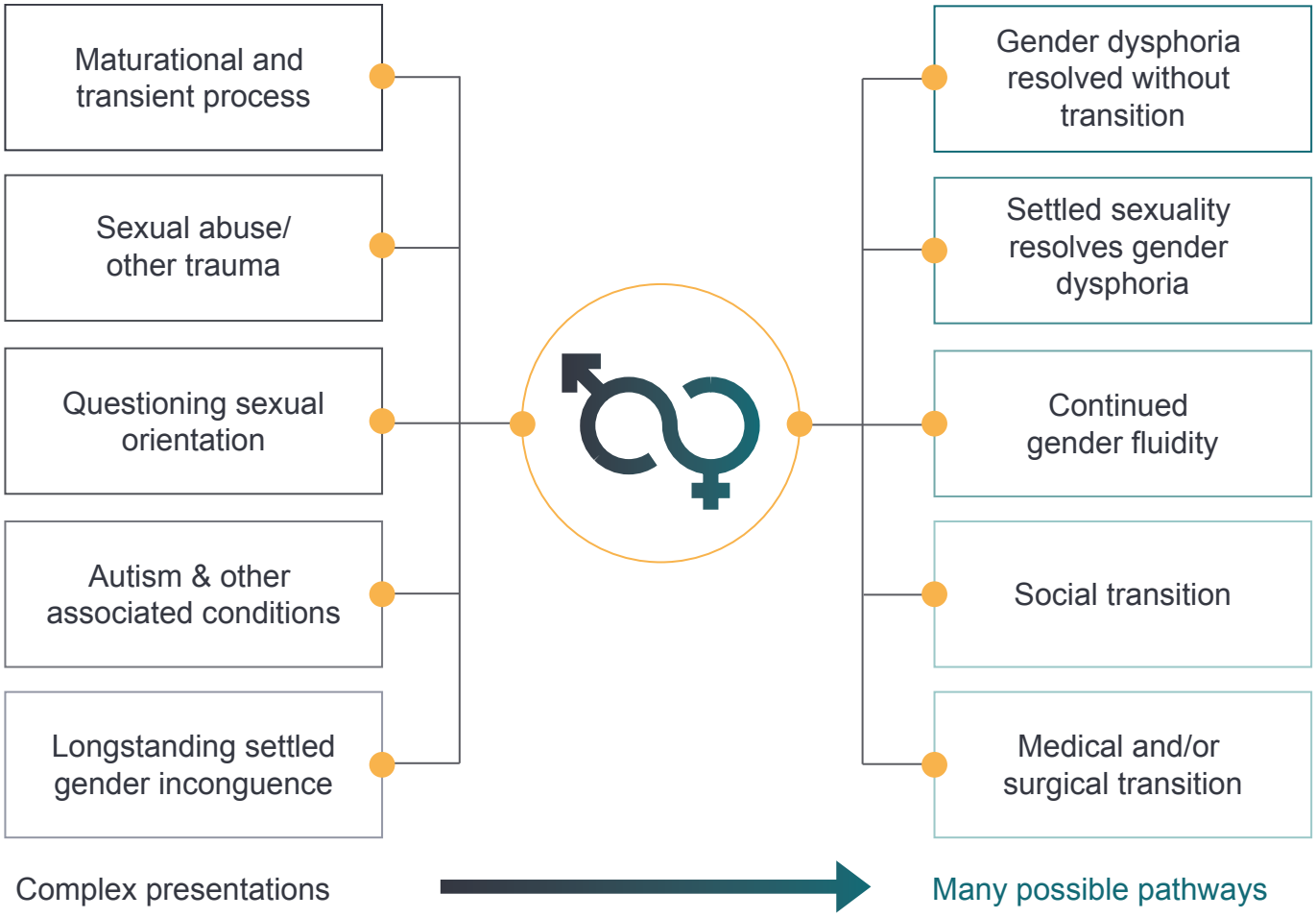
5.7. There is extensive literature discussing the possible aetiology of gender incongruence. Based on the available evidence, many authors would suggest that it is likely that biological, cultural, social and psychological factors all contribute. The examples in **Appendix 4** show that this is not an uncommon situation; many conditions do not have a single clear causation – they are in other words 'multifactorial'.

5.8. Regardless of aetiology, the more contentious and important question is how fixed or fluid gender incongruence is at different ages and stages of development, and whether, regardless of aetiology, can be an inherent characteristic of the individual concerned. There is a spectrum of academic, clinical and societal opinion on this. At one end are those who believe that gender identity can fluctuate over time and be highly mutable and that, because gender incongruence or gender-related distress may be a response to many psychosocial factors, identity may sometimes change or the distress may resolve in later adolescence or early adulthood, even in those whose early incongruence or distress was quite marked. At the other end are those who believe that gender incongruence or dysphoria in childhood or adolescence is generally a clear indicator of that child or young person being transgender and question the methodology of some of the desistance studies. Previous literature has indicated that if gender incongruence continues into puberty, desistance is unlikely.[54,55] However, it should be noted that these older studies were not based on the current changed case-mix or the different socio-cultural climate of recent years, which may have led to different outcomes. Having an open discussion about these questions is essential if a shared understanding of how to provide appropriate assessment and treatment is to be reached.

---

[51] Zucker KJ (2017). Epidemiology of gender dysphoria and transgender identity. Sex Health 14(5): 404–11. DOI:10.1071/SH1.
[52] Zucker KJ, Lawrence AA (2009). Epidemiology of gender identity disorder: recommendations for the Standards of Care of the World Professional Association for Transgender Health. Int J Transgend 11(1): 8-18. DOI: 10.1080/15532730902799946.
[53] de Graaf NM, Giovanardi G, Zitz C, Carmichael P (2018). Sex ratio in children and adolescents referred to the gender identity development service in the UK (2009-2016). Arch Sex Behav 47(5): 1301–4.
[54] Steensma TD, Biemond R, de Boer F, Cohen-Kettenis PT (2011). Desisting and persisting gender dysphoria after childhood: a qualitative follow-up study. Clin Child Psychol Psychiatry 16(4): 485-97. DOI: 10.1177/135910451037803.
[55] Steensma TD, McGuire JK, Kreukels BPC, Beekman AJ, Cohen-Kettenis PT (2013). Factors associated with desistence and persistence of childhood gender dysphoria: a quantitative follow-up study. J Am Acad Child Adolesc Psychiatry 52: 582-590. DOI: 10.1016/j.jaac.2013.03.016.

Complex presentations and complex pathways – exemplars, not comprehensive lists



Epidemiology: Epidemiologists collect data to find out how common a condition is, who is most likely to be affected, what the age distribution is and so on. This allows health service planners to work out how many services are needed, where they should be established, and what staff are needed.

They also report on changes in who is most affected, which may mean that either the disease is changing, or the susceptibility of the population is changing.

5.9. As previously indicated, the epidemiology of gender dysphoria is changing, with an increase in the numbers of birth-registered females presenting in early teens.[56,57] In addition, the majority of children and young people presenting to GIDS have other complex mental health issues and/or neurodiversity.[58] There is also an over-representation of looked after children.[59]

5.10. There are several implications arising from the change in epidemiology:

- Firstly, the speed of change in the numbers presenting means that services have not kept pace with demand.
- Secondly, the cohort that the original Dutch Approach was based on is different from the current more complex NHS cohort, and also from the current case-mix internationally, and therefore it is difficult to extrapolate from older literature to this current group.
- Thirdly, different subgroups may have quite different needs and outcomes, and these must be built into any service design, so that it works for all children and young people.

5.11. At present we have the least information for the largest group of patients – birth-registered females first presenting in early teen years. Since the rapid increase in this group began around 2015, they will not reach late 20s for another 5+ years, which would be the best time to assess longer-term wellbeing.

---

[56] Steensma TD, Cohen-Kettenis PT, Zucker KJ (2018). Evidence for a change in the sex ratio of children referred for gender dysphoria: Data from the Center of Expertise on Gender Dysphoria in Amsterdam (1988-2016). Journal of Sex & Marital Therapy 44(7): 713–5. DOI: 10.1080/0092623X.2018.1437580.
[57] de Graaf NM, Carmichael P, Steensma TD, Zucker KJ (2018). Evidence for a change in the sex ratio of children referred for Gender Dysphoria: Data from the Gender Identity Development Service in London (2000–2017). J Sex Med 15(10): 1381–3. DOI: 10.1016/j.jsxm.2018.08.002.
[58] Van Der Miesen AIR, Hurley H, De Vries ALC (2016). Gender dysphoria and autism spectrum disorder: A narrative review. Int Rev Psychiatry 28: 70-80. DOI: 10.3109/09540261.2015.1111199.
[59] Matthews T, Holt V, Sahin S, Taylor A, Griksaitis (2019). Gender Dysphoria in looked-after and adopted young people in a gender identity development service. Clinical Child Psychol Psychiatry 24: 112-128. DOI: 10.1177/1359104518791657.

Assessment and diagnosis: Clinicians will usually take a history from (that is, of their symptoms) and examine the patient (that is, for signs and symptoms), and where appropriate undertake a series of investigations or tests, to help them reach an accurate diagnosis.

Sometimes the whole process of making a diagnosis through talking to the patient and asking them to complete formal questionnaires, examining them and/or undertaking investigations is called 'clinical assessment'.

As well as diagnosing and ruling out a particular condition, clinicians often need to consider and exclude other, sometimes more serious, conditions that present in a similar way but may need quite different treatment – this process is called 'differential diagnosis'.

5.12. For children and young people with gender-related distress, many people would dispute the notion that 'making a diagnosis' is a meaningful concept, arguing that gender identity is a personal, internal perception of oneself. However, there are several reasons to why a diagnostic framework is used:

- Firstly, the clinician will seek to determine whether the child or young person has a stable transgender identity, or whether there might be other causes for the gender-related distress.
- Secondly, the clinician will determine whether there are other issues or diagnoses that might be having an impact on the young person's mental health. The Dutch Approach suggesting that these should be addressed prior to or alongside initiation of any medical treatments.
- Thirdly, in any situation where life-altering treatments are being administered, the clinician holds the responsibility for ensuring that they are being administered based on an appropriate decision making process. Therefore, it is usual practice for a diagnosis of gender dysphoria to be made prior to referring for any physical treatments.

5.13. When the word 'diagnosis' is used, people often associate this with the use of blood tests, X-rays, or other laboratory tests. As set out in the **Appendix 4**, the public is very familiar with diagnosis of Covid-19 and understands that there need to be tests that give a high degree of certainty about whether an individual is Covid-19 positive or not. False positive lateral flow tests are rare, but caused problems for schools, while PCR has been treated as the 'gold standard' test for accuracy.

5.14. When it comes to gender dysphoria, there are no blood tests or other laboratory tests, so assessment and diagnosis in children and young people with gender-related distress is reliant on the judgements of experienced clinicians. Because medical, and subsequently possibly surgical treatments will follow, it may be argued that a highly sensitive and specific assessment process is required. The assessment should be able to accurately identify those children or young people for whom physical intervention is going to be the best course of action, but it is equally important that it identifies those who need an alternative pathway or treatment.

5.15. The formal criteria for diagnosing gender dysphoria (DSM-5) are listed in **Appendix 3**. However, there are two problems associated with the use of these criteria:

- Firstly, several of the criteria are based on gender stereotyping which may not be deemed relevant in current society, although the core criteria remain valid.
- Secondly, and more importantly, these criteria give a basis on which to make a diagnosis that a young person is clinically distressed by the incongruence between their birth-registered and their experienced gender, but they do not help in determining which factors may have led to this distress and how they might best be resolved.

5.16. At present, the assessment process varies considerably, dependent on the perceptions, experience and beliefs of different clinicians. There are some existing measurement tools, but it is suggested that these have substantial limitations.[60]

5.17. The challenges are similar to the early difficulties in diagnosing autism, as set out in **Appendix 4**. As with autism, the framework for assessment needs to become formalised so there are clearer criteria for diagnosis and treatment pathways which are shared more widely. These should incorporate not just whether the child or young person meets DSM-5 criteria for gender dysphoria, but how a broader psychosocial assessment should be conducted and evaluated, and what other factors need to be considered to gain a holistic understanding of the child or young person's experience. Professional judgement and experience will still be important, but if the frameworks and criteria for assessment and diagnosis were more consistent and reproducible, there would be a greater likelihood that two different people seeing the same child or young person would come to the same conclusion. This would also mean that any research on interventions or long-term outcomes would be more reliable because the criteria on which a diagnosis was made, and hence the patients within the sample, would have the same characteristics.

[60] Bloom TM, Nguyen TP, Lami F, Pace CC, Poulakis Z, Telfer N (2021). Measurement tools for gender identity, gender expression, and gender dysphoria in transgender and gender-diverse children and adolescents: a systematic review. Lancet Child Adolescent Health. 5: 582-588. DOI: 10.1016/s2352-4642(21)00098-5.

5.18. As outlined above, it is standard clinical practice to undertake a process called differential diagnosis. This involves summarising the main points of the clinical assessment, the most likely diagnosis, other possible diagnoses and the reasons for including or excluding them, as well as any further assessments that may be required to clarify the diagnosis and the treatment options and plan. This is important when a medical intervention is being provided on the basis of the assessment, so the process is robust, explicit and reproducible. These considerations need to be applied to the assessment of children and young people presenting with gender-related distress. In mental health services, practitioners may also undertake a diagnostic or psychological formulation, which is a holistic summary of how the patient is feeling and why, and how to make sense of it, and a plan for moving forward with management or treatment.

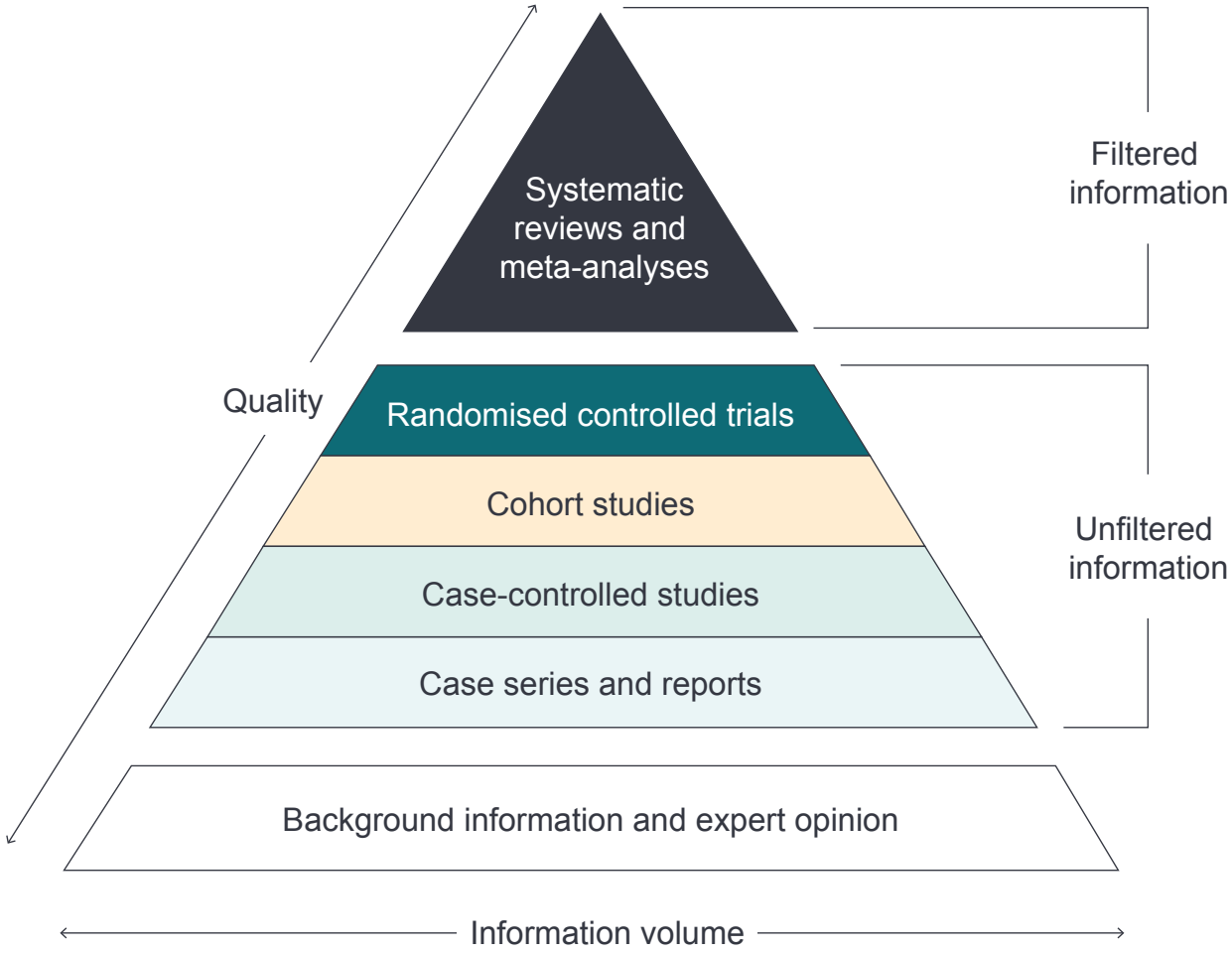
Developing and implementing new treatments: Clinicians and scientists work on developing treatments. This involves clinical trials and, where there are new treatments, comparing them to any existing treatments. Questions include: What are the intended outcomes or benefits of treatment? What are the complications or side effects? What are the costs? To initiate a new treatment, it must be both safe and effective. Questions of affordability can sometimes become controversial.

The best type of single study is considered to be the randomised controlled trial (RCT), but sometimes this is not feasible. Even where RCTs are not available, it is usual to at least have data on the outcomes of sufficient cases or cohorts to understand the risk/benefit of the treatment under consideration. As demonstrated in Fig. 4, the highest level of evidence is when the results of several different studies are pooled, but this is only useful if the individual studies themselves are of high quality.

In many instances, evidence is not perfect and difficult decisions have to be made. Where treatments are innovative or life-changing, the whole multi-disciplinary team will usually meet to consider the available options, and how to advise the child or young person and family so that a shared decision can be made. Sometimes an ethics committee is involved. This is one of the most challenging areas of medicine and is underpinned by GMC guidance.[61,62]

[61] General Medical Council (2020). Decision making and consent.
[62] National Institute for Health and Care Excellence (2021). Shared decision making.

Figure 3: Pyramid of standards of evidence



Source: Levels of evidence pyramid, OpenMD. Reproduced with permission[63]

5.19. There are three types of intervention or treatment for children and young people with gender-related distress, which may be introduced individually or in combination with one another:

- **Social transition** – this may not be thought of as an intervention or treatment, because it is not something that happens within health services. However, it is important to view it as an active intervention because it may have significant effects on the child or young person in terms of their psychological functioning.[64,65] There are different views on the benefits versus the harms of early social transition. Whatever position one

[63] OpenMD (2021). New Evidence in Medical Research.
[64] Sievert EDC, Schweizer K, Barkmann C, Fahrenkrug S, Becker-Hebly I (2020). Not social transition status, but peer relations and family functioning predict psychological functioning in a German clinical sample of children with Gender Dysphoria.Clin Child Psychol Psychiatry 26(1): 79–95. DOI: 10.1177/1359104520964530
[65] Ehrensaft D, Giammattei SV, Storck K, Tishelman AC, Colton K-M (2018). Prepubertal social gender transitions: What we know; what we can learn—A view from a gender affirmative lens. Int J Transgend 19(2): 251–68. DOI: 10.1080/15532739.2017.1414649.

takes, it is important to acknowledge that it is not a neutral act, and better information is needed about outcomes.

- **Counselling, social or psychological interventions** – these may be offered before, instead of, or alongside physical interventions. Again, they should be viewed as active interventions which require robust evaluation in their own right.
- **Physical treatments** – these comprise puberty blockers and feminising/masculinising hormones (administered by endocrinologists) and surgery. The latter is not considered as part of this Review since it is not available to those under age 18.

5.20. It should also be recognised that 'doing nothing' cannot be considered a neutral act.

5.21. The lack of available high-level evidence was reflected in the recent NICE review into the use of puberty blockers and feminising/masculinising hormones commissioned by NHS England, with the evidence being too inconclusive to form the basis of a policy position.[66,67] Assessing treatments for gender dysphoria has many of the same problems as assessing treatment for children with autism – it can take many years to get a full appreciation of outcomes and there may be other complicating factors in the child or young person's life during this period. However, this of itself is not an adequate reason for the major gaps in the international literature.

5.22. It is still common that drugs are not specifically licensed for children because the trials have only taken place on adults. This does not preclude their use or make their use inherently unsafe, particularly if they are used very commonly in children. However, where their use is innovative, patients receiving the drug should ideally do so under trial conditions.

5.23. The same considerations apply to 'off-label' drugs, where the drug is used for a condition different to the one for which it was licensed. This is the case for puberty blockers, which are licensed for use in precocious puberty, but not for puberty suppression in gender dysphoria. Again, it is important that it is not assumed that outcomes for, and side effects in, children treated for precocious puberty will necessarily be the same in children or young people with gender dysphoria.

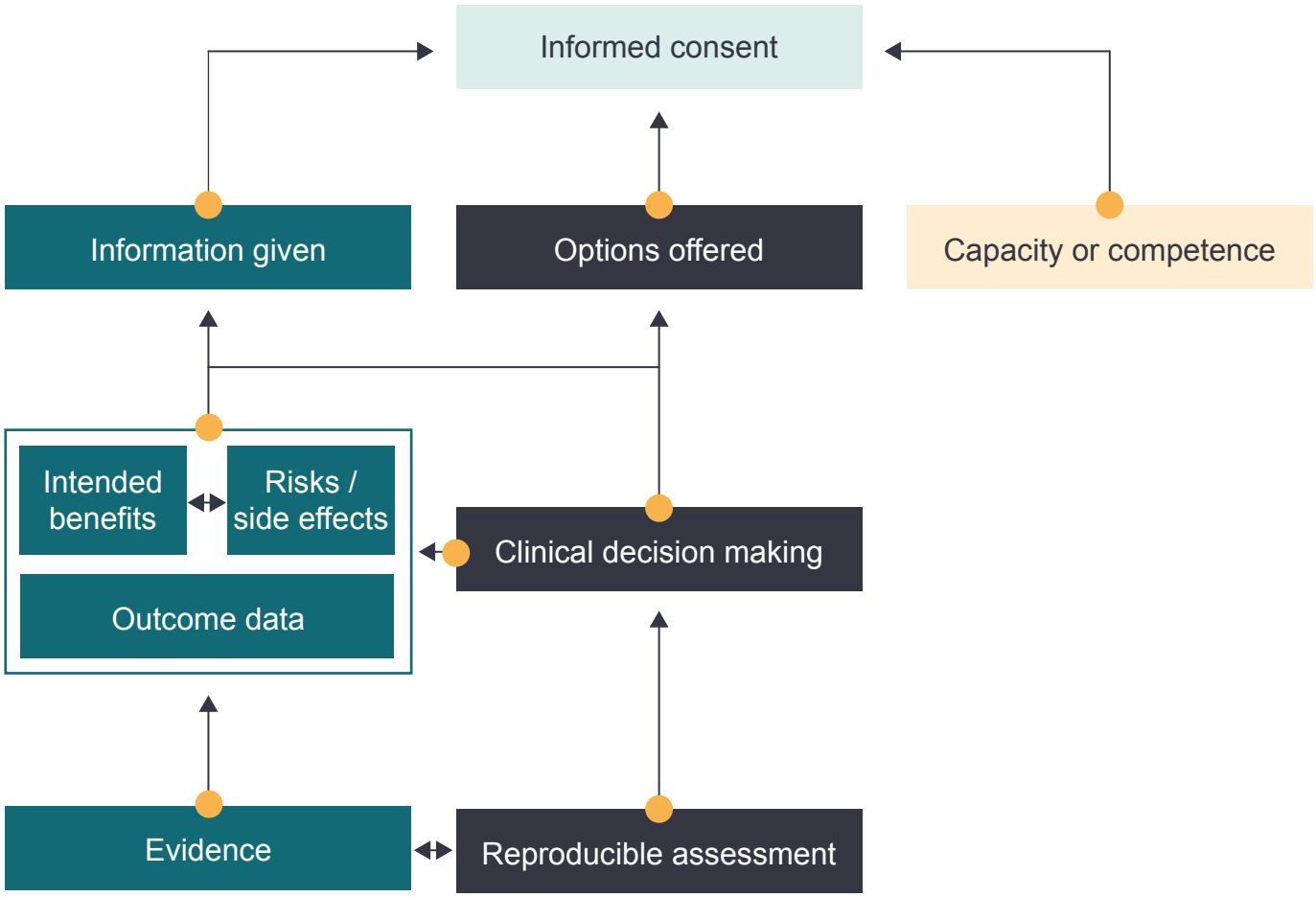
5.24. As outlined above, in other areas of practice where complex or potentially life-altering treatment is being considered for a child or young person, it is usual for the case to be discussed by an MDT including all professionals involved in their care. In gender services for children and young people in the Netherlands, as well as a number of other countries, there are full

[66] National Institute for Health and Care Excellence (2020). Evidence Review: Gonadotrophin Releasing Hormone Analogues for Children and Adolescents with Gender Dysphoria
[67] National Institute for Health and Care Excellence (2020). Evidence review: gender-affirming hormones for children and adolescents with gender dysphoria.

MDT meetings, including psychiatrists and endocrinologists, to make decisions about suitability for hormone intervention and to review progress.[68,69]

5.25. Recent legal proceedings have examined the question of the competence and capacity of children and young people to consent to hormone treatment. However, there are some essential components that underpin informed consent; the robustness of the options offered to the patient, the information provided to them about those options, and their competence and capacity to consider them. The courts have given consideration to competence and capacity, and it is incumbent on this Review to consider the soundness of the decision making which underpins the options offered, and the quality and accuracy of the information provided about those options.

Elements of informed consent



[68] Kyriakou A, Nicolaides NC, Skordis N (2020). Current approach to the clinical care of adolescents with gender dysphoria. Acta Biomed 91(1): 165–75. DOI: 10.23750/abm.v91i1.9244.
[69] Cohen-Kettenis PT, Steensma TD, de Vries ALC. Treatment of adolescents with gender dysphoria in the Netherlands. Child Adolesc Psychiatr Clin N Am 20. 689–700. 2001. DOI: 10.1016/j.chc.2011.08.001.

Service development and service improvement: Central to any service improvement is the systematic and consistent collection of data on outcomes of treatment. There is a process of continuous service improvement as new presentations or variations on the original condition are recognised, diagnosis or screening improves and/or trials on new treatments or variations on existing treatments are ongoing.

There should be consistent treatment protocols or guidelines in place, in order to make sense of variations in outcomes. Where possible, these should be compared between and across multiple different centres.

As time passes, services need to be changed or extended based on patient need, and on what resources are needed to deliver the available treatments. They need to be accessible where the prevalence of the condition is highest. The relevant workforce to deliver the service needs to be recruited and trained, contingent on the type of treatments or therapy that is required.

5.26. When a pioneering treatment or specialist service starts, it is often delivered in a single centre. Thereafter, additional centres take on the work as increasing numbers of patients need to access the treatment. Current provision of NHS specialist gender identity services for children and young people has remained concentrated within a single organisation, but demand has grown dramatically.

5.27. The situation has been exacerbated because there are not many local services seeing gender-questioning children at an earlier stage in their journey, which means that GIDS is carrying an unsustainable workload of increasingly complex young people.

5.28. As a condition evolves, rigorous data collection and quantitative research is an essential prerequisite to refining understanding and treatment. Historically, The Tavistock and Portman NHS Foundation Trust built its international reputation as the home of psychoanalysis, psychotherapy and family therapy, with a strong track record of publishing qualitative rather than quantitative research; consequently its approach to quantitative data collection about this important group of children and young people has been weak.

5.29. A further anomaly is a public perception that The Tavistock and Portman NHS Foundation Trust is the responsible organisation for leading the management of children receiving hormone treatment for their gender dysphoria. In reality, the hormone treatment is delivered by paediatric services in University College London Hospitals NHS Foundation Trust and The Leeds Teaching Hospitals NHS Trust.

5.30. In practice, it is important that for children and young people who need physical intervention, paediatric and mental health services are seen as equal partners, with seamless joint working and shared responsibility. When there were very small numbers of patients, it was easier for this to be achieved, but cross-site working with a very large caseload has made this more difficult to achieve, despite the best intentions of the staff.

5.31. Over the last two years there have been strong efforts on the part of The Tavistock and Portman NHS Foundation Trust to make practice within GIDS more consistent, with tighter procedures for case management, consent, and safeguarding. However, although this has resulted in better documentation, variations and inconsistencies in clinical decision making remain. In responding to a changing legal framework, some processes have become more cumbersome and complex, and the team are working hard to streamline the process.

5.32. Overall, GIDS faces a daunting task as a single provider in managing risk on the waiting list, seeing new referrals, reviewing and supporting those on hormone treatment, undertaking an ongoing transformation programme, recruiting and training new staff and trying to retain existing staff. This suggests that the current model is not sustainable and that another model is needed.

6. Interim advice, research programme and next steps

Dealing with uncertainty

6.1. As outlined throughout this report, there are major gaps in the research base underpinning the clinical management of children and young people with gender incongruence and gender dysphoria, including the appropriate approaches to assessment and treatment.

6.2. As with any other area of medicine, where there are gaps in the evidence base and uncertainties about the correct clinical approach, three tasks must be undertaken:

- Clinical services must be run as safely and effectively as possible, within the constraints of current knowledge; treatment options must be weighed carefully; and treatment decisions must be made in partnership between the clinicians and the children, young people and their families and carers, based on our current understanding about outcomes.
- Consistent data must be collected by clinical services, for both audit and research purposes so that knowledge gaps can be filled, alongside an active research programme.
- Where there is not an immediate prospect of filling research gaps, professional consensus should be developed on the correct way to proceed pending clearer research evidence, supported by input from service users.

6.3. The additional problem with the current service model is that safety and access are further compromised by the pace at which referrals have grown and outstripped capacity at tertiary level, and the lack of service availability at local level.

6.4. The Review's approach to these tasks is as follows:

- Our **interim advice** focuses on the issues of capacity, safety, and standards around treatment decisions, as well as data and audit.
- Our **research streams** will provide the Review with an independent collation of published evidence relevant to epidemiology, clinical management, models of care, and outcomes, as well as delivering qualitative and quantitative research relevant to the Terms of Reference of the Review. This offers a real opportunity to contribute to the international evidence base for this service area.
- There will be an ongoing and wide-ranging **programme of engagement** to address areas on which we will not be able to obtain definitive evidence during the lifetime of the Review.

Interim advice

6.5. The Review considers that there are some areas where there is sufficient clarity about the way forward and we are therefore offering some specific observations and interim advice. The Review will work with NHS England, providers and the broader stakeholder community to progress action in these areas.

Service model

6.6. It has become increasingly clear that a single specialist provider model is not a safe or viable long-term option in view of concerns about lack of peer review and the ability to respond to the increasing demand.

6.7. Additionally, children and young people with gender-related distress have been inadvertently disadvantaged because local services have not felt adequately equipped to see them. It is essential that they can access the same level of psychological and social support as any other child or young person in distress, from their first encounter with the NHS and at every level within the service.

6.8. A fundamentally different service model is needed which is more in line with other paediatric provision, to provide timely and appropriate care for children and young people needing support around their gender identity. This must include support for any other clinical presentations that they may have.

6.9. The Review supports NHS England's plan to establish regional services, and welcomes the move from a single highly specialist service to regional hubs.

6.10. Expanding the number of providers will have the advantages of:

- creating networks within each area to improve early access and support;
- reducing waiting times for specialist care;
- building capacity and training opportunities within the workforce;
- developing a specialist network to ensure peer review and shared standards of care; and
- providing opportunities to establish a more formalised service improvement strategy.

Service provision

6.11. The primary remit of NHS England's proposed model is for the regional hubs to provide support and advice to referrers and professionals. However, it includes limited provision for direct contact with children and young people and their families.

1: The Review advises that the regional centres should be developed, as soon as feasibly possible, to become direct service providers, assessing and treating children and young people who may need specialist care, as part of a wider pathway. The Review team will work with NHS England and stakeholders to further define the proposed model and workforce implications.

2: Each regional centre will need to develop links and work collaboratively with a range of local services within their geography to ensure that appropriate clinical, psychological and social support is made available to children and young people who are in early stages of experiencing gender distress.

3: Clear criteria will be needed for referral to services along the pathway from primary to tertiary care so that gender-questioning children and young people who seek help from the NHS have equitable access to services.

4: Regional training programmes should be run for clinical practitioners at all levels, alongside the online training modules developed by Health Education England (HEE). In the longer-term, clearer mapping of the required workforce, and a series of competency frameworks will need to be developed in collaboration with relevant professional organisations.

Data, audit and research

6.12. A lack of routine and consistent data collection means that it is not possible to accurately track the outcomes and pathways children and young people take through the service. Standardised data collection is required in order to audit service standards and inform understanding of the epidemiology, assessment and treatment of this group. This, alongside a national network which brings providers together, will help build knowledge and improve outcomes through shared clinical standards and systematic data collection. In the longer-term, formalisation of such a network into a learning health system[70] with an academic host would mean that there was systematised use of data to produce a continuing research programme with rapid translation into clinical practice and a focus on training.

5: The regional services should have regular co-ordinated national provider meetings and operate to shared standards and operating procedures with a view to establishing a formal learning health system.

6: Existing and future services should have standardised data collection in order to audit standards and inform understanding of the epidemiology, assessment and treatment of this group of children and young people.

[70] Scobie S, Castle-Clarke S (2019). Implementing learning health systems in the UK NHS: Policy actions to improve collaboration and transparency and support innovation and better use of analytics. Learning Health Systems 4(1): e10209. DOI:10.1002/lrh2.10209.

7: Prospective consent of children and young people should be sought for their data to be used for continuous service development, to track outcomes, and for research purposes. Within this model, children and young people put on hormone treatment should be formally followed up into adult services, ideally as part of an agreed research protocol, to improve outcome data.

Clinical approach

Assessment processes

6.13. We have heard that there are inconsistencies and gaps in the assessment process. Our work to date has also demonstrated that clinical staff have different views about the purpose of assessment and where responsibility lies for different components of the process within the pathway of care. The Review team has commenced discussions with clinical staff across primary, secondary and tertiary care to develop a framework for these processes.

8: There needs to be agreement and guidance about the appropriate clinical assessment processes that should take place at primary, secondary and tertiary level.

9: Assessments should be respectful of the experience of the child or young person and be developmentally informed. Clinicians should remain open and explore the patient's experience and the range of support and treatment options that may best address their needs, including any specific needs of neurodiverse children and young people.

Hormone treatment

6.14. The issues raised by the Multi-Professional Review Group echo several of the problems highlighted by the CQC. It is essential that principles of the General Medical Council's Good Practice in Prescribing and Managing Medicine's and Devices[71] are closely followed, particularly given the gaps in the evidence base regarding hormone treatment. Standards for decision making regarding endocrine treatment should also be consistent with international best practice.[72,73,74]

---

[71] General Medical Council (2021). Good practice in prescribing and managing medicines and devices (76-78).
[72] Hembree WC, Cohen-Kettenis PT, Gooren L, Hannema SE, Meyer WJ, Murad MH, et al (2017). Endocrine treatment of gender-dysphoric/gender-incongruent persons: an Endocrine Society clinical practice guideline. J Clin Endocrinol Metab 102(11): 3869–903. DOI: 10.1210/jc.2017-01658.
[73] Cohen-Kettenis PT, Steensma TD, de Vries ALC (2001). Treatment of adolescents with gender dysphoria in the Netherlands. Child Adolesc Psychiatr Clin N Am 20: 689–700. DOI: 10.1016/j.chc.2011.08.001.
[74] Kyriakou A, Nicolaides NC, Skordis N (2020). Current approach to the clinical care of adolescents with gender dysphoria. Acta Biomed 91(1): 165–75. DOI: 10.23750/abm.v91i1.9244.

10: Any child or young person being considered for hormone treatment should have a formal diagnosis and formulation, which addresses the full range of factors affecting their physical, mental, developmental and psychosocial wellbeing. This formulation should then inform what options for support and intervention might be helpful for that child or young person.

11: Currently paediatric endocrinologists have sole responsibility for treatment, but where a life-changing intervention is given there should also be additional medical responsibility for the differential diagnosis leading up to the treatment decision.

6.15. Paediatric endocrinologists develop a wide range of knowledge within their paediatric training, including safeguarding, child mental health, and adolescent development. Being party to the discussions and deliberations that have led up to the decision for medical intervention supports them in carrying out their legal responsibility for consent to treatment and the prescription of hormones.

12: Paediatric endocrinologists should become active partners in the decision making process leading up to referral for hormone treatment by participating in the multidisciplinary team meeting where children being considered for hormone treatment are discussed.

6.16. Given the uncertainties regarding puberty blockers, it is particularly important to demonstrate that consent under this circumstance has been fully informed and to follow GMC guidance[75] by keeping an accurate record of the exchange of information leading to a decision in order to inform their future care and to help explain and justify the clinician's decisions and actions.

13: Within clinical notes, the stated purpose of puberty blockers as explained to the child or young person and parent should be made clear. There should be clear documentation of what information has been provided to each child or young person on likely outcomes and side effects of all hormone treatment, as well as uncertainties about longer-term outcomes.

[75] General Medical Council (2020). Decision making and consent.

14: In the immediate term the Multi-Professional Review Group (MPRG) established by NHS England should continue to review cases being referred by GIDS to endocrine services.

Research programme

6.17. The Review's formal academic research programme, comprising a literature review, quantitative analysis and primary qualitative research, has been based on the identified gaps in the evidence and the feasibility of filling them within the lifetime of the Review.

6.18. Initial work has identified the existing evidence base on epidemiology, natural history, and the treatment and outcomes of children and young people with gender dysphoria/gender-related distress. It has also assessed the feasibility of linking data between local, regional or national datasets in order to assess intermediate and longer-term outcomes.

Literature review

6.19. A literature review is being undertaken, which will interface with evidence gathering from the professional community (see qualitative research section below). Its aim is to systematically identify, collate and synthesise the existing evidence on the changing epidemiology of gender-related distress in children and young people and the appropriate social, clinical, psychological and medical management of that distress.

6.20. The literature review will capture primary studies of any design, including experimental, observational, survey and qualitative, and is looking to answer the following questions:

1. How has the population of children and young people presenting with gender dysphoria and/or gender-related distress changed over time?
2. What are the appropriate referral, assessment and treatment pathways for children and young people with gender dysphoria and/or gender-related distress?
3. What are the short-, medium- and long-term outcomes for children and young people with gender dysphoria and/or gender-related distress?
4. How do children and young people and their families negotiate distress, present this distress to services, and what are their expectations, following presentation?
5. How do children, young people and their families/carers experience referral, assessment and treatment? And how are these negotiated among children and young people, parents/carers, families and healthcare professionals?

6.21. A separate synthesis for each question will be undertaken. The systematic review has been registered on PROSPERO [ID:289659].

Quantitative research

6.22. The National Institute for Health and Care Excellence (NICE) recently published two evidence reviews.[76,77] These highlight shortcomings in the follow-up data collected about children and young people, when they are referred to a specialist gender identity service. The quantitative research will therefore focus on the collection and analysis of data to uncover patterns and quantify problems, thereby helping the Review to address some of these shortcomings.

6.23. The aim of the quantitative study is to supplement the material collected by the literature review, further examining the changing epidemiology of gender-related distress in children and young people, in addition to exploring the appropriate social, clinical, psychological and medical management. Its objectives are to:

a) describe the clinical and demographic characteristics of this population of children and young people and their clinical management in the GIDS service; and

b) assess the intermediate and longer-term outcomes of this population of children and young people utilising national healthcare data.

6.24. This research will provide an evidence base to facilitate informed decision making among children and young people and their families. It will also provide an evidence base for those responsible for commissioning, delivering and managing services.

Qualitative research

6.25. The qualitative research will capture a diverse range of trajectories experienced by gender-questioning children and young people, exploring a range of different experiences and outcomes. This will include talking to children and young people and their families/carers who are currently negotiating gender-related distress, young adults who have gone through the process of resolving their distress and care professionals.

[76] National Institute for Health and Care Excellence (2020). Evidence Review: Gonadotrophin Releasing Hormone Analogues for Children and Adolescents with Gender Dysphoria.
[77] National Institute for Health and Care Excellence (2020). Evidence review: gender-affirming hormones for children and adolescents with gender dysphoria.

The objectives of the qualitative research are to:

1. Explore how children and young people understand, respond and negotiate gender-related distress within the context of their social networks, alongside the perspectives of young adults who experienced gender distress as children.
2. Examine the perspectives, understandings and responses of parents (or carers), including how they support their child.
3. Investigate how children, young people, young adults and their families experience(d) and negotiate(d) referral, assessment and possible treatment and intervention options.
4. Understand the role and experiences of care professionals who offer support, including identifying shared and potentially divergent views among care professionals, children and young people, and parents of what constitutes optimal care.

Progress

6.26. The literature review is already underway and is identifying relevant studies. Initial meetings have also taken place with voluntary organisations and other researchers working in the area to ensure there is no duplication and in recognition of research fatigue among this population.

6.27. Children and young people and young adults who have experienced gender-related distress are involved in the research programme. Their advice has been, and will continue to be, sought throughout this work, including in relation to the focus of the research and interpretation of findings and the design and content of dissemination materials.

6.28. Three research protocols have been produced setting out how the research will be undertaken, and the research team is currently gaining the necessary ethical and governance approvals to progress the study. The systematic review is published on the PROSPERO website and will be published on the Review website in due course, along with the qualitative and quantitative research proposals once ethical and governance approvals have been received.

6.29. The research findings will be subject to peer review through the publication process and various summaries, aimed at different audiences, will be available on the project website and distributed via support organisations. These summaries will also be made available on the Review website.

Ongoing engagement

6.30. In recognition that not all the published evidence is likely to be of high enough quality to form the sole basis for our recommendations, a consensus development approach will be used to synthesise the published evidence and research outputs of the academic work with stakeholder submissions and expert opinion.

6.31. Over the coming months, the Review will build on its engagement to date and, alongside the academic research programme, will continue informal and structured engagement with service users, their families, support and advocacy groups and professionals to test emerging thinking, provide opportunities for challenge and further develop the evidence base.

6.32. This review is an iterative process and we will share important findings when they become available. For the latest updates, please visit our website: https://cass.independent-review.uk/

6.33. We thank those who have participated in the Review to date and welcome engagement with us as work progresses towards final recommendations.

Glossary

Glossary

There is sometimes no consensus on the best language to use relating to this subject. The language surrounding this area has also changed rapidly and young people have developed varied ways of describing their experiences using different terms and constructs that are relevant to them.

The Review tries as far as possible to use language and terms that are respectful and acknowledge diversity, but that also accurately illustrate the complexity of what we are trying to describe and articulate.

The terms we have used may not always feel right to some; nevertheless, it is important to emphasise that the language used is not an indication of a position being taken by the Review. The glossary below sets out a description of some of the terms we have used in the Review.

Term		Description
Affirmative model		A model of gender healthcare that originated in the USA[78,79,80,81] which affirms a young person's subjective gender experience while remaining open to fluidity and changes over time. This approach is used in some key child and adolescent clinics across the Western world.
Assent		To agree to or approve of something (idea, plan or request), especially after thoughtful consideration.
Autonomy		Personal autonomy is the ability of a person to make their own decisions. In health this refers specifically to decisions about their care.

[78] Hidalgo MA, Ehrensaft D, Tishelman AC, Clark LF, Garofalo R, Rosenthal SM, et al (2013). The gender affirmative model: What we know and what we aim to learn [Editorial]. Human Dev 56(5): 285–290. DOI:10.1159/000355235.
[79] Chen D, Abrams M, Clark L, Ehrensaft D, Tishelman AC, Chan YM, et al (2021). Psychosocial characteristics of transgender youth seeking gender-affirming medical treatment: baseline findings from the trans youth care study. J Adol Health 68(6): 1104–11.
[80] Olson-Kennedy J, Chan YM, Rosenthal S, Hidalgo MA, Chen D, Clark L, et al (2019). Creating the Trans Youth Research Network: A collaborative research endeavor. Transgend Health 4(12): 304–12. DOI: 10.1089/trgh.2019.0024.
[81] Ehrensaft D, Giammattei SV, Storck K, Tishelman AC, Colton K-M (2018). Prepubertal social gender transitions: What we know; what we can learn—A view from a gender affirmative lens. Int J Transgend 19(2): 251–68. DOI: 10.1080/15532739.2017.1414649.

Term		Description
Best interests		Clinicians and the courts seek to act in the best interests of children and young people. For the Mental Capacity Act (MCA) 2005, decisions for someone who cannot decide for themselves must be made in their best interests. Under the Children Act 1989, in any decision of the court about a child (under 18), the welfare of the child must be paramount. For these purposes, there is little or no material difference between the welfare and best interests, and we have used "best interests" throughout the report. Although there is no standard definition of "best interests of the child," the General Medical Council advises that an assessment of best interests will include what is clinically indicated as well as additional factors such as the child or young person's views, the views of parents and others close to the child or young person and cultural, religious and other beliefs and values of the child or young person.[82]
		The MCA s4,[83] and extensive Court of Protection case law, deals with the approach to best interests under that legislation. Whether in the Court of Protection or the High Court, when the court is asked to make an assessment of a child or young person's best interests, it will consider their welfare/best interests in the widest sense. This will include not just medical factors but also social and psychological factors.
Case-mix		The mix of patients within a particular group.
Child and adolescent mental health services	CAMHS	NHS children and young people's mental health services.[84]

[82] General Medical Council (2018). 0-18 years – guidance for all doctors.
[83] Mental Health Law Online. MCA 2005 s4.
[84] Young Minds. Guide to CAMHS: a guide for young people.

Term		Description
Child and/or young person		In law, everyone under 18 years of age is a child (Children Act 1989) but we recognise that it may be more appropriate to refer to those approaching the age of 18 as a young person, and that such young people may not recognise themselves as a "child". In places, we have referred only to "young person", or only to "child", for example where treatment in question is only given towards the later stages of childhood, closer to the age of 18, or in reference to the parent/child relationship, in which they remain the parents' child, regardless of their age. Otherwise, we have used the phrase "child and/or young person" throughout the report for this reason only, and do not intend there to be a material difference between them other than that.
Cognitive		Relating to, or involving, the process of thinking and reasoning.
Consent		Permission for a clinical intervention (such as an examination, test or treatment) to happen. For consent to be 'informed', information must be disclosed to the person about relevant risks, benefits and alternatives (including the option to take no action), and efforts made to ensure that the information is understood. In legal terms, consent is seen as needing: 1 – capacity (or *Gillick* competence under 16) to make the relevant decision; 2 – to be fully informed (ie the information provided about the available options, the material risks and benefits of each option, and of doing nothing, "material" meaning (per the Montgomery Supreme Court judgment in 2015) what a reasonable patient would want to know, and what this patient actually wants to know, NOT what a reasonable doctor would tell them); and 3 – to be freely given (that is,without coercion).
Contraindications		A condition or circumstance that suggests or indicates that a particular technique or drug should not be used in the case in question.

Term		Description
Court of Appeal		(England and Wales) The Court of Appeal hears appeals against both civil and criminal judgments from the Crown Courts, High Court and County Court. It is second only to the Supreme Court.
Detransition/ detransitioners		Population of individuals who experienced gender dysphoria, chose to undergo medical and/or surgical transition and then detransitioned by discontinuing medications, having surgery to reverse the effects of transition, or both.[85]
Diagnostic and Statistical Manual of Mental Disorders Fifth edition	DSM-5	The American diagnostic manual used to diagnose mental health disorders, and commonly used in UK practice. See **Appendix 3**.
Diagnostic formulation		The comprehensive assessment that includes a patient's history, results of psychological tests, and diagnosis of mental health difficulties.
Divisional Court		(England and Wales) When the High Court of Justice of England and Wales hears a case with at least two judges sitting, it is referred to as the Divisional Court. This is typically the case for certain judicial review cases (as well as some criminal cases).
Dutch Approach		Protocol published in 1998 by the Amsterdam child and adolescent gender identity clinic.[86]
Endocrine treatment		In relation to this clinical area, this term is used to describe the use of gonadotropin-releasing hormones (see below) and feminising and masculinising hormones (see below).
Endocrinologist		An endocrinologist is a medical doctor specialising in diagnosing and treating disorders relating to problems with the body's hormones.
Endocrinology		The study of hormones.

[85] Littman L (2021). Individuals treated for gender dysphoria with medical and/or surgical transition who subsequently detransitioned: a survey of 100 detransitioners. Arch Sex Abuse 50: 3353–69. DOI: 10.1007/s10508-021-02163-w

[86] de Vries ALC, Cohen-Kettenis PT (2012). Clinical management of gender dysphoria in children and adolescents: The Dutch approach. J Homosex 59: 301-320. DOI: 10.1080/00918369.2012.653300.

Term		Description
Epidemiology		Epidemiology is the study of the distribution and determinants of health-related states or events in specified populations, and the application of this study to the control of health problems.[87]
Exploratory approaches		Therapeutic approaches that acknowledge the young person's subjective gender experience, whilst also engaging in an open, curious, non-directive exploration of the meaning of a range of experiences that may connect to gender and broader self-identity.[88,89,90,91]
Feminising and masculinising hormones (also known as cross-sex hormones, and gender affirming hormones).		Hormones given as part of a medical transition for gender dysphoric individuals, where sex hormones (testosterone for transgender males and oestrogen for transgender females).
Gender dysphoria		Diagnostic term used in DSM-5.[92] Gender dysphoria describes "a marked incongruence between one's experienced/expressed gender and assigned gender of at least 6 months duration" which must be manifested by a number of criterion – see **Appendix 3** for further detail.
Gender fluid		An experience of gender that is not fixed, but changes between two or more identities.
Gender identity		This term is used to describe an individual's internal sense of being male or female or something else.
Gender identity development		The developmental experience of a child or young person in seeking to understand their gender identity over time.
Gender Identity Development Service	GIDS	The service that NHS England commissions for children and adolescents with gender dysphoria.

[87] Centers for Disease Control and Prevention (2012). Principles of Epidemiology in Public Health Practice: An introduction to Applied Epidemiology and Biostatistics, 3rd ed.
[88] Di Ceglie D (2009). Engaging young people with atypical gender identity development in therapeutic work: A developmental approach. J Child Psychother 35(1): 3–12. DOI: 10.1080/00754170902764868.
[89] Spiliadis A (2019). Towards a gender exploratory model: Slowing things down, opening things up and exploring identity development. Metalogos Systemic Ther J 35: 1–9.
[90] Churcher Clarke A, Spiliadis A (2019). 'Taking the lid off the box': The value of extended clinical assessment for adolescents presenting with gender identity difficulties. Clin Child Psychol Psychiatry 24(2): 338–52. DOI:10.1177/1359104518825288.
[91] Bonfatto M, Crasnow E (2018). Gender/ed identities: an overview of our current work as child psychotherapists in the Gender Identity Development Service. J Child Psychother 44(1): 29–46. DOI:10.1080/007541 7X.2018.1443150.
[92] American Psychiatric Association (2013). Diagnostic and Statistical Manual of Mental Health Disorders: DSM-5™, 5th ed.

Term		Description
Gender incongruence		Diagnostic term used in ICD-11.[93] Gender incongruence is characterised by "a marked and persistent incongruence between an individual's experienced gender and the assigned sex". See **Appendix 3** for further detail.
Gender-questioning		A broader term that might describe children and young people who are in a process of working out how they want to present in relation to their gender.
Gender-related distress		A way of describing distress that may arise from a broad range of experiences connected to a child or young person's gender identity development. Often used for young people whereby any formal diagnosis of gender dysphoria has not yet been made.
Gillick competence/ Fraser guidelines		A term derived from *Gillick v West Norfolk And Wisbech AHA*, 1984 that is used to decide whether a child or young person up to the age of 16 years is able to consent to their own medical treatment, without the need for parental permission or knowledge. A child or young person will be 'Gillick competent' for that decision if they have the necessary maturity and understanding to make the decision.
Gonadotropin-releasing hormone analogues (also known as the hormone blocker/s and puberty blocker/s)	GnRH	GnRH analogues competitively block GnRH receptors to prevent the spontaneous release of two gonadotropin hormones, Follicular Stimulating Hormone (FSH) and Luteinising Hormone (LH) from the pituitary gland. This arrests the progress of puberty.
General Practitioner	GP	GPs deal with a whole range of health problems and manage the care of their patients, referring onto specialists as appropriate.[94]
High Court		The third highest court in the UK. It deals with all high value and high importance civil law (non-criminal) cases and appeals of decisions made in lower courts. When the High Court sits with more than one judge, as required for certain kinds of cases, it is called the Divisional Court.
International Classification of Diseases, Version 11	ICD-11	ICD-11[95] is the World Health Organization (WHO) mandated health data standard used for medical diagnosis.

[93] World Health Organization (2022). International Classification of Diseases Eleventh Revision.
[94] NHS. GP services.
[95] World Health Organization (2022). International Classification of Diseases Eleventh Revision.

Term		Description
Looked after children		Children who are in the care of their Local Authority who may be living with foster parents or in a residential care setting.
Multi-disciplinary-team	MDT	The identified group of professional staff who provide a clinical service.
Neurodiverse		Displaying or characterised by autistic or other neurologically atypical patterns of thought or behaviour; not neurotypical.
Non-binary		A gender identity that does not fit into the traditional gender binary of male and female.[96]
Paediatrics		The branch of medicine dealing with children and their medical conditions.
Pass/passing		A person's gender being seen and read in the way they identify.
Precocious puberty		This is when a child's body begins changing into that of an adult (puberty) too soon – before age 8 in girls and before age 9 in boys.
Primary care		Primary care includes general practice, community pharmacy, dental and optometry (eye health) services. This tends to be the first point of access to healthcare.
Psychological formulation		A structured approach to understanding the factors underlying distressing states in a way that informs the changes needed and the therapeutic intervention for these changes to occur.
Psychosocial		Describes the psychological and social factors that encompass broader wellbeing.
Puberty blockers		See gonadotropin-releasing hormone above.
Secondary care		Hospital and community health care services that do not provide specialist care and are usually relatively close to the patient. For children this will include Child and Adolescent Mental Health Services (CAMHS), child development and general paediatric services.
Tanner Stage		Classification of puberty by stage of development. This ranges from Stage 1, before physical signs of puberty appear, to Stage 5 at full maturity.

[96] Twist J, de Graaf NM (2019). Gender diversity and non-binary presentations in young people attending the United Kingdom's National Gender Identity Development Service. Clin Child Psychol Psychiatry 24(2): 277–90. DOI: 10.1177/1359104518804311.

Term		Description
Tertiary care		Tertiary care is the specialist end of the NHS. These services relate to complex or rare conditions. Services are usually delivered in a number of hospitals/centres.
Transgender	trans	This is an umbrella term that includes a range of people whose gender identity is different from the sex they were registered at birth.
Transition		These are the steps a person may take to live in the gender in which they identify. This may involve different things, such as changing elements of social presentation and role and/or medical intervention for some.

Appendix 1

Terms of reference

TERMS OF REFERENCE FOR REVIEW OF GENDER IDENTITY DEVELOPMENT SERVICE FOR CHILDREN AND ADOLESCENTS

INTRODUCTION

1. NHS England is the responsible commissioner for specialised gender identity services for children and adolescents. The Gender Identity Development Service for children and adolescents is currently managed by the Tavistock and Portman NHS Foundation Trust.

2. In recent years there has been a significant increase in the number of referrals to the Gender Identity Development Service, and this has occurred at a time when the service has moved from a psychosocial and psychotherapeutic model to one that also prescribes medical interventions by way of hormone drugs. This has contributed to growing interest in how the NHS should most appropriately assess, diagnose and care for children and young people who present with gender incongruence and gender identity issues.

3. It is in this context that NHS England and NHS Improvement's Quality and Innovation Committee has asked Dr Hilary Cass to chair an independent review, and to make recommendations on how to improve services for children and young people experiencing issues with their gender identity or gender incongruence, and ensure that the best model/s for safe and effective services are commissioned.

REVIEW SCOPE

The independent review, led by Dr Cass, will be wide ranging in scope and will conduct extensive engagement with all interested stakeholders. The review is expected to set out findings and make recommendations in relation to:

i. Pathways of care into local services, including clinical management approaches for individuals with less complex expressions of gender incongruence who do not need specialist gender identity services;
ii. Pathways of care into specialist gender identity services, including referral criteria into a specialist gender identity service; and referral criteria into other appropriate specialist services;
iii. Clinical models and clinical management approaches at each point of the specialised pathway of care from assessment to discharge, including a description of objectives, expected benefits and expected outcomes for each clinical intervention in the pathway;
iv. Best clinical approach for individuals with other complex presentations.
v. The use of gonadotropin-releasing hormone analogues and gender affirming drugs, supported by a review of the available evidence by the National Institute for Health and Care Excellence; any treatment recommendations will include a description of treatment objectives, expected benefits and expected outcomes, and potential risks, harms and effects to the individual;
vi. Ongoing clinical audit, long term follow-up, data reporting and future research priorities;
vii. Current and future workforce requirements;
viii. Exploration of the reasons for the increase in referrals and why the increase has disproportionately been of natal females, and the implications of these matters; and,

TERMS OF REFERENCE FOR REVIEW OF GENDER IDENTITY DEVELOPMENT SERVICE FOR CHILDREN AND ADOLESCENTS

ix. Any other relevant matters that arise during the course of the review

4. In addition, and with support from the Royal College of Paediatrics and Child Health and other relevant professional associations, the Chair will review current clinical practice concerning individuals referred to the specialist endocrine service. It is expected that findings and any recommendations on this aspect of the review will be reported early in 2021 with the review's wider findings and recommendations delivered later in 2021.

5. The review will not immediately consider issues around informed consent as these are the subject of an ongoing judicial review. However, any implications that might arise from the legal ruling could be considered by the review if appropriate or necessary.

Appendix 2

Letter to NHS England from Dr Cass – May 2021

Dr Hilary Cass
Chair
Review of GIDS for Children and Young People

John Stewart
National Director
Specialised Commissioning
NHS England and NHS Improvement

Sent by email

10 May 2021

Dear John

INDEPENDENT REVIEW INTO GENDER IDENTITY SERVICES FOR CHILDREN AND YOUNG PEOPLE

I am writing to update you on my current approach to the work of the independent review into gender identity services for children and young people. However, the most pressing issue is how we augment the immediate support for children and young people currently needing assessment and treatment, some of whom have already been waiting for an extended period for an appointment. I will therefore also make some suggestions about interim arrangements and ways in which the review team could help to support and strengthen these.

Commissioned research programme

As you know, a key principle of the review is that it should be evidence-based, and that final conclusions will be developed through a consensus development process contingent on the synthesised evidence.

I am pleased to see that the National Institute for Health and Care Excellence (NICE) evidence reviews of gonadotrophin releasing hormone analogues and gender affirming hormones for children and adolescents with gender dysphoria have now been published. Although this is a helpful starting point, despite following a standard and robust process the NICE review findings are not conclusive enough to inform policy decisions. As part of my review, I am therefore exploring other methodologies to give increased confidence and clarity about the optimal treatment approaches.

My team is commissioning a broader literature review of the existing evidence base on the epidemiology, management and outcomes of children with gender dysphoria. We are also commissioning qualitative and quantitative research, including considering other approaches which might be employed to understand the intermediate and longer-term outcomes of children with gender dysphoria. We intend to include a review of international models and data in this programme of work.

Addressing the immediate situation

Recognising that the outcome of the review is going to take some time, I have been reflecting on the recent court rulings on puberty blockers and consent and the Care Quality Commission (CQC) report on the Gender Identity Development Service (GIDS) run by the Tavistock and Portman NHS Foundation Trust. These significant developments have changed the context in which the review is taking place, and further added to the service pressures.

I note the proposal to establish an independent multidisciplinary professional review group to confirm decision-making has followed a robust process, which seems an appropriate interim measure pending further clarification of the legal situation.

I know that everyone concerned with the delivery of services – both commissioners and providers – are worried about the increasing number of children on the waiting list for assessment by the GIDS service and the resulting distress for the children and young people and their families. The difficulty in managing risk for those on the waiting list is exacerbated by the staff vacancies at GIDS, the increasing volume of new referrals, and the fact that the support and engagement from local services is highly variable and, in some cases, very limited.

Having a single provider may have been a logical position when the GIDS service was first set up, given that this is a highly specialised service that was seeing a relatively small number of cases each year. As the epidemiology has changed and there has been an exponential increase in numbers of children with gender incongruence or dysphoria, concentration of expertise within a single service has become unsustainable. At the same time, local services have not developed the skills and competencies to provide support for children on the waiting list and those with lesser degrees of gender incongruence who may not wish to pursue specialist medical intervention, and / or to provide help for children with additional complex needs.

I know from discussions we have had that your team is working hard to find some practical alternative arrangements, and that you have been in discussion with relevant professional bodies to come up with creative interim solutions while awaiting the outcome of my review.

The review team has also been in discussion with CQC, with the Tavistock and Portman NHS Foundation Trust and with colleagues within and external to NHS England and NHS Improvement to consider which aspects of this situation we can help with in the short to medium term, whilst keeping our focus on the longer-term questions of the appropriate clinical management and whole care pathway for these children and young people. In the past months I have also met with many groups and individuals with expertise and lived experience relevant to the review, including charities and support groups, Royal Colleges and healthcare professionals.

Recommendations to NHS England and NHS Improvement

I would encourage you to consider the following when developing an interim pathway for children and young people experiencing gender dysphoria:

- **Access and referral:** Children and young people need ready access to services. However, it is unusual for a specialist service to take direct referrals. The risk of having a national service as the first point of access is that assessment and treatment of children and young people who have the greatest need for specialist care is delayed because of the lack of differentiation of those on the waiting list. In addition, many children and

young people have complex needs, but once they are identified as having gender dysphoria, other important healthcare issues which would normally be managed by local services can sometimes be overlooked.

- **Assessment and management:** All children and young people who are referred to specialist services should have a competent local multi-disciplinary assessment and should remain under active holistic local management until they are seen at a specialist centre.

  I recognise that developing capacity and capability outside of the existing GIDS service to provide such initial assessment and support will be difficult to achieve at speed and will be incremental. This means that there will likely be a range of different models and options around the country, dependent on local resources, with some of the work being delivered through existing secondary service teams, and some being delivered at regional level. The support of wider services is vital.

- **Data:** The lack of systematic data collection is a significant issue. Therefore, when employing interim measures, I would suggest that particular attention is paid to the gathering of good quality data, which can then be used to inform the evidence base and future model of provision.

Actions for the review team

I would like to suggest how the review team might help with the challenging problem of growing an infrastructure outside of GIDS. From my conversations to date, I believe there are three barriers to the involvement of local services:

- **Capacity** – the staff most appropriately trained to be involved in initial assessment are those who are already most stretched within Child and Adolescent Mental Health Services (CAMHS) and paediatric services, and this situation has been significantly worsened through the impact of the Covid-19 pandemic on children's mental health. However, I know that there is substantial investment in CAMHS services, so close engagement with the relevant national policy teams at NHS England and NHS Improvement and at Health Education England (HEE) will be crucial.

- **Capability and confidence** – clinical teams outside of GIDS do not feel confident in initial assessment and support of children and young people with gender incongruence and dysphoria, in large part because they have not had the necessary training and experience, but also because of the societal polarisation and tensions surrounding the management of this group.

- **Lack of an explicit assessment framework** – currently expertise in assessment of children and young people presenting to GIDS is held in a small body of clinicians and their assessment processes have not been made explicit. The CQC report drew attention to the lack of structured assessment in the GIDS notes, and this is something that the Tavistock and Portman NHS Foundation Trust is already working to address internally. However, it is equally important to develop an initial assessment approach that can be used by first contact professionals, not just those working in the specialist service.

In the first instance, it is important that we test these assumptions with a range of clinical staff and ascertain whether there are other barriers that are preventing local engagement in this work. Then we would plan to prioritise a series of workshops, in collaboration with relevant professional groups, service users and close engagement with HEE. The purpose of these workshops would be to address identified barriers and develop:

- A framework for initial assessment of children and young people presenting with gender dysphoria.
- An approach to training for professionals at local and regional level.
- Some preliminary workforce recommendations, which will be particularly important in meeting the timelines of the three-year Comprehensive Spending Review.

These workshops will serve multiple purposes – firstly to support NHS England and NHS Improvement in the establishment of local and / or regional teams; secondly as an essential component of the work needed to inform the questions that the review is tackling; and thirdly to form the professional networks that will be needed to underpin future service and research networks.

Timelines

As you will recognise, setting up a complex national review is difficult and time consuming at the best of times. It requires a team to support the work and mechanisms for stakeholders to engage safely and with confidence. Starting a review in the midst of a pandemic is even more challenging.

I have committed to a review approach which is participative, consensus-based, evidence-based, transparent, and informed by lived and professional experience. This requires extensive engagement. Pending the appointment of our research team, the review has now launched its website and I have been proactively engaging with the stakeholder community.

It is critical that we get the approach right, particularly the engagement, the evidence review and the quantitative research given the gaps in the evidence highlighted through the NICE review, and this will take time.

My intention is that an interim report will be delivered in the summer, with a report next year setting out my final recommendations.

Yours sincerely

Dr Hilary Cass
Chair, Independent Review into Gender Identity Services for Children and Young People

Cc: Care Quality Commission
Health Education England
Tavistock and Portman NHS Foundation Trust

Appendix 3

Diagnostic criteria for gender dysphoria

DSM-5 diagnostic criteria for gender dysphoria

Gender Dysphoria in Children

A. A marked incongruence between one's experienced/expressed gender and assigned gender, of at least 6 months' duration, as manifested by at least six of the following (one of which must be Criterion A1):

1. A strong desire to be of the other gender or an insistence that one is the other gender (or some alternative gender different from one's assigned gender).
2. In boys (assigned gender), a strong preference for cross-dressing or simulating female attire; or in girls (assigned gender), a strong preference for wearing only typical masculine clothing and a strong resistance to the wearing of typical feminine clothing.
3. A strong preference for cross-gender roles in make-believe play or fantasy play.
4. A strong preference for the toys, games, or activities stereotypically used or engaged in by the other gender.
5. A strong preference for playmates of the other gender.
6. In boys (assigned gender), a strong rejection of typically masculine toys, games, and activities and a strong avoidance of rough-and-tumble play; or in girls (assigned gender), a strong rejection of typically feminine toys, games, and activities.
7. A strong dislike of one's sexual anatomy.
8. A strong desire for the primary and/or secondary sex characteristics that match one's experienced gender.

B. The condition is associated with clinically significant distress or impairment in social, school, or other important areas of functioning.

Specify if:

With a disorder of sex development (e.g., a congenital adrenogenital disorder such as congenital adrenal hyperplasia or androgen insensitivity syndrome).

Gender Dysphoria in Adolescents and Adults

A. A marked incongruence between one's experienced/expressed gender and assigned gender, of at least 6 months' duration, as manifested by at least two of the following:

1. 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).
2. 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).

3. A strong desire for the primary and/or secondary sex characteristics of the other gender.
4. A strong desire to be of the other gender (or some alternative gender different from one's assigned gender).
5. A strong desire to be treated as the other gender (or some alternative gender different from one's assigned gender).
6. 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).

B. The condition is associated with clinically significant distress or impairment in social, occupational, or other important areas of functioning.

Specify if:

With a disorder of sex development (e.g., a congenital adrenogenital disorder such as congenital adrenal hyperplasia or androgen insensitivity syndrome).

Specify if:

Post transition: the individual has transitioned to full-time living in the desired gender (with or without legalization of gender change) and has undergone (or is preparing to have) at least one cross-sex medical procedure or treatment regimen – namely, regular cross-sex hormone treatment or gender reassignment surgery confirming the desired gender (e.g., penectomy, vaginoplasty in a natal male; mastectomy or phalloplasty in a natal female).

ICD-11: HA60 Gender incongruence of adolescence or adulthood

Gender Incongruence of Adolescence and Adulthood is characterised by a marked and persistent incongruence between an individual´s experienced gender and the assigned sex, which often leads to a desire to 'transition', in order to live and be accepted as a person of the experienced gender, through hormonal treatment, surgery or other health care services to make the individual´s body align, as much as desired and to the extent possible, with the experienced gender. The diagnosis cannot be assigned prior the onset of puberty. Gender variant behaviour and preferences alone are not a basis for assigning the diagnosis.

Exclusions:

Paraphilic disorders.

ICD-11: HA61 Gender incongruence of childhood

Gender incongruence of childhood is characterised by a marked incongruence between an individual's experienced/expressed gender and the assigned sex in pre-pubertal children. It includes a strong desire to be a different gender than the assigned sex; a strong dislike on the child's part of his or her sexual anatomy or anticipated secondary sex characteristics and/or a strong desire for the primary and/or anticipated secondary sex characteristics that match the experienced gender; and make-believe or fantasy play, toys, games, or activities and playmates that are typical of the experienced gender rather than the assigned sex. The incongruence must have persisted for about 2 years. Gender variant behaviour and preferences alone are not a basis for assigning the diagnosis.

Exclusions:

Paraphilic disorders.

Appendix 4

The standard approach to clinical service development

The standard approach to clinical service development

The three examples below illustrate the usual process of developing a clinical service: Covid-19 is included because this is a new condition that everyone is familiar with; childhood epilepsy because it is a complex condition with physical manifestations; and autism because it is a condition with neuro-behavioural manifestations.

By comparing these examples of clinical service development, it is possible to demonstrate some of the challenges in developing services for children and young people with gender incongruence or dysphoria, and to identify where there are gaps and questions that need to be addressed for this population, in order to ensure any future service model delivers the highest possible standards of care.

The stages below may proceed in a different sequence for different conditions, but each stage is important in the development of evidence based care.

Stage	Covid-19	Childhood Epilepsy	Autism
New condition is observed This often begins with a few case reports and then clinicians begin to recognise a recurring pattern and key clinical features, and to develop fuller descriptions of the condition.	Covid-19 is an example of a recent new condition that we all recognise, and this started with a few unusual cases of respiratory illness being described in Wuhan.	Childhood epilepsy has been recognised for centuries, but over the last century there has been growing understanding of the many different subtypes.	Individuals with autism have probably also existed for an indefinite period, but it wasn't until 1943 and 1944 that Leo Kanner and Hans Asperger wrote the first scientific accounts about the condition.

Stage	Covid-19	Childhood Epilepsy	Autism
Aetiology Clinicians and scientists try to work out the cause of the condition or the underlying physical or biological basis. Sometimes the answers to this are never found.	The cause of Covid-19 was identified at a very early stage as being due to a novel coronavirus, although it remains unclear where and how this originated.	It is now known that there are numerous different types of epilepsy, with many different causes – for example, epilepsy can be caused by specific epilepsy genes, by birth trauma, by metabolic conditions, by brain tumours and many other mechanisms. Epilepsies due to a change in the brain structure which occur after birth are called 'symptomatic' – they are a symptom of something else. Epilepsies for which there is no identified cause are called 'idiopathic'.	The first theory about the aetiology of autism was that it was caused by so called 'refrigerator parents'. This was inaccurate and damaging. It has subsequently been shown that there are many complex genetic and physical or chemical brain changes underpinning this condition.
Natural history and prognosis It is important to understand how a condition usually evolves over time, with or without treatment. The latter is important if treatment has limited efficacy and the condition is 'self-limiting' (that is, it resolves without treatment), because otherwise there is a risk that treatments create more difficulties than the condition itself.	Covid-19 is an example of a condition where there are quite polarised views about management based on its prognosis and natural history. A relatively small proportion of people are seriously affected and need treatment, and for the majority the natural history is that it will get better by itself. This has led some people to question the need for lockdowns, vaccinations and other measures which they see as impacting personal freedoms.	In epilepsy the natural history is very important. Some epilepsies get better through puberty and into adulthood, and some can get worse with hormonal changes. This is important to know when monitoring and reviewing drug treatment.	

Stage	Covid-19	Childhood Epilepsy	Autism
Epidemiology Epidemiologists collect data to find out how common a condition is, who is most likely to be affected, what the age distribution is and so on. This allows health service planners to work out how many services are needed, where they should be established, and what staff are needed. They also report on changes in who is most affected, which may mean that either the disease is changing, or the susceptibility of the population is changing.	Epidemiologists have been crucial in supporting the management of Covid-19 because they have extracted and analysed the data on which patients are at greater risk from the virus. This has been fundamental to planning a vaccination strategy and other protective measures.		The epidemiology of autism has changed considerably, with a dramatic increase in the numbers of children diagnosed over the last 20 years. This has had major implications for service provision. There is ongoing debate about the cause of the increase – whether it is because of greater awareness and better diagnosis, or because there are more children with autism. Current opinion favours the first option.

Stage	Covid-19	Childhood Epilepsy	Autism
Assessment and diagnosis Clinicians will usually take a history from (that is, of their symptoms) and examine the patient (that is, for signs and symptoms), and where appropriate undertake a series of investigations or tests, to help them reach an accurate diagnosis. Sometimes the whole process of making a diagnosis through talking to the patient and asking them to complete formal questionnaires, examining them and/or undertaking investigations is called 'clinical assessment'. As well as diagnosing and ruling out a particular condition, clinicians often need to consider and exclude other, sometimes more serious, conditions that present in a similar way but may need quite different treatment – this process is called 'differential diagnosis'.	PCR has been used as a 'gold standard' test for diagnosis of Covid-19 since the beginning of the pandemic. Lateral flow testing was developed to provide a quicker and cheaper option, but it demonstrates the limitations of testing; it is 99.68% specific, which is a very high specificity. This means there are only a tiny number of false positives. It has lower sensitivity at 76.8%, which means it will miss about a quarter of all cases, so giving many more false negatives, BUT it will only miss 5% of cases with high viral load.	Epilepsy can only be definitively diagnosed by either getting a really clear description of the events from a parent or carer, or seeing the child or young person having a seizure on a video. An EEG (brain wave tracing) and other tests can provide information about the type of epilepsy, but unless a seizure happens during the recording, it does not demonstrate that they actually have seizures – only that they may be susceptible to seizures.	In autism there are no blood tests or X-rays to make the diagnosis. It is a 'clinical' diagnosis, which means it is dependent on taking a standardised history from the parents, and performing standardised assessments on the child or young person to distinguish between autism and other possible diagnoses (for example, language disorder, social anxiety). In the early days, these standardised measures did not exist; the diagnosis was very dependent on experts who were used to diagnosing autism by making a clinical judgement about each child. This made it difficult to teach new people how to do this without a long apprenticeship, and also made it difficult to know whether two different experts would come to the same conclusion about the same child or young person. Standardisation of the questions and process made diagnosis more reliable and consistent, as did an improved evidence base. At the same time, because children with autism all present differently, the assessment had to be flexible enough to accommodate, for example, non-verbal children with severe learning disability, as well as high-functioning children with strong verbal skills.

Stage	Covid-19	Childhood Epilepsy	Autism
Differential diagnosis As well as making a positive diagnosis, clinicians often need to exclude other, sometimes more serious conditions that present in a similar way, but may need quite different treatment.		There are conditions that can be mistaken for epilepsy, so it is important to accurately diagnose whether seizures are happening and exclude other conditions (differential diagnoses) by carrying out relevant tests.	There are many conditions that may be mistaken for autism – for example, children who have language disorders, learning disability, severe social anxiety for other reasons, or ADHD can all appear to have autism. It is important to exclude these other conditions as well as making a positive diagnosis of autism. Sometimes these conditions can exist alongside autism, and management must then be planned to address all the child's difficulties.

Stage	Covid-19	Childhood Epilepsy	Autism
Developing and implementing new treatments Clinicians and scientists work on developing treatments. This involves clinical trials and, where there are new treatments, comparing them to any existing treatments. Questions include: What are the intended outcomes or benefits of treatment? What are the complications or side effects? What are the costs? To initiate a new treatment, it must be both safe and effective. Questions of affordability can sometimes become controversial. The best type of single study is considered to be the randomised controlled trial (RCT), but sometimes this is not feasible. Even where RCTs are not available, it is usual to at least have data on the outcomes of sufficient cases or cohorts to understand the risk/benefit of the treatment under consideration. As demonstrated in Fig. 3, the highest level of evidence is when the results of several different studies are pooled, but this is only useful if the individual studies themselves are of high quality.	Developing treatments for Covid-19 has been possible at speed because of the large numbers of patients, and the fact that outcomes can be observed on each patient within a matter of days to weeks. Because Covid-19 was a new condition, clinicians also started in a position of 'equipoise' which means that they did not have reason to believe any one treatment might be more effective than another; this made it ethical to have one group having a treatment and another group having a different treatment or a placebo. There are also really clear outcome measures, such as whether or not patients survive or need hospitalisation. This has facilitated a high level of evidence through randomised controlled trials (see diagram below).	Similar considerations apply to the treatment of epilepsy in that there are 'hard' outcome measures (for example, frequency of seizures), but it can take several months to determine whether a new drug is better than an existing one for any one patient, and some side effects may be longer-term, so trials can take several years. In addition, children with epilepsy may have very different conditions causing their seizures which can also make trials more challenging. In the most severe cases of epilepsy, surgery may be the best option for controlling seizures. This can be very radical in certain cases and have lifelong implications for how they function. These options, which have a cost as well as a benefit to the child, will only be offered after a multi-disciplinary team meeting, including the paediatricians, therapists, neuropsychologists, radiologists, neurophysiologists and neurosurgeons have all discussed whether the benefits will outweigh the costs.	Evaluating interventions for autism is the most difficult of these three examples. This is because it can take many years to see developmental outcomes; it is hard to get uniform groups of children; outcomes are extremely sensitive to the social (and historical) response of others; and many other things happen in children's lives (such as changes of school, other medications, new diets). Isolating the effect of the target treatment is therefore challenging.

Stage	Covid-19	Childhood Epilepsy	Autism
In many instances, evidence is not perfect and difficult decisions have to be made. Where treatments are innovative or life-changing, the whole multi-disciplinary team will usually meet to consider the available options, and how to advise the child or young person and family so that a shared decision can be made. Sometimes an ethics committee is involved. This is one of the most challenging areas of medicine and is underpinned by GMC guidance.[97, 98]	The UK has been internationally recognised for its Recovery Trial, led by Oxford University. This has recruited over 46,000 participants, and resulted in several treatments being approved. A key factor in this success was the willingness of patients to participate in these studies – with over 46,000 being recruited and consented.		

[97] General Medical Council (2020). Decision making and consent.
[98] National Institute for Health and Care Excellence (2021). Shared decision making.

Stage	Covid-19	Childhood Epilepsy	Autism
Service development and service improvement Central to any service improvement is the systematic and consistent collection of data on outcomes of treatment. There is a process of continuous service improvement as new presentations or variations on the original condition are recognised, diagnosis or screening improves and/or trials on new treatments or variations on existing treatments are ongoing. There should be consistent treatment protocols or guidelines in place, in order to make sense of variations in outcomes. Where possible, these should be compared between and across multiple different centres. As time passes, services need to be changed or extended based on patient need, and on what resources are needed to deliver the available treatments. They need to be accessible where the prevalence of the condition is highest. The relevant workforce to deliver the service needs to be recruited and trained, contingent on the type of treatments or therapy that is required.	Service development to manage Covid-19 has been on a scale unlike any normal new service development ever experienced. It has also demonstrated how other non-Covid services have had to evolve alongside, including the need for isolation, and/or PCR testing prior to routine clinical appointments, use of remote consultation and an array of other changes across the NHS. Continuous audit and monitoring of outcomes has resulted in major improvements in survival – for example, changing ventilation approach to include 'proning' (putting patients on their front while on the ventilator) and delaying fully intubated ventilation by giving mask ventilation for as long as possible.	Paediatric epilepsy is a good example of how a national approach can be taken to service improvement through the Epilepy12 programme.[99] This is a nationally co-ordinated audit which collects a standardised dataset, incorporating NICE standards, and is used to drive up standards of care for children and young people with epilepsy.	Improvement in autism services has been driven by the changing epidemiology, NICE standards, extensive training of the workforce and attempts to improve public understanding. Where previously diagnosis was undertaken in a few specialist centres, the rising waiting times and NICE standards on access, assessment and appropriate multi-professional provision have led to almost every community child development service having an autism assessment clinic or team. Services are able to self-assess against national standards to inform local improvement strategies.

[99] Royal College of Paediatrics and Child Health (2021). Epilepsy 12 – national organisational audit and clinical audit.

References

AB v CD & Ors [2021] EWHC 741.

American Psychiatric Association (2013). Diagnostic and Statistical Manual of Mental Health Disorders: DSM-5TM, 5th ed.

Bell v Tavistock [2020] EWHC 3274 (Admin).

Bloom TM, Nguyen TP, Lami F, Pace CC, Poulakis Z, Telfer N (2021). Measurement tools for gender identity, gender expression, and gender dysphoria in transgender and gender-diverse children and adolescents: a systematic review. Lancet Child Adolescent Health. 5: 582-588. DOI: 10.1016/s2352-4642(21)00098-5.

Bonfatto M, Crasnow E (2018). Gender/ed identities: an overview of our current work as child psychotherapists in the Gender Identity Development Service. J Child Psychother 44(1): 29–46. DOI:10.1080/0075417X.2018.1443150.

Brik T, Vrouenraets LJJJ, de Vries MC, Hannema SE (2020). Trajectories of adolescents treated with gonadotropin-releasing hormone analogues for gender dysphoria. Arch Sex Behav 49: 2611–8. DOI: 10.1007/s10508-020-01660-8.

Care Quality Commission (2021). The Tavistock and Portman NHS Foundation Trust Gender Identity Service Inspection Report. London: CQC.

Carmichael P, Butler G, Masic U, Cole TJ, De Stavola BL, Davidson S, et al (2021). Short-term outcomes of pubertal suppression in a selected cohort of 12 to 15 year old young people with persistent gender dysphoria in the UK. PLoS One. 16(2):e0243894. DOI:10.1371/journal.pone.0243894.

Centers for Disease Control and Prevention (2012). Principles of Epidemiology in Public Health Practice: An introduction to Applied Epidemiology and Biostatistics, 3rd ed.

Chen D, Abrams M, Clark L, Ehrensaft D, Tishelman AC, Chan YM, et al (2021). Psychosocial characteristics of transgender youth seeking gender-affirming medical treatment: baseline findings from the trans youth care study. J Adol Health 68(6): 1104–11.

Chen D, Strang JF, Kolbuck VD, Rosenthal SM, Wallen K, Waber DP, et al (2020). Consensus parameter: research methodologies to evaluate neurodevelopmental effects of pubertal suppression in transgender youth. Transgender Health 5(4). DOI: 10.1089/trgh.2020.0006.

Churcher Clarke A, Spiliadis A (2019). 'Taking the lid off the box': The value of extended clinical assessment for adolescents presenting with gender identity difficulties. Clin Child Psychol Psychiatry 24(2): 338–52. DOI:10.1177/1359104518825288.

Cohen-Kettenis PT, Steensma TD, de Vries ALC (2001). Treatment of adolescents with gender dysphoria in the Netherlands. Child Adolesc Psychiatr Clin N Am 20: 689–700. DOI: 10.1016/j.chc.2011.08.001.

Cohen-Kettenis PT, Van Goozen S (1998). Pubertal delay as an aid in diagnosis and treatment of a transsexual adolescent. Eur Child Adolesc Psychiatry 7: 246–8. DOI: 10.1007/s007870050073.

de Graaf NM, Carmichael P, Steensma TD, Zucker KJ (2018). Evidence for a change in the sex ratio of children referred for Gender Dysphoria: Data from the Gender Identity Development Service in London (2000–2017). J Sex Med 15(10): 1381–3. DOI: 10.1016/j.jsxm.2018.08.002.

de Graaf NM, Giovanardi G, Zitz C, Carmichael P (2018). Sex ratio in children and adolescents referred to the gender identity development service in the UK (2009-2016). Arch Sex Behav 47(5): 1301–4.

de Vries ALC, Cohen-Kettenis PT (2012). Clinical management of gender dysphoria in children and adolescents: the Dutch approach. J Homosex 59: 301–320. DOI: 10.1080/00918369.2012.653300.

Delemarre-van de Wall HA, Cohen-Kettinis PT (2006). Clinical management of gender identity disorder in adolescents: a protocol on psychological and paediatric endocrinology aspects. Eur J Endocrinol 155 (Suppl 1): S131–7. DOI: 10.1530/eje.1.02231.

Delevichab K, Klinger M, Nana OJ, Wilbrecht L (2021). Coming of age in the frontal cortex: The role of puberty in cortical maturation. Semin Cell Dev Biol 118: 64–72. DOI: 10.1016/j.semcdb.2021.04.021.

Di Ceglie D (2009). Engaging young people with atypical gender identity development in therapeutic work: A developmental approach. J Child Psychother 35(1): 3–12. DOI: 10.1080/00754170902764868.

Ehrensaft D, Giammattei SV, Storck K, Tishelman AC, Colton K-M (2018). Prepubertal social gender transitions: What we know; what we can learn—A view from a gender affirmative lens. Int J Transgend 19(2): 251–68. DOI: 10.1080/15532739.2017.1414649.

EWCA [2021] Civ 1363.

Gender Identity Clinic, The Tavistock and Portman NHS Foundation Trust. Waiting times.

Gender Identity Development Service. Referrals to GIDS, financial years 2010-11 to 2020-21.

General Medical Council (2018). 0-18 years – guidance for all doctors.

General Medical Council (2020). Decision making and consent.

General Medical Council (2021). Good practice in prescribing and managing medicines and devices (76-78).

Gillick v West Norfolk and Wisbech AHA [1986] AC 112.

Goddings A-L, Beltz A, Jiska S, Crone EA, Braams BR (2019). Understanding the role of puberty in structural and functional development of the adolescent brain. J Res Adolesc 29(1): 32–53. DOI: 10.1111/jora.12408.

Hembree WC, Cohen-Kettenis PT, Gooren L, Hannema SE, Meyer WJ, Murad MH, et al (2017). Endocrine treatment of gender-dysphoric/gender-incongruent persons: an Endocrine Society clinical practice guideline. J Clin Endocrinol Metab 102(11): 3869–903. DOI: 10.1210/jc.2017-01658.

Hidalgo MA, Ehrensaft D, Tishelman AC, Clark LF, Garofalo R, Rosenthal SM, et al (2013). The gender affirmative model: What we know and what we aim to learn [Editorial]. Human Dev 56(5): 285–290. DOI:10.1159/000355235.

Kyriakou A, Nicolaides NC, Skordis N (2020). Current approach to the clinical care of adolescents with gender dysphoria. Acta Biomed 91(1): 165–75. DOI: 10.23750/abm.v91i1.9244.

Littman L (2021). Individuals treated for gender dysphoria with medical and/or surgical transition who subsequently detransitioned: a survey of 100 detransitioners. Arch Sex Abuse 50: 3353–69. DOI: 10.1007/s10508-021-02163-w

Matthews T, Holt V, Sahin S, Taylor A, Griksaitis (2019). Gender Dysphoria in looked-after and adopted young people in a gender identity development service. Clinical Child Psychol Psychiatry 24: 112-128. DOI: 10.1177/1359104518791657.

Mental Health Law Online. MCA 2005 s4.

National Health Service Commissioning Board and Clinical Commissioning Groups (Responsibilities and Standing Rules) Regulations 2012.

National Institute for Health and Care Excellence (2020). Evidence Review: Gonadotrophin Releasing Hormone Analogues for Children and Adolescents with Gender Dysphoria.

National Institute for Health and Care Excellence (2020). Evidence review: gender-affirming hormones for children and adolescents with gender dysphoria.

National Institute for Health and Care Excellence (2021). Shared decision making.

NHS. GP services.

Office for National Statistics (2019). What is the difference between sex and gender?

Olson-Kennedy J, Chan YM, Rosenthal S, Hidalgo MA, Chen D, Clark L, et al (2019). Creating the Trans Youth Research Network: A collaborative research endeavor. Transgend Health 4(12): 304–12. DOI: 10.1089/trgh.2019.0024.

OpenMD (2021). New Evidence in Medical Research.

Royal College of Paediatrics and Child Health (2021). Epilepsy 12 – national organisational audit and clinical audit – 2021.

Scobie S, Castle-Clarke S (2019). Implementing learning health systems in the UK NHS: Policy actions to improve collaboration and transparency and support innovation and better use of analytics. Learning Health Systems 4(1): e10209. DOI:10.1002/lrh2.10209.

Sievert EDC, Schweizer K, Barkmann C, Fahrenkrug S, Becker-Hebly I (2020). Not social transition status, but peer relations and family functioning predict psychological functioning in a German clinical sample of children with Gender Dysphoria.Clin Child Psychol Psychiatry 26(1): 79–95. DOI: 10.1177/1359104520964530

Spiliadis A (2019). Towards a gender exploratory model: Slowing things down, opening things up and exploring identity development. Metalogos Systemic Ther J 35: 1–9.

Steensma TD, Biemond R, de Boer F, Cohen-Kettenis PT (2011). Desisting and persisting gender dysphoria after childhood: a qualitative follow-up study. Clin Child Psychol Psychiatry 16(4): 485-97. DOI: 10.1177/135910451037803.

Steensma TD, Cohen-Kettenis PT, Zucker KJ (2018). Evidence for a change in the sex ratio of children referred for gender dysphoria: Data from the Center of Expertise on Gender Dysphoria in Amsterdam (1988-2016). Journal of Sex & Marital Therapy 44(7): 713–5. DOI: 10.1080/0092623X.2018.1437580.

Steensma TD, McGuire JK, Kreukels BPC, Beekman AJ, Cohen-Kettenis PT (2013). Factors associated with desistence and persistence of childhood gender dysphoria: a quantitative follow-up study. J Am Acad Child Adolesc Psychiatry 52: 582-590. DOI: 10.1016/j.jaac.2013.03.016.

Turban JL, King D, Carswell JM, et al (2020). Pubertal suppression for transgender youth and risk of suicidal ideation. Pediatrics 145 (2): e20191725. DOI: 10.1542/peds.2019-1725.

Twist J, de Graaf NM (2019). Gender diversity and non-binary presentations in young people attending the United Kingdom's National Gender Identity Development Service. Clin Child Psychol Psychiatry 24(2): 277–90. DOI: 10.1177/1359104518804311.

Van Der Miesen AIR, Hurley H, De Vries ALC (2016). Gender dysphoria and autism spectrum disorder: A narrative review. Int Rev Psychiatry 28: 70-80. DOI: 10.3109/09540261.2015.1111199.

Vrouenraets LJ, Fredriks AM, Hannema SE, Cohen-Kettenis PT, de Vries MC (2015). Early medical treatment of children and adolescents with gender dysphoria: an empirical ethical study. J Adolesc Health 57(4): 367-73. DOI: 10.1016/.2015.04.004.

World Health Organization (2022). International Classification of Diseases Eleventh Revision.

Wren B (2019). Notes on a crisis of meaning in the care of gender-diverse children. In: Hertzmann L, Newbigin J (eds) Sexuality and Gender Now: Moving Beyond Heteronormativity. Routledge.

Young Minds. Guide to CAMHS: a guide for young people.

Zucker KJ (2017). Epidemiology of gender dysphoria and transgender identity. Sex Health 14(5): 404–11. DOI:10.1071/SH1.

Zucker KJ, Lawrence AA (2009). Epidemiology of gender identity disorder: recommendations for the Standards of Care of the World Professional Association for Transgender Health. Int J Transgend 11(1): 8-18. DOI: 10.1080/15532730902799946.

EXHIBIT 25

Summary of a recommendation by COHERE Finland 16.6.2020

Medical treatment methods for dysphoria associated with variations in gender identity in minors – recommendation

In its meeting on 11 June 2020, the Council for Choices in Health Care in Finland (COHERE Finland) adopted a recommendation on medical treatment methods for dysphoria associated with variations in the gender identity of minors

The recommendation clarifies the roles of different healthcare operators in a situation where a minor is uncertain about their gender identity. The recommendation presents the medical treatment methods that fall within the range of public healthcare services when it comes to the medical treatment of gender dysphoria in minors.
In COHERE's view, psychosocial support should be provided in school and student healthcare and in primary healthcare for the treatment of gender dysphoria due to variations in gender identity in minors, and there must be sufficient competency to provide such support. Consultation with a child or youth psychiatrist and the necessary psychiatric treatment and psychotherapy should be arranged locally according to the level of treatment needed. If a child or young person experiencing gender-related anxiety has other simultaneous psychiatric symptoms requiring specialised medical care, treatment according to the nature and severity of the disorder must be arranged within the services of their own region, as no conclusions can be drawn on the stability of gender identity during the period of disorder caused by a psychiatric illness with symptoms that hamper development.

In Finland, the diagnostics of gender identity variation, the assessment of the need for medical treatments and the planning of their implementation are centralised by law in the multi-professional research clinics of Helsinki University Central Hospital (HUS) and Tampere University Hospital (TAYS). The consultation, evaluation periods and treatments provided by the TAYS or HUS working group on the gender identity of minors shall be carried out in accordance with the following principles.

Children who have not started puberty and are experiencing persistent, severe anxiety related to gender conflict and/or identification as the other sex may be sent for a consultation visit to the research group on the gender identity of minors at TAYS or HUS. Any need for support beyond the consultation visit or need for other psychiatric treatment should be addressed by local services according to the nature and severity of the problem.

If a child is diagnosed prior to the onset of puberty with a persistent experience of identifying as the other sex and shows symptoms of gender-related anxiety, which increases in severity in puberty, the child can be guided at the onset of puberty to the research group on the gender identity of minors at TAYS or HUS for an assessment of the need for treatment to suppress puberty. Based on these assessments, puberty suppression treatment may be initiated on a case-by-case basis after careful consideration and appropriate diagnostic examinations if the medical indications for the treatment are present and there are no contraindications. Therapeutic amenorrhea, i.e. prevention of menstruation, is also medically possible.

A young person who has already undergone puberty can be sent to the research clinic on the gender identity of minors at TAYS or HUS for extensive gender identity studies if the variation in gender identity and related dysphoria do not reflect the temporary search for identity typical of the development stage of adolescence and do not subside once the young person has had the opportunity to reflect on their identity but rather their identity and personality development appear to be stable.

PO BOX 33 (Meritullinkatu 8, Helsinki)
00023 Government, Finland
www.palveluvalikoima.fi
Tel 0295 16001
e-mail:
palveluvalikoima@stm.fi

Summary of a recommendation by COHERE Finland 16.6.2020

Based on thorough, case-by-case consideration, the initiation of hormonal interventions that alter sex characteristics may be considered before the person is 18 years of age only if it can be ascertained that their identity as the other sex is of a permanent nature and causes severe dysphoria. In addition, it must be confirmed that the young person is able to understand the significance of irreversible treatments and the benefits and disadvantages associated with lifelong hormone therapy, and that no contraindications are present.

If a young person experiencing gender-related anxiety has experienced or is simultaneously experiencing psychiatric symptoms requiring specialised medical care, a gender identity assessment may be considered if the need for it continues after the other psychiatric symptoms have ceased and adolescent development is progressing normally. In this case, a young person can be sent by the specialised youth psychiatric care in their region for an extensive gender identity study by the TAYS or HUS research group on the gender identity of minors, which will begin the diagnostic studies. Based on the results of the studies, the need for and timeliness of medically justified treatments will be assessed individually.

Surgical treatments are not part of the treatment methods for dysphoria caused by gender-related conflicts in minors. The initiation and monitoring of hormonal treatments must be centralised at the research clinics on gender identity at HUS and TAYS.
Research data on the treatment of dysphoria due to gender identity conflicts in minors is limited. COHERE considers that, moving forward, multi-professional clinics specialising in the diagnostics and treatment of gender identity conflicts at HUS and TAYS should collect extensive information on the diagnostic process and the effects of different treatment methods on the mental wellbeing, social capacity and quality of life of children and youth. There is also a need for more information on the disadvantages of procedures and on people who regret them.

Link to the COHERE website: https://palveluvalikoima.fi/en/frontpage

The Council for Choices in Health Care in Finland (COHERE Finland) works in conjunction with the Ministry of Social Affairs and Health, and its task is to issue recommendations on services that should be included in the range of public health services. Further information: www.palveluvalikoima.fi.

PO BOX 33 (Meritullinkatu 8, Helsinki)
00023 Government, Finland
www.palveluvalikoima.fi
Tel 0295 16001
e-mail:
palveluvalikoima@stm.fi

EXHIBIT 26

STM038:00/2020

Recommendation of the Council for Choices in Health Care in Finland (PALKO / COHERE Finland)

Medical Treatment Methods for Dysphoria Related to Gender Variance In Minors

Concepts

Suppression treatment	Pubertal suppression with GnRH analogues (drugs that inhibit gonadotropin-releasing hormone activity) to halt the development of secondary sex characteristics of the biological sex.
Cisgender/Cis person	A person whose gender identity matches the sex determined at birth (identifies, and is satisfied with, the sex determined at birth and generally expresses his/her gender accordingly).
Other gender identity	A person who does not identify as a man or a woman, but rather somewhere along the continuum or outside of it; genderless, nonbinary, or multigendered.
Transgender	A person whose gender identity differs from the legal and biological sex determined at birth but instead aligns with the opposite sex.

STM038:00/2020

Content

1. Basis for Preparing These Recommendations

As the number of patients, including minors, referred to the Helsinki University Hospital (HUS) and the Tampere University Hospital (TAYS) multidisciplinary outpatient clinics for assessment and treatment of gender dysphoria has increased, PALKO (Council for Choices in Healthcare in Finland / COHERE Finland) decided to prepare recommendations for medical treatments of gender dysphoria, i.e., distress which is associated with a minor's gender variance and impairs function. Gender variance refers to a spectrum of gender experience anywhere on the male-female identity continuum or outside it, and is not exclusively confined to the dichotomized male/female conception of gender. Not all patients with gender variance experience significant suffering or functional impairments, and not all seek medical treatment.

These recommendations are based on the legislation in force at the time of the adoption of the recommendation, the available research evidence, and the clinical experience of multidisciplinary teams with expertise in gender dysphoria assessment and treatment at HUS and TAYS. The knowledge base supporting these recommendations is detailed in a separate Preparatory Memorandum and appendices and includes a description of planning and implementation of medical treatments, a literature review of medical treatments, an extensive ethical analysis, and feedback following meetings with patients and the advocacy groups who represent them.

Finnish legislation defines the requirements for the legal gender recognition of transsexuals (Act on Legal Recognition of the Gender of Transsexuals (Trans Act) 536/2002). The detailed requirements for providing the assessment and treatment to enable legal gender recognition are spelled out further in a Decree of the Ministry of Social Affairs and Health (1053/2002). The Trans Act and the related Decree apply to adults. For those who are not of legal age, there are no laws governing the provision and needs of transgender healthcare; however, these are subject to the Health Care Act of Finland (1326/2010), in particular section 7 (criteria for integrated care), section 7a (criteria for treatment options), section 8 (evidence-based, high quality, safe and appropriate care) and section 10 (rationale for centralization); and also to the Constitution of Finland (731/1999)'s section 6 on equality and section 19 on the right to adequate social and healthcare services. Finland's Act on the Status and Rights of Patients, (785/1992), and especially sections 5, 6, and 7, are also relevant.

2. Recommendations' Target Population

These recommendations apply to minors suffering from dysphoria related to gender variance who are seeking a consultation regarding an evaluation of medical examination and treatment needs; the children and adolescents may identify with the opposite sex (transgender), or may identify as genderless, non-binary, or anywhere along or outside the male/female gender identity continuum (other gender).

3. Procedures Assessed

These recommendations focus on medical treatment procedures that aim to decrease suffering and functional impairment of gender-dysphoric minors.

4. Current Care

Cross-sex identification in childhood, even in extreme cases, generally disappears during puberty. However, in some cases, it persists or even intensifies. Gender dysphoria may also emerge or intensify at the onset of puberty. There is considerable variation in the timing of the onset of puberty in both sexes. The first-line treatment for gender dysphoria is psychosocial support and, as necessary, psychotherapy and treatment of possible comorbid psychiatric disorders.

Consultation appointments (for parents / caregivers) regarding pre-pubescent children's cross-sex identification or gender dysphoria are provided by the research group on the gender identity of minors at TAYS or HUS. However, ongoing support or other treatment of psychiatric disorders are provided through the local municipal services.

In clear cases of pre-pubertal onset of gender dysphoria that intensified during puberty, a referral can be made for an assessment by the research group at TAYS or HUS regarding the appropriateness for puberty suppression. If no contraindications to early intervention are identified, pubertal suppression with GnRH analogues (to suppress the effect of gonadotropin-releasing hormone) may be considered to prevent further development of secondary sex characteristics of the biological sex.

Adolescents who have already undergone puberty, whose gender dysphoria occurs in the absence of co-occurring symptoms requiring psychiatric treatment, and whose experience of transgender identity failed to resolve following a period of reflection, can be referred for assessment by the research group on the gender identity of minors at TAYS or HUS. Hormone therapy (testosterone/estrogen and anti-androgen) can be started after the diagnostic evaluations, but no earlier than age 16. Additionally, patients under 18 receive three to six months of GnRH analogue treatment prior to the initiation of cross-sex hormones in order to suppress the hormonal activity of the gonads. No gender confirmation surgeries are performed on minors.

5. Risks, Benefits and Uncertainty

The literature review identified two studies with the total of 271 persons diagnosed with childhood-onset gender identity disorder and associated gender or body dysphoria that intensified after the onset of puberty (Preparatory Memorandum Appendix 1, Tables 15 and 16, pages 46-48).

In a smaller study of 70 adolescents, puberty was suppressed with the GnRH analogue at the average age of 14.8 (12-18 years) and puberty blockade continued for an average of 2 years. During the treatment period, the adolescents' mood improved, and the risk of behavioral disorders diminished, but gender dysphoria itself did not diminish, and there were no changes in body image. In a larger study consisting of 201 adolescents, 101 patients with the average age of 15.5 (12-18 years) started an 18-month psychological supportive intervention, and, additionally at six months, pubertal development was suppressed by starting GnRH analogue treatment. The other cohort of 100 only received psychological supportive intervention for 18 months. In both groups, statistically significant increases in global psychosocial functioning were found at 12 and 18 months; among those having received psychological intervention alone, the improvement in global functioning was already significant at the 6-month mark. Both studies lack long-term treatment follow-up into adulthood.

A recent Finnish study, published after the completion of this literature review, reported on the effect of initiating cross-sex hormone therapy on functioning, progression of developmental tasks of adolescence, and psychiatric symptoms. This study found that during cross-sex hormone therapy, problems in these areas did not decrease.

Potential risks of GnRH therapy include disruption in bone mineralization and the as yet unknown effects on the central nervous system. In trans girls, early pubertal suppression inhibits penile growth, requiring the use of alternative sources of tissue grafts for a potential future vaginoplasty. The effect of pubertal suppression and cross-sex hormones on fertility is not yet known.

6. Ethical Assessment

Although the ethics analysis did not systematically address the issues pertaining to children and adolescents, they have been discussed in several areas in the related documents (Preparatory Memorandum pages 52-62; Appendix 5).

According to the Health Care Act (section 8), healthcare services must be based on evidence and recognized treatment and operational practices. As far as minors are concerned, there are no medical treatment that can be considered evidence-based. At the same time, the numbers of minors developing gender dysphoria has increased. In this situation, it is vital to assure that children and young people are able to talk about their feelings, and that their feelings are acknowledged. The opportunity to reflect on one's experience should be easily accessible through the local health system (i.e., school or student health care, primary care). A young

person's feelings should not be interpreted as immediately requiring specialized medical examinations or treatments.

In cases of children and adolescents, ethical issues are concerned with the natural process of adolescent identity development, and the possibility that medical interventions may interfere with this process. It has been suggested that hormone therapy (e.g., pubertal suppression) alters the course of gender identity development; i.e., it may consolidate a gender identity that would have otherwise changed in some of the treated adolescents. The reliability of the existing studies with no control groups is highly uncertain, and because of this uncertainty, no decisions should be made that can permanently alter a still-maturing minor's mental and physical development.

From the point of view of patient advocacy groups, halting puberty is providing young people with a period of reflection, rather than consolidating their gender identity. This is based on the premise that halting the development of one's permanent sex characteristics will improve the minor's social interactions, while allowing more time for diagnostic evaluations. Additionally, patient advocacy groups assert that early intervention with hormonal treatments will lead to improved outcomes for the patients who do eventually pursue gender reassignment. Professionals, for their part, consider it important to ensure that irreversible interventions, which may also have significant adverse effects, both physical and mental, are only performed on individuals who are able to understand the permanence of the changes and the potential for harm, and who are unlikely to regret such interventions. It is not known how the hormonal suppression of puberty affects young people's judgement and decision-making.

The Act on the Status and Rights of Patients (1992/785) states that the patient shall be provided with information about his/her state of health, the significance of the treatment, various alternative forms of treatment and their effects, and about other factors concerning treatment that have an effect on treatment decision-making. In a situation where a minor's identification with the opposite sex causes long-term and severe dysphoria, it is important to make sure that he/she understands the realistic potential of gender reassignment treatments to alter secondary sex characteristics, the reality of a lifelong commitment to medical therapy, the permanence of the effects, and the possible physical and mental adverse effects of the treatments. Although patients may experience regret, after reassignment treatments, there is no going back to the non-reassigned body and its normal functions. Brain development continues until early adulthood – about age 25, which also affects young people's ability to assess the consequences of their decisions on their own future selves for rest of their lives.

A lack of recognition of comorbid psychiatric disorders common among gender-dysphoric adolescents can also be detrimental. Since reduction of psychiatric symptoms cannot be achieved with hormonal and surgical interventions, it is not a valid justification for gender reassignment. A young person's identity and personality development must be stable so that they can genuinely face and discuss their gender dysphoria, the significance of their own feelings, and the need for various treatment options.

For children and adolescents, these factors are key reasons for postponing any interventions until adulthood.

7. Conclusions

The first-line intervention for gender variance during childhood and adolescent years is psychosocial support and, as necessary, gender-explorative therapy and treatment for comorbid psychiatric disorders. Uncertainty related to gender identity should be dealt with according to the severity of symptoms and the need for treatment and should be handled at the school / student health care, primary health care at the local level, or in specialty care.

In adolescents, psychiatric disorders and developmental difficulties may predispose a young person to the onset of gender dysphoria. These young people should receive treatment for their mental and behavioral health issues, and their mental health must be stable prior to the determination of their gender identity.

Clinical experience reveals that autistic spectrum disorders (ASD) are overrepresented among adolescents suffering from gender dysphoria; even if such adolescents are presenting with gender dysphoria, rehabilitative interventions for ASD must be properly addressed.

In light of available evidence, gender reassignment of minors is an experimental practice. Based on studies examining gender identity in minors, hormonal interventions may be considered before reaching adulthood in those with firmly established transgender identities, but it must be done with a great deal of caution, and no irreversible treatment should be initiated. Information about the potential harms of hormone therapies is accumulating slowly and is not systematically reported. It is critical to obtain information on the benefits and risks of these treatments in rigorous research settings.

At a minimum, a consultation for a pre- pubescent child at the specialist setting at the TAYS includes an extensive assessment appointment costing EUR 369. If necessary, a day-long outpatient consultation can be arranged, costing EUR 1,408.

The consultation and assessment process for minors at the specialist settings of TAYS or HUS costs EUR 4,300. If it is determined that this process would be untimely, the minimum cost is EUR 640. An initial assessment / consultation by phone costs EUR 100.

The planning and monitoring costs for pubertal suppression are EUR 2,000 for the first year, and EUR 1,200 for subsequent years. The costs for the planning and monitoring of hormone treatments are a minimum of EUR 400 per year.

These costs do not take into account the additional costs of psychosocial support provided in the local level, the possible need for psychiatric treatment, or hormone treatment medication costs.

8. Summary of the Recommendations

PALKO / COHERE maintains the following:

1. For the treatment of gender dysphoria due to variations in gender identity in minors, psychosocial support should be provided in school and student healthcare and in primary healthcare, and there must be sufficient competency to provide such support.
2. Consultation with a child or youth psychiatrist and the necessary psychiatric treatment and psychotherapy should be arranged locally according to the level of treatment needed.
3. If a child or young person experiencing gender-related anxiety has other simultaneous psychiatric symptoms requiring specialised medical care, treatment according to the nature and severity of the disorder must be arranged within the services of their own region, as no conclusions can be drawn on the stability of gender identity during the period of disorder caused by a psychiatric illness with symptoms that hamper development.

PALKO / COHERE considers that the consultation, periods of assessment, and treatments by the research group on the gender identity of minors at TAYS or HUS must be carried out according to the following principles:

1. Children who have not started puberty and are experiencing persistent, severe anxiety related to gender conflict and/or identification as the other sex may be sent for a consultation visit to the research group on the gender identity of minors at TAYS or HUS. Any need for support beyond the consultation visit or need for other psychiatric treatment should be addressed by local services according to the nature and severity of the problem.

2. If a child is diagnosed prior to the onset of puberty with a persistent experience of identifying as the other sex and shows symptoms of gender-related anxiety, which increases in severity in puberty, the child can be guided at the onset of puberty to the research group on the gender identity of minors at TAYS or HUS for an assessment of the need for treatment to suppress puberty. Based on these assessments, puberty suppression treatment may be initiated on a case-by-case basis after careful consideration and appropriate diagnostic examinations if the medical indications for the treatment are present and there are no contraindications. Therapeutic amenorrhea, i.e. prevention of menstruation, is also medically possible.

3. A young person who has already undergone puberty can be sent to the research clinic on the gender identity of minors at TAYS or HUS for extensive gender identity studies if the variation in gender identity and related dysphoria do not reflect the temporary search for identity typical of the development stage of adolescence and do not subside once the young person has had the opportunity to reflect on their identity but rather their identity and personality development appear to be stable.

4. Based on thorough, case-by-case consideration, the initiation of hormonal interventions that alter sex characteristics may be considered before the person is 18 years of age only if it can be ascertained that their identity as the other sex is of a permanent nature and causes severe dysphoria. In addition, it must be confirmed that the young person is able to understand the significance of irreversible treatments and the

benefits and disadvantages associated with lifelong hormone therapy, and that no contraindications are present.

5. If a young person experiencing gender-related anxiety has experienced or is simultaneously experiencing psychiatric symptoms requiring specialized medical care, a gender identity assessment may be considered if the need for it continues after the other psychiatric symptoms have ceased and adolescent development is progressing normally. In this case, a young person can be sent by the specialized youth psychiatric care in their region for an extensive gender identity study by the TAYS or HUS research group on the gender identity of minors, which will begin the diagnostic studies. Based on the results of the studies, the need for and timeliness of medically justified treatments will be assessed individually.

Surgical treatments are not part of the treatment methods for dysphoria caused by gender-related conflicts in minors. The initiation and monitoring of hormonal treatments must be centralized at the research clinics on gender identity at HUS and TAYS.

9. Additional Evidence Gathering and Monitoring the Effectiveness of Recommendations

Moving forward, the following information must be obtained about the patients diagnosed and receiving treatments in Finland before re-evaluating these recommendations:

- Number of new patient referrals
- Number of patients starting the assessment period, and numbers of new transgender (F64.0) vs “other gender” (F64.8) diagnoses
- Whether the diagnosis remains stable or changes during the assessment phase
- Number of patients discontinuing the assessment period and the reasons for the discontinuation
- Adverse effects of treatments (especially long-term effects and effect on fertility)
- Number of patients regretting hormone therapy
- Analysis of the effects of the assessment and the treatment period on gender dysphoria outcomes, as measured by the Gender Congruence and Life Satisfaction Scale (GCLS)
- Analysis of the effects of the assessment and the treatment period on functional capacity and quality of life
- The prevalence of co-occurring psychiatric diagnoses (especially neurodevelopmental diagnoses F80-F90) among those diagnosed with / seeking treatment for gender dysphoria, and whether the presence of these co-occurring diagnoses impacts the ability to achieve the desired outcome (e.g. decreased dysphoria) in the assessment or the treatment phase.
- Whether the assessment and treatment periods lead to a reduction of suicide attempts
- Whether the assessment and treatment periods lead to a reduction in depression and distress

10. Appendices

Preparatory Memorandum, with Appendices 1-5.

EXHIBIT 27

Received: 20 January 2023 | Revised: 5 April 2023 | Accepted: 17 April 2023

DOI: 10.1111/apa.16791

ACTA PÆDIATRICA NURTURING THE CHILD WILEY

REVIEW ARTICLE

A systematic review of hormone treatment for children with gender dysphoria and recommendations for research

Jonas F. Ludvigsson[1,2,3] | Jan Adolfsson[4,5] | Malin Höistad[5] | Per-Anders Rydelius[6,†] | Berit Kriström[7] | Mikael Landén[1,8]

[1]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden

[2]Department of Paediatrics, Örebro University Hospital, Örebro, Sweden

[3]Division of Digestive and Liver Disease, Department of Medicine, Columbia University Medical Center, New York, New York, USA

[4]Department of Clinical Science, Intervention and Technology, Karolinska Institutet, Stockholm, Sweden

[5]The Swedish Agency for Health Technology Assessment and Assessment of Social Services, Stockholm, Sweden

[6]Department of Women's and Children's health, Karolinska Institutet, Stockholm, Sweden

[7]Department of Clinical Sciences/Paediatrics, Umeå University, Umeå, Sweden

[8]Department of Neuroscience and Physiology, Sahlgrenska Academy at the University of Gothenburg, Gothenburg, Sweden

Correspondence
Mikael Landén, Section of Psychiatry and Neurochemistry, Sahlgrenska University Hospital, Blå Stråket 15, S-431 45 Gothenburg, Sweden.
Email: mikael.landen@neuro.gu.se

Berit Kriström, Department of Clinical Sciences/Paediatrics, Umeå University, S-90185 Umeå, Sweden.
Email: berit.kristrom@umu.se

Funding information
Swedish Agency for Health Technology Assessment and Assessment of Social Services

Abstract

Aim: The aim of this systematic review was to assess the effects on psychosocial and mental health, cognition, body composition, and metabolic markers of hormone treatment in children with gender dysphoria.

Methods: Systematic review essentially follows PRISMA. We searched PubMed, EMBASE and thirteen other databases until 9 November 2021 for English-language studies of hormone therapy in children with gender dysphoria. Of 9934 potential studies identified with abstracts reviewed, 195 were assessed in full text, and 24 were relevant.

Results: In 21 studies, adolescents were given gonadotropin-releasing hormone analogues (GnRHa) treatment. In three studies, cross-sex hormone treatment (CSHT) was given without previous GnRHa treatment. No randomised controlled trials were identified. The few longitudinal observational studies were hampered by small numbers and high attrition rates. Hence, the long-term effects of hormone therapy on psychosocial health could not be evaluated. Concerning bone health, GnRHa treatment delays bone maturation and bone mineral density gain, which, however, was found to partially recover during CSHT when studied at age 22 years.

Abbreviations: BMD, bone mineral density; CSHT, cross-sex hormone treatment; DXA, dual-energy X-ray absorptiometry; GnRHa, gonadotropin-releasing hormone agonist (analogues); GRADE, grades of recommendation, assessment, development and evaluation; ICD, International Classification of Diseases; MRI, magnetic resonance imaging; SBU, Swedish Agency for Health Technology Assessment and Assessment of Social Services.

Berit Kriström and Mikael Landén have equal contrbution.

[†]Part of the original study group but deceased in December 2021.

Conclusion: Evidence to assess the effects of hormone treatment on the above fields in children with gender dysphoria is insufficient. To improve future research, we present the GENDHOR checklist, a checklist for studies in gender dysphoria.

KEYWORDS
adolescent, bone density, gender dysphoria, gonadotropin-releasing hormone agonist, psychosocial functioning

1 | INTRODUCTION

Gender incongruence refers to a mismatch between the biological sex and perceived gender identity. When gender incongruence causes significant discomfort, it is called gender dysphoria. When gender dysphoria causes clinically significant distress, the condition might meet the diagnostic criteria for transsexualism according to the (international classification of disease) ICD-10 guidelines,[1] or gender dysphoria according to the DSM-5.[2] Gender identity-affirming health care is provided to ease gender dysphoria.[3] The treatment aims to align bodily characteristics with the individual's gender identity, and usually includes cross-sex hormone treatment (CSHT), as well as chest and genital surgery.

In youth with gender dysphoria, gonadotropin-releasing hormone analogues (GnRHa) have been used to inhibit spontaneous puberty development. The rationale is to prevent irreversible bodily changes and give young individuals time to explore their gender identity. Following the first case report in which a GnRHa was used to suppress puberty in a female-to-male transsexual individual,[4] the '"Dutch protocol" was developed.[5] According to this protocol, young pubertal people presenting with gender dysphoria should first undergo a thorough psychological evaluation. If the diagnosis gender dysphoria is confirmed, GnRHa treatment is recommended to start during the early stages of puberty (Tanner stages 2–3). If gender dysphoria subsides, the individual may discontinue GnRHa treatment, at which point spontaneous puberty will restart. If gender dysphoria persists, CSHT might start at age 16 years and sex-reassignment surgery at 18 years. Gender dysphoria in youth was a rare phenomenon when the Dutch multidisciplinary protocol for the treatment of gender dysphoria was introduced. Seeking care for gender dysphoria has since become increasingly common in younger people in many parts of the western world,[6,7] with an exponential rise among children born female.[8] Although not all children with gender dysphoria receive gender identity affirming treatment, there has been an ensuing increase in hormones to treat children with gender dysphoria, of which data on the effects and side effects are limited. There is no previous systematic review or meta-analysis of hormone treatment for children with gender dysphoria.

This systematic review aimed at assessing (a) psychosocial effects, (b) effects on bone health, (c) effects on body composition and metabolism, and (d) satisfaction and therapy persistence in children aged <18 years with gender dysphoria undergoing hormone therapy.

Key Notes

- This systematic review assessed psychosocial effects, bone health, body composition and metabolism, and therapy persistence in children (<18 years of age) with gender dysphoria undergoing treatment with gonadotropin-releasing hormone analogues (GnRHa).
- Long-term effects of hormone therapy on psychosocial health are unknown. GnRHa treatment delays bone maturation and gain in bone mineral density.
- GnRHa treatment in children with gender dysphoria should be considered experimental treatment of individual cases rather than standard procedure.

In this review, trans women are referred to as male-to-female and trans men as female-to-male.

2 | METHODS

2.1 | Preregistration

This systematic review originated from a 2-year commissioned work from the governmental body the Swedish Agency for Health Technology Assessment and Assessment of Social Services (SBU). Ongoing SBU reviews are registered on the SBU website (https://www.sbu.se/en/ongoing-projects/) but not recorded in external databases.

2.2 | Selection criteria

The search was restricted to children aged <18 years with reported gender dysphoria. We included observational studies, randomised controlled trials, and systematic reviews according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.[9] Case reports, editorials, and non-human studies were excluded from further review. The search was limited to English-language publications.

2.3 | Search strategy

Two professional information specialists at the Swedish Agency for Health Technology Assessment and Assessment for Social Services (SBU) performed a comprehensive search of the following medical databases up until 9 November 2021: CINAHL (EBSCO), Cochrane Library (Wiley), EMBASE (Embase.com), PsycINFO (EBSCO), PubMed (NLM), Scopus (Elsevier), and SocINDEX (EBSCO). They also searched the Campbell Library, Epistemonikos, Evidence Search, International HTA database, as well as three NIHR Centre for Reviews and Dissemination (CRD) databases: Database of Abstracts of Reviews of Effects (DARE), Health, and Technology Assessment (HTA), and NHS Economic Evaluation Database (EED). Finally, we searched PROSPERO, an international prospective register for systematic reviews, to identify any relevant ongoing systematic reviews but found none. The search, selection, and assessment were conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines.[9] The search and selection processes are outlined in Figure 1. Only studies of low or moderate bias were eligible for this review. Full literature search strategy is provided at the SBU web page (https://www.sbu.se/contentassets/4062b596a35c4e1383405766b7365076/bilaga-1-litteratursokning.pdf).

2.4 | Relevance, risk of bias, and quality of evidence

Two independent experts checked all hits for relevance. Relevant studies (based on a pre-defined PICO) were then evaluated for risk of bias, also by two independent experts, according to ROBINS-I (Risk of bias in non-randomised studies of interventions).[10,11] Robins-I assesses possible bias in seven domains: confounding; bias due to selection, measurement classification of interventions, deviations from intended interventions, missing data, measurement of outcomes, and selection of the reported result.

If the two reviewers did not agree on content or quality, the paper was discussed in the larger research team of four experts (JFL, PR, BK, ML). Randomised controlled trials were planned to be assessed by RoB–2.[10,11] To rate the quality of evidence for specific outcomes, we used the Grades of Recommendation, Assessment, Development and Evaluation (GRADE) system.[12] GRADE has four levels of evidence (very low, low, moderate, high) and considers five domains that can decrease the level of certainty one or two levels (risk of bias, imprecision, inconsistency, indirectness (similar to 'external validity'), and publication bias).

2.5 | Data extraction

Two reviewers (MH, JA) retrieved data from the included studies. The data extracted included the outcomes mental and psychosocial health including suicidality, anthropometric measures and metabolism, bone health, adverse events, and the characteristics of each study including age at referral or intake, age at start of GnRHa treatment, age at start of CSHT, number of participants enrolled in study, number of transgender participants, number of hormone treated transgender participants, number of non-transgender participants, number of participants evaluated, treatment type (drugs, dosages, type of administration, treatment frequency), total treatment duration, and total follow-up time. The full data extraction of included studies is provided at the SBU web page (https://www.sbu.se/contentassets/4062b596a35c4e1383405766b7365076/bilaga-3-tabellverk-over-inkluderade-studier.pdf).

2.6 | Statistics

No statistical analyses were performed.

2.7 | Ethics

Ethical approval is not applicable for this systematic review.

3 | RESULTS

3.1 | Identified studies

After duplicate removal, the search yielded 9934 potential studies (Figure 1). Of these, 195 were selected for thorough reading. Of these, 36 were relevant and assessed for risk of bias. Twelve studies were excluded because of high risk for bias, leaving 24 studies with low to moderate, moderate, or moderate to high risk of bias reviewed in this paper. A list of excluded studies is provided at the SBU web page (https://www.sbu.se/contentassets/4062b596a35c4e1383405766b7365076/bilaga-2-exkluderade-studier-med-hog-risk-for-bias.pdf).

3.2 | Characteristics of the 24 studies

All 24 relevant studies had been published since 2014 (Table 1). Study participant age at the start of GnRHa therapy was typically between 11 and 15 years (range 9–18.6 years), with CSHT rarely being introduced before age 15. Except for the Hisle-Gorman et al.[6] ($n=3754$ participants) and Mullins et al.[13] ($n=611$) papers, few studies included >200 individuals. GnRHa treatment often continued for around 2 years, sometimes up to 4 years, and similar treatment durations were observed or reported for CSHT as observations were usually not reported after age 18 years. Full details of included studies are given at the SBU web page. Overall, there were eight studies on GnRH alone, 13 studies on GnRH + CSHT, and three studies on CSHT alone.

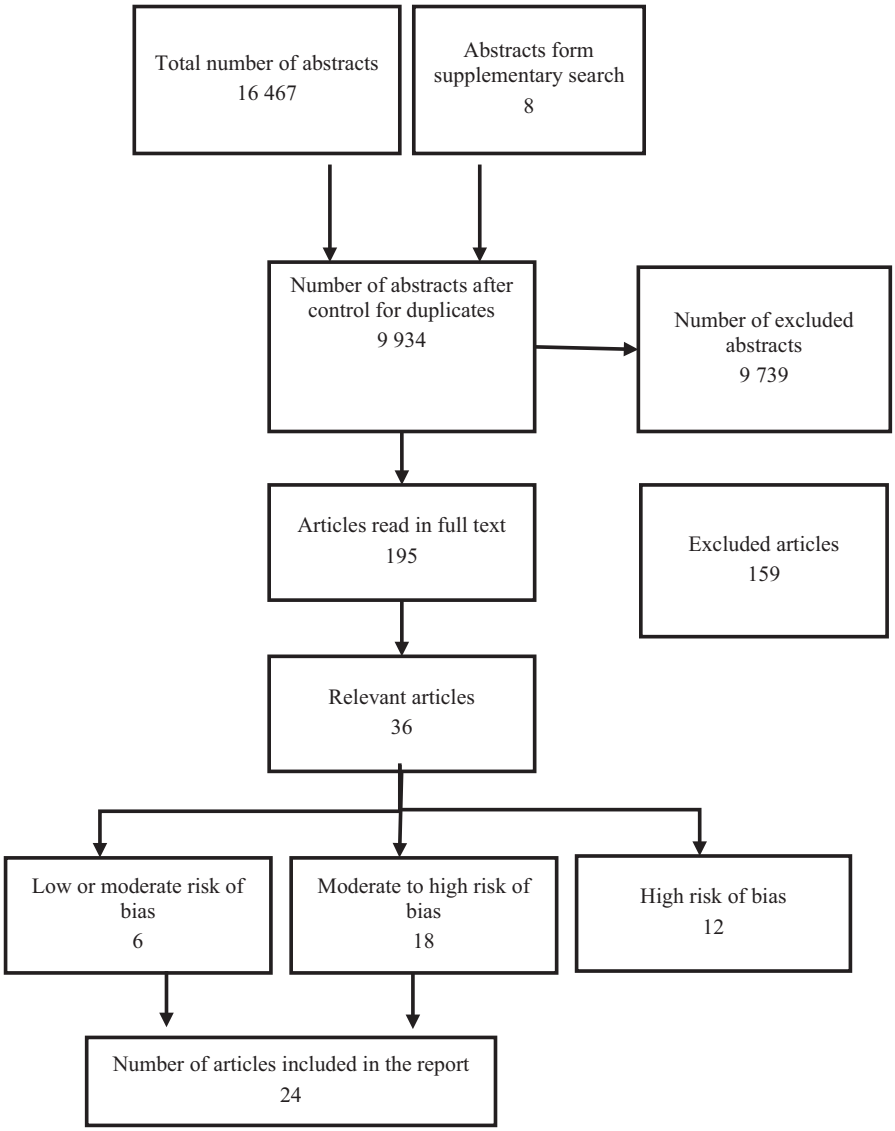


FIGURE 1 PRISMA flow diagram.

3.3 | Psychosocial and mental health

Table 2 outlines the six studies that examined psychosocial outcomes and cognitive effects.[14–19] Three of these studies found significantly improved overall psychosocial function after GnRHa treatment as measured by the Children's Global Assessment Scale (CGAS).[14–16] Two of these studies observed no statistically significant change in gender dysphoria.[15,16] Two of these studies reported significantly improved self-rated quality of life after treatment measured through Kidscreen-27, Short Form-8 (SF-8), Child Behaviour Checklist (CBCL) (parent report), and Youth Self Report (YSR),[16,17] while another study reported no statistically significant differences in anxiety and depression between those who started and not started hormone therapy.[18]

Because these studies were hampered by small number of participants and substantial risk of selection bias, the long-term effects of hormone treatment on psychosocial health could not be evaluated. Of note, the above studies do not allow separation of potential effects of psychological intervention independent of hormonal effects.

3.4 | Cognitive outcomes

We could only identify one study of low-moderate bias on cognitive outcomes in children with gender dysphoria receiving GnRHa therapy.[19] This cross-sectional study from the USA comprised 20 treated (8 male-to-female and 12 female-to-male) and 20 untreated (10 male-to-female and 10 female-to-male) young transgender persons and a control group ($n = 45$). Controls were identified from age-matched family members and friends. The Tower of London task was administered to assess executive functioning. The study neither found differences in cognitive function between treated and untreated transgender persons, nor between treated transgender persons and controls. However, because no before-after GnRHa therapy analyses were performed, the study

TABLE 1 Overview of 24 included studies.

Reference	Ages of patients (years)			Numbers of patients						Interventions			Time: duration and follow-up			Outcomes extracted
	Age at intake range (mean)	Age at start of GnRH range (mean)	Age at start of CSHT range (mean)	*n* referred	*n* TG enrolled	*n* TG HT	*n* TG non-HT	*n* non-TG	*n* TG HT at last FU	GnRH	CSHT	Surgery[b]	GnRH duration range (mean)	CSHT duration range (mean)	Follow-Up time range (mean)	Mental health Bone health Anthropometrics Metabolism
Mental health																
de Vries 2014[14]	11–17 (13.6)	11.5–18.5 (14.8)	13.9–19 (16.7)	196	111	55			32	x	x	x	1 year[a]	4 years[a]		UGDS, global functioning (CGAS), depression (BDI), anxiety (STAI), anger (TPI)
Costa 2015[15]	12–17 (15.5)	13–17 (16.5)		436	201	101	100		35	x			1 year		1.5 years	UGDS, psychosocial functioning (CGAS)
Becker-Hebly 2020[17]		11–17 (15.5)	13–17 (15.5)	434	75	54	21		54	x	x	x	0.5–4 years[a]	0.5–4 years[a]	7–49 months	Global functioning (CGAS), psychosocial functioning (YSR/ASR)
Cantu 2020[18]		11–xx (15)	xx–18 (15)		80	42	38		28	x	x		NR	NR	1–11 months (5 months)	Psychosocial functioning (PHQ-9, GAD-7), acute distress, suicidality
Carmichael 2021[16]		12.0–15.3 (13.6)			44	44			14	x			12–59 months (31 months)		12–36 months	UGDS, CGAS, psychological functioning (CBCL, YSR), Self-harm, BIS, HRQoL (Kidscreen52)
Hisle-Gorman 2021[6]	8–13 (10)		16.6–19.8 (18.2)		3754	963		6603	963	x	x		0.7–2.7 years (1.5)	0.7–2.7 years (1.5)	8.5 years	Mental health diagnosis, psychotropic medication use, medication days, service use
Staphorsius 2015[19]		min 12			41	20	20	45		x			0.6–2.6 years (1.6)			Psychological functioning (CBCL), cognitive function (executive function task)

Reference	Ages of patients (years)			Numbers of patients						Interventions			Time: duration and follow-up			Outcomes extracted
	Age at intake range (mean)	Age at start of GnRH range (mean)	Age at start of CSHT range (mean)	*n* referred	*n* TG enrolled	*n* TG HT	*n* TG non-HT	*n* non-TG	*n* TG HT at last FU	GnRH	CSHT	Surgery[b]	GnRH duration range (mean)	CSHT duration range (mean)	Follow-Up time range (mean)	Mental health Bone health Anthropometrics Metabolism
Bone health																
Joseph 2019[23]		12–14 (13)				70			70	x			1–xx years		up to 2.8 years	Height, weight, BMI BMD, BMAD, Z-score (hip, spine)
Klink 2015[21]		11.4–18.3 (15)	15.6–19 (16)			34			34	x	x	x	0.25–8 years	xx–8 years	up to age 22	Height, BMD, aBMD, Z-score, T-score (femoral neck, lumbar spine)
Vlot 2017[22]		11.5–18.6 (14)	14.0–19.5 (16)		215	70			57	x	x		1–xx years		up to 2 years	Height, BMAD, Z-score (hip, lumbar spine), bone markers (P1NP, OC, ICTP)
Schagen 2020[20]		12.2–16.5 (14)	15.0–17.9 (16)			127			121	x	x		1.5–4 years	3 years		aBMD, Z-score (hip)
Stoffers 2019[24]		11.8–18.0 (16)	14.9–18.4 (17.2)		64	62			15	x	x		3 months–3 years	5 months–3 years	2 years	Height, BP, BMD, Z-score (femoral neck, lumbar spine)
Navabi 2021[25]		13.4–17.4 (15)			198	172			116	x			6 months–2 years		1.5 years	BMD, aBMAD, Z-score (hip, lumbar spine)
van der Loos 2021[26]		11–17	15–17			322			322	x	x	x	1–3 years	2–6 years	up to 4 years	Subperiostal width, endocortical diameter
Lee 2020[27]		9.6–13.4 (11.5)			95	63			63	x			2 months			BMD, aBMAD, Z-score (hip, lumbar spine)
Anthropometrics and metabolism																
Schagen 2016[28]		11.1–18.6 (14)			138	116			77	x			3–12 months		1 year	Height, weight, BMI, lean body mass, liver enzymes, creatinine
Klaver 2018[31]		12.7–17.3[a] (15)	15.3–17.8[a] (16)	489	192	192			192	x	x	x	0.5–2.9 years (1.5[a])	1.6–3.4 years (2.9[a])	age 22	Weight, BMI, total body %, WHR

Reference	Ages of patients (years)			Numbers of patients						Interventions			Time: duration and follow-up			Outcomes extracted
	Age at intake range (mean)	Age at start of GnRH range (mean)	Age at start of CSHT range (mean)	*n* referred	*n* TG enrolled	*n* TG HT	*n* TG non-HT	*n* non-TG	*n* TG HT at last FU	GnRH	CSHT	Surgery[b]	GnRH duration range (mean)	CSHT duration range (mean)	Follow-Up time range (mean)	Mental health Bone health Anthropometrics Metabolism
Klaver 2020[32]		12.8–17.2[a] (14.9)	15.3–17.8[a] (16.6)		192	192			192	x	x	x	0.5–2.9 years (1.5)[a]	1.1–3.4 years (2.5[a])	age 22	BMI, SBP, DBP, glucose, insulin, HOMA-IR, cholesterol, triglycerides
Perl 2020[33]		13.4–15.4 (14)	14.2–16.0 (15)		48	15			15	x	x		2–4 months	2–6 months		BMI, BP
Schulmeister 2021[29]		9.0–14.5 (11.5)			92	55	226		55	x			10–14 months		1 year	Height velocity, BMI, z-score
Nokoff 2021[30]		10.2–14.1 (12)			17	17	31		17	x			0.5–5.8 years			Insulin, glucose HbA1c, HOMA-IR, body fat, % lean mass
Tack 2016[34]			NR (15–17)		45	43			43		x			6–18 months (12)	1.5 years	Height, weight, BMI, triglycerides, cholesterol, suicide, side effects
Jarin 2017[35]		103–xx	xx–25 (16–18)		116	116			116	(x)	x				2 years	BMI, BP, haematocrit, Hb, cholesterol
Mullins 2021[13]			13–24 (17)	1406	611	611			611		x			0.8–2.8 years (1.5 years)	3 years	Haematology, thrombosis, BMI

Note: Number of patients:*n* referred = number of patients referred to gender clinic for evaluation of gender dysphoria (not same at number of patients receiving GD diagnosis)*n* TG enrolled = number of patients enrolled in the study at start*n* TG = number of patients with gender dysphoria*n* TG HT = number of patients with gender dysphoria treated with hormones (GnRH alone, GnRH + CSHT, or CSHT only)*n* TG non-HT = number of patients with gender dysphoria treated NOT with hormones*n* TG HT at last FU = number of patients with gender dysphoria treated with hormones (GnRH alone, GnRH + CSHT, or CSHT only) evaluated at last follow-up time*n* non-TG = number of subjects in study without gender dysphoria (reference population).

Abbreviations: BDI, Beck Depression Inventory; BIS, Body Image Scale; BMAD, Bone Mineral Apparent Density; BMD, Bone Mineral Density; BMI, Body Mass Index; BP, Blood pressure; CBCL, Child Behaviour Checklist; CGAS, Global functioning Children's Global Assessment Scale, [higher scores (>80) indicating better global functioning]; CSHT, Cross-Sex Hormone Treatment/ gender-affirming treatment, testosterone, oestradiol, cyproterone acetate (CA), spironolactone, lynestrenol; GAD-7, Generalised Anxiety Disorder-7; GnRH, Gonadotropin Releasing Hormone analogue: triptorelin; HRQoL, Health Related Quality of Life; HT Hormone treatment, either GnRH, CSHT, or both; PHQ-9, Patient Health Questionnaire-9; SF-8, Short Form-8: (<18 years); STAI, Spielberger's Trait Anxiety; TG, Transgender; TPI, Anger Spielberger's Trait Anger; UGDS, Utrecht Gender Dysphoria Scale, score range 12–60 points [high score = high level of GD]; WHR, Waist-hip ratio; YSR, Youth Self Report: YSR (ages 11–18 years); Adult version (ASR, >18 years), [higher scores reflect higher degree of problems]; NR, not reported.

[a]Calculated by SBU.

[b]Surgery = any kind of gender reassignment surgery (gonadectomy, mastectomy, hysterectomy, laryngeal surgery, hair removal, phalloplasty, vaginoplasty).

TABLE 2 Summary of findings on psychosocial outcomes of puberty-blocking treatment (GnRHa) treatment in children with gender dysphoria.[14–19]

Outcome measures	Number of study participants, description of studies	Main result	"Certainty of evidence"	Deduction in GRADE[a]
Global function	*n* on hormones = 254 *n* evaluated = 113 Four observational cohort studies: one prospective and three retrospective studies[14–17]	Improved global function as assessed with the CGAS	Cannot be assessed	−2 risk of overall bias[b] −2 precision[c]
Suicide ideation	*n* on hormones = 42 *n* evaluated = 28 One prospective observational cohort study with mixed treatment (38 subjects with no pharmacological treatment)[18]	No change in suicide ideation	Cannot be assessed	−2 risk of overall bias[b] −2 precision[c]
Gender dysphoria	*n* on hormones = 145 *n* evaluated = 49 Two prospective observational cohort studies[15,16]	No change in gender dysphoria	Cannot be assessed	−2 risk of overall bias[b] −2 precision[c]
Depression	*n* on hormones = 97 *n* evaluated = 60 Two prospective observational cohort studies of which one included mixed treatment[14,18]	No change in depression	Cannot be assessed	−2 risk of overall bias[b] −2 precision[c]
Anxiety	n on hormones = 97 n evaluated = 60 Two prospective observational cohort studies[14,18]	No change in anxiety	Cannot be assessed	−2 risk of overall bias[a] −2 precision[b]
Cognition	*n* on hormones = 20 *n* evaluated = 20 One study[19]	No change in cognition compared with matched controls	Cannot be assessed	−2 risk of overall bias[b] −2 precision[c]
Quality of life	*n* on hormones = 98 *n* evaluated = 46 Two observational cohort studies, whereof one retrospective[16,17]	1. Improvement in quality of life most pronounced in subjects receiving puberty-blocking hormones, followed by gender-affirming hormone treatment[17] 2. Some improvement[16]	Cannot be assessed	−2 risk of overall bias[b] −2 precision[c]

Abbreviation: CGAS, Children's Global Assessment Scale.

[a]Starting at 4 for optimal studies in each study type.

[b]Selection of study participants is difficult to assess, analysis not based on stage in puberty development.

[c]Few study subjects in each study, heterogeneity in outcome and analyses.

could not investigate potential cognitive effects of hormone therapy.

3.5 | Bone health outcomes

Six longitudinal studies used dual-energy X-ray absorptiometry (DXA) scan technology to explore bone health before and again after some time with GnRHa treatment (Table 3). The second DXA scan usually coincided with CSHT initiation leading to different follow-up durations. The third DXA scan was performed after variable time with CSHT, performed with variable dosing and administration. The lumbar spine and hip were most often examined. One study investigated bone geometry.[20] Six studies were retrospective[21–26] and one study was prospective.[20] An additional study was cross-sectional where study participants in early puberty (Tanner stages 2–3) were examined only once, before the start of GnRHa therapy.[27]

Three studies reported a lower bone mineral density (BMD) in patients before or at start of GnRHa treatment compared with the general population of the same biological sex and age.[21,23,27] During GnRHa treatment, BMD estimated through area or volume, and expressed in z-scores increased less compared with general population reference values. However, the mean absolute BMD remained unchanged up to 2–3 years of GnRHa treatment.[20,23] The initiation of CSHT stimulated bone maturation and mineral accrual, increasing BMD.[21,22] After a median CSHT duration of 5.4 years in in female-to-male and 5.8 years in male-to-female, the lumbar spine mean areal BMD z-score was still significantly lower than at the start of GnRH therapy, while the other volume BMD and femoral neck estimates had normalised.[21] In another study, female-to-male receiving testosterone replacement therapy for 1–2 years had not regained their group mean BMD z-score registered at the start of GnRHa therapy.[24]

Bone geometry, estimated as subperiosteal width and endocortical diameter, was studied on DXA scans before start of GnRHa

TABLE 3 Summary of effects on bone development by puberty-blocking treatment (GnRHa) followed by CSHT in children with gender dysphoria.[20–25]

Outcome measures	Number of study participants, description of studies	Main Result	"Certainty of Evidence"	Deduction in GRADE[a]
Bone density during puberty-blocking hormonal treatment (g/cm^2, g/cm^3)	*n* on hormones = 363 *n* evaluated = 297 Five observational cohort studies (four retrospective and one prospective)[20–24]	Unchanged bone density (DXA measurement)	⊕⊕◯◯ Low certainty	–1 risk of overall bias[b] –1 precision
Bone density during puberty blocking hormonal treatment in relation to reference data in the literature (z-score)	*n* on hormones = 408 *n* evaluated = 292 Five observational cohort studies (four retrospective, and one prospective)[21–25]	Decreased increase in bone density over time	⊕⊕◯◯ Low certainty	–1 risk of overall bias[b] –1 precision
Bone density after 1–3 years (up to 22 years of age) of CSHT, which had been preceded by puberty-blocking hormonal treatment in relation to reference data in the literature	*n* on hormones = 268 *n* evaluated = 165 Three observational cohort studies (two retrospective and one prospective)[21,24,25]	After group median five years with CSHT, bone density recovered in hip but not in lumbar spine compared to data at start of treatment (z-score)	⊕⊕◯◯ Low certainty	–1 risk of overall bias[b] –1 precision

Abbreviations: CSHT, Cross-sex hormone treatment; DXA, Dual-Energy X-ray Absorptiometry.

[a]Starting at 4 for optimal studies in each study type.

[b]Analysis not based on stage in puberty development.

treatment and after at least two years on CSHT and compared with reference values of the general population: the bone geometry resembled the reference curve for the experienced sex only when GnRHa was started during early puberty. Bone geometry estimates in those who started GnRHa treatment during mid and late puberty remained within the reference curve of the biological sex.[26]

3.6 | Body composition and metabolic markers

GnRHa treatment effectively reduced endogenous sex hormone serum levels (Table 4). DXA scans after 1 year of GnRHa treatment revealed increased fat mass and reduced lean body mass.[28] Longitudinal growth depends on bone maturity (bone age) of those in the study group. Ongoing pubertal growth spurt will be arrested when GnRHa therapy is started, reducing the growth velocity to the prepubertal rate.[29]

Nokoff et al studied body composition and insulin sensitivity during 1 year of GnRHa therapy.[30] In addition to body composition, metabolic effects as insulin sensitivity during CSHT, and changes in blood pressure during testosterone therapy were examined.[31–33] Of these studies, three originated from Amsterdam.[29,32,33] The Amsterdam studies included observations during GnRHa therapy,[28] 1 year after starting CSHT,[32] as well as after a group median >5 years with CSHT in a cohort of 22-year-old adolescents.[31,33] The studies from Amsterdam were generally larger than the other studies. CSHT changed body composition towards the affirmed sex.[31,32] Obesity (defined as BMI >30 at age 22 years) was more prevalent in the transgender population[33] (Table 4).

3.7 | CSHT in children without prior GnRHa treatment

We were able to identify three studies of low-to-moderate bias examining CSHT in children without prior GnRHa treatment.[13,34,35] All were retrospective longitudinal studies. Because the number of study participants was small, studies were deemed to have low external validity, and because the studies examined different outcomes (e.g., lipid serum levels, Hb, blood pressure, metrorrhagia), it was not possible to draw any overall conclusions from these studies. Although the Mullins et al. paper[13] included several individuals at elevated risk of arterial or venous thrombosis, no cases of thrombosis were reported.

4 | DISCUSSION

We performed an extensive literature search to examine psychosocial and cognitive outcomes as well as metabolic and bone health in children with gender dysphoria taking hormone therapy. No randomised controlled trials were found, but we could identify 24 relevant observational studies. However, these were limited by

TABLE 4 Summary of findings of puberty-blocking (GnRHa) hormone treatment on anthropometric measures, body composition, and metabolism in children with gender dysphoria.[28–33]

Outcome measures	Number of study participants, description of studies	Main result	"Certainty of Evidence"	Deduction in GRADE[a]
Anthropometric measures	*n* on hormones = 192 *n* evaluated = 192 One retrospective observational cohort study[31]	Increased weight and body mass index	Cannot be assessed	−2 risk for overall bias[b] −1 precision[c] −1 indirectness[d]
Body composition	*n* on hormones = 325 *n* evaluated = 286 Two prospective observational cohort studies and one controlled cross-sectional study[28,30,31]	Decreased lean body mass	Cannot be assessed	−2 risk for overall bias[b] −1 precision[c] −1 indirectness[d]
Metabolic measures	*n* on hormones = 209 *n* evaluated = 209 One retrospective observational cohort study and one controlled cross-sectional study[30,32]	No change in serum lipids or blood pressure Increased insulin level in MtF Decreased insulin sensitivity	Cannot be assessed	−2 risk for overall bias[b] −1 precision[c] −1 indirectness[d]
Blood pressure	*n* on hormones = 15 *n* evaluated =15 One retrospective observational cohort study[33]	Change in blood pressure	Cannot be assessed	−2 risk for overall bias[b] −1 precision[c] −1 indirectness[d]
Growth (cm/year)	*n* on hormones = 55 *n* evaluated = 55 One prospective multicentre observational GnRHa treatment cohort study[29]	Reduced growth velocity	Cannot be assessed	−2 risk for overall bias[b] −1 precision[c] −1 indirectness[d]

[a]Starting at 4 for optimal studies in each study type.

[b]Selection of study participants is difficult to assess. Analysis not based on stage in puberty development.

[c]Few study subjects in each study, hence there is heterogeneity in outcome and analyses.

[d]Single study. In this context, 'indirectness' is similar to 'external validity'.

methodological weaknesses, for instance lack of or inappropriate control group, lack of intra-individual analyses, high attrition rates that precluded conclusion to be drawn. The exception being that children with gender dysphoria often had lower group mean values for BMD already prior to GnRHa treatment, and that GnRHa treatment delays the physiologically occurring BMD gain during pubertal sex hormone stimulation. However, this GnRHa-induced delay in BMD gain is almost fully compensated for by later ensuing CSHT. Although study participants were followed up to 22 years of age, the observed remaining deficit may depend on the limited study group size or on too short observation time.[21]

Our review highlights several specific knowledge gaps in gender dysphoria that are important to bridge not least given the recent increased incidence in many countries.[6,7] First, randomised controlled trials are lacking in gender dysphoria research. We call for such studies, which may be the only way to address biases that we have noted in the field. Given the current lack of evidence for hormonal therapy improving gender dysphoria, another ethically feasible option would be to randomise individuals to hormone therapy with all study participants, independent of intervention status, receiving psychological and psychosocial support. However, controlled trials do not necessarily require placebo treatment, but could for example build on the date or time of starting hormonal therapy to generate comparison groups. However, it should also be noted that this is a highly vulnerable population.

A second limitation concerns the statistical management of data. In the reviewed studies, observational data have frequently been analysed at a group level where intra-individual changes would have been more appropriate. Intra-individual analyses would allow for a better understanding of how subgroups of individuals respond (both positively and negatively) to hormone therapy. Group-level analyses are sensitive to selection bias because of high drop-out rates: The group studied at the end of the study is a selection of the group studied at baseline, which increases indirectness (reduces external validity). Moreover, it is important to analyse the distribution of individual data to be able to identify outliers who may be at risk for severe consequences of treatment.

Third, many studies only present data on chronological age but fail to account for puberty stage and biological age. This is a concern because the main purpose of GnRHa treatment is to suppress puberty and, with that, biological ageing.

Fourth, long-term studies are lacking. The duration of GnRHa treatment and CSHT was rarely >4 years. The absence of long-term studies is worrying because many individuals start treatment as minors (<18 years) and CSHT is lifelong. Fifth, individuals who stop GnRHa treatment before the start of CSHT need to be described and followed up. Sixth, some of the findings underlying this review are old, and studies reflecting the changing demographics of individuals seeking care for gender dysphoria are warranted.

TABLE 5 The GEnder Dysphoria HORmone treatment (GENDHOR) checklist.

	Recommendations
Aim	Describe the aim of the study
Study participants:	
Cases/exposed	Define gender dysphoria in your study, including the assessment tools used. Define eligibility criteria for your study (including chronological age, bone age or puberty stage, according to Tanner or Prader (when study concerns adolescents), biological sex, perceived gender identity, psychiatric and somatic comorbidities, medications at baseline). List exclusion criteria (diagnoses). List ages of participants at the start of each treatment (including absolute age ranges).
Comparators/ unexposed	Clarify how controls were selected (were controls recruited from the general population?) or whether national/ regional reference data (for instance, Z-scores) were used instead of individual controls.
Study design	Describe the study design: Cross-sectional, retrospective, prospective; case–control (and if nested), cohort study, randomised clinical trial.
Setting	Describe the setting of the study. Were study participants included at a tertiary centre or from the general population? Describe the catchment area/population of participating centres.
Intervention	Hormone treatment Describe whether GnRHa, anti-androgens, CSHT, or a combination was used. List generic names, mode of administration, and dosages of all treatments. Specify the treatment duration of each treatment. If hormone serum concentrations are studied, include the standard procedure for the timing of blood samples to hormone intake. If patients undergo surgery, clarify the type of surgery and number of participants undergoing each surgical procedure (gonadectomy, mastectomy, laryngeal surgery, vaginoplasty/phalloplasty, etc.). Clarify if any participant received psychiatric counselling before, or during the study, including total duration and frequency of counselling.
Variables	Define each variable (including co-variates) and its source. If possible, mention any effort to validate the variables.
Data measurement	Clarify who collected the data on study participants. Present time between first and second measurements if your study is longitudinal and includes "before-after" measurements in relation to the intervention. Mention if study participants had previously been included in other studies with a different aim or examining other outcomes.
Blinding	Describe if the data collectors were blinded to participant status/treatment or not.
Loss to follow-up	Indicate the number of participants discontinuing GnRHa/ CSHT and the reason(s) for discontinuation, including no longer wish to pursue gender reassignment treatment. Describe loss to follow-up/missing data
Statistical methods	Describe statistics according to a relevant checklist. Consider when applicable: Intra-individual changes (mean, SD, median, range) vs. between-group differences.
Descriptive data	In addition to usual demographic, clinical, social/socioeconomic information, report body mass index (BMI), smoking, use of oral contraceptives (type) or other hormonal treatment, puberty stage. Report any psychiatric illness at baseline, as well as the use of psychotropic medication. Describe other comorbidities, including disorders that could be considered contraindications for either hormone treatment or surgery. Specify follow-up time (median, mean) since the start of the intervention and since start of hormone treatment (define intervention start).
Outcome data	Specify main outcome of the study. Indicate all secondary outcomes, including adverse events.
Adverse events/ complications	Describe all adverse events.
Main results	Present absolute numbers. Calculate absolute and relative risks/Intraindividual effects/change and group mean/ median. Present incidence data. Describe any adjustment for potential confounders.
Limitations	Discuss limitations of your study, including limitations of the measurements used (e.g., DXA) and sources of potential bias or imprecision.
Generalisability/ external validity	Can data be generalised to individuals with gender dysphoria outside your study centre and the study country?
Conflict of interest	Report any conflict of interest.

Note: Based on our literature review, we created a *GEnder Dysphoria HORmone treatment checklist* (GENDHOR).

This list consists of recommendations that researchers may consider when planning a study of gender dysphoria, whether observational or interventional.

Abbreviations: CSHT, Cross-sex hormone treatment; DXA, Dual-Energy X-ray Absorptiometry; GnRHa, Gonadotropin-releasing hormone agonist (analogues).

Finally, we could not evaluate the frequency of individuals who drop out from GnRHa treatment and no longer wish to continue with gender transition. However, a follow up study was published after our literature search.[36] Of 720 children (31% born male and 69% born female) who started GnRHa treatment in adolescence, 98% continued to use hormone treatment into adulthood, which suggests that children generally continue with gender transition once they have started GnRHa treatment. We know from internet-based surveys that detransitioning exists,[37] but such studies cannot provide reliable estimates of detransitioning frequency because of selection bias. Studies that closely follow individuals who start GnRHa therapy and/or CSHT until at least age 30 are urgently needed. We also acknowledge there are other potential side effects from GnRHa therapy or CSHT that were not included in our review such as alopecia and abscesses from injections.[38]

Due to limitations in reporting of data, previous published studies in this field repeatedly contain insufficient details on drug administration and dosages, treatment duration, and the type of surgery performed. Some of these limitations will be partly remedied by the introduction of the new ICD version 11, and the Utrecht criteria,[39] but the field also urgently needs high quality longitudinal studies that not only assess medical outcomes but also those outcomes that matter most for affected individuals. Building on the identified limitations in previous research, we compiled a checklist to improve gender dysphoria research ("GENDHOR", Table 5). The aim of this checklist is not to replace existing research guidelines, but using it together with existing guidelines might support researchers and peer reviewers, and ultimately benefit patients and their families.

Last, there have been studies in this field published after the date of our literature search (9 November 2021). These have not been added to this study in order to not depart from the systematic approach. We nevertheless wish to comment on some of the publications. First, the National Institute for Health and Care Excellence in England (NICE) conducted evidence reviews of GnRHa[40] as well as CSHT[41] for children with gender dysphoria, which were independent from our work. The conclusions generally align with our findings. Second, Chien et al.[42] recently published a prospective study of psychosocial functioning during 2 years after initiation of CSHT in youths (12–20 years of age) with gender dysphoria. Of 315 participants, 162 completed that study. Life satisfaction increased, and depression and anxiety scores decreased, among biological females but not biological males. The strongest finding was a moderately improved appearance congruence. No information on concomitant psychological or psychopharmacological therapy was provided.

5 | CONCLUSION

This systematic review of almost 10 000 screened abstracts suggests that long-term effects of hormone therapy on psychosocial and somatic health are unknown, except that GnRHa treatment seems to delay bone maturation and gain in bone mineral density.

AUTHOR CONTRIBUTIONS

Study concept and design: All authors. Acquisition of data: Malin Höistad, Jan Adolfsson. Drafting of the manuscript: All authors. Interpretation of data and critical revision of the manuscript for important intellectual content: All authors. Administrative, technical, or material support: Jan Adolfsson, Malin Höistad. Funding acquisition: the Swedish agency for technology assessment and assessment for social services.

ACKNOWLEDGEMENTS

Information specialists Klas Moberg and Hanna Olofsson designed and performed the literature search.

FUNDING INFORMATION

This study received funding from the Swedish Agency for Health Technology Assessment and Assessment of Social Services.

CONFLICT OF INTEREST STATEMENT

JFL coordinated an unrelated study on behalf of the Swedish inflammatory bowel disease quality register (SWIBREG) that received funding from the Janssen Corporation. JFL has also received financial support from Merck Sharp & Dohme developing a paper reviewing national healthcare registers in China. JFL is currently discussing potential research collaboration with Takeda. ML has received lecture honoraria for Lundbeck pharmaceuticals and served as consultant for AstraZeneca. The other authors report no conflict of interest.

The data collection of this study was funded by the Swedish agency for technology assessment and assessment for social services. JA and MH are employees at this agency.

ORCID

Jonas F. Ludvigsson https://orcid.org/0000-0003-1024-5602
Jan Adolfsson https://orcid.org/0000-0003-2992-1869
Malin Höistad https://orcid.org/0000-0002-2312-3885
Per-Anders Rydelius https://orcid.org/0000-0002-1923-0282
Berit Kriström https://orcid.org/0000-0002-5456-2514
Mikael Landén https://orcid.org/0000-0002-4496-6451

REFERENCES

1. World Health Organization. The ICD-10 Classification of Mental and Behavioural Disorders. Diagnostic Criteria for Research. WHO; 1993.
2. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 5th ed. American Psychiatric Association; 2013.
3. 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 Transgend. 2012;13(4):165-232. doi:10.1080/15532739.2011.700873
4. Cohen-Kettenis PT, van Goozen SH. Pubertal delay as an aid in diagnosis and treatment of a transsexual adolescent. Eur Child Adolesc Psychiatry. 1998;7(4):246-248. doi:10.1007/s007870050073
5. de Vries AL, Cohen-Kettenis PT. Clinical management of gender dysphoria in children and adolescents: the Dutch approach. J Homosex. 2012;59(3):301-320. doi:10.1080/00918369.2012.653300

6. Hisle-Gorman E, Schvey NA, Adirim TA, et al. Mental healthcare utilization of transgender youth before and after affirming treatment. J Sex Med. 2021;18(8):1444-1454. doi:10.1016/j.jsxm.2021.05.014
7. Landén M. Dramatic increase in adolescent gender dysphoria requires careful consideration. Lakartidningen. 2019;116:FSMH.
8. Thompson L, Sarovic D, Wilson P, Sämfjord A, Gillberg C. A PRISMA systematic review of adolescent gender dysphoria literature: (1) Epidemiology. PLOS Glob Public Health. 2022;2(3):e0000245. doi:10.1371/journal.pgph.0000245
9. Moher D, Shamseer L, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev. 2015;4:1. doi:10.1186/2046-4053-4-1
10. Sterne JA, Hernan MA, Reeves BC, 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
11. SBU. Utvärdering av metoder i hälso- och sjukvården och insatser i socialtjänsten: en metodbok. Stockholm. Statens beredning för medicinsk och social utvärdering (SBU); 2020. Available from: https://www.sbu.se/metodbok
12. Guyatt GH, Oxman AD, Schunemann HJ, et al. GRADE guidelines: a new series of articles in the journal of clinical epidemiology. J Clin Epidemiol. 2011;64(4):380-382. doi:10.1016/j.jclinepi.2010.09.011
13. Mullins ES, Geer R, Metcalf M, et al. Thrombosis risk in transgender adolescents receiving gender-affirming hormone therapy. Pediatrics. 2021;147(4):e2020023549. doi:10.1542/peds.2020-023549
14. de Vries AL, McGuire JK, Steensma TD, et al. Young adult psychological outcome after puberty suppression and gender reassignment. Pediatrics. 2014;134(4):696-704. doi:10.1542/peds.2013-2958
15. 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(11):2206-2214. doi:10.1111/jsm.13034
16. Carmichael P, Butler G, Masic U, 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(2):e0243894. doi:10.1371/journal.pone.0243894
17. 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(11):1755-1767. doi:10.1007/s00787-020-01640-2
18. Cantu AL, Moyer DN, Connelly KJ, Holley AL. Changes in anxiety and depression from intake to first follow-up among transgender youth in a pediatric endocrinology clinic. Transgend Health. 2020;5(3):196-200. doi:10.1089/trgh.2019.0077
19. Staphorsius AS, Kreukels BP, Cohen-Kettenis PT, et al. Puberty suppression and executive functioning: an fMRI-study in adolescents with gender dysphoria. Psychoneuroendocrinology. 2015;56:190-199. doi:10.1016/j.psyneuen.2015.03.007
20. 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. doi:10.1210/clinem/dgaa604
21. 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(2):E270-E275. doi:10.1210/jc.2014-2439
22. 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;95:11-19. doi:10.1016/j.bone.2016.11.008
23. 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(10):1077-1081. doi:10.1515/jpem-2019-0046
24. 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):1459-1468. doi:10.1016/j.jsxm.2019.06.014
25. Navabi B, Tang K, Khatchadourian K, Lawson ML. Pubertal suppression, bone mass, and body composition in youth with gender dysphoria. Pediatrics. 2021;148(4):e2020039339. doi:10.1542/peds.2020-039339
26. van der Loos MA, Hellinga I, Vlot MC, et al. 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):931-941. doi:10.1002/jbmr.4262
27. Lee JY, Finlayson C, Olson-Kennedy J, 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. doi:10.1210/jendso/bvaa065
28. Schagen SE, Cohen-Kettenis PT, Delemarre-van de Waal HA, et al. Efficacy and safety of gonadotropin-releasing hormone agonist treatment to suppress puberty in gender dysphoric adolescents. J Sex Med. 2016;13(7):1125-1132. doi:10.1016/j.jsxm.2016.05.004
29. Schulmeister C, Millington K, Kaufman M, et al. Growth in transgender/gender-diverse youth in the first year of treatment with gonadotropin-releasing hormone agonists. J Adolesc Health. 2022;70(1):108-113. doi:10.1016/j.jadohealth.2021.06.022
30. Nokoff NJ, Scarbro SL, Moreau KL, et al. Body composition and markers of Cardiometabolic health in transgender youth compared with cisgender youth. J Clin Endocrinol Metab. 2020;105(3):e704-e714. doi:10.1210/clinem/dgz029
31. Klaver M, de Mutsert R, Wiepjes CM, et al. Early hormonal treatment affects body composition and body shape in young transgender adolescents. J Sex Med. 2018;15(2):251-260. doi:10.1016/j.jsxm.2017.12.009
32. Klaver M, de Mutsert R, van der Loos M, et al. Hormonal treatment and cardiovascular risk profile in transgender adolescents. Pediatrics. 2020;145(3):e20190741. doi:10.1542/peds.2019-0741
33. Perl L, Segev-Becker A, Israeli G, Elkon-Tamir E, Oren A. Blood pressure dynamics after pubertal suppression with gonadotropin-releasing hormone analogs followed by testosterone treatment in transgender male adolescents: a pilot study. LGBT Health. 2020;7(6):340-344. doi:10.1089/lgbt.2020.0026
34. Tack LJ, Craen M, Dhondt K, et al. Consecutive lynestrenol and cross-sex hormone treatment in biological female adolescents with gender dysphoria: a retrospective analysis. Biol Sex Differ. 2016;7:14. doi:10.1186/s13293-016-0067-9
35. Jarin J, Pine-Twaddell E, Trotman G, et al. Cross-sex hormones and metabolic parameters in adolescents with gender dysphoria. Pediatrics. 2017;139(5):e20163173. doi:10.1542/peds.2016-3173
36. van der Loos M, Hannema SE, Klink DT, et al. Continuation of gender-affirming hormones in transgender people starting puberty suppression in adolescence: a cohort study in The Netherlands. Lancet Child Adolesc Health. 2022;6:869-875. doi:10.1016/S2352-4642(22)00254-1
37. 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):3353-3369. doi:10.1007/s10508-021-02163-w

38. Khatchadourian K, Amed S, Metzger DL. Clinical management of youth with gender dysphoria in Vancouver. J Pediatr. 2014;164(4):906-911. doi:10.1016/j.jpeds.2013.10.068
39. McGuire JK, Berg D, Catalpa JM, et al. Utrecht gender dysphoria scale - gender Spectrum (UGDS-GS): construct validity among transgender, nonbinary, and LGBQ samples. Int J Transgend Health. 2020;21(2):194-208. doi:10.1080/26895269.2020.1723460
40. Evidence review: gonadotrophin releasing hormone analogues for children and adolescents with gender dysphoria (NICE. National Institute for Health and Care Excellence). 2020 https://cassindependent-reviewuk/nice-evidence-reviews/
41. Evidence review: gender-affirming hormones for children and adolescents with gender dysphoria (NICE. National Institute for Health and Care Excellence). 2020 https://cass.independent-review.uk/nice-evidence-reviews/
42. Chen D, Berona J, Chan YM, et al. Psychosocial functioning in transgender youth after 2 years of hormones. N Engl J Med. 2023;388(3):240-250. doi:10.1056/NEJMoa2206297

How to cite this article: Ludvigsson JF, Adolfsson J, Höistad M, Rydelius P-A, Kriström B, Landén M. A systematic review of hormone treatment for children with gender dysphoria and recommendations for research. Acta Paediatr. 2023;00:1–14. https://doi.org/10.1111/apa.16791