Classification: Official

NHS England Evidence Review:

Masculinising medicines comprising testosterone monotherapy for children and young people with gender incongruence who identify as a male gender and wish a binary physical transition

NHS England URN: 2417j

Publication reference: PRN02421_ix

NHS England Evidence Review

Masculinising medicines comprising testosterone monotherapy for children and young people with gender incongruence who identify as a male gender and wish a binary physical transition

Completed: January 2026

Prepared by Solutions for Public Health (SPH) on behalf of NHS England Specialised Commissioning

Contents

1. Introduction

This evidence review examines the clinical effectiveness, safety, and cost-effectiveness of masculinising medicines comprising testosterone monotherapy with or without psychological and psychosocial support compared with one or a combination of psychological support or social transitioning to the desired gender or no intervention, for children and young people (CYP) with gender incongruence who identify as male and wish binary physical transition.

The International Classification of Diseases (ICD)-11 (WHO, 2025) splits gender incongruence into that identified in childhood and that identified in adolescents and adults. 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. The incongruence must have persisted for about two years. Gender incongruence of adolescence and adulthood is 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. The diagnosis cannot be assigned prior to the onset of puberty. Gender variant behaviour and preferences alone are not a basis for assigning the diagnosis.

Although the diagnosis of gender incongruence includes both adolescence and adulthood, this evidence review refers specifically to CYP up to their 18th birthday.

Treatment for gender incongruence aims to help people live the way they want to, in their preferred gender identity, whilst aiming to improve mental health and quality of life outcomes. Masculinising medicines are used to help treat gender incongruence and make the patient's body more congruent with their gender identity. Treatment includes testosterone which will result in the patient's body developing a more male physical appearance. These treatments will be used in combination with a number of other interventions. This evidence review focusses on individuals that use testosterone monotherapy.

Studies in which gonadotropin-releasing hormone (GnRH) analogues are used in the context of puberty suppression or used as puberty suppressing hormones are outside of the scope of this evidence review. NHS England and the National Institute of Health and Care Research (NIHR) are working together to set up a study into the potential benefits and harms of puberty suppressing hormones as a treatment option for CYP with gender incongruence.

In addition, the review scope included the identification of possible subgroups of CYP within the included studies who might benefit from treatment with testosterone monotherapy more than the wider population, the criteria used by the research studies to define gender

incongruence, the starting criteria, formulation, duration and dose of testosterone monotherapy for those aged 16 years up to their 18th birthday, the circumstances under which any children aged 15 years or younger received testosterone monotherapy, monitoring arrangements and study exclusion criteria.

2. Executive summary of the review

This evidence review examines the clinical effectiveness, safety, and cost-effectiveness of masculinising medicines comprising testosterone monotherapy compared with one or a combination of psychological support or social transitioning to the desired gender, or no intervention in CYP with gender incongruence who identify as male gender and wish a binary physical transition.

The terminology in this topic area is continually evolving and is different depending on stakeholder perspectives. In this evidence review we have used the phrase 'CYP who identify as male gender and wish a binary physical transition' rather than saying natal or biological sex and 'cross-sex hormones' are now referred to as 'masculinising medicines'. The studies referenced in this review may use historical terms which are no longer considered appropriate (Table 1, Appendix E: Evidence Table, Appendix G: GRADE profiles).

The searches for evidence published since 01 January 2005 were conducted on 04 June 2025 and identified 1,233 references. The titles and abstracts were screened, and 102 full text papers were obtained and assessed for relevance against the criteria defined in the Population, Intervention, Comparator, Outcome (PICO) for this review.

Eleven studies were identified for inclusion in this review (Allen et al 2019, Baines et al 2023, Grannis et al 2023, Grimstad et al 2021, Kramer et al 2024, Laurenzano et al 2021, Millington et al 2024, Moussaoui et al 2024, Mullins et al 2021, Persky et al 2024, Valentine et al 2022). Of the 11 studies, three provided comparator evidence in cross-sectional studies (Grannis et al 2023, Kramer et al 2024, Valentine et al 2022). The remaining eight studies (Allen et al 2019, Baines et al 2023, Grimstad et al 2021, Laurenzano et al 2021, Millington et al 2024, Moussaoui et al 2024, Mullins et al 2021, Persky et al 2024) did not include an in-scope comparator group. The number of individuals in the studies ranged from 26 in Baines et al (2023) to 2,766 in Valentine et al (2022). Four studies reported mean and median duration of treatment (Allen et al 2019, median 349 days; Grannis et al 2023, mean 12.87 months, Grimstad et al 2021, mean 28.5 and 37.3 months, Mullins et al 2021, median 577 days).

Evidence was available for all the critical outcomes of interest. One cross-sectional study (Grannis et al 2023) reported the impact on gender incongruence, two cross-sectional studies (Grannis et al 2023, Kramer et al 2024) and one retrospective case series (Allen et al 2019) provided evidence for impact on mental health outcomes (including eating disorder behaviours and suicidality). Allen et al (2019) and one prospective case series (Baines et al

2023) reported on wellbeing and quality of life (QoL) in individuals who received testosterone monotherapy.

For the important outcomes of interest, three case series provided evidence on perceived physical changes in individuals who received testosterone monotherapy. Baines et al (2023) reported perceived masculinising changes (including skin oiliness/acne, facial hair, body hair, increased muscle mass/strength, menstrual cessation, and clitoral enlargement). Grimstad et al (2021) reported on the number of individuals experiencing breakthrough bleeding whilst taking testosterone monotherapy, and Persky et al (2024) reported on final adult height and final adult height Z-scores. One retrospective case series (Laurenzano et al 2021) reported discontinuation of treatment. No evidence was identified for the important outcomes 'psychosocial impact',' fertility', 'feasibility of masculinising genital surgery', 'cognitive outcomes', or 'regret after receipt of masculinising medicines'.

Safety was reported in six studies (one comparator study and five non-comparator studies), across six major themes: cardiometabolic-related diagnoses such as hypertension (Valentine et al 2022, Laurenzano et al 2021); venous thromboembolism (VTE) or arterial thrombosis (Mullins et al 2021); change in laboratory measures (Millington et al 2024, Laurenzano et al 2021); adverse events (Baines et al 2023); pelvic pain, abdomino-pelvic pain or pain in the lower part of the abdomen (Moussaoui et al 2024); and other safety concerns, including injection site reactions and progression of acne (Laurenzano et al 2021).

No evidence was identified for cost-effectiveness or any subgroups of interest.

Most of the evidence was from children and young people with gender incongruence attending gender incongruence clinics at paediatric hospitals in the United States of America (10 studies), with the remaining study was conducted in children and young people with gender incongruence attending a single paediatric gender clinic in Australia.

The included studies all provided very low certainty evidence for all the outcomes reported when assessed using modified GRADE.

In terms of clinical effectiveness:

Critical outcomes

Impact on gender incongruence

- *Testosterone monotherapy vs no hormones*

One cross-sectional study reported that dissatisfaction with body image[1] was *statistically significantly lower* in CYP who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy compared to individuals not receiving testosterone: mean 92.29 vs mean 103.14, respectively. The mean duration of treatment with testosterone monotherapy was 12.87 months.

- *Testosterone monotherapy (no comparator)*
  No evidence was identified for this outcome.

Impact on mental health

- *Testosterone monotherapy vs no hormones*
  One cross-sectional study reported *statistically significantly lower* general anxiety symptoms[2] (mean 38.75 vs 50.25), depressive symptoms[3] (mean 13.38 vs 18.34), and social anxiety[4] (mean 50.21 vs 78.62) in CYP who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy compared to individuals not receiving testosterone, respectively. The same study reported *no statistically significant difference* in suicidality[5] between the two groups. The mean treatment duration was 12.87 months. The second cross-sectional study reported *no statistically significant difference* in eating disorder behaviours (subjective binge eating, self-induced vomiting, laxative use and compensatory exercise) between CYP who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy compared to individuals not receiving testosterone. The treatment duration/follow-up was not reported.

- *Testosterone monotherapy (no comparator)*
  One retrospective case series reported a reduction in suicidal ideation[6] from baseline to at least three months follow-up (mean 1.01 vs mean 0.29, respectively) in CYP who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy, but no statistical measures were reported.

Impact on Quality of Life (QoL)

- *Testosterone monotherapy vs no hormones*
  No evidence was identified for this outcome.

[1] Higher values indicate greater dissatisfaction with one's body image.
[2] Higher values indicate greater generalised anxiety symptoms.
[3] A higher score indicates a higher level of depressive symptoms.
[4] Higher values indicate greater social anxiety.
[5] Suicidality was measured by counting the frequency of suicidal ideation and/or attempts in the past year.
[6] Higher scores indicate greater levels of suicidal ideation.

- *Testosterone monotherapy (no comparator)*
  One prospective case series and one retrospective case series provided evidence relating to the impact on QoL in CYP who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy. The prospective case series reported QoL in CYP who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy at baseline (median 66.08) and at three months (median 64.31), six months (median 61.41), and nine months follow-up (median 59.78). The difference was *not statistically significant* at six months, and no statistical measures were reported at three and nine months follow-up. The retrospective case series reported higher levels of wellbeing in CYP who identify as a male gender and wish a binary physical transition following at least three months of testosterone monotherapy (mean 70.94) compared to baseline (mean 64.95), but no statistical measures were reported.

Important outcomes

Masculinising physical changes

- *Testosterone monotherapy vs no hormones*
  No evidence was identified for this outcome.

- *Testosterone monotherapy (no comparator)*
  One prospective case series reported a *statistically significant increase* in perceived physical changes in CYP who identify as a male gender and wish a binary physical transition and receiving testosterone monotherapy from three to six months follow-up (median 16.0 vs 19.5, respectively). The same study reported perceived physical changes at nine months follow-up, but statistical measures were not reported. One retrospective case series reported breakthrough bleeding in 58 (25%) CYP who identify as a male gender and wish a binary physical transition and receiving testosterone monotherapy compared to 174 (75%) individuals reporting no breakthrough bleeding whilst taking testosterone monotherapy, but no statistical measures were reported. The mean treatment duration was 28.5 months and 37.3 months, respectively. The remaining retrospective case series indicated *no statistically significant difference* in final adult height or final adult height Z-scores in CYP who identify as a male gender and wish a binary physical transition and receiving testosterone monotherapy at an unknown treatment duration/follow-up.

Psychosocial impact

- No evidence was identified for psychosocial impact.

Fertility

- No evidence was identified for fertility.

Feasibility of masculinising genital surgery

- No evidence was identified for masculinising genital surgery.

Cognitive outcomes

- No evidence was identified for cognitive outcomes.

Detransition after receipt of masculinising medicines

- *Testosterone monotherapy vs no hormones*
  No evidence was identified for this outcome.

- *Testosterone monotherapy (no comparator)*
  One retrospective case series reported that of the 119 CYP who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy, three discontinued treatment with testosterone monotherapy: *"two stopped because they were satisfied with effects of SC-T* [subcutaneous testosterone] *achieved and one stopped within 6 months of starting after reassessing their gender identity…. due to concerns about body changes and potential impact on fertility"*. The median follow-up duration was 1.9 years (range 6 months to 5.5 years). No statistical measures were reported.

Regret after receipt of masculinising medicines

- No evidence was identified for regret after receipt of masculinising medicines.

In terms of safety:

Frequency of adverse events

- *Testosterone monotherapy vs no hormones*
  One cross-sectional study reported that CYP who identify as a male gender and wish a binary physical transition and receiving testosterone monotherapy reported *statistically significantly higher* odds of overweight/obesity (odds ratio [OR] 1.8, 95% confidence interval [CI] 1.5 to 2.1), dyslipidaemia (OR 1.7, 95% CI 1.3 to 2.3), liver dysfunction (OR 1.5, 95% CI 1.1 to 1.9), and hypertension (OR 1.6, 95% CI 1.2 to 2.2) compared to individuals not receiving testosterone, but these were *not statistically significant* after adjusting for confounders. The same study reported *no statistically significant difference* in the odds of dysglycaemia or polycystic ovary

syndrome (PCOS) between individuals receiving vs not receiving testosterone. The duration of treatment/follow-up was not reported.

- *Testosterone monotherapy (no comparator)*
  One prospective case series reported that three (11.5%) CYP who identify as a male gender and wish a binary physical transition experienced skin irritation at initiation of testosterone monotherapy. The same study reported that two participants experienced elevated Alanine transaminase (ALT) or Aspartate aminotransferase (AST) levels at three or six months follow-up, but no participants developed elevated haematocrit, or significant dyslipidaemia at nine months follow-up. No statistical measures were reported for any safety outcome. One retrospective case series reported no VTE or arterial thrombosis events (including stroke) at a median follow-up of 577 days. One retrospective case series reported a *statistically significant increase* in haemoglobin and haematocrit levels from baseline to six months follow-up, from six to 12 months follow-up, and from 12 to 24 months follow-up, but *no statistically significant difference* in glycated haemoglobin (HbA1c), ALT or AST levels for the same time periods. One retrospective case series reported that 37 (23.4%) participants experienced pelvic pain compared to 121 (76.6%) who did not experience pelvic pain whilst receiving testosterone monotherapy at a median follow-up of 22.1 months. No statistical measures were reported. The remaining retrospective case series reported a *statistically significant increase* in haematocrit from baseline to final dose of testosterone monotherapy, but *no statistically significant difference* in total cholesterol, ALT, or AST levels from baseline to final dose at a median follow-up of 1.9 years (range six months to 5.5 years). The same study reported that 14 (11.8%) CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy experienced mild injection site reactions, 77 (64.7%) experienced progression of acne, and one (0.8%) experienced dyslipidaemia, while no participants experienced hypertension, transaminitis, or haematocrit >55%.

In terms of cost-effectiveness:

- No evidence was identified for cost-effectiveness.

In terms of subgroups:

- No evidence was identified for subgroups of interest.

In terms of criteria used by the research studies to define gender incongruence:

- The included studies used different terminology to describe the study populations. Terminology included transgender and gender diverse (TGD) adolescents and young adults or youth, transgender and non-binary (TNB) youth, transmasculine and gender diverse (TM/GD) adolescents and young adults, transgender (TG) adolescents and young adults, or transmasculine (TM) youth. Only two of the included studies described the criteria used to diagnose gender incongruence: one study used World Professional Association of Transgender Health (WPATH) Standards of Care guidelines and the second study used ICD-10 F.64.0, F64.9, and E34.9 or ICD-9 302.85.

In terms of the starting criteria, formulation, duration and dose of testosterone treatment:

- One prospective case series and six retrospective case series reported on testosterone schedules/doses. The first study reported that all participants initiated treatment with testosterone cypionate 200 milligrams per millilitre (mg/mL), the dosing protocol included either subcutaneous injection at a starting dose of 14 mg every week, a dose of 26 mg every week at three to six months, and 100 mg every two weeks at six to nine months, or intramuscular injection at a starting dose of 50 mg every two weeks, a dose of 75 mg every two weeks at three to six months, and 40 mg every week at six to nine months. The second study reported use of testosterone formulations including injectable, topical gel, injectable then subcutaneous pellets, topical gel then injectable, injectable then topical gel then subcutaneous pellets, or injectable then topical gel. The third study reported that *"Nearly all subjects were started on 50–100 mg SC-T monthly divided into every 2 weeks (biweekly) doses; two subjects misunderstood instructions and started on 120–140 mg SC-T monthly"*. The final follow-up dose of monthly subcutaneous testosterone range from 100 to 200 mg or 240 to 320 mg. The fourth study reported that at baseline, the median dose for subcutaneous testosterone was 25.0 mg per week and for transdermal testosterone 25.0 mg per day. At the 12 month follow-up visit, the median subcutaneous dose was 50.0 mg per week and transdermal dose was 40.5 mg per day. At the 24 month visit, the median subcutaneous dose was 50.0 mg per week and the transdermal dose was 40.5 mg per day. The fifth study reported that participants received topical testosterone, 3-weekly intramuscular testosterone undecanoate, or 3-monthly intramuscular testosterone enanthate or mixed testosterone esters. The dose at initiation was defined as low (defined as topical testosterone 12.5 to 25 mg, intramuscular testosterone enanthate or mixed

testosterone esters 125 mg or intramuscular testosterone undecanoate 500 mg) or high (defined as topical testosterone 32.5 to 50 mg, intramuscular testosterone enanthate or mixed testosterone esters 250 mg or intramuscular testosterone undecanoate 1,000 mg [doses equivalent to those used for standard maintenance dosing in adult men]). The remaining study administered testosterone subcutaneously, intramuscularly, as a gel, or transdermally. The median dose was 70 mg. The remaining study reported that participants received testosterone cypionate/enanthate intramuscular injection, testosterone cypionate/enanthate subcutaneous injection, testosterone gel (topical). The median initial starting dose for testosterone cypionate/enanthate intramuscular injection was 25 mg, and for testosterone cypionate/enanthate subcutaneous injection 26 mg. The median frequency of administration for both formulations was 7 days. Only two of these studies reported treatment duration; the first study reported a mean duration of 28.5 months in participants without breakthrough bleeding, and 37.3 months in participants with breakthrough bleeding. The remaining study reported a median follow-up duration of 577 days.

In terms of CYP aged 15 years and younger that received testosterone monotherapy:

- Two retrospective case series reported the circumstances under which children aged 15 years or younger received testosterone monotherapy for gender transition. The first study stated that their centre initiates subcutaneous testosterone around the age of 14 years or older, following multidisciplinary assessment of readiness in youth with gender dysphoria and who have no contraindications. The second study stated that the minimum age for inclusion in the study was eight years in order to ensure that potential participants who might be eligible for hormones based on their Tanner stage would not be excluded due to age alone. Although the remaining studies included some children aged 15 years or younger who received testosterone monotherapy, they did not provide details on the circumstances.

In terms of the monitoring arrangements that were in place for CYP with gender incongruence who identify as male and wish a binary transition receiving masculinising medicines, comprising testosterone monotherapy:

- Two retrospective case series reported on monitoring for CYP who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy. The first study reported that TG youth were seen once a year in clinic by a multidisciplinary team, but during the interim, could be seen by endocrinologists,

nurses, and psychologists individually for follow-up care. The second study stated that all participants were followed by medical and mental health clinicians, but no further details were provided.

In terms of the exclusion criteria of the studies:

- One prospective case series and five retrospective case series reported study exclusion criteria. The first study reported that TG youth without a second follow-up at least three months after initiation of treatment were not eligible for inclusion in the study, one study excluded TGD adolescents AFAB with prior testosterone use, and another study excluded TG adolescents aged less than 13 years at the initiation of GAHT. The second study excluded TGD adolescents and young adults AFAB who had received testosterone for less than one year, had no uterine bleeding documented in their medical records, had Mayer-Rokintansky-Kuster-Hauser Syndrome or hypogonadal hypogonadism. The third study reported that individuals without were excluded from the study if they had a history of GAHT use, presence of serious psychiatric symptoms (eg active hallucinations, thought disorder), visibly distraught (eg suicidal), or were intoxicated or under the influence of alcohol or other substances. The remaining study excluded TM youth on growth- altering medications, such as systemic glucocorticoids, or with any history of a growth-altering disorder, such as precocious puberty, growth hormone deficiency, or an advanced or delayed baseline bone age.

Please see the results table (section 5) in the review for further details of outcomes and definitions.

Limitations

Evidence was identified for the critical outcomes ‘impact on gender incongruence’, ‘impact on mental health’, and ‘impact on quality of life’, and the important outcomes ‘masculinising physical changes’, and ‘detransition after receipt of masculinising medicines’, and safety. No evidence was identified for the important outcomes ‘psychosocial impact’,’ fertility’, ‘feasibility of masculinising genital surgery’, ‘cognitive outcomes’, or ‘regret after receipt of masculinising medicines’. The outcomes reported were primarily assessed using standard assessment tools (with the exception of perceived physical changes and pelvic pain outcomes which used unvalidated questionnaires). The use of standardised outcome measures allows some interpretation of the level of burden associated with specific scores; however, it was not clear how clinically significant the changes observed were. No specific detail about what the minimal clinically important thresholds or differences might be was reported for the outcomes considered.

All of the included studies were judged to be at high risk of bias and the outcomes reported were assessed as very low certainty evidence when evaluated using modified GRADE. Of particular note was the lack of reporting of psychological or psychosocial support / interventions in both the intervention and control populations; this increases the risk of bias through potential confounding. Eight of the papers included in this evidence review were uncontrolled observational studies, which are subject to bias and confounding. Furthermore, three of the included studies (Grannis et al 2023, Kramer et al 2024 and Valentine et al 2022), all of the studies with comparator evidence, were cross-sectional studies which limit the assessment of causality; since both the outcomes and the exposure (in this case testosterone monotherapy) are measured at the same point in time, it is impossible to determine temporality, and these types of studies can only demonstrate associations. Factors reducing confidence in the outcomes reported in the papers include limited reporting of study eligibility criteria, lack of clinical information for participants and/or measurement of gender incongruence, use of subjective outcome measures (ie self-reported) or use of an unvalidated questionnaire, and/or lack of statistical measures. In addition, there was a high degree of indirectness in all of the included studies due to the inclusion of out-of-scope participants and/or interventions for this evidence review. The majority of studies did not report treatment duration, and some studies did not involve or report a follow-up duration, which means that it is not possible to determine the impact of testosterone over time or whether these were sufficient to capture some of the outcomes measured. Further limitations include the loss to follow-up in some studies (where this was reported) and the use of electronic data in some studies which may have limited the reliability of the data collected, and may have meant that not all patients, treatments or outcomes were identified. In addition, most of the studies included at least some (or all) data from more than 10 years ago, and all of the included studies were conducted outside the UK (10 studies were based in paediatric gender clinics in the USA and one study was conducted in a single paediatric gender clinic in Australia). It is therefore unclear whether the populations and aspects of gender affirming care in the included studies reflect that seen in clinical practice in England, and the generalisability of the findings to the NHS may therefore be limited and should be interpreted with caution. No evidence on cost-effectiveness was identified and none of the included studies reported relevant subgroup analyses. No evidence was identified comparing testosterone monotherapy to psychological and psychosocial support or social transitioning to the gender with which the individual identifies.

Conclusion

This evidence review included 11 studies. One study excluded non-binary CYP while the remaining studies included both binary and non-binary CYP who identify as male and wish a

binary transition, or did not report this information. Three cross-sectional studies provided comparator evidence. The remaining eight studies did not include an in-scope comparator.

All evidence was of very low certainty evidence when assessed using modified GRADE. No studies were identified that reported the important outcomes ‘psychosocial impact’, ‘fertility’, ‘feasibility of masculinising genital surgery’, ‘cognitive outcomes’, or ‘regret after receipt of masculinising medicines’.

One comparator study provided evidence that there was a statistically significant improvement in gender incongruence and mental health (including anxiety and depression) in individuals who received testosterone monotherapy compared to individuals not receiving testosterone, but there was no statistically significant difference in suicidality. One comparator study provided evidence that there was no statistically significant difference in eating disorder behaviours (including subject and objective binge episodes, self-induced vomiting, laxative use, or compensatory exercise) among individuals receiving testosterone monotherapy compared to those not receiving treatment. Only two non-comparator studies reported statistical measures for critical and important outcomes: one study reported no statistically significant change in quality of life from three to six months follow-up, but a statistically significant improvement in perceived physical changes at six months follow-up; the second study reported no statistically significant difference in final adult height at an unknown timepoint.

Safety was reported across six studies (one comparator study and five non-comparator studies). The comparator study reported statistically significant higher odds of overweight/obesity, dyslipidaemia, hypertension, and liver function in individuals receiving testosterone compared to those not receiving testosterone, but no statistically significant difference in the odds of dysglycaemia or PCOS. One non-comparator study reported no occurrence of venous thromboembolism or arterial thrombosis (including stroke) among 429 individuals receiving testosterone monotherapy. One prospective case series reported that a small proportion of individuals experienced skin irritation or elevated liver enzymes, but none of the participants developed elevated haemoglobin or haematocrit, or significant dyslipidaemia. One retrospective case series reported statistically significant increases in haemoglobin and haematocrit levels in 136 individuals (the direction of effect was unclear), but no statistically significant differences in HbA1c, ALT and AST at any timepoint. Another retrospective case series reported that 23.4% of individuals receiving testosterone monotherapy experienced pelvic pain. The remaining retrospective case series reported a statistically significant increase in haematocrit from baseline to the final dose of testosterone monotherapy in 119 individuals (the direction of effect was unclear), but no statistically

significant differences in total cholesterol, ALT, or AST. The same study reported that 64.7% of individuals experienced progression of acne, a small proportion of participants experienced mild injection site reactions or dyslipidaemia, while none of the participants reported experiencing hypertension, transaminitis, or haematocrit greater than 55%.

Overall, there is very low certainty evidence with inconsistent results for the selected outcomes in CYP who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy. There is also a lack of direct evidence due to the mixed populations and mixed interventions which further limits the conclusions that can be drawn. No conclusions can be drawn about cost-effectiveness as no evidence was identified. Published studies which allow conclusions to be drawn about the effectiveness of testosterone monotherapy for this population are needed.

3. Methodology

Review questions

The review questions for this evidence review are:

1. For CYP with gender incongruence who identify as a male gender and wish a binary physical transition, what is the clinical effectiveness of treatment with testosterone monotherapy with or without psychological and psychosocial support compared with one or a combination of psychological support or social transitioning to the desired gender or with no intervention?
2. For CYP with gender incongruence who identify as a male gender and wish to undergo a binary physical transition, what is the short-term and long-term safety of testosterone monotherapy with or without psychological and psychosocial support compared with one or a combination of psychological support or social transitioning to the desired gender or with no intervention?
3. For CYP with gender incongruence who identify as a male gender and wish to undergo a binary physical transition, what is the cost-effectiveness of testosterone monotherapy with or without psychological and psychosocial support compared to one or a combination of psychological support or social transitioning to the desired gender or with no intervention?
4. From the evidence selected, are there particular sub-groups of CYP with gender incongruence who identify as a male gender and wish to undergo a binary physical transition that may benefit more from treatment with testosterone monotherapy than the wider population?
5. From the evidence selected:
    a) What were the criteria used by the research studies to define gender incongruence?
    b) What were the starting criteria, formulation, duration and dose of testosterone monotherapy for those aged 16 years up to their 18th birthday?
    c) Did any children aged 15 years or younger receive testosterone monotherapy for gender transition? If so, in what circumstances?
    d) What monitoring was in place for CYP with gender incongruence who identify as a male gender and wish to undergo a binary physical transition receiving testosterone monotherapy?
    e) What were the exclusion criteria in the studies?

See Appendix A for the full PICO document.

Review process

The methodology to undertake this review is specified by NHS England in its '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 04 June 2025.

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 document. Full text of potentially relevant studies 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.

This review prioritised data from comparator studies and studies with larger sample sizes in accordance with the NHS England methodology for conducting these rapid evidence reviews. Smaller, non-comparator studies were only selected when no higher-level evidence was identified for specific PICO outcomes. In such instances, only the PICO outcomes absent from higher-level evidence were extracted from these studies to avoid redundancy and maintain the overall coherence of the review.

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. In studies where only one population is in-scope of the PICO, the case series checklist has been selected to assess quality.

The available evidence was assessed by outcome for certainty using modified GRADE. See Appendix G for GRADE profiles.

4. Summary of included studies

Eleven papers were identified for inclusion in this evidence review (Allen et al 2019, Baines et al 2023, Grannis et al 2023, Grimstad et al 2021, Kramer et al 2024, Laurenzano et al 2021, Millington et al 2024, Moussaoui et al 2024, Mullins et al 2021, Persky et al 2024, Valentine et al 2022). Of the 11 included studies, three were cross-sectional studies providing comparator evidence (Grannis et al 2023, Kramer et al 2024, Valentine et al 2022). The remaining eight studies did not include an in-scope comparator group.

Evidence identified in relation to the critical outcomes included the following: for impact on gender incongruence, one comparator cross-sectional study; for impact on mental health, two comparator cross-sectional studies and one non-comparator retrospective case series; for impact on quality of life, two non-comparator case series (one prospective and one retrospective).

In relation to the important outcomes, evidence was identified for inclusion in relation to masculinising physical changes: three non-comparator case series (one prospective and two retrospective). Detransition after receipt of masculinising medicines was reported in one non-comparator retrospective case series.

Safety was reported in six studies, across six major themes: cardiometabolic-related diagnoses such as hypertension (one comparator cross-sectional study and one non-comparator retrospective case series); VTE or arterial thrombosis (one non-comparator retrospective case series); change in laboratory measures (two non-comparator retrospective case series); adverse events (one non-comparator prospective case series); pelvic pain, abdomino-pelvic pain or pain in the lower part of the abdomen (one non-comparator retrospective case series); and other safety concerns, including injection site reactions and progression of acne (one non-comparator retrospective case series).

No studies were identified that reported on the important outcomes for ‘psychosocial impact’, ‘fertility’, ‘feasibility of masculinising genital surgery’, ‘cognitive outcomes’, or ‘regret after receipt of masculinising medicines’. No evidence was identified comparing testosterone monotherapy to psychological and psychosocial support or social transitioning to the gender with which the individual identifies. No cost-effectiveness studies were identified for inclusion in this review.

Gender incongruence is described differently and the papers included in this review report on TGD adolescents and young adults or youth, TNB youth, TM/GD adolescents and young adults, TG adolescents and young adults, or TM youth. For the purposes of this review, the

terms referred to in the individual papers are used when referring to an individual paper, and CYP with gender incongruence who identify as a male gender and wish a binary physical transition is used when referring to a group of studies or when summarising the results for an outcome.

The terminology in this topic area is continually evolving and is different depending on stakeholder perspectives. In this evidence review we have used the phrase 'CYP with gender incongruence who identify as male gender and wish a binary physical transition' rather than saying natal or biological sex and 'cross-sex hormones' are now referred to as 'masculinising medicines'. The data extracted from studies into Table 1, Appendix E: Evidence Table and Appendix G: GRADE profiles may use historical terms which are no longer considered appropriate.

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

Table 1: Summary of included studies

Study	Population	Intervention and comparison	Outcomes reported

Allen et al 2019
Retrospective cohort study[a]
USA (single centre)
	N=47 TG youth (n=33 AFAB)
	Baseline characteristics are for the whole population (AMAB and AFAB)
	Age (years) at GAHT administration, mean (SD; range)
	16.59[b] (1.19; 13.73 to 19.04)[c]
	Race/Ethnicity, n (%)
	White: 39 (83)
	Biracial or multiracial: 2 (4.3)
	Latinx or Hispanic: 3 (6.4)
	Black or African American: 1 (2.1)
	American Indian or Alaskan Native: 1 (2.1)
	Asian: 1 (2.1)
	No relevant subgroups reported
		Interventions
		Testosterone[d]
		Comparators
		No relevant comparator
		Use of concomitant treatments: NR
		Treatment duration (days) (whole population; AMAB and AFAB), mean (SD): 349 (193)
			Follow-up: at least 3 months
			Important outcomes
			• Impact on mental health
			  • Suicidality measured using the ASQ at baseline and final assessment[e]
			• Impact on QoL
			  • Wellbeing[f]

Baines et al 2023 RCT[g] USA (single centre)	N=26 TGD adolescents AFAB Age (years), median (IQR) 15.5 (15.0 to 17.0)	**Interventions** Testosterone method,[h] n (%) Subcutaneous injection: starting dose (n=16) 14	Follow-up: 3, 6 and 9 months **Critical outcomes** • Impact on QoL • Change in PedsQL, score

Study	Population	Intervention and comparison	Outcomes reported
	Race, n (%) Caucasian: 21 (81) Hispanic or Latino: 4 (15) Unknown/NR: 1 (4) Weight (kg), median (IQR) 68.25 (58.5 to 82.2) BMI (kg/m^2), median (IQR) 24.7 (20.6 to 30.1) No relevant subgroups reported	mg every week; at 3 (n=14) to 6 (n=13) months 26 mg every week; at 6 to 9 months (n=3) 100 mg every 2 weeks Intramuscular injection: starting dose (n=10) 50 mg every 2 weeks; at 3 (n=9) to 6 (n=6) months 75 mg every 2 weeks; at 6 to 9 months (n=2) 40 mg every week **Comparators** No relevant comparator Use of concomitant treatments, n (%): 5 (19.2%) participants received GnRH analogues treatment (leuprolide-depot) during the study period Treatment duration: NR	**Important outcomes** • Masculinising physical changes • Perceived physical changes (non-validated measure)[i] **Safety** • Number of individuals reporting adverse effects

Grannis et al 2023
Retrospective cross-sectional study
USA (single centre)
	N=82 TNB youth (n=50 AFAB)
	Baseline characteristics are for TNB youth AFAB
	Age (years), mean (SD)
	Testosterone: 17.04 (1.18)
	No testosterone: 15.24 (1.72)
	Race, n (%)
	Black or African American: Testosterone 3 (14.29); no testosterone 1 (3.45)
	Multiracial: Testosterone 3 (14.29); no testosterone 3 (10.34)
	Native American or American Indian: Testosterone 1 (4.76); no testosterone 1 (3.45)
	White: Testosterone 14 (66.67); no testosterone 22 (75.86)
	Prefer not to answer: Testosterone 0 (0); no testosterone 2 (6.90)
		Interventions
		Testosterone injections: n=21
		Comparators
		No testosterone: n=29
		Use of concomitant treatments: NR
		Treatment duration (months), mean (SD): 12.87 (9.94)
			Follow-up: none
			Critical outcomes
			• Impact on gender incongruence
			  • Body image dissatisfaction measured using the BID scale
			• Impact on mental health
			  • Generalised anxiety symptoms measured using the SCARED scale
			  • Depression symptoms measured using the CDI scale
			  • Social anxiety symptoms measured using the LSAS scale
			  • Frequency of suicidal ideation and/or attempts in the past year

Study	Population	Intervention and comparison	Outcomes reported
	No relevant subgroups reported		
Grimstad et al 2021 Retrospective case series USA (single centre)	N=232 TGD adolescents and young adults AFAB *Baseline characteristics are for participants with no breakthrough bleeding (n=174) or breakthrough bleeding (n=58)* Age (years) at initiation of testosterone, mean (SD) No breakthrough bleeding: 16.3 (1.8) Breakthrough bleeding: 16.3 (2.2) Race/ethnicity, n (%) White (non-Hispanic): no breakthrough bleeding 130 (74.7); breakthrough bleeding 42 (72.4) White (Hispanic): no breakthrough bleeding 6 (3.4); breakthrough bleeding 8 (13.8) Asian: no breakthrough bleeding 4 (2.3); breakthrough bleeding 1 (1.7) Black: no breakthrough bleeding 3 (1.7); breakthrough bleeding 3 (5.2) American Indian: no breakthrough bleeding 2 (1.1); breakthrough bleeding 0 (0) Additional race: no breakthrough bleeding 3 (1.7); breakthrough bleeding 0 (0) Declined to answer: no breakthrough bleeding 26 (14.9); breakthrough bleeding 4 (6.9) Comorbid conditions, n (%) Endometriosis: no breakthrough bleeding 1	**Interventions** Type of testosterone used, n (%) Injectable:[j] no breakthrough bleeding 170 (97.7); breakthrough bleeding 51 (87.9) Topical gel: no breakthrough bleeding 1 (0.6); breakthrough bleeding 0 (0) Injectable then subcutaneous pellets: no breakthrough bleeding 1 (0.6); breakthrough bleeding 3 (5.2) Topical gel then injectable: no breakthrough bleeding 1 (0.6); breakthrough bleeding 1 (1.7) Injectable then topical gel then subcutaneous pellets: no breakthrough bleeding 1 (0.6); breakthrough bleeding 0 (0) Injectable then topical gel: no breakthrough bleeding 0 (0); breakthrough bleeding 3 (5.2) **Comparators** No relevant comparator Use of concomitant treatments, n On menstrual suppression but discontinued during study: 74 On menstrual suppression throughout study: 43 Never on menstrual suppression or GnRH analogues: 106 *Menstrual suppression use, n (%)*	Follow-up: NR[k] **Important outcomes** • Masculinising physical changes • Number of participants reporting breakthrough bleeding[l]

Study	Population	Intervention and comparison	Outcomes reported
	(0.6); breakthrough bleeding 3 (5.2) Hypothyroidism: no breakthrough bleeding 2 (1.1); breakthrough bleeding 0 (0)	No breakthrough bleeding 93 (53.4) Breakthrough bleeding 33 (56.9) Treatment duration (months), mean (SD) No breakthrough bleeding 28.5 (14.6) Breakthrough bleeding: 37.3 (16.9)	
Kramer et al 2024 Retrospective cross-sectional study USA (single centre)	N=251 TGD youth (n=181 [72.1%] binary AFAB) *Baseline characteristics are for TGD youth AFAB* Age (years), mean (SD) 16.77 (SD 2.74) BMI, mean (SD) 26.87 (7.88) EDE-Q score, mean (SD) 1.34 (1.17) Race, n (%) White: 122 (89.7) Black: 3 (2.2) Asian: 3 (2.2) Other: 8 (5.9) No relevant subgroups reported	**Interventions** Testosterone, n (%): 39 (21.5) **Comparators** No testosterone, n (%): 142 (78.5) Use of concomitant treatments: NR Treatment duration: NR	Follow-up: none[m] **Crticial outcomes** • Impact on mental health • Number of individuals reporting eating disorder behaviours[n]
Laurenzano et al 2021 Retrospective case series USA (single centre)	N=119 TM/GD adolescents and young adults AFAB (n=110 [92.4%] binary) *Baseline characteristics are for TM/GD adolescents and young adults AFAB* Age at presentation (years), mean (range) 16 (10.1 to 19.8) Age at start of testosterone (years), mean (range) 16.5 (13 to 19.9) Race, n (%)	**Interventions** Initial subcutaneous testosterone dose The authors stated that *"Nearly all subjects were started on 50–100 mg SC-T monthly divided into every 2 weeks (biweekly) doses; two subjects misunderstood instructions and started on 120–140 mg SC-T monthly"* Final follow-up dose of monthly subcutaneous testosterone, n (%)	Follow-up: median 1.9 years (range 6 months to 5.5 years) **Important outcomes** • Detransition after receipt of masculinising medicines • Number of participants discontinuing treatment **Safety** • Change in laboratory values • Number of adverse events during treatment with testosterone

Study	Population	Intervention and comparison	Outcomes reported
	Caucasian: 79 (66.4) Asian: 7 (5.9) Native American/Alaska Native: 4 (3.4) African American: 2 (1.7) Native Hawaiian/Pacific Islander: 1 (0.8) Unknown/unavailable: 20 (16.8) Other:[o] 17 (14.3) BMI Z-score category, n (%) Obese: 24 (20.2) Overweight: 19 (16) Underweight: 3 (2.5) No relevant subgroups reported	100 to 200 mg: 94 (79) 240 to 320 mg: 21 (18) **Comparators** No relevant comparator Use of concomitant treatments: NR Treatment duration: NR	

Millington et al 2024
Retrospective cohort study[p]
USA (four centres)
	N=293 TG adolescents (n=200 AFAB)
	Baseline characteristics are for TG adolescents AFAB
	Age at presentation (years), median (IQR)
	16.2 (15.1 to 17.6)
	Gender identity, n (%)
	Male: 81 (41)
	TG male: 106 (53)
	Gender fluid: 1 (1)
	Gender queer: 1 (0.5)
	Non-binary: 10 (5)
	Tanner stage at baseline, n (%)
	3: 1 (0.5)
	4: 17 (9)
	5: 168 (91)
	No relevant subgroups reported
		Interventions
		Testosterone formulation at baseline, n (%)
		Subcutaneous: 195 (97)
		Transdermal gel: 5 (3)
		Testosterone dose at baseline visit, median (IQR)
		Subcutaneous administration, mg per week: 25.0 (25.0 to 26.0)
		Transdermal administration, mg per day: 25.0 (20.25 to 25.0)
		Testosterone dose at 24-month visit, median (IQR)
		Subcutaneous administration, mg per week: 50.0 (50.0 to 60.0)
		Transdermal administration, mg per day: 40.5 (20.25 to 60.75)
		Comparators
		No relevant comparator
			Follow-up: 6, 12 and 24 months
			Safety
			• Change in laboratory values from baseline to follow-up
			  • Change in haemoglobin
			  • Change in haematocrit
			  • Change in HbA1c
			  • Change in ALT
			  • Change in AST

Study	Population	Intervention and comparison	Outcomes reported
		Use of concomitant treatments, n (%) Progesterone use: Norethindrone acetate: 31 (16) Medroxyprogesterone acetate: 6 (3) Etonogestrel implant: 1 (0.5) Combined oestrogen and progesterone oral contraceptive: 22 (11) Treatment duration: NR	
Moussaoui et al 2024 Retrospective case series Australia (one centre)	N=158 TGD AFAB children and adolescents Age at testosterone initiation (years), median (range) 16.6 (13.9 to 18.6) No relevant subgroups reported	**Interventions** Testosterone, n (%) Topical testosterone: 17 (10.8) 3-weekly intramuscular testosterone undecanoate: 63 (39.9) 3- monthly intramuscular testosterone enanthate or mixed testosterone esters: 78 (49.4) Low dose:[q] 74 (46.8) High dose:[r] 84 (53.2) **Comparators** No relevant comparator Use of concomitant treatments, n (%): hormonal medication to suppress menstruation 137 (86.7)[s] Treatment duration: NR	Follow-up: 22.1 (IQR median value 15.4) months **Safety** • Number of individuals reporting pelvic pain[t]
Mullins et al 2021 Retrospective case series USA (single centre)	N=611 TG adolescents and young adults (n=428 AFAB;[u] n=12 [2.8%] non-binary) *Baseline characteristics are for the whole population (AMAB and AFAB)* Age at presentation (years), median (IQR) 17 (15 to 19)	**Interventions** Testosterone (n=429): median dose 70 mg (IQR 60 to 80) Testosterone formulation, n (%) Subcutaneous: 312 (72.7) Intramuscular: 105 (24.4) Gel: 11 (2.8)	Follow-up: median 577 days (IQR 283 to 923) **Safety** • Incidence of thrombosis during treatment with GAHT

Study	Population	Intervention and comparison	Outcomes reported
	Race, n (%) White: 544 (89) African American: 50 (8.2) Asian American: 8 (1.3) Other: 12 (1.9) Not documented: 8 (1.3) Ethnicity, n (%) Hispanic: 14 (2.3) Non-Hispanic: 595 (97.4) Not documented: 2 (0.3) BMI, kg/m^2, n (%) 18.5: 40 (6.5) 18.5 to 25: 212 (34.7) 25 to 30 148 (24.2) 30: 211 (34.5) Family history of thrombosis, n (%) Yes: 49 (8.0) No: 388 (63.5) Not documented: 174 (28.5) No relevant subgroups reported	Transdermal: 1 (0.7) **Comparators** No relevant comparator Use of concomitant treatments, n (%) Anticoagulation (rivaroxaban): 1 (0.2)[v] Treatment duration, median (IQR): 577 days (283 to 923)	

Persky et al 2024
Retrospective cohort study[w]
USA (five centres)
	N=94 TM youth
	Baseline characteristics are for TM youth receiving testosterone only
	Age at start of testosterone (years), median (IQR)
	16.8 (15.8 to 17.8)
	Race, n (%)
	White: 49 (79)
	Other: 13 (21) [Black/African American: 3; Asian or Pacific Islander: 1; American Indian or Alaskan: 1; Other/Unknown: 8]
	Ethnicity, n (%)
	Hispanic: 5 (8)
		Interventions
		Formulation of testosterone at initiation, n (%)
		Testosterone cypionate/enanthate intramuscular injection: 9/62 (14.5)
		Testosterone cypionate/enanthate subcutaneous injection: 53/62 (85.5)
		Testosterone gel (topical): 0/62 (0)
		Initial starting dose of injectable testosterone, median (IQR)
		Testosterone cypionate/enanthate
			Follow-up: NR
			Important outcomes
			• Masculinising physical changes
			  • Final adult height from baseline to follow-up
			  • Final adult height Z-scores[x] from baseline to follow-up

Study	Population	Intervention and comparison	Outcomes reported
	BMI (at initial visit; Z-score), mean (SD) 0.71 (1.1) No relevant subgroups reported	intramuscular injection (mg), median (IQR) (n=9): 25 (25 to 25) Testosterone cypionate/enanthate subcutaneous injection (mg), median (IQR) (n=53): 26 (25 to 50) **Comparators** No relevant comparator Use of concomitant treatments: the authors reported that participants on stimulant medications were included, but no further details were provided Treatment duration: NR	
Valentine et al 2022 Retrospective cross-sectional study USA (PEDSnet study; six centres)	N=4,172[y] TGD youth (n=2,766 AFAB) *Baseline characteristics are reported for the whole population (AMAB and AFAB)* Age (years) at first visit, median (IQR) 10.0 (4.4 to 14.6) Age (years) at last visit, median (IQR) 16.7 (14.6 to 18.3) Race, n (%) White: 3,027 (72.5) Unknown: 401 (9.6) Other: 390 (9.3) Black: 257 (6.2) Asian: 98 (2.3) Ethnicity, n (%) Non-Hispanic: 3,538 (84.8) Hispanic: 354 (8.5) Unknown: 281 (6.7)	**Interventions** Testosterone: n=832 **Comparators** No testosterone: n=1,934 Use of concomitant treatments, n (%) Progestin norethindrone and medroxyprogesterone: 112 (4.1) COCP: 199 (7.2) The authors mentioned the use of antipsychotics, but no further details were provided Treatment duration: NR	Follow-up: none **Safety** • Odds of cardiometabolic-related diagnoses (overweight/obese, dyslipidaemia, liver dysfunction, dysglycaemia, hypertension, PCOS)

Abbreviations
AFAB: assigned female at birth; ALT: Alanine aminotransferase; AMAB: assigned male at birth; ASQ: Ask Suicide-Screening Questions; AST: aspartate aminotransferase; BID: Body Image Dissatisfaction; BMI: body mass index; CDI: Children's Depression Inventory; COCP: combined oral contraceptive pill; GAHT: gender

Study	Population	Intervention and comparison	Outcomes reported
affirming hormones; GnRH: gonadotropin-releasing hormone; HbA1c: glycated haemoglobin; IQR: interquartile range; kg: kilogram; kg/m^2: kilogram per square metre; LSAS: Leibowitz Social Anxiety Scale; mg: milligram; n: number; NR: not reported; PCOS: polycystic ovary syndrome; PEDSnet: Paediatric Learning Health System network; QoL: quality of life; RCT: randomised controlled trial; SCARED: Screen for Child Anxiety Related Emotional Disorders; SC-T: subcutaneous testosterone; SD: standard deviation; TG: transgender; TGD: transgender and gender diverse; TM: transmasculine; TM/GD: transmasculine and gender diverse; TNB: transgender and non-binary; USA: United States of America.			
Footnotes a. Although this study was a retrospective cohort study, as the comparator groups were out-of-scope for this evidence review, it was treated as a case series. b. There was a discrepancy in the mean age reported in the main text (16.59) and table (16.50) and we have reported the figure presented in the main text. c. The authors reported that most participants (90%) were at or below the age of 18.01 years at initiation of GAHT. d. The authors reported that the endocrinologists in their clinic sometimes begin participants at hormone levels lower than the recommended protocol, and typically, patients' doses are gradually increased every three to six months so that the dosage levels recommended by suggested protocols are reached by the end of treatment. e. ASQ is a four-item measure used to identify patients who are at risk of attempting suicide. Questions include: In the past few weeks have you… "…wished you were dead?", "…felt that you or your family would be better off if you were dead?", "…been having thoughts about harming or killing yourself?", or "…done anything to hurt yourself or to end your life?" A response of "no" was scored as 0 and a response of "yes" was scored as 1; with an overall score for suicidality on a scale ranging from 0 to 4, with higher scores indicating greater levels of suicidal ideation. f. Wellbeing was measured using the PedsQL [General Wellbeing scale] which is a 5-point response scale, containing seven items, and measures "general well-being" and "general health". The general well-being subscale includes six items (eg "I feel happy" and "I think my health will be good in the future"). Participants are asked to consider each item over the past month and rate responses from 0 (never) to 4 (almost always). The general health subscale contains one item, "In general, how is your health?" ranging from 0 (Bad) to 4 (Excellent). All items are scored and linearly transformed to a 0 to 100 scale (initial score of 0 = 0, 1 = 25, 2 = 50, 3 = 75, and 4 = 100). High scores indicate perceptions of minimal problems, high wellbeing. g. Although this was an RCT, it was treated as a case series as no relevant comparator group (as stated in the PICO document) was included in the study. h. All participants initiated treatment with testosterone cypionate 200 mg/mL (transitioned to testosterone enanthate if skin irritation occurred). i. Masculinising effects questionnaire assessed perceived physical changes after starting testosterone, including skin oiliness/acne, facial hair, body hair, increased muscle mass/strength, menstrual cessation, and clitoral enlargement. j. Includes intramuscular and subcutaneous testosterone cypionate and testosterone enanthate [reported as ethanate in the paper]. k. The authors stated that to observe breakthrough bleeding effects after a medication change, they followed participants up three months after the change, but no further details were reported. l. Breakthrough bleeding was defined as any bleeding presumed to originate in the uterus while on testosterone. m. The presence of eating disorders was self-reported by individuals at their first visit to the clinic and no follow-up data were presented. n. Eating disorder behaviours were measured using the Eating Disorder Examination-Questionnaire (EDE-Q) subscales, including Restraint, Weight Concern, Shape Concern, and Eating Concern. Individual EDE-Q items reflect specific eating disorder behaviours, ie objective and subjective binge eating episodes, self-induced vomiting, laxative use, and compensatory exercise, and these items were used in the study analyses. o. Other self-reported races included: Mexican (n=1), Peruvian (n=1), Puerto Rican (n=1), and Hispanic not otherwise specified (n=14). p. Participants were recruited for the Trans Youth Care-United States (TYCUS) Study. q. 'Low' dose of testosterone was defined as topical testosterone 12.5 to 25 mg, intramuscular testosterone enanthate or mixed testosterone esters 125 mg or intramuscular testosterone undecanoate 500 mg.			

Study	Population	Intervention and comparison	Outcomes reported
r. ‘High’ dose of testosterone was defined as topical testosterone 32.5 to 50 mg, intramuscular testosterone enanthate or mixed testosterone esters 250 mg or intramuscular testosterone undecanoate 1,000 mg (doses equivalent to those used for standard maintenance dosing in adult men). s. Menstrual suppression included: levonorgestrel intrauterine device: 10 (7.3%); oral combined contraceptive pill: 2 (1.5%); etonogestrel implant: 3 (2.2%); norethisterone: 106 (77.4%); medroxyprogesterone acetate injection: 13 (9.5%); oral medroxyprogesterone acetate: 3 (2.2%). t. Pelvic pain was defined as the timing of onset as reported by the individual, or as the date of first documentation in medial chart if no mention of timing of onset. Intensity of pain was self-reported by participants based on a score of 1 to 10, with 10 being most severe pain. u. There was an unexplained discrepancy between the number of TG men (reported as n=428 [70%]) and the number of TG men receiving testosterone (n=429 [70.2%]) and we have reported the number as n=429 for the purposes of reporting outcomes. v. One TG male had a thrombotic event of VTE prior to starting testosterone and was maintained on rivaroxaban during treatment with testosterone. w. Although this was a cohort study, it was treated as a case series as there was no relevant in-scope comparator group. x. Final adult height Z-scores were calculated based on the CDC 20 year old growth chart for girls. y. There was a discrepancy in the total sample size which was reported as n=4,172, but the number of AMAB and AFAB individuals totals n=4,173.			

5. Results

For CYP with gender incongruence who identify as a male gender and wish a binary physical transition, what is the clinical effectiveness and safety of treatment with testosterone monotherapy with or without psychological and psychosocial support compared with one or a combination of psychological support or social transitioning to the desired gender or with no intervention?

Outcome	Evidence statement
Clinical Effectiveness	
Critical outcomes	
Impact on gender incongruence **Certainty of evidence:** Very low	*This outcome is important to patients because gender incongruence is associated with significant distress and problems functioning.* One cross-sectional study provided comparator evidence relating to the impact of testosterone monotherapy on gender dysphoria (measured using the BID) in CYP with gender incongruence who identify as a male gender and wish a binary physical transition compared to individuals not receiving testosterone. The mean treatment duration was 12.87 (SD 9.94) months. *Testosterone monotherapy vs no hormones* **At mean 12.87 (SD 9.94) months treatment duration** • One cross-sectional study (Grannis et al 2023; n=50) reported that dissatisfaction with body image (measured using the BID)[7] was *statistically significantly lower* in CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy (mean 92.29 [SD 19.74]) compared to individuals not receiving testosterone (mean 103.14 [SD 18.99]): ANOVA F-value 7.46 (df 1, 47), p<0.01. The mean treatment duration was 12.87 (SD 9.94) months. **(VERY LOW)** *Testosterone monotherapy (no comparator)* No evidence was identified for this outcome. ***Testosterone monotherapy vs no hormones*** **One cross-sectional study reported that dissatisfaction with body image was *statistically significantly lower* in CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy compared to individuals not receiving testosterone: mean 92.29 (SD 19.74) vs mean 103.14 (SD 18.99), respectively. The mean duration of treatment with testosterone monotherapy was 12.87 (SD 9.94) months. This study provided very low certainty evidence.** ***Testosterone monotherapy (no comparator)*** **No evidence was identified for this outcome.**

[7] Higher values indicate greater dissatisfaction with one's body image.

Outcome	Evidence statement
Impact on mental health **Certainty of evidence:** Very low	*This outcome is important to patients because gender incongruence is associated with psychological distress which can lead to the development of mental health problems.* Two cross-sectional studies and one retrospective case series provided evidence relating to the impact of masculinising medicines on mental health (measured using various questionnaires) in CYP with gender incongruence who identify as a male gender and wish a binary physical transition. Two cross-sectional studies provided comparator evidence relating to anxiety and depression or eating disorders in CYP taking testosterone monotherapy vs individuals not taking testosterone. The mean treatment duration was 12.87 (SD 9.94) months in one study and at an unknown treatment duration/follow-up in the second study. One retrospective case series provided non-comparator evidence relating to suicidality in CYP receiving testosterone monotherapy to at least three months follow-up. *Testosterone monotherapy vs no hormones* **At mean 12.87 (SD 9.94) months treatment duration** • One cross-sectional study (Grannis et al 2023; n=50) reported *statistically significantly less* generalised anxiety symptoms[8] in CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy (n=21; mean 38.75 [SD 17.41]) compared to those not receiving testosterone (n=29; mean 50.25 [SD 14.12]); ANOVA F-value 7.76 (df 1, 45), p<0.01. The mean treatment duration was 12.87 (SD 9.94) months. **(VERY LOW)** • One cross-sectional study (Grannis et al 2023; n=50) reported *statistically significantly less* symptoms of depression[9] in CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy (n=21; mean 13.38 [SD 7.81]) compared to those not receiving testosterone (n=29; mean 18.34 [SD 7.02]); ANOVA F-value 5.62 (df 1, 47), p<0.05. The mean treatment duration was 12.87 (SD 9.94) months. **(VERY LOW)** • One cross-sectional study (Grannis et al 2023; n=50) reported *statistically significantly less* symptoms of social anxiety[10] in CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy (n=21; mean 50.21 [SD 22.34]) compared to those not receiving testosterone (n=29; mean 78.62 [SD 31.41]); ANOVA F-value 14.8 (df 1, 42), p<0.001. The mean treatment duration was 12.87 (SD 9.94) months. **(VERY LOW)** • One cross-sectional study (Grannis et al 2023; n=50) reported *no statistically significant difference* in suicidality[11] in CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy (n=21; mean 1.95 [SD 0.92]) compared to those not receiving testosterone (n=29; mean 2.86 [SD 1.46]); ANOVA F-value 3.05 (df 1, 47), p=NS. The mean treatment duration was 12.87 (SD 9.94) months. **(VERY LOW)**

[8] Measured using Screen for Child Anxiety Related Emotional Disorders (SCARED); higher values indicate greater generalised anxiety symptoms.
[9] Measured using the Children's Depression Inventory (CDI); higher values indicate greater depression symptoms.
[10] Measured using the Leibowitz Social Anxiety Scale (LSAS); higher values indicate greater social anxiety.
[11] Defined as frequency of suicidal ideation and/or attempts in the past year.

Outcome	Evidence statement
	At unknown treatment duration/follow-up • One cross-sectional study (Kramer et al 2024; n=181) reported *no statistically significant difference* in the presence of eating disorder behaviours in the form of subjective binge episodes[12] among CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy (n=5/39, 13.5%) compared to those not receiving testosterone (n=16/142, 18.0%), p>0.05. The treatment duration/follow-up was not reported. **(VERY LOW)** • One cross-sectional study (Kramer et al 2024; n=181) reported *no statistically significant difference* in the presence of eating disorder behaviours in the form of objective binge episodes among CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy (n=4/39, 10.8%) compared to those not receiving testosterone (n=19/142, 21.1%), p>0.05. The treatment duration/follow-up was not reported. **(VERY LOW)** • One cross-sectional study (Kramer et al 2024; n=181) reported *no statistically significant difference* in the presence of eating disorder behaviours in the form of self-induced vomiting episodes among CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy (n=1/39, 2.7%) compared to those not receiving testosterone (n=1/142, 1.1%), respectively, p>0.05. The treatment duration/follow-up was not reported. **(VERY LOW)** • One cross-sectional study (Kramer et al 2024; n=181) reported *no statistically significant difference* in laxative use reported by CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy (n=39) compared to those not receiving testosterone (n=142): 2 (5.1%) vs 6 (6.3%), p>0.05. The treatment duration/follow-up was not reported. **(VERY LOW)** • One cross-sectional study (Kramer et al 2024; n=181) reported *no statistically significant difference* in the presence of eating disorder behaviours in the form of compensatory exercise behaviour among CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy (n=7/39, 18.9%) compared to those not receiving testosterone (n=16/142, 17.8%), p>0.05. The treatment duration/follow-up was not reported. **(VERY LOW)** *Testosterone monotherapy (no comparator)* **To at least 3 months follow-up** • One retrospective case series (Allen et al 2019; n=33) reported a reduction in suicidal ideation[13] from baseline (mean 1.01 [SE 0.23]) to

[12] Eating disorder behaviours were measured using the Eating Disorder Examination-Questionnaire (EDE-Q) subscales, including Restraint, Weight Concern, Shape Concern, and Eating Concern. Individual EDE-Q items reflect specific eating disorder behaviours, ie objective and subjective binge eating episodes, self-induced vomiting, laxative use, and compensatory exercise, and these items were used in the study analyses.
[13] Measured using the Ask Suicide-Screening Questions (ASQ) which is a four-item measure used to identify patients who are at risk of attempting suicide. Questions include: In the past few weeks have you… "…wished you were dead?", "…felt

Outcome	Evidence statement
	at least three months follow-up (mean 0.29 [SE 0.13]) in CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy, but no statistical measures were reported. **(VERY LOW)** ***Testosterone monotherapy vs no hormones*** **One cross-sectional study reported *statistically significantly lower* general anxiety symptoms (mean 38.75 vs 50.25), depressive symptoms (mean 13.38 vs 18.34), and social anxiety (mean 50.21 vs 78.62) in CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy compared to individuals not receiving testosterone, respectively. The same study reported *no statistically significant difference* in suicidality between the two groups. The mean treatment duration was 12.87 months. The second cross-sectional study reported *no statistically significant difference* in eating disorder behaviours between CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy compared to individuals not receiving testosterone. The treatment duration/follow-up was not reported.** ***Testosterone monotherapy (no comparator)*** **One retrospective case series reported a reduction in suicidal ideation from baseline (mean 1.01) to at least three months follow-up (mean 0.29) in CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy, but no statistical measures were reported.** **All of the above studies provided very low certainty evidence.**
Impact on Quality of life (QoL) **Certainty of evidence:** Very low	*This outcome is important to patients because gender incongruence may be associated with a significant reduction in health-related quality of life.* One prospective case series and one retrospective case series provided non-comparator evidence relating to the impact on QoL in CYP with gender incongruence who identify as a male gender and wish a binary physical transition and receiving testosterone monotherapy. The studies assessed wellbeing to at least three months follow-up or QoL at three, six, and nine months follow-up. *Testosterone vs no hormones* No evidence was identified for this outcome. *Testosterone (no comparator)* **At 3 months follow-up** • One prospective case series (Baines et al 2023) reported median QoL (measured using PedsQL) at baseline (n=24; 66.08, IQR 60.87 to 80.98) and at three months follow-up (n=20; 64.31, IQR 54.97 to 76.63) in CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy, but no statistical measures were reported. **(VERY LOW)** **At least 3 months follow-up**

that you or your family would be better off if you were dead?", "…been having thoughts about harming or killing yourself?", or "…done anything to hurt yourself or to end your life?" A response of "no" was scored as 0 and a response of "yes" was scored as 1; with an overall score for suicidality on a scale ranging from 0 to 4, with higher scores indicating greater levels of suicidal ideation.

Outcome	Evidence statement
	• One retrospective case series (Allen et al 2019; n=33) reported wellbeing [14] at baseline (64.95, SE 2.66) to at least three months follow-up (70.94, SE 2.35) in CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy, but no statistical measures were reported. **(VERY LOW)** **At 6 months follow-up** • One prospective case series (Baines et al 2023) reported *no statistically significant difference* in median QoL (measured using PedsQL) at baseline (n=24; 66.08, IQR 60.87 to 80.98) and six months follow-up (n=18; 61.41, IQR 56.52 to 85.87) in CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy, p=0.71. **(VERY LOW)** **At 9 months follow-up** • One prospective case series (Baines et al 2023) reported median QoL (measured using PedsQL) at baseline (n=24; 66.08, IQR 60.87 to 80.98) and nine months follow-up (n=5; 59.78, IQR 42.39 to 66.30) in CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy, but no statistical measures were reported. **(VERY LOW)** ***Testosterone monotherapy vs no hormones*** **No evidence was identified for this outcome.** ***Testosterone monotherapy (no comparator)*** **One prospective case series reported median QoL (measured using PedsQL) in CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy at baseline (66.08) and at three months (64.31), six months (61.41), and nine months follow-up (59.78). The difference was *not statistically significant* at six months, but no statistical measures were reported for three and nine months follow-up. One retrospective case series reported greater wellbeing in CYP with gender incongruence who identify as a male gender and wish a binary physical transition following at least three months on testosterone monotherapy (mean 70.94)compared to baseline (mean 64.95), but no statistical measures were reported. These two studies provided very low certainty evidence.**
Important outcomes	
Masculinising physical changes **Certainty of evidence:** Very low	*This outcome is important because most patients with gender incongruence wish 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.* One prospective case series and two retrospective case series provided non-comparator evidence relating to masculinising physical changes in CYP with gender incongruence who identify as a male gender and wish a binary physical

[14] Measured using the Paediatric Quality of Life Inventory General Wellbeing Scale (PedsQL GWBS) which is a 5-point response scale, containing seven items, and measures “general well-being” and “general health”. The general well-being subscale includes six items (eg “I feel happy” and “I think my health will be good in the future”). Participants are asked to consider each item over the past month and rate responses from 0 (never) to 4 (almost always). The general health subscale contains one item, “In general, how is your health?” ranging from 0 (Bad) to 4 (Excellent). All items are scored and linearly transformed to a 0 to 100 scale (initial score of 0 = 0, 1 = 25, 2 = 50, 3 = 75, and 4 = 100). High scores indicate perceptions of minimal problems, high wellbeing.

Outcome	Evidence statement
	transition and receiving testosterone monotherapy. The prospective case series reported outcomes at three, six and nine months follow-up. One retrospective case series reported outcomes at mean treatment durations of 28.5 (SD 14.6) and 37.3 (SD 16.9) months. The remaining retrospective case series reported outcomes at an unknown treatment duration/follow-up. *Testosterone vs no hormones* No evidence was identified for this outcome. *Testosterone (no comparator)* **At 3 months follow-up** • One prospective case series (Baines et al 2023) reported perceived physical changes (using an unvalidated masculinising effects questionnaire)[15] in CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy at three months follow-up (n=20): median 16.0 (IQR 12.5 to 20.0). **(VERY LOW)** **At 6 months follow-up** • One prospective case series (Baines et al 2023) reported a *statistically significant increase* in perceived physical changes (using an unvalidated masculinising effects questionnaire) in CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy from three (n=20) to six months follow-up (n=18): median 16.0 (IQR 12.5 to 20.0) vs 19.5 (IQR 17.0 to 23.0), respectively, p<0.001. **(VERY LOW)** **At 9 months follow-up** • One prospective case series (Baines et al 2023) reported a perceived physical changes (using an unvalidated masculinising effects questionnaire) in CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy at nine months follow-up (n=5): median 21.0 (IQR 20.0 to 22.0). No statistical measures were reported. **(VERY LOW)** **At mean 28.5 and 37.3 months treatment duration** • One retrospective case series (Grimstad et al 2021; n=232) reported breakthrough bleeding in 58 (25%)[16] CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy compared to 174 (75%) individuals reporting no breakthrough bleeding whilst taking testosterone monotherapy. Mean treatment duration was 28.5 (SD 14.6) months and 37.3 (SD 16.9) months, respectively. No statistical measures were reported. **(VERY LOW)**

[15] Masculinising effects questionnaire assessed perceived physical changes after starting testosterone, including skin oiliness/acne, facial hair, body hair, increased muscle mass/strength, menstrual cessation, and clitoral enlargement. Higher scores indicate greater perceived physical changes.
[16] For participants with breakthrough bleeding, bleeding started at a mean of 24.3 (SD 7.2) months after initiation of testosterone. Mean age at time of first breakthrough bleeding was 18.4 (SD 2.8) years. Eight (13.8%) of these participants had never become amenorrhoeic in the first year and continued to bleed beyond the first year. A total of 48 patients (82.8%) had more than one episode of bleeding after one year on testosterone.

Outcome	Evidence statement
	Unknown treatment duration/follow-up • One retrospective case series (Persky et al 2024; n=62) indicated *no statistically significant difference* in final adult height[17] in CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy at an unknown treatment duration/follow-up compared to baseline: mean 164.1 (SD 6.8) cm vs 164.7 (SD 6.7) cm, respectively. No statistical measures were reported. **(VERY LOW)** • One retrospective case series (Persky et al 2024; n=62) indicated *no statistically significant difference* in final adult height Z-scores[18] in CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy at an unknown treatment duration/follow-up compared to baseline: mean 0.21 (SD 1.0) vs 21.0 (SD 1.0), respectively. No statistical measures were reported. **(VERY LOW)** ***Testosterone monotherapy vs no hormones*** **No evidence was identified for this outcome.** ***Testosterone monotherapy (no comparator)*** **One prospective case series reported a *statistically significant increase* in perceived physical changes (using an unvalidated masculinising effects questionnaire) in CYP with gender incongruence who identify as a male gender and wish a binary physical transition and receiving testosterone monotherapy from three to six months follow-up (median 16.0 vs 19.5, respectively). The same study reported perceived physical changes at nine months follow-up, but statistical measures were not reported. One retrospective case series reported breakthrough bleeding in 58 (25%) CYP with gender incongruence who identify as a male gender and wish a binary physical transition and receiving testosterone monotherapy compared to 174 (75%) individuals reporting no breakthrough bleeding whilst taking testosterone monotherapy, but no statistical measures were reported. The mean treatment duration was 28.5 months and 37.3 months, respectively. The remaining retrospective case series indicated *no statistically significant difference* in final adult height or final adult height Z-scores in CYP with gender incongruence who identify as a male gender and wish a binary physical transition and receiving testosterone monotherapy at an unknown treatment duration/follow-up.** **All of the above studies provided very low certainty evidence.**
Psychosocial impact **Certainty of evidence:** Not applicable	*This outcome is important to patients because gender incongruence is associated with internalising and externalising behaviours and emotional and behavioural problems which may impact on social and occupational functioning.* **No evidence was identified for psychosocial impact.**
Fertility **Certainty of evidence:** Not applicable	*This outcome is important to patients because masculinising medicines can reduce fertility. Prior to commencing masculinising medicines patients should be counselled on the impact of treatment on their fertility and offered fertility preservation options.*

[17] The authors stated that due to the variability in height measurements reported in the medical records, the average of all heights recorded after the final adult height was reached was calculated and this was used for data analysis purposes.
[18] Final adult height Z-scores were calculated based on the CDC 20 year old growth chart for girls.

Outcome	Evidence statement
	No evidence was identified for fertility.
Feasibility of masculinising genital surgery **Certainty of evidence:** Not applicable	*This outcome is important to patients because masculinising medicines can have an impact on surgical outcomes. Treatment may alter the amount of genital tissue available for phalloplasty, metoidioplasty, hysterectomy and bilateral salpingo-oophorectomy.* **No evidence was identified for feasibility of masculinising genital surgery.**
Cognitive outcomes **Certainty of evidence:** Not applicable	*This outcome is important to patients because masculinising medicines can negatively impact cognitive processes such as concentration, memory, and executive function.* **No evidence was identified for cognitive outcomes.**
Detransition after receipt of masculinising medicines **Certainty of evidence:** Very low	*Medical detransition is a complex experience encompassing medical, psychological, social implications and is important to patients because they may choose to discontinue treatment. The decision to detransition may or may not be associated with regret.* One retrospective case series provided non-comparator evidence relating to detransition after receipt of masculinising medicines in CYP with gender incongruence who identify as a male gender and wish a binary physical transition taking testosterone monotherapy. The median follow-up was 1.9 years (range 6 months to 5.5 years). *Testosterone monotherapy vs no hormones* No evidence was identified for this outcome. *Testosterone monotherapy (no comparator)* **At median 1.9 years (range 6 months to 5.5 years) follow-up** • One retrospective case series (Laurenzano et al 2021; n=119) reported that three CYP with gender incongruence who identify as a male gender and wish a binary physical transition discontinued testosterone monotherapy at a median follow-up duration of 1.9 years (range 6 months to 5.5 years): *"two stopped because they were satisfied with effects of SC-T achieved and one stopped within 6 months of starting after reassessing their gender identity…. due to concerns about body changes and potential impact on fertility".* No statistical measures were reported. **(VERY LOW)** ***Testosterone monotherapy vs no hormones*** **No evidence was identified for this outcome.** ***Testosterone monotherapy (no comparator)*** **One retrospective case series reported that of the 119 CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy *"two stopped because they were satisfied with effects of SC-T achieved and one stopped within 6 months of starting after reassessing their gender identity…. due to concerns about body changes and potential impact on fertility"*. The median follow-up duration was 1.9 years (range 6 months to 5.5 years). No statistical measures were reported.** **This study provided very low certainty evidence.**

Outcome	Evidence statement
Regret after receipt of masculinising medicines **Certainty of evidence:** Not applicable	*This outcome is important to patients because some patients who choose to take masculinising medicines may regret this decision. Regret may or may not be associated with detransition.* **No evidence was identified for regret after receipt of masculinising medicines.**
Safety	
Frequency of adverse events **Certainty of evidence:** Very low	*It 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*. One cross-sectional study, one prospective case series, and four retrospective case series provided evidence relating to safety in CYP with gender incongruence who identify as a male gender and wish a binary physical transition. The cross-sectional study provided comparator evidence on cardiometabolic-related diagnoses in individuals taking testosterone monotherapy vs no testosterone at an unknown treatment duration/follow-up. One prospective case series reported adverse events at treatment initiation and at nine months follow-up. The four retrospective case series reported safety at a median follow-up of 577 days, 22.1 months, or 1.9 years, or at six, nine, 12 and 24 months follow-up. *Testosterone monotherapy vs no hormones* **Unknown treatment duration/follow-up** • One cross-sectional study (Valentine et al 2022) reported *statistically significantly higher* odds of overweight/obesity in CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy compared to individuals not receiving testosterone (52.3% vs 39.4%, respectively): OR 1.8 (95% CI 1.5 to 2.1), p<0.0001; this was *not statistically significant* after adjusting for confounders (results not presented in the paper).[19] The treatment duration/follow-up was not reported. **(VERY LOW)** • One cross-sectional study (Valentine et al 2022) reported *statistically significantly higher* odds of dyslipidaemia in CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy compared to individuals not receiving testosterone (14.2% vs 6.7%, respectively): OR 1.7 (95% CI 1.3 to 2.3), p<0.01; this was *not statistically significant* after adjusting for confounders (results not presented in the paper).[16] The treatment duration/follow-up was not reported. **(VERY LOW)** • One cross-sectional study (Valentine et al 2022) reported *statistically significantly higher* odds of liver dysfunction in CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy compared to individuals not receiving testosterone (14.3% vs 10.1%, respectively): OR 1.5 (95% CI 1.1 to 1.9), p≤0.01; this was *not statistically significant* after adjusting for confounders (results not presented in the paper).[16] The treatment duration/follow-up was not reported. **(VERY LOW)** • One cross-sectional study (Valentine et al 2022) reported *no statistically significant difference* in the odds of dysglycaemia in CYP with gender incongruence who identify as a male gender and wish a

[19] Analyses were adjusted for electronic health record (EHR) recorded sex/sex assigned at birth, age at last visit, duration in PEDSnet/EHR, overweight/obesity status, depression status and antipsychotic prescription.

Outcome	Evidence statement
	binary physical transition receiving testosterone monotherapy compared to individuals not receiving testosterone (1.9% vs 2.2%, respectively): OR (95% CI) NR, p=NS. The treatment duration/follow-up was not reported. **(VERY LOW)** • One cross-sectional study (Valentine et al 2022) reported *statistically significantly higher* odds of hypertension in CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy compared to individuals not receiving testosterone (11.2% vs 8.3%, respectively): OR 1.6 (95% CI 1.2 to 2.2), p<0.01; this was *not statistically significant* after adjusting for confounders (results not presented in the paper).[16] The treatment duration/follow-up was not reported. **(VERY LOW)** • One cross-sectional study (Valentine et al 2022) reported *no statistically significant difference* in the odds of PCOS in CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy compared to individuals not receiving testosterone (1.0% vs 1.0%, respectively): OR (95% CI) NR, p=NS. The treatment duration/follow-up was not reported. **(VERY LOW)** *Testosterone monotherapy (no comparator)* **At treatment initiation** • One prospective case series (Baines et al 2023; n=26) reported that three (11.5%) CYP with gender incongruence who identify as a male gender and wish a binary physical transition experienced skin irritation at initiation of testosterone monotherapy. No statistical measures were reported. **(VERY LOW)** **At median 577 (IQR 283 to 923) days follow-up** • One retrospective case series (Mullins et al 2021; n=429) reported no VTE or arterial thrombosis events (including stroke) in CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy at a median follow-up of 577 (IQR 283 to 923) days. **(VERY LOW)** **At 3 months follow-up** • One prospective case series (Baines et al 2023; n=26) reported that one (3.8%) CYP with gender incongruence who identifies as a male gender and wishes a binary physical transition receiving testosterone monotherapy experienced elevated AST (≤36 U/L) at three months follow-up. No statistical measures were reported. **(VERY LOW)** **At 6 months follow-up** • One retrospective case series (Millington et al 2024) reported a *statistically significant increase* in haemoglobin levels from baseline (n=189) to six months follow-up (n=155) in CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy: median (IQR) 13.2 mg/dL (12.5 to 13.9) vs 14.2 mg/dL (13.3 to 15.1), respectively, p<0.001. The direction of benefit was unclear. **(VERY LOW)** • One retrospective case series (Millington et al 2024) reported a *statistically significant increase* in haematocrit levels from baseline (n=191) to six months follow-up (n=157) in CYP with gender

Outcome	Evidence statement
	incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy: median (IQR) 39.9% (37.8 to 41.6) vs 43.8% (41.6 to 45.8), respectively, p<0.001. The direction of benefit was unclear. **(VERY LOW)** • One retrospective case series (Millington et al 2024) reported *no statistically significant difference* in HbA1c from baseline (n=105)[20] to six months follow-up (n=66) in CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy: median (IQR) 5.2% (5.0 to 5.4) vs 5.1% (5.0 to 5.4), respectively, p=NS. The direction of benefit was unclear. **(VERY LOW)** • One retrospective case series (Millington et al 2024) reported *no statistically significant difference* in ALT[21] from baseline (n=77) to six months follow-up (n=58) in CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy: median (IQR) 17 U/L (11 to 25) vs 19.5 U/L(14 to 28), respectively, p=NS. The direction of benefit was unclear. **(VERY LOW)** • One prospective case series (Baines et al 2023; n=26) reported that one (3.8%) CYP with gender incongruence who identifies as a male gender and wishes a binary physical transition receiving testosterone monotherapy experienced elevated ALT (≤60 U/L) at three months follow-up. No statistical measures were reported. **(VERY LOW)** • One retrospective case series (Millington et al 2024) reported *no statistically significant difference* in AST from baseline (n=77) to six months follow-up (n=58) in CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy: median (IQR) 20 U/L (17 to 25) vs 23 U/L (19 to 29), respectively, p=NS. The direction of benefit was unclear. **(VERY LOW)** **At 9 months follow-up** • One prospective case series (Baines et al 2023; n=26) reported that no CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy developed elevated haemoglobin at nine months follow-up. **(VERY LOW)** • One prospective case series (Baines et al 2023; n=26) reported that no CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy developed elevated haematocrit at nine months follow-up. **(VERY LOW)**

[20] Five participants with pre-existing diabetes mellitus were excluded from the HbA1c analysis. There were 10 (5.1%) participants with baseline HbA1c measurements in the prediabetes range (5.7% to 6.4%). Of these, three participants did not have additional follow-up measurements, five had subsequently normal measurements, and two had persistently elevated HbA1c measurements of 5.7%.

[21] Four (2%) participants had elevations in liver enzymes >2 times the upper limit of normal during testosterone therapy (range 118 to 263 U/L). The authors stated that all four participants were receiving concurrent treatment with other medications that have been associated with abnormal liver function tests (ie isotretinoin, quetiapine, lamotrigine, bupropion). Three of these four participants had subsequent ALT and AST measurements that returned to normal during the study visit. The fourth participant, who was taking isotretinoin and sertraline, continued to have elevated ALT at the 24-month follow-up visit.

Outcome	Evidence statement
	• One prospective case series (Baines et al 2023; n=26) reported that no CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy developed significant dyslipidaemia at nine months follow-up. **(VERY LOW)** **At 12 months follow-up** • One retrospective case series (Millington et al 2024) reported a *statistically significant increase* in haemoglobin levels from six months (n=155) to 12 months follow-up (n=136) in CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy: median (IQR) 14.2 mg/dL (13.3 to 15.1) vs 14.7 mg/dL (13.6 to 15.6), respectively, p<0.001. The direction of benefit was unclear. **(VERY LOW)** • One retrospective case series (Millington et al 2024) reported a *statistically significant increase* in haematocrit levels from six months (n=157) to 12 months follow-up (n=136) in CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy: median (IQR) 43.8% (41.6 to 45.8) vs 44.6% (41.6 to 47.0), respectively, p<0.001. The direction of benefit was unclear. **(VERY LOW)** • One retrospective case series (Millington et al 2024) reported *no statistically significant difference* in HbA1c[22] from six months (n=66) to 12 months follow-up (n=64) in CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy: median (IQR) 5.1% (5.0 to 5.4) vs 5.1% (4.9 to 5.3), respectively, p=NS. The direction of benefit was unclear. **(VERY LOW)** • One retrospective case series (Millington et al 2024) reported *no statistically significant difference* in ALT[23] from six months (n=58) to 12 months follow-up (n=57) in CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy: median (IQR) 19.5 U/L (14 to 28) vs 18 U/L (13 to 26), respectively, p=NS. The direction of benefit was unclear. **(VERY LOW)** • One retrospective case series (Millington et al 2024) reported *no statistically significant difference* in AST from six months (n=58) to 12 months follow-up (n=57) in CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy: median (IQR) 23 U/L (19 to 29) vs 22 U/L (18 to 27), respectively, p=NS. The direction of benefit was unclear. **(VERY LOW)** **At median 22.1 (IQR median value 15.4) months follow-up**

[22] Five participants with pre-existing diabetes mellitus were excluded from the HbA1c analysis. There were 10 (5.1%) participants with baseline HbA1c measurements in the prediabetes range (5.7% to 6.4%). Of these, three participants did not have additional follow-up measurements, five had subsequently normal measurements, and two had persistently elevated HbA1c measurements of 5.7%.

[23] Four (2%) participants had elevations in liver enzymes >2 times the upper limit of normal during testosterone therapy (range 118 to 263 U/L). The authors stated that all four participants were receiving concurrent treatment with other medications that have been associated with abnormal liver function tests (ie isotretinoin, quetiapine, lamotrigine, bupropion). Three of these four participants had subsequent ALT and AST measurements that returned to normal during the study visit. The fourth participant, who was taking isotretinoin and sertraline, continued to have elevated ALT at the 24-month follow-up visit.

Outcome	Evidence statement
	• One retrospective case series (Moussaoui et al 2024; n=158) reported that pelvic pain[24] was experienced by 37 (23.4%)[25] CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy compared to 121 (76.6%) individuals who did not experience pelvic pain at a median follow-up of 22.1 (IQR median value 15.4) months. No statistical measures were reported. **(VERY LOW)** **At median 1.9 years (range 6 months to 5.5 years) follow-up** • One retrospective case series (Laurenzano et al 2021; n=119) reported *no statistically significant difference* in total cholesterol from baseline (mean 158.0 mg/dL, SD 29.8) to final dose of testosterone monotherapy (mean 154.6 mg/dL, SD 31.3) in CYP with gender incongruence who identify as a male gender and wish a binary physical transition: mean difference -3.5 mg/dL (SD 0.194), p=NS.[26] The median follow-up duration was 1.9 years (range 6 months to 5.5 years). **(VERY LOW)** • One retrospective case series (Laurenzano et al 2021; n=119) reported a *statistically significant increase* in haematocrit from baseline (mean 39.2%, SD 2.6) to final dose of testosterone monotherapy (mean 44.1%, SD 3.3) in CYP with gender incongruence who identify as a male gender and wish a binary physical transition: mean difference 4.9% (SD not reported), p<0.001.[27] The median follow-up duration was 1.9 years (range 6 months to 5.5 years). **(VERY LOW)** • One retrospective case series (Laurenzano et al 2021; n=119) reported *no statistically significant difference* in ALT from baseline (mean 20.8 U/L, SD 10.0) to final dose of testosterone monotherapy (mean 21.4 U/L, SD 12.9) in CYP with gender incongruence who identify as a male gender and wish a binary physical transition: mean difference 0.7 U/L (SD 0.60), p=NS.[28] The median follow-up duration was 1.9 years (range 6 months to 5.5 years). **(VERY LOW)** • One retrospective case series (Laurenzano et al 2021; n=119) reported *no statistically significant difference* in AST from baseline (mean 21.8 U/L, SD 8.0) to final dose of testosterone monotherapy (mean 23.0 U/L, SD 8.8) in CYP with gender incongruence who identify as a male gender and wish a binary physical transition: mean difference 1.1 U/L

[24] Pelvic pain was defined as the timing of onset as reported by the individual, or as the date of first documentation in medical chart if no mention of timing of onset. The median interval between testosterone initiation and onset of pain was 1.6 months (range 0.3 to 6.4).
[25] Pelvic pain description: cramps 17 (45.9%); similar to period pain 8 (21.6%); sex-related pain 10 (27%); early morning/or waking up at night 5 (13.5%); suspicion of pelvic floor spasms 5 (13.5%); associated with breakthrough bleeding 11 (29.7%); associated with nausea and/vomiting 4 (10.8%); pain radiating into lower limbs 2 (5.4%). Thirty six of 37 participants reporting pelvic pain had not received past puberty blockers. Pain intensity was reported for n=11 adolescents: n=1 mild (self-reported score of 1 to 3 out of 10, with a score of 10 being most severe); n=10 severe (self-reported score of 7 to 10 out of 10, with a score of 10 being most severe).
[26] There was no statistically significant difference in mean total cholesterol at baseline or follow-up across three testosterone dose ranges (<160 mg, 160 to 240 mg and >240 mg).
[27] There was a statistically significant increase in mean final haematocrit levels as testosterone dose increased (<160 mg, 160 to 240 mg and >240 mg); p=0.024, but no statistically significant difference in baseline haematocrit or difference between baseline and final haematocrit.
[28] The authors reported that there was a borderline statistically significant increase in mean final ALT as testosterone dose increased (<160 mg, 160 to 240 mg and >240 mg); p=0.05, but no statistically significant difference in baseline ALT or difference between baseline and final ALT.

Outcome	Evidence statement
	(SD 0.29), p=NS.[29] The median follow-up duration was 1.9 years (range 6 months to 5.5 years). **(VERY LOW)** • One retrospective case series (Laurenzano et al 2021; n=119) reported that 14 (11.8%) CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy experienced mild injection site reactions at a median follow-up duration was 1.9 years (range 6 months to 5.5 years). No statistical measures were reported. **(VERY LOW)** • One retrospective case series (Laurenzano et al 2021; n=119) reported that no CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy experienced hypertension at a median follow-up duration was 1.9 years (range 6 months to 5.5 years). **(VERY LOW)** • One retrospective case series (Laurenzano et al 2021; n=119) reported that 77 (64.7%) CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy experienced progression of acne[30] at a median follow-up duration was 1.9 years (range 6 months to 5.5 years). No statistical measures were reported. **(VERY LOW)** • One retrospective case series (Laurenzano et al 2021; n=119) reported that no CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy experienced transaminitis at a median follow-up duration was 1.9 years (range 6 months to 5.5 years). **(VERY LOW)** • One retrospective case series (Laurenzano et al 2021; n=119) reported that one (0.8%) CYP with gender incongruence who identifies as a male gender and wishes a binary physical transition receiving testosterone monotherapy experienced dyslipidaemia at a median follow-up duration was 1.9 years (range 6 months to 5.5 years). No statistical measures were reported. **(VERY LOW)** • One retrospective case series (Laurenzano et al 2021; n=119) reported that no CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy experienced haematocrit >55% at a median follow-up duration was 1.9 years (range 6 months to 5.5 years). **(VERY LOW)** **At 24 months follow-up** • One retrospective case series (Millington et al 2024) reported a *statistically significant increase* in haemoglobin levels from 12 months follow-up (n=136) to 24 months follow-up (n=119) in CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy: median (IQR) 14.7 mg/dL (13.6 to 15.6) vs 15.0 mg/dL (14.1 to 15.8, respectively, p=0.01.The direction of benefit was unclear. **(VERY LOW)** • One retrospective case series (Millington et al 2024) reported a *statistically significant increase* in haematocrit levels from 12 months

[29] There was a statistically significant increase in the difference between baseline and final AST as testosterone dose increased (<160 mg, 160 to 240 mg and >240 mg); p=0.01, but no statistically significant difference in baseline or final AST.
[30] Advanced acne management, ie oral treatment and/or referral to dermatology, was reported in 23 of 77 participants.

Outcome	Evidence statement
	follow-up (n=136) to 24 months follow-up (n=119)[31] in CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy: median (IQR) 44.6% (41.6 to 47.0) vs 45.4% (42.9 to 47.6), respectively, p=0.03. The direction of benefit was unclear. **(VERY LOW)** • One retrospective case series (Millington et al 2024) reported *no statistically significant difference* in HbA1c[32] from 12 months follow-up (n=64) to 24 months follow-up (n=59)[33] in CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy: median (IQR) 5.1% (4.9 to 5.3) vs 5.1% (4.9 to 5.3), respectively, p=NS. The direction of benefit was unclear. **(VERY LOW)** • One retrospective case series (Millington et al 2024) reported *no statistically significant difference* in ALT[34] from 12 months follow-up (n=57) to 24 months follow-up (n=42) in CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy: median (IQR) 18 U/L (13 to 26) vs 19 U/L (13 to 27), respectively, p=NS. The direction of benefit was unclear. **(VERY LOW)** • One retrospective case series (Millington et al 2024) reported *no statistically significant difference* in AST from 12 months follow-up (n=57) to 24 months follow-up (n=42) in CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy: median (IQR) 22 U/L (18 to 27) vs 22 U/L (18 to 28), respectively, p=NS. The direction of benefit was unclear. **(VERY LOW)** ***Testosterone monotherapy vs no hormones*** **One cross-sectional study reported that CYP with gender incongruence who identify as a male gender and wish a binary physical transition and receiving testosterone monotherapy reported *statistically significant higher* odds of overweight/obesity (OR 1.8, 95% CI 1.5 to 2.1), dyslipidaemia (OR 1.7, 95% CI 1.3 to 2.3), liver dysfunction (OR 1.5, 95% CI 1.1 to 1.9), and hypertension (OR 1.6, 95% CI 1.2 to 2.2) compared to individuals not receiving testosterone, but these were *not statistically significant* after adjusting for confounders. The same study reported *no statistically significant difference* in the odds of dysglycaemia or PCOS**

[31] There were 13 participants (6.5%) DFAB, one of whom had recent tobacco use, who had haematocrit above the typical cisgender male range (haematocrit 40% to 50%) during treatment with testosterone, ranging from 50.1% to 53.1%. Apart from a decrease in the testosterone dose, there were no other clinical interventions 9eg therapeutic phlebotomy) required to address the increased haematocrit.
[32] Five participants with pre-existing diabetes mellitus were excluded from the HbA1c analysis. There were 10 (5.1%) participants with baseline HbA1c measurements in the prediabetes range (5.7% to 6.4%). Of these, three participants did not have additional follow-up measurements, five had subsequently normal measurements, and two had persistently elevated HbA1c measurements of 5.7%.
[33] One participant had an increase in HbA1c from 5.4% to 5.7% over the 24-month treatment period. This participant also had an increase in BMI from 24.5 to 27.8 kg/m2. Another participant had an increase in HbA1c from 5.6% to 6.1% at the 12-month follow-up visit, but HbA1c was in the normal range at the 24-month follow-up visit (4.3%). This participant also had a baseline BMI in the obese range (32.0 kg/m2). In total, there were 15 (7.7%) participants who had an elevated HbA1c measurement at any point during the study period.
[34] Four (2%) participants had elevations in liver enzymes >2 times the upper limit of normal during testosterone therapy (range 118 to 263 U/L). The authors stated that all four participants were receiving concurrent treatment with other medications that have been associated with abnormal liver function tests (ie isotretinoin, quetiapine, lamotrigine, bupropion). Three of these four participants had subsequent ALT and AST measurements that returned to normal during the study visit. The fourth participant, who was taking isotretinoin and sertraline, continued to have elevated ALT at the 24-month follow-up visit.

Outcome	Evidence statement
	between individuals receiving vs not receiving testosterone. The duration of treatment/follow-up was not reported. ***Testosterone monotherapy (no comparator)*** **One prospective case series reported that three (11.5%) CYP with gender incongruence who identify as a male gender and wish a binary physical transition experienced skin irritation at initiation of testosterone monotherapy. The same study reported that two participants experienced elevated ALT or AST levels at three or six months follow-up, but no participants developed elevated haematocrit, or significant dyslipidaemia at nine months follow-up. No statistical measures were reported for any safety outcome. One retrospective case series reported no VTE or arterial thrombosis events (including stroke) at a median follow-up of 577 days. One retrospective case series reported a *statistically significant increase* in haemoglobin and haematocrit levels from baseline to six months follow-up, from six to 12 months follow-up, and from 12 to 24 months follow-up, but *no statistically significant difference* in HbA1c, ALT or AST levels for the same time periods. One retrospective case series reported that 37 (23.4%) participants experienced pelvic pain compared to 121 (76.6%) who did not experience pelvic pain whilst receiving testosterone monotherapy at a median follow-up of 22.1 months. No statistical measures were reported. The remaining retrospective case series reported a *statistically significant increase* in haematocrit from baseline to final dose of testosterone monotherapy, but *no statistically significant difference* in total cholesterol, ALT, or AST levels from baseline to final dose at a median follow-up of 1.9 years. The same study reported that 14 (11.8%) CYP with gender incongruence who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy experienced mild injection site reactions, 77 (64.7%) experienced progression of acne, and one (0.8%) experienced dyslipidaemia, while no participants experienced hypertension, transaminitis, or haematocrit >55%.** **All of the above studies provided very low certainty evidence.**
Abbreviations ALT: alanine transaminase; ASQ: Ask Suicide-Screening Questions; AST: aspartate aminotransferase; BID: body image dissatisfaction; CDI: Children's Depression Inventory; CI: confidence interval; CYP: children and young people; EDE-Q: Eating Disorders Examination-Questionnaire; HbA1c: glycated haemoglobin; IQR: interquartile range; LSAS: Leibowitz Social Anxiety Scale; n: number; mg/dL: milligrams per decilitre; NR: not reported: NS: not significant; OR: odds ratio; PCOS: polycystic ovary syndrome; PedsQL: Paediatric Quality of Life Inventory; QoL: quality of life; SCARED: Screen for Child Anxiety Related Emotional Disorders; SC-T: subcutaneous testosterone; SD: standard deviation; SE: standard error; U/L: units per litre; VTE: venous thromboembolism.	

For CYP with gender incongruence who identify as a male gender and wish to undergo a binary physical transition, what is the cost-effectiveness of testosterone monotherapy with or without psychological and psychosocial support compared to one or a combination of psychological support or social transitioning to the desired gender or with no intervention?

Outcome	Evidence statement
Cost-effectiveness	No evidence was identified for cost-effectiveness.

From the evidence selected, are there particular sub-groups of CYP with gender incongruence who identify as a male gender and wish to undergo a binary physical transition that may benefit more from treatment with testosterone monotherapy than the wider population?

Subgroups	Evidence statement
	No evidence was identified for any subgroups.

From the evidence selected:

a) **What were the criteria used by the research studies to define gender incongruence?**
b) **What were the starting criteria, formulation, duration and dose of testosterone monotherapy for those aged 16 years up to their 18th birthday?**
c) **Did any children aged 15 years or younger receive testosterone monotherapy for gender transition? If so, in what circumstances?**
d) **What monitoring was in place for CYP with gender incongruence who identify as a male gender and wish to undergo a binary physical transition receiving testosterone monotherapy?**
e) **What were the exclusion criteria in the studies?**

Outcome	Evidence statement
Definitions of gender incongruence	Gender incongruence descriptions were provided by all the included studies. • One retrospective case series (Allen et al 2019; n=33) did not define gender incongruence but included TG youth with a history of gender nonconformity or gender dysphoria, and identified as binary or non-binary. • One prospective case series (Baines et al 2023; n=26) included transgender male adolescents AFAB but did not describe the criteria used for diagnosing gender incongruence. • One cross-sectional study (Grannis et al 2023; n=50) included TNB youth who met a diagnostic classification of gender dysphoria, based on assessment by a specialist mental health provider, but did not

Outcome	Evidence statement
	describe the criteria for gender dysphoria diagnoses. Gender identify was self-reported and expressed as male, female, transgender, female to male transgender/FTM, male to female transgender/MTF, trans male/trans man, trans female/trans woman, trans*r/trans asterisk, genderqueer, gender nonconforming, gender variant, gender fluid, gender expansive, intersex, androgynous, non-binary, transsexual, cross-dresser, two-spirited, third gender, agender, not sure, other. • One retrospective case series (Grimstad et al 2021; n=232) included TGD adolescents and young adults AFAB but did not describe the criteria used. • One cross-sectional study (Kramer et al 2024; n=181) included TGD youth but did not describe the criteria used for diagnosing gender incongruence. • One retrospective case series (Laurenzano et al 2021; n=119) included TM/GD adolescents and young adults AFAB if they met diagnostic criteria for gender dysphoria in accordance with WPATH Standards of Care guidelines. • One retrospective case series (Millington et al 2024; n=200) included TGD adolescents if they met diagnostic criteria for gender dysphoria as determined by a clinician but did not describe the criteria used. • One retrospective case series (Moussaoui et al 2024; n=158) included TGD children and adolescents AFAB but did not describe the criteria used for diagnosing gender incongruence. • One retrospective case series (Mullins et al 2021; n=428) included TG adolescents and young adults but did not describe the criteria used for diagnosing gender incongruence. • One retrospective case series (Persky et al 2024; n=62) included TM youth if they met diagnostic criteria for ICD-10 F.64.0, F64.9, and E34.9 or ICD-9 302.85. • One cross-sectional study (Valentine et al 2022; n= n=2,766) included TGD youth but did not describe the criteria used for diagnosing gender incongruence.
Testosterone dosing	Testosterone schedules/doses were reported in one prospective case series and six retrospective case series. Three cross-sectional studies and one retrospective case series did not provide this information. • One prospective case series (Baines et al 2023; n=26) reported that all participants (aged 15 to 17 years) initiated treatment with testosterone cypionate 200 mg/mL (participants transitioned to testosterone enanthate if skin irritation occurred). Sixteen participants used subcutaneous injection at a starting dose of 14 mg every week; a dose of 26 mg every week at three to six months; and 100 mg every two weeks at six to nine months. Ten participants used intramuscular injection at a starting dose of 50 mg every two weeks; a dose of 75 mg every two weeks at three to six months; and 40 mg every week at six to nine months. Treatment duration was not reported. • One retrospective case series (Grimstad et al 2021; n=232) reported the different testosterone formulations used: injectable, topical gel,

Outcome	Evidence statement
	injectable then subcutaneous pellets, topical gel then injectable, injectable then topical gel then subcutaneous pellets, or injectable then topical gel. Mean treatment duration was 28.5 (SD 14.6) months in participants without breakthrough bleeding, and 37.3 (SD 16.9) months in participants with breakthrough bleeding. • One retrospective case series (Laurenzano et al 2021; n=119) reported that *"Nearly all subjects were started on 50–100 mg SC-T monthly divided into every 2 weeks (biweekly) doses; two subjects misunderstood instructions and started on 120–140 mg SC-T monthly"*. The final follow-up dose of monthly subcutaneous testosterone range from 100 to 200 mg or 240 to 320 mg. Treatment duration was not reported. • One retrospective case series (Millington et al 2024; n=200) reported that at baseline visit, the median (IQR) dose for subcutaneous testosterone was 25.0 (25.0 to 26.0) mg per week and for transdermal testosterone 25.0 (20.25 to 25.0) mg per day. At the 12 month visit, the median (IQR) subcutaneous dose was 50.0 (40.0 to 50.0) mg per week and transdermal dose was 40.5 (37.5 to 40.5) mg per day. At the 24 month visit, the median (IQR) subcutaneous dose was 50.0 (50.0 to 60.0) mg per week and the transdermal dose was 40.5 (20.25 to 60.75) mg per day. The treatment duration was not reported. • One retrospective case series (Moussaoui et al 2024; n=158) reported that participants received topical testosterone, 3-weekly intramuscular testosterone undecanoate, or 3-monthly intramuscular testosterone enanthate or mixed testosterone esters. The dose at initiation was defined as low (defined as topical testosterone 12.5 to 25 mg, intramuscular testosterone enanthate or mixed testosterone esters 125 mg or intramuscular testosterone undecanoate 500 mg) or high (defined as topical testosterone 32.5 to 50 mg, intramuscular testosterone enanthate or mixed testosterone esters 250 mg or intramuscular testosterone undecanoate 1,000 mg [doses equivalent to those used for standard maintenance dosing in adult men]). The treatment duration was not reported. • One retrospective case series (Mullins et al 2021; n=428) administered testosterone subcutaneously, intramuscularly, as a gel, or transdermally. The median dose was 70 mg (IQR 60 to 80). The median treatment duration was 577 days (IQR 283 to 923). • One retrospective case series (Persky et al 2024; n=62) reported that participants received testosterone cypionate/enanthate intramuscular injection, testosterone cypionate/enanthate subcutaneous injection, testosterone gel (topical). The median (IQR) initial starting dose for testosterone cypionate/enanthate intramuscular injection was 25 mg (25 to 25), and for testosterone cypionate/enanthate subcutaneous injection 26 mg (25 to 50). The median frequency of administration for both formulations was 7 days (range 7 to 14). The treatment duration was not reported.
Testosterone monotherapy for those <15 years	Two retrospective case series reported the number of children aged 15 years or young who received testosterone monotherapy for gender transition. Three cross-sectional studies, one prospective case series and five retrospective case series did not provide this information.

Outcome	Evidence statement
	• One retrospective case series (Laurenzano et al 2021; n=119) stated that their centre typically initiates subcutaneous testosterone around the age of 14 years or older, following multidisciplinary assessment of readiness in youth with gender dysphoria and who have no contraindications. Six participants commenced subcutaneous testosterone younger than the age of 14 years. • One retrospective case series (Millington et al 2024; n=200) stated that the minimum age for inclusion in the study was eight years in order to ensure that potential participants who might be eligible for hormones based on their Tanner stage would not be excluded due to age alone.
Monitoring arrangements	Monitoring for CYP with gender incongruence who identify as a male gender was reported in two retrospective case series. Three cross-sectional studies, one prospective case series, and four retrospective case series did not provide details on monitoring. • One retrospective case series (Allen et al 2019; n=33) reported that TG youth were seen once a year in clinic by a multidisciplinary team, but during the interim, could be seen by endocrinologists, nurses, and psychologists individually for follow-up care. • One retrospective case series (Grimstad et al 2021; n=232) stated that all participants were followed by medical and mental health clinicians, but no further details were provided.
Study exclusion criteria	Study exclusion criteria were reported in one prospective case series and five retrospective case series. Three cross-sectional studies and one retrospective case series did not report study exclusion criteria. • One retrospective case series (Allen et al 2019; n=33) reported that TG youth without a second follow-up at least three months after initiation of treatment were not eligible for inclusion in the study. • One prospective case series (Baines et al 2023; n=26) excluded TGD adolescents AFAB with prior testosterone use. • One retrospective case series (Grimstad et al 2021; n=232) excluded TGD adolescents and young adults AFAB who had received testosterone for less than one year, had no uterine bleeding documented in their medical records, had Mayer-Rokintansky-Kuster-Hauser Syndrome or hypogonadal hypogonadism. • One retrospective case series (Millington et al 2024; n=200) reported that individuals without were excluded from the study if they had a history of GAHT use, presence of serious psychiatric symptoms (eg active hallucinations, thought disorder), visibly distraught (eg suicidal), or were intoxicated or under the influence of alcohol or other substances. • One retrospective case series (Mullins et al 2021; n=428) excluded TG adolescents if they were aged less than 13 years at the initiation of GAHT. • One retrospective case series (Persky et al 2024; n=62) excluded TM youth on growth- altering medications, such as systemic glucocorticoids, or with any history of a growth-altering disorder, such as precocious puberty, growth hormone deficiency, or an advanced or delayed baseline bone age.

Outcome	Evidence statement
Abbreviations AFAB: assigned female at birth; CYP: children and young people; FTM: female to male; GAHT: gender-affirming hormone therapy; ICD: International Classification of Diseases; IQR: interquartile range; mg: milligram; m/mL: milligrams per millilitre; MTF: male to female; n: number; SC-T: subcutaneous testosterone; SD: standard deviation; TG: transgender; TGD: transgender and diverse; TM/GD trans masculine/gender diverse; TNB: transgender and non-binary; WPATH: World Professional Association for Transgender Health.	

6. Discussion

This evidence review examines the clinical effectiveness, safety, and cost-effectiveness of masculinising medicines comprising testosterone monotherapy compared with one or a combination of psychological support or social transitioning to the desired gender, or no intervention in CYP with gender incongruence who identify as male gender and wish a binary physical transition. The critical outcomes of interest were 'impact on gender incongruence', 'impact on mental health', and 'impact on quality of life' (QoL). The important outcomes of interest were 'masculinising physical changes', 'psychosocial impact', 'fertility', 'feasibility of masculinising genital surgery', 'cognitive outcomes', 'detransition after receipt of masculinising medicines', 'regret after receipt of masculinising medicines', and safety.

Evidence for the clinical effectiveness and safety of masculinising medicines was available from eleven papers (Allen et al 2019, Baines et al 2023, Grannis et al 2023, Grimstad et al 2021, Kramer et al 2024, Laurenzano et al 2021, Millington et al 2024, Moussaoui et al 2024, Mullins et al 2021, Persky et al 2024, Valentine et al 2022). Three studies provided comparator evidence (Grannis et al 2023, Kramer et al 2024, Valentine et al 2022); all three were cross-sectional studies. The remaining eight studies did not include an in-scope comparator group. The included studies were conducted in Australia (one study) or the US (10 studies) and, where reported, were conducted between 2007 and 2022. It is therefore unclear how generalisable the findings might be to current NHS settings.

Baseline characteristics of all CYP with gender incongruence who identify as a male gender and wish a binary physical transition are included in Appendix E. Three studies stated that CYP were eligible for inclusion if they had a history of gender nonconformity or gender dysphoria, but six studies did not provide any definitions. The remaining two studies assessed participants based on the WPATH guidelines or ICD-10 and ICD-11 criteria. Sample sizes were small for most studies, ranging from 26 to 2,766 individuals. The mean age of participants ranged from 15.24 to 17.04 years, or a median age between 15.5 and 17.0 years. Some studies included an unknown proportion of out-of-scope participants (ie older than 18 years of age), and although most of the studies included a proportion of participants aged 15 years or younger, only two studies reported the circumstances under which these children received testosterone monotherapy for gender transition. One study stated that subcutaneous testosterone is initiated around the age of 14 years or older, following multidisciplinary assessment of readiness in youth with gender dysphoria and who have no contraindications. The second study stated that the minimum age for inclusion in the study was eight years in order to ensure that potential participants who might be eligible for hormones based on their Tanner stage would not be excluded due to age alone.

Some studies included a small proportion of individuals who identified as non-binary (2.5% to 10%), other studies did not clearly report this information, whilst one study reported that non-binary CYP were excluded from the analysis (Kramer et al 2024). Six studies mentioned the use of GnRH analogues prior to commencing testosterone in a small proportion of participants (6.9% to 17%). Details on other demographic characteristics indicated a higher proportion of individuals (for the whole study population) identifying as white (ranging between 72% and 89.7%). Limited details were provided by the studies in relation to comorbidities.

Seven studies provided details on testosterone formulations and/or doses, which varied considerably within and across studies. Where reported, participants received testosterone cypionate, enanthate, undecanoate, or mixed esters and methods included subcutaneous or intramuscular injections, topical gel, or transdermal testosterone. Only two studies reported the duration of treatment, Grimstad et al (2021) reported a mean treatment duration of 28.5 (standard deviation [SD] 14.6) months in participants without breakthrough bleeding, and 37.3 (SD 16.9) months in participants with breakthrough bleeding. Mullins et al (2021) reported a median treatment duration of 577 days (interquartile range [IQR] 283 to 923). Two studies provided limited details on monitoring of CYP receiving testosterone monotherapy. One study reported that individuals were seen once a year in clinic by a multidisciplinary team, but during the interim, could be seen by endocrinologists, nurses, and psychologists individually for follow-up care. The second study stated that all participants were followed by medical and mental health clinicians, but no further details were provided.

The use of concomitant treatments was reported in eight studies. Baines et al (2023) reported that a small proportion of participants received GnRH analogues during the study period, but it was not clear whether this was administered for menstrual suppression or other purposes. Three studies (Grimstad et al 2021, Millington et al 2024, Moussaoui et al 2024) reported the use of menstrual suppression and/or contraception in varying forms. Mullins et al (2021) stated that one participant with a history of venous thromboembolism prior to taking testosterone monotherapy received anticoagulation (rivaroxaban) during the study period. One study each reported the use of stimulation medication or antipsychotics but did not provide further details. The remaining studies did not report on concurrent treatments. It is therefore not possible to determine what impact treatments such as psychological or psychosocial interventions for gender incongruence, or specific treatments for depression, anxiety or eating disorders (where these were the outcomes reported), or familial support, may have had on the outcomes. The study reporting on breakthrough bleeding (Grimstad et al 2021) reported the number of participants who underwent hysterectomy after the age of 18 years, and the study reporting cardiometabolic-related diagnoses (Valentine et al 2022)

stated that gender-affirming surgical procedures had not been evaluated, but the remaining studies did not mention gender-affirming surgery. It was therefore unclear whether study participants may have undergone surgery in these remaining studies, but for the purposes of inclusion in this review, it was assumed to be a minority of the population.

Comparator evidence for clinical effectiveness was available for two critical outcomes. Only one study (Grannis et al 2023) reported reasons for participants not receiving testosterone, which included lack of parental consent, deferring treatment in favour of fertility preservation, uncertainty of interest in pursuing treatment, not interested in treatment, or interested but not yet started treatment. One cross-sectional study (Grannis et al 2023) compared the impact on gender incongruence and mental health (using validated questionnaires) at a mean treatment duration of 12.87 (SD 9.94) months among 21 individuals treated with testosterone monotherapy compared to 29 individuals who had not received treatment with testosterone. The findings indicated statistically significant improvements in gender incongruence and symptoms of anxiety and depression in individuals treated with testosterone monotherapy, but no statistically significant difference in suicidality. A second cross-sectional study (Kramer et al 2024) reported no statistically significant differences in eating disorder behaviours (including subjective and objective binge episodes, self-induced vomiting, laxative use, and compensatory exercise) in 39 individuals receiving testosterone monotherapy compared to 142 individuals not receiving testosterone. The treatment/follow-up duration was not reported for this study. No comparator evidence was identified for the remaining critical outcomes or for any important outcomes. Comparator evidence for safety was provided by one cross-sectional study (Valentine et al 2022), which reported a statistically significant higher odds of overweight/obesity, dyslipidaemia, hypertension, and liver function in TGD youth receiving testosterone compared to those not receiving testosterone, but no statistically significant difference in the odds of dysglycaemia or PCOS. It was unclear whether these cardiometabolic-related diagnoses preceded or followed the initial testosterone prescription. In addition, it was unclear how many individuals were included in each group and the treatment/follow-up duration was not reported.

Non-comparator evidence for clinical effectiveness was available for three critical outcomes (impact on gender incongruence, mental health and QoL) and two important outcomes (masculinising physical changes and detransition after receipt of masculinising medicines). One retrospective case series (Allen et al 2019) reported a reduction in suicidality and an increase in wellbeing from baseline to at least three months follow-up among 33 individuals receiving testosterone monotherapy, but no statistical measures were reported. One prospective case series (Baines et al 2023) indicated a reduction in QoL in individuals who received testosterone monotherapy at three, six months, and nine months follow-up, but

statistical measures were only reported for six months follow-up, which indicated no statistically significant difference compared to baseline. Baines et al (2023) also reported perceived physical changes at three, six and nine months follow-up, but statistical measures were only reported at six months, which indicated a statistically significant improvement compared to three months follow-up. Two retrospective case series also reported on masculinising physical changes. Grimstad et al (2021) reported that, of 232 individuals receiving testosterone monotherapy, a greater proportion (75%) did not experience breakthrough bleeding. Persky et al (2024) reported no statistically significant difference in final adult height or final adult height Z-scores from baseline to an unknown treatment/follow-up duration in 62 individuals receiving testosterone monotherapy. One retrospective case series (Laurenzano et al 2021) reported that of 119 individuals, three discontinued treatment with testosterone; two discontinued treatment due to satisfaction with outcomes and one discontinued due to concerns about body change and potential impact on fertility.

Five non-comparator studies provided evidence on safety. Mullins et al (2021) reported that none of the 429 individuals receiving testosterone monotherapy experienced venous thromboembolism or arterial thrombosis (including stroke) at a mean follow-up duration of 577 days (IQR 283 to 923). Although the authors acknowledged that not all participants had reached target physiologic hormone levels during the study which means that not all safety events may have been captured. Baines et al (2023) reported that a small proportion of individuals experienced skin irritation at initiation of testosterone monotherapy (11.5%) and elevated liver enzymes (7.7%) up to nine months follow-up, but none of the participants developed elevated haemoglobin or haematocrit, or significant dyslipidaemia. Millington et al (2024) reported a statistically significant increase in haemoglobin and haematocrit levels from baseline to six, 12 and 24 months follow-up in 136 individuals, but no statistically significant differences in HbA1c, ALT AST at any timepoint. A small proportion of participants included in Millington et al (2024) were reported to have elevated laboratory measures or diabetes at baseline. Moussaoui et al (2024) reported that 23.4% of individuals receiving testosterone monotherapy reported pelvic pain, but it was unclear whether this pre-existed prior to testosterone use and the authors acknowledged that the follow-up duration may not have been sufficient to capture the onset and development of pelvic pain. Laurenzano et al (2021) reported a statistically significant increase in haematocrit levels from baseline to the final dose of testosterone monotherapy in 119 individuals, but no statistically significant difference in total cholesterol, ALT or AST. The same study reported that 64.7% of individuals experienced progression of acne, a small proportion of participants experienced mild injection site reactions (11.8%) or dyslipidaemia (0.8%), while none of the participants reported experiencing hypertension, transaminitis, or haematocrit greater than 55%. Laurenzano et al (2021) was the only study to assess the impact of testosterone dose (<160

mg, 160 to 240 mg and >240 mg) on safety outcomes (total cholesterol, haematocrit, ALT and AST levels). Statistically significant increases were indicated for final haematocrit (p=0.024), and the difference between baseline and final AST (p=0.01) as the testosterone dose increased, and a borderline statistically significant difference was reported for final ALT (p=0.05).

The outcomes reported were primarily assessed using standard assessment tools (with the exception of perceived physical changes and pelvic pain outcomes which used unvalidated questionnaires). The use of standardised outcome measures allows some interpretation of the level of burden associated with specific scores; however, it was not clear how clinically significant the changes observed were. No specific detail about what the minimal clinically important thresholds or differences might be was reported for the outcomes considered.

All of the included studies were judged to be at high risk of bias and the outcomes reported were assessed as very low certainty evidence when evaluated using modified GRADE. Of particular note was the lack of reporting of psychological or psychosocial support / interventions in both the intervention and control populations; this increases the risk of bias through potential confounding. Eight of the papers included in this evidence review were uncontrolled observational studies, which are subject to bias and confounding. Furthermore, three of the included studies (Grannis et al 2023, Kramer et al 2024 and Valentine et al 2022), were cross-sectional studies which limit the assessment of causality as both the outcomes and the exposure (in this case testosterone monotherapy) are measured at the same point in time, it is impossible to determine temporality, and these types of studies can only demonstrate associations; the cross-sectional studies were the only studies to include comparator evidence. Additionally, for the three comparator studies, there was very serious risk of bias due to limited reporting of study eligibility criteria, unclear measurement of the exposure, or the use of subjective outcome measures (ie self-reported) which can result in bias. There was very serious risk of bias in the three non-comparator studies due to the lack of clinical information for participants and/or measurement of gender incongruence, use of subjective outcome measures (ie self-reported), use of an unvalidated questionnaire in one study, and/or lack of statistical measures.

There was a high degree of indirectness in all of the included studies due to the inclusion of out-of-scope participants for this evidence review (ie non-binary individuals or aged greater than 18 years when commencing masculinising medicines), and/or inclusion of out-of-scope interventions (ie GnRH analogues or gender affirming surgery). The majority of studies did not report treatment duration, and some studies did not involve or report a follow-up duration, which means that it is not possible to determine the impact of testosterone over time. Where

follow-up durations were reported, these ranged between three and 24 months, which may not have been sufficient to capture some of the outcomes measured. Further limitations include the use of electronic data in some studies which may have limited reliability and may have meant that not all patients, treatments or outcomes were identified, and loss to follow-up in some studies, where this was reported. In addition, most of the studies included at least some (or all) data from more than 10 years ago, and all of the included studies were conducted outside the UK (10 studies were conducted in paediatric gender clinics in the United States of America and one study was conducted in a single paediatric gender clinic in Australia). It is therefore unclear whether the populations and aspects of gender affirming care in the included studies reflect that seen in clinical practice in England, and the generalisability of the findings to the NHS may therefore be limited and should be interpreted with caution.

No evidence on cost-effectiveness was identified and none of the included studies reported relevant subgroup analyses.

7. Conclusion

This evidence review includes 11 studies. One study excluded non-binary CYP while the remaining studies included both binary and non-binary CYP or did not report this information. Three cross-sectional studies provided comparator evidence for individuals taking testosterone monotherapy versus individuals not taking testosterone. The remaining eight studies did not include an in-scope comparator. All evidence was of very low certainty.

Included studies provided evidence for the critical outcomes 'impact on gender incongruence', 'impact on mental health', 'impact on quality of life', and the important outcomes 'masculinising physical changes', 'detransition after receipt of masculinising medicines', and safety. No studies were identified that reported the important outcomes 'psychosocial impact', 'fertility', 'feasibility of masculinising genital surgery', 'cognitive outcomes', or 'regret after receipt of masculinising medicines'.

One comparator study provided evidence that there was a statistically significant improvement in gender incongruence and mental health (including anxiety and depression) in individuals who received testosterone monotherapy compared to individuals not receiving testosterone, but there was no statistically significant difference in suicidality. One comparator study provided evidence that there was no statistically significant difference in eating disorder behaviours (including subject and objective binge episodes, self-induced vomiting, laxative use, or compensatory exercise) among individuals receiving testosterone monotherapy compared to those not receiving treatment. Only two non-comparator studies reported statistical measures for critical and important outcomes: one study reported no statistically significant change in quality of life from three to six months follow-up, but a statistically significant improvement in perceived physical changes at six months follow-up; the second study reported no statistically significant difference in final adult height at an unknown timepoint.

Safety was reported across six studies (one comparator study and five non-comparator studies). The comparator study reported statistically significant higher odds of overweight/obesity, dyslipidaemia, hypertension, and liver function in individuals receiving testosterone compared to those not receiving testosterone, but no statistically significant difference in the odds of dysglycaemia or PCOS. One non-comparator study reported no occurrence of venous thromboembolism or arterial thrombosis (including stroke) among 429 individuals receiving testosterone monotherapy. One prospective case series reported that a small proportion of individuals experienced skin irritation or elevated liver enzymes, but none of the participants developed elevated haemoglobin or haematocrit, or significant dyslipidaemia. One retrospective case series reported statistically significant increases in

haemoglobin and haematocrit levels in 136 individuals (the direction of effect was unclear), but no statistically significant differences in HbA1c, ALT or AST at any timepoint. Another retrospective case series reported that 23.4% of individuals receiving testosterone monotherapy experienced pelvic pain. The remaining retrospective case series reported a statistically significant increase in haematocrit from baseline to the final dose of testosterone monotherapy in 119 individuals (the direction of effect was unclear), but no statistically significant differences in total cholesterol, ALT, or AST. The same study reported that 64.7% of individuals experienced progression of acne, a small proportion of participants experienced mild injection site reactions or dyslipidaemia, while none of the participants reported experiencing hypertension, transaminitis, or haematocrit greater than 55%.

All outcomes reported were assessed as very low certainty evidence when evaluated using modified GRADE. Of particular note was the lack of reporting of psychological or psychosocial support / interventions in both the intervention and control populations; this increases the risk of bias through potential confounding. Three of the included studies, all of the studies with comparator evidence, were cross-sectional studies which limit the assessment of causality; since both the outcomes and the exposure (in this case testosterone monotherapy) are measured at the same point in time, it is impossible to determine temporality, and these type of studies can only demonstrate associations. Further limitations include the limited reporting on treatment and/or follow-up duration, the use of electronic data, and loss to follow-up in some studies. In addition, most of the studies included at least some data from more than 10 years ago, and all of the included studies were conducted outside the UK (10 studies were conducted in paediatric gender clinics in the United States of America and one study was conducted in a paediatric gender clinic in Australia). It is therefore unclear whether the populations and aspects of gender affirming care in the included studies reflect that seen in clinical practice in England, and the generalisability of the findings to the NHS may therefore be limited and should be interpreted with caution.

No evidence on cost-effectiveness was identified and none of the included studies reported relevant subgroup analyses.

Overall, there is very low certainty evidence with inconsistent results for the selected outcomes in CYP who identify as a male gender and wish a binary physical transition receiving testosterone monotherapy. There is also a lack of direct evidence due to the mixed populations and mixed interventions which limits the conclusions that can be drawn. No conclusions can be drawn about cost-effectiveness as no evidence was identified. Published

studies which allow conclusions to be drawn about the effectiveness of testosterone monotherapy for this population are needed.

Appendix A PICO document

The review questions for this evidence review are:

1. For CYP with gender incongruence who identify as a male gender and wish a binary physical transition, what is the clinical effectiveness of treatment with testosterone monotherapy with or without psychological and psychosocial support compared with one or a combination of psychological support or social transitioning to the desired gender or with no intervention?
2. For CYP with gender incongruence who identify as a male gender and wish to undergo a binary physical transition, what is the short-term and long-term safety of testosterone monotherapy with or without psychological and psychosocial support compared with one or a combination of psychological support or social transitioning to the desired gender or with no intervention?
3. For CYP with gender incongruence who identify as a male gender and wish to undergo a binary physical transition, what is the cost-effectiveness of testosterone monotherapy with or without psychological and psychosocial support compared to one or a combination of psychological support or social transitioning to the desired gender or with no intervention?
4. From the evidence selected, are there particular sub-groups of CYP with gender incongruence who identify as a male gender and wish to undergo a binary physical transition that may benefit more from treatment with testosterone monotherapy than the wider population?
5. From the evidence selected:
    a. What were the criteria used by the research studies to define gender incongruence?
    b. What were the starting criteria, formulation, duration and dose of testosterone monotherapy for those aged 16 years up to their 18th birthday?
    c. Did any children aged 15 years or younger receive testosterone monotherapy for gender transition? If so, in what circumstances?
    d. What monitoring was in place for CYP with gender incongruence who identify as a male gender and wish to undergo a binary physical transition receiving testosterone monotherapy?
    e. What were the exclusion criteria in the studies?

P –Population and Indication	Children and young people (up to their 18th birthday) who have gender incongruence as defined by the study and identify as a male gender and wish to undergo a binary physical transition. [Some terms used to describe this population include, but are not limited to, female to male (FTM; F2M), gender queer, transperson, transmasculine, transmale, transmasc, transman, transgender, transsexual, trans-sex, trans*, cross gender or cross-sex (alternate spellings may be considered). The term gender incongruence may also be referred to as, but is not limited to, gender dysphoria, gender identity disorder, gender dysfunction, gender diverse, gender questioning or transsexualism. 'Gender incongruence of childhood' is a diagnostic term used by health professionals, found in the WHO International Classification of Diseases ICD-11 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 (WHO, 2024). Gender variant behaviour and preferences alone are not a basis for assigning the diagnosis. 'Gender incongruence of adolescence or adulthood' is a diagnostic term used by health professionals, found in the WHO International Classification of Diseases ICD-11. Gender incongruence is characterised by "a marked and persistent incongruence between an individual's experienced gender and the assigned sex". It is important to note that it has been moved out of the "Mental and behavioural disorders" chapter and into the "Conditions related to sexual health" chapter so that it is not perceived as a mental health disorder. It does not include references to dysphoria or dysfunction. Gender dysphoria, within the section of gender identity disorders, is the term used in Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, Text Revision (DSM-5-TR) (American Psychiatric Association, 2022). In the DSM-5-TR definition, gender dysphoria has to be associated with clinically significant distress or impairment of function. Gender dysphoria is the more commonly used term clinically and among research papers. It is also most likely to be familiar to the lay public since it has been used widely in mainstream and social media. It is a label that is used colloquially to describe feelings, as well as being a formal diagnosis.] The following subgroups of CYP with gender incongruence are of interest: • Peri-pubertal vs post-pubertal • The stated duration of gender incongruence is either less than 6 months, 6- 24 months or more than 24 months at time of assessment and/or treatment • The age of onset of gender incongruence • The age of onset of puberty • The age/ Tanner stage at which treatment was initiated with testosterone monotherapy

	• CYP with gender incongruence who have a preexisting diagnosis of neurodiversity • CYP with gender incongruence who have a preexisting diagnosis of a learning disability • CYP with gender incongruence with a history of severe enduring mental disorder including anxiety, depression (with or without a history of self-harm and suicidality), psychosis, personality disorder, and eating disorders. •
I – Intervention	Masculinising medicines comprising testosterone monotherapy. Individuals taking masculinising medicines may also be receiving psychological or psychosocial support. [Masculinising medicines may be referred to as gender affirming hormones, cross sex hormones, sex reassignment, sex change, sex transformation, sex hormones, gender reassignment, gender change, gender transformation or gender hormones. Testosterone can be given an as intramuscular injection (IM), oral tablet or applied as a gel. Examples include: testosterone gel (Tostran, Testogel, Testim, Testavan), short-acting intramuscular injections such as testosterone propionate, phenypropionate, isocaproate and decanoate (Sustanon), testosterone enantate (Delatestryl) and long-acting injection testosterone undecanoate (Nebido, Roxadin, Aveed), oral testosterone capsules in the form of testosterone undecanoate (Restandol Testocaps, Andriol testocaps, Jatenzo, Kyzatrex, Tlando). Individuals may also have experienced a period of time or process known as ‘real-life experience’ (RLE), sometimes historically called ‘real-life test’ (RLT) where they have lived full-time in their identified gender role in order to be eligible for masculinising medicines. This PICO excludes individuals who are receiving or have received GnRH analogues for the indication of puberty suppression or gender affirmation.]
C – Comparator(s)	One or a combination of: 1. Psychological and psychosocial support 2. Social transitioning to the gender with which the individual identifies **OR** 3. No intervention [Psychological and psychosocial support include cognitive behavioural therapy (CBT), Psychoanalytic and Psychodynamic therapies, Humanistic and Existential Therapies, Interpersonal and Relational Therapies, Trauma-Focused Therapies, Arts and Expressive Therapies, mindfulness and self-compassion, attachment-based family therapy, attachment therapy, psychoeducation, gender exploratory therapy, exploratory therapy. • Examples of Cognitive and Behavioural Therapies include: Cognitive Behavioural Therapy (CBT), Dialectical Behaviour Therapy (DBT), Acceptance and Commitment Therapy (ACT), Exposure Therapy, Behaviour Therapy Examples of Psychoanalytic and Psychodynamic Therapies include: Psychoanalysis, Psychodynamic Therapy, Intensive short-

	term dynamic psychotherapy (ISTDP), sensorimotor psychotherapy • Examples of Humanistic and Existential Therapies include: Person-Centered Therapy (Carl Rogers), Gestalt Therapy, Existential Therapy • Examples of Interpersonal, Relational and Systemic Therapies include: Interpersonal Therapy (IPT), Couples Therapy, Family Therapy, Group Therapy, Narrative Therapy, Mentalisation-based Therapy, Dyadic Developmental Psychotherapy (DDP), Narrative exposure therapy • Examples of Trauma-Focused Therapies include: Eye Movement Desensitization and Reprocessing (EMDR), Trauma-Focused CBT (TF-CBT) • Examples of Mindfulness-Based Therapies include: Mindfulness-Based Stress Reduction (MBSR), Mindfulness-Based Cognitive Therapy (MBCT) • Examples of Arts and Expressive Therapies include: Art Therapy, Music Therapy, Drama Therapy, Play-based Therapy, Theraplay • Examples of Integrative and Holistic Therapies include: Integrative Therapy, integrative counselling • Examples of Specialised Therapies include: Compassion-Focused Therapy (CFT), Schema Therapy, Solution-Focused Brief Therapy (SFBT). Psychosocial support also includes: assessment, extended assessment, therapeutic assessment. These longer assessments allow exploration at a deeper level to seek understanding. Interventions can be delivered by psychological practitioners including Clinical and Counselling Psychologists, Psychotherapists, other healthcare professionals with additional training and supervision (e.g., specialist nurse or therapeutic social worker), trained facilitators or counsellors. Interventions can be delivered face to face or online, individually or in groups. Duration of intervention can range from a single session to having no fixed duration or number of sessions. No intervention may include individuals who actively choose not to take any interventions.]
O – Outcomes	**<u>Clinical Effectiveness</u>** *There are no known minimal clinically important differences and there are no preferred timepoints for the outcome measures selected.* *<u>Critical to decision-making:</u>* • **Impact on gender incongruence** *This outcome is important to patients because gender incongruence is associated with significant distress and problems functioning.* [This outcome may be measured using the Utrecht Gender Dysphoria Scale (UGDS), Gender Dysphoria Questionnaire, Gender Identity Interview for Adolescents and Adults, Gender Identity Interview for Children, Gender Distress Scale (TYC-GDS), Self-reported

	satisfaction. Other measures (including self-reported) may be used as an alternative to the stated measures.] • **Impact on mental health** *This outcome is important to patients because gender incongruence is associated with psychological distress which can lead to the development of mental health problems.* [Examples of mental health problems include self-harm, thoughts of suicide, suicide attempts, suicide, eating disorders, depression/low mood, anxiety, psychotic symptoms/psychosis, substance abuse, minority stress and trauma. This outcome may be measured using Child Behaviour Checklist (CBCL), Youth Self Report (YSR), Childhood Global Assessment Scale (CGAS), Revised Children's Anxiety and Depression Scale (and Subscales) (RCADS), The Child and Adolescent Psychiatric Assessment (CAPA), ED-15-Y eating disorder measure, Depression Anxiety Stress Scales (DASS-Y), Patient health questionnaire (PHQ-9) Modified for Teens, Beck Depression Inventory for Youth (BDI-Y), Beck Depression Inventory-II (BDI-II), Quick Inventory of Depressive Symptoms [QIDS], Generalised Anxiety Disorder Questionnaire (GAD-7), Hospital Anxiety and Depression Scale (HADS), Screen for Child Anxiety Related Emotional Disorders (SCARED), Ask Suicide Screening Questions (ASQ), Suicide Ideation Questionnaire Junior, Children's Rosenberg Self-Esteem Scale (CRSES), Clinical Outcomes in Routine Evaluation (CORE), Child Revised Impact of Events Scale 8 or 13 (CRIES 8 or 13), Dissociative Experiences Scale (DES), Assessment Checklist for Adolescents (ACA), Assessment Checklist for Children (ACC). Other measures (including self-reported) may be used as an alternative to the stated measures.] • **Impact on Quality of Life** *This outcome is important to patients because gender incongruence may be associated with a significant reduction in health-related quality of life.* [Quality of life can be measured using a recognised quality of life score for example KINDL questionnaire, Kidscreen 10/27/52, Pediatric Quality of Life Inventory (PedsQL), EuroQuality of Life Five Dimensions Youth (EQ-5D-Y/EQ-5D-3L/EQ-5D-5L), Satisfaction with Life Scale for Children (SWLS-C), Quality of Life Enjoyment and Satisfaction Questionnaire (QLES-Q-SF), General Well-Being Scale (GWBS). Other measures (including self-reported) may be used as an alternative to the stated measures.] *<u>Important to decision making:</u>* • **Masculinising physical changes** *This outcome is important because most patients with gender incongruence wish 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.* [Masculinising physical changes can include: menstrual cycling, facial/body/head hair, body shape, voice changes, sexual and genital effects.

	Measures can include The Children's Body Image Scale (CBIS), Body Image Scale for Children (BISC), Body Dysmorphia scale YBOCS, Yale-Brown Obsessive Compulsive Scale Modified for Body Dysmorphic Disorder (BD D-YBO CS). Other measures (including self-reported) may be used as an alternative to the stated measures.] • **Psychosocial impact** *This outcome is important to patients because gender incongruence is associated with internalising and externalising behaviours and emotional and behavioural problems which may impact on social and occupational functioning.* [Examples of psychosocial impact are coping mechanisms (such as substance misuse) which may impact on family relationships; peer relationships, living arrangements, educational attendance, work participation, romantic involvement, prosocial skills. Measures that may be used are The Work and Social Adjustment Scale – Youth versions (WSAS-Y), Strengths and Difficulties Questionnaire (SDQ), Multidimensional Scale of Perceived Social Support (MSPSS), Inventory of Interpersonal Problems (IIP32), Family Adaptability, Partnership, Growth, Affection and Resolve test. Other measures (including self-reported) may be used as an alternative to the stated measures.] • **Fertility** *This outcome is important to patients because masculinising medicines can reduce fertility. Prior to commencing masculinising medicines patients should be counselled on the impact of treatment on their fertility and offered fertility preservation options.* [Examples of fertility outcomes include, but are not limited to ovulation, pregnancy as well as pregnancy outcomes.] • **Feasibility of masculinising genital surgery** *This outcome is important to patients because masculinising medicines can have an impact on surgical outcomes. Treatment may alter the amount of genital tissue available for phalloplasty, metoidioplasty, hysterectomy and bilateral salpingo-oophorectomy.* • **Cognitive outcomes** *This outcome is important to patients because feminising medicines can negatively impact cognitive processes such as concentration, memory, and executive function.* [Observations and cognitive testing are performed by a trained professional which may include a key worker, support worker, social care, social worker or through school observations. This might include assessment of visuospatial ability, verbal memory, verbal fluency, verbal reasoning, verbal comprehension, visual memory, working memory, processing speed, computation, motor coordination, executive functioning, timed task completion or cognitive flexibility. Measures can include Wechsler Intelligence Scale for Children (WISC), Wechsler Adult Intelligence Scale (WAIS), Adaptive

	Behaviours Assessment System (ABAS) or Wechsler Preschool and Primary Scale of Intelligence (WPPSI).] • **Detransition after receipt of masculinising medicines** *Medical detransition is a complex experience encompassing medical, psychological, social implications and is important to patients because they may choose to discontinue treatment. The decision to detransition may or may not be associated with regret.* [Detransitioning is a concept that has evolved over time. Older studies may incorporate terminology relating to retransition. Relevant terms in the literature may include: detransitioner, desistence, discontinuation, cessation, termination, reversion, reversal, disidentification, reidentification.] • **Regret after receipt of masculinising medicines** *This outcome is important to patients because some patients who choose to take masculinising medicines may regret this decision. Regret may or may not be associated with detransition.* [This may be expressed as a proportion of the study population or other measures such as documentation of regret or semi-structured interviews.] **<u>Safety</u>** *It 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.* • Aspects to be reported could include: o Of most importance: Thromboembolic disease, cardiovascular events, polycythaemia, pulmonary oil microembolism. o Pre-diabetes (glycosylated haemoglobin (HbA1c) 42mmol/mol – 47mmol/mol, 6% vs 6.4%) or diabetes (HbA1c ≥48mmol/mol, ≥6.5%) and for those with diabetes, worsening control e.g. increase in HbA1c despite treatment or as defined in study, anaemia, breast, ovarian or endometrial cancer, migraine, seizures, impaired liver function, sleep apnoea, sexually transmitted infections, gynaecomastia, skin reactions, severe acne. **<u>Cost-effectiveness</u>**
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	Up to 18 years

Date limits	2005-2025
Exclusion criteria	
Publication type	Conference abstracts, non-systematic reviews, narrative reviews, commentaries, letters, editorials, pre-prints and guidelines
Study design	Case reports, resource utilisation studies

Appendix B Search strategy

Medline, Embase, PsycINFO, and the Cochrane Library were searched limiting the search to papers published in English language in the last 20 years. Searches were not limited by hormone type (masculinising / feminising) or final transition goals (binary transition or non-binary transition); this was to ensure that the widest selection of papers were included in the search. Conference abstracts, non-systematic reviews, narrative reviews, commentaries, letters, editorials, pre-prints, guidelines, and case reports were excluded.

Search dates: 01 January 2005 to 04 June 2025

Medline search strategy

1	adolescent/ or young adult/ or child/
2	adolescent health/ or child health/
3	Transition to Adult Care/
4	Pediatrics/
5	Puberty/
6	(child* or school* or p?ediatric* or adolescen* or preadolescen* or teen* or preteen* or young or youth? or girl? or boy? or puberty or pubescen*).ti,ab,kf.
7	or/1-6
8	Gender Dysphoria/
9	gender identity/ or transsexualism/
10	gender-nonconforming persons/ or transgender persons/
11	(gender adj2 (incongruen* or dysphoria* or dysfunction* or identit* or divers* or question*)).ti,ab,kf.
12	(trans or transgender* or transsex* or transperson* or transwom?n or transfem* or crossgender* or cross gender* or cross sex* or crosssex* or mtf or m2f or queer*).ti,ab,kf.
13	or/8-12
14	(femini?ing adj2 (drug? or medicine? or medication? or agent? or hormone?)).ti,ab,kf.
15	((gender* adj2 (affirm* or reassign* or re-assign* or transform* or transition* or chang*)) and (drug? or medicine? or medication? or agent? or hormone?)).ti,ab,kf.
16	(gender adj2 (drug? or medicine? or medication? or agent? or hormone?)).ti,ab,kf.
17	((sex adj2 (affirm* or reassign* or re-assign* or transform* or transition* or chang*)) and (drug? or medicine? or medication? or agent? or hormone?)).ti,ab,kf.
18	((cross-sex adj hormon*) or (hormon* adj (therap* or treatment? or "use" or usage or supplement*))).ti,ab,kf.
19	Hormone Replacement Therapy/ or Estrogen Replacement Therapy/

20 Estrogens/tu

21 estradiol/tu

22 Ethinyl Estradiol/

23 (oestrogens or estrogens).ti,kf.

24 ((oestrogen? or estrogen?) adj3 (drug? or medicine? or medication? or agent? or therap* or treatment? or "use" or usage or supplement*)).ti,ab,kf.

25 ((oestrogen? or estrogen?) adj3 (oral* or buccal* or sublingual* or sub-lingual* or pellet? or implant* or patch* or spray* or gel? or cream? or dermal* or transdermal or subcutaneous or sub-cutaneous or inject* or intramuscular or intra-muscular)).ti,ab,kf.

26 (oestradiols or estradiols or ethinylestradiols or oestriols or estriols).ti,kf.

27 ((oestradiol or estradiol or ethinylestradiol or oestriol or estriol) adj3 (drug? or medicine? or medication? or agent? or therap* or treatment? or "use" or supplement*)).ti,ab,kf.

28 ((oestradiol or estradiol or ethinylestradiol or oestriol or estriol) adj3 (oral* or buccal* or sublingual* or sub-lingual* or pellet? or implant* or patch* or spray* or gel? or cream? or dermal* or transdermal or subcutaneous or sub-cutaneous or inject* or intramuscular or intra-muscular)).ti,ab,kf.

29 (zumenon or delestrogen* or sandrena or oestrogel or evorel or estradot or oestraderm or estraderm or progynova or ts patch* or femseven or fem seven or lenzetto or estraor or Elleste Solo or Bedol).ti,ab,kf.

30 or/14-29

31 7 and 13 and 30

32 (animal or rat or rats or mice or mouse or murine or rodent? or cows or heifers or sheep or ewes or goats or pigs or cats or dogs).ti.

33 31 not 32

34 limit 33 to (english language and yr="2005 -Current")

35 (comment or editorial or letter or preprint or review).pt. or case report.ti.

36 34 not 35

37 ("systematic review" or scoping review).pt. or "Systematic Reviews as Topic"/ or ("Cochrane Database of Systematic Reviews" or evidence report technology assessment or evidence report technology assessment summary).jn. or (((((comprehensive or comprehensively) adj (analysis or review or reviewed)) or ((literature or scoping) adj (search or searches))).ti,ab,kf,kw. not "narrative review".ti.) and (database or databases or cinahl or cochrane or embase or psycinfo or pubmed or medline or scopus or (web adj1 science) or ((bibliographic or literature) adj (review or reviews)) or (((electronic adj (database or databases)) or (databases adj3 searched)) and (eligibility or excluded or exclusion or included or inclusion))).ti,ab,kf,kw.) or (((comparative adj effectiveness) and (effectiveness adj review)) or ((critical adj interpretive) and ((interpretive adj review) or (interpretive adj synthesis)))).ti,ab,kf,kw. or ((diagnostic adj test) and ((accuracy adj review) or (accuracy adj reviews) or (accuracy adj studies) or (accuracy adj study)) and (meta-analysis or scoping or systematic)).ti,ab,kf,kw. or ((evidence adj assessment) and

GRADE).ti,ab,kf,kw. or ((evidence adj2 gap) and (gap adj map)).ti,ab,kf,kw. or ((evidence adj mapping) or (evidence adj review) or (exploratory adj review) or (framework adj synthesis) or (mapping adj review)).ti,ab,kf,kw. or ((meta adj (epidemiological or ethnographic or ethnography or interpretation or narrative or review or study or synthesis or summary or theory)) or metaethnographic or metaethnography or metasynthesis).ti,ab,kf,kw. or ((methodological or methodology) adj1 review).ti,ab,kf,kw. or ((mixed adj methods) and (methods adj1 (review or synthesis))).ti,ab,kf,kw. or ((narrative adj1 synthesis) or (overview adj4 reviews) or ("PRISMA" adj4 (guideline or guidelines or preferred or reporting or requirements)) or (PRISMA adj "P")).ti,ab,kf,kw. or (((prognostic or psychometric) adj1 review) or ((qualitative adj (evidence or research)) and ((evidence or research) adj synthesis))).ti,ab,kf,kw. or (((rapid adj evidence) and (evidence adj assessment)) or (rapid adj realist) or (rapid adj2 (review or reviews)) or (realist adj2 (review or reviews or syntheses or synthesis))).ti,ab,kf,kw. or (((review adj economic) and (economic adj1 (evaluation or evaluations))) or ((scoping or systematic) adj2 (review or reviews or studies or study))).ti,ab,kf,kw. or ((review adj1 reviews) or ((systematic adj evidence) and (evidence adj map)) or (systematic adj2 mapping) or (systematic adj2 literature) or (systematic adj2 (Embase or Medline or PsycInfo or PubMed)) or (systematic adj2 (review or reviews)) or ((systematical or systematically) adj2 (review or reviewed reviews)) or (systematically adj identified) or (systematized adj review) or (umbrella adj (review or reviews))).ti,ab,kf,kw. or "Meta-Analysis".pt. or "meta-analysis as topic"/ or (meta adj2 (analyse or analyser or analyses or analysis or analytic or analytical or analytics or analyze or analyzed or analyzes)).ti,ab,kf,kw. or (metaanalyse or Metaanalysen or metaanalyser or metaanalyses or metaanalysis* or metaanalytic or metaanalytical or metaanalytics or metaanalyze or metaanalyzed or metaanalyzes).ti,ab,kf,kw. or "network meta-analysis"/ or (network adj1 (meta or metaanalyses or metaanalysis or metaregression)).ti,ab,kf,kw. or (systematic and ((meta adj regression) or metagression)).ti,ab,kf,kw.

38 34 and 37

39 36 or 38

40 Gender Dysphoria/

41 gender identity/ or transsexualism/

42 gender-nonconforming persons/ or transgender persons/

43 (gender adj2 (incongruen* or dysphoria* or dysfunction* or identit* or divers* or question*)).ti,ab,kf.

44 (trans or transgender* or transsex* or transperson* or transm?n or transmale? or transmasc* or crossgender* or cross gender* or cross sex* or crosssex* or ftm or f2m or queer*).ti,ab,kf.

45 or/40-44

46 (masculini?ing adj2 (drug? or medicine? or medication? or agent? or hormone?)).ti,ab,kf.

47 ((gender* adj2 (affirm* or reassign* or re-assign* or transform* or transition* or chang*)) and (drug? or medicine? or medication? or agent? or hormone?)).ti,ab,kf.

48 (gender adj2 (drug? or medicine? or medication? or agent? or hormone?)).ti,ab,kf.

49 ((sex adj2 (affirm* or reassign* or re-assign* or transform* or transition* or chang*)) and (drug? or medicine? or medication? or agent? or hormone?)).ti,ab,kf.

50 ((cross-sex adj hormon*) or (hormon* adj (therap* or treatment? or "use" or usage or supplement*))).ti,ab,kf.

51 Hormone Replacement Therapy/

52 exp Testosterone/tu

53 (testosterone adj3 (drug? or medicine? or medication? or agent? or therap* or treatment? or "use" or usage or supplement*)).ti,ab,kf.

54 (testosterone adj3 (capsule? or tablet? or oral* or buccal* or sublingual* or sub-lingual* or pellet? or implant* or patch* or spray* or gel? or cream? or dermal* or transdermal or subcutaneous or sub-cutaneous or inject* or intramuscular or intra-muscular)).ti,ab,kf.

55 (testosterone adj (isocaproate or undecanoate or enantate)).ti,ab,kf.

56 (tostran or testogel or testavan or sustanon or Testim or Delatestryl or Nebido or Roxadin or Aveed or Restandol Testocaps or Andriol testocaps or Jatenzo or Kyzatrex or Tlando).ti,ab,kf.

57 or/46-56

58 7 and 45 and 57

59 (animal or rat or rats or mice or mouse or murine or rodent? or cows or heifers or sheep or ewes or goats or pigs or cats or dogs).ti.

60 58 not 59

61 limit 60 to (english language and yr="2005 -Current")

62 (comment or editorial or letter or preprint or review).pt. or case report.ti.

63 61 not 62

64 ("systematic review" or scoping review).pt. or "Systematic Reviews as Topic"/ or ("Cochrane Database of Systematic Reviews" or evidence report technology assessment or evidence report technology assessment summary).jn. or (((((comprehensive or comprehensively) adj (analysis or review or reviewed)) or ((literature or scoping) adj (search or searches))).ti,ab,kf,kw. not "narrative review".ti.) and (database or databases or cinahl or cochrane or embase or psycinfo or pubmed or medline or scopus or (web adj1 science) or ((bibliographic or literature) adj (review or reviews)) or (((electronic adj (database or databases)) or (databases adj3 searched)) and (eligibility or excluded or exclusion or included or inclusion))).ti,ab,kf,kw.) or (((comparative adj effectiveness) and (effectiveness adj review)) or ((critical adj interpretive) and ((interpretive adj review) or (interpretive adj synthesis)))).ti,ab,kf,kw. or ((diagnostic adj test) and ((accuracy adj review) or (accuracy adj reviews) or (accuracy adj studies) or (accuracy adj study)) and (meta-analysis or scoping or systematic)).ti,ab,kf,kw. or ((evidence adj assessment) and GRADE).ti,ab,kf,kw. or ((evidence adj2 gap) and (gap adj map)).ti,ab,kf,kw. or ((evidence adj mapping) or (evidence adj review) or (exploratory adj review) or (framework adj synthesis) or (mapping adj review)).ti,ab,kf,kw. or ((meta adj (epidemiological or ethnographic or ethnography or interpretation or narrative or review or study or synthesis or summary or theory)) or metaethnographic or metaethnography or metasynthesis).ti,ab,kf,kw. or ((methodological or methodology) adj1 review).ti,ab,kf,kw. or ((mixed adj methods) and (methods adj1 (review or synthesis))).ti,ab,kf,kw. or ((narrative adj1 synthesis) or (overview adj4 reviews) or ("PRISMA" adj4 (guideline or guidelines or preferred or reporting or

requirements)) or (PRISMA adj "P")).ti,ab,kf,kw. or (((prognostic or psychometric) adj1 review) or ((qualitative adj (evidence or research)) and ((evidence or research) adj synthesis))).ti,ab,kf,kw. or (((rapid adj evidence) and (evidence adj assessment)) or (rapid adj realist) or (rapid adj2 (review or reviews)) or (realist adj2 (review or reviews or syntheses or synthesis))).ti,ab,kf,kw. or (((review adj economic) and (economic adj1 (evaluation or evaluations))) or ((scoping or systematic) adj2 (review or reviews or studies or study))).ti,ab,kf,kw. or ((review adj1 reviews) or ((systematic adj evidence) and (evidence adj map)) or (systematic adj2 mapping) or (systematic adj2 literature) or (systematic adj2 (Embase or Medline or PsycInfo or PubMed)) or (systematic adj2 (review or reviews)) or ((systematical or systematically) adj2 (review or reviewed reviews)) or (systematically adj identified) or (systematized adj review) or (umbrella adj (review or reviews))).ti,ab,kf,kw. or "Meta-Analysis".pt. or "meta-analysis as topic"/ or (meta adj2 (analyse or analyser or analyses or analysis or analytic or analytical or analytics or analyze or analyzed or analyzes)).ti,ab,kf,kw. or (metaanalyse or Metaanalysen or metaanalyser or metaanalyses or metaanalysis* or metaanalytic or metaanalytical or metaanalytics or metaanalyze or metaanalyzed or metaanalyzes).ti,ab,kf,kw. or "network meta-analysis"/ or (network adj1 (meta or metaanalyses or metaanalysis or metaregression)).ti,ab,kf,kw. or (systematic and ((meta adj regression) or metagression)).ti,ab,kf,kw.

65 61 and 64

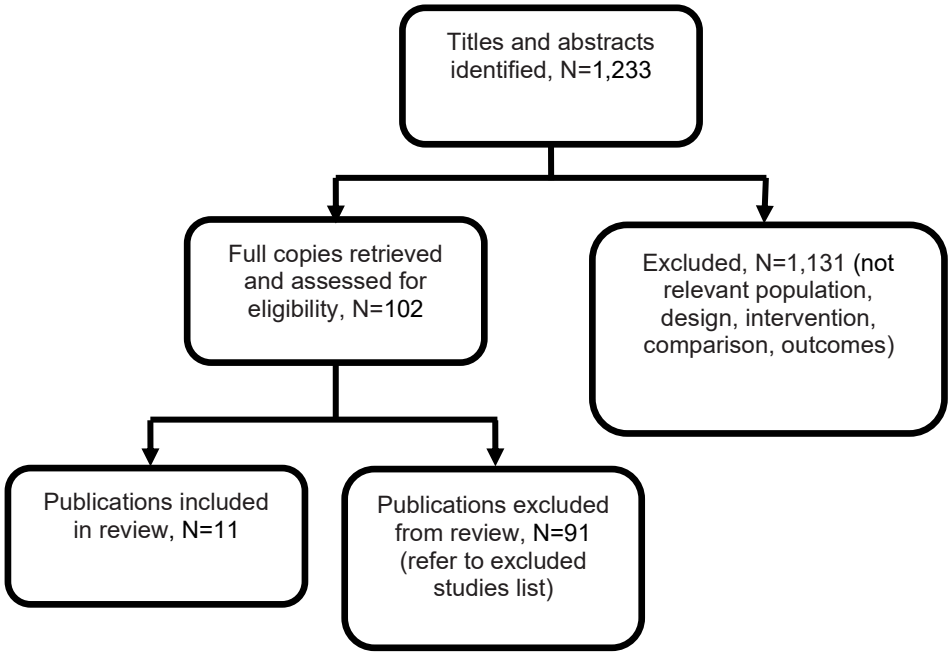
66 63 or 65

67 39 or 66

Appendix C Evidence selection

The literature searches identified 1,233 references. These were screened using their titles and abstracts and 102 references were obtained in full text and assessed for relevance. Of these, 11 references are included in the evidence summary. The remaining 91 references were excluded and are listed in Appendix D.

Figure 1- Study selection flow diagram



References submitted with Preliminary Policy Proposal

Not applicable.

Appendix D Excluded studies table

Study reference	Reason for exclusion
Achille C, Taggart T, Eaton NR, Osipoff J, Tafuri K, Lane A, et al. Longitudinal impact of gender-affirming endocrine intervention on the mental health and well-being of transgender youths: preliminary results. Int J Pediatr Endocrinol. 2020;2020:8.	Higher level evidence identified reporting on QoL.
Alaniz VI, Sheeder JL, Whitmore GT, Wilde MD, Hutchens KJ, Nokoff NJ, et al. Menstrual Suppression in Adolescent and Young Adult Transgender Males. J Pediatr Adolesc Gynecol. 2023;36(2):116-21.	Retrospective cohort study reporting median time to cessation of menses in individuals receiving testosterone monotherapy. Number (%) of participants experiencing cessation of menses not reported. Outcome out-of-scope.
Arnoldussen M, Hooijman EC, Kreukels BP, de Vries AL. Association between pre-treatment IQ and educational achievement after gender-affirming treatment including puberty suppression in transgender adolescents. Clinical Child Psychology and Psychiatry. 2022;27(4):1069-76.	Intervention out-of-scope as all participants received GnRH analogues, followed by GAHT, followed by surgery.
Avila JT, Golden NH, Aye T. Eating Disorder Screening in Transgender Youth. J Adolesc Health. 2019;65(6):815-7.	Insufficient information to confirm whether outcomes are reported for the in-scope population receiving testosterone monotherapy; unclear whether the four of seven individuals AFAB receiving hormonal treatment received GnRH analogues or testosterone.
Baker KE, Wilson LM, Sharma R, Dukhanin V, McArthur K, Robinson KA. Hormone Therapy, Mental Health, and Quality of Life Among Transgender People: A Systematic Review. J. 2021;5(4):bvab011.	Systematic review including in-scope and out-of-scope studies. No meta-analysis or pooled results. In-scope studies have been included separately in the evidence review.
Baldassarre M, Giannone FA, Foschini MP, Battaglia C, Busacchi P, Venturoli S, et al. Effects of long-term high dose testosterone administration on vaginal epithelium structure and estrogen receptor-alpha and -beta expression of young women. Int J Impot Res. 2013;25(5):172-7.	Mean age of population >18 years. Population out-of-scope.
Baram S, Myers SA, Yee S, Librach CL. Fertility preservation for transgender adolescents and young adults: a systematic review. Hum Reprod Update. 2019;25(6):694-716.	Systematic review including in-scope and out-of-scope studies. No meta-analysis or pooled results. In-scope studies have been included separately in the evidence review.
Baskaran C, Roberts SA, Barrera E, Pilcher S, Kumar R. Venous thromboembolism in transgender and gender non-binary youth is rare and occurs in the setting of secondary risk factors: A retrospective cohort study. Pediatr Blood Cancer. 2024;71(11):e31284.	Population experiencing thrombotic events out-of-scope due to age (ie >18 years) or not receiving testosterone monotherapy.

Study reference	Reason for exclusion
Becker I, Auer M, Barkmann C, Fuss J, Moller B, Nieder TO, et al. A Cross-Sectional Multicenter Study of Multidimensional Body Image in Adolescents and Adults with Gender Dysphoria Before and After Transition-Related Medical Interventions. Arch Sex Behav. 2018;47(8):2335-47.	Outcomes not reported separately for in-scope population receiving testosterone alone.
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):1755-67.	Intervention out-of-scope as participants received GnRH analogues with or without hormones and surgery, or no treatment.
Borger O, Perl L, Yackobovitch-Gavan M, Sides R, Brener A, Segev-Becker A, et al. Body Composition and Metabolic Syndrome Components in Transgender/Gender Diverse Adolescents and Young Adults. LGBT health. 2024;11(5):359-69.	Intervention out-of-scope as all participants had received previous treatment with GnRH analogues.
Boskey ER, Scheffey KL, Pilcher S, Barerra EP, McGregor K, Carswell JM, et al. A Retrospective Cohort Study of Transgender Adolescents' Gender-Affirming Hormone Discontinuation. J Adolesc Health. 2025;76(4):584-91.	Insufficient information to be able to confirm that the intervention was testosterone monotherapy without prior use of GnRH analogues.
Burke SM, Kreukels BP, Cohen-Kettenis PT, Veltman DJ, Klink DT, Bakker J. Male-typical visuospatial functioning in gynephilic girls with gender dysphoria - organizational and activational effects of testosterone. J Psychiatry Neurosci. 2016;41(6):395-404.	Intervention out-of-scope as all AFAB individuals had received GnRH analogues prior to receiving testosterone.
Butler G, Adu-Gyamfi K, Clarkson K, El Khairi R, Kleczewski S, Roberts A, et al. Discharge outcome analysis of 1089 transgender young people referred to paediatric endocrine clinics in England 2008-2021. Arch Dis Child. 2022;107(11):1018-22.	Insufficient information to confirm the proportion of in-scope participants receiving in-scope intervention. Outcomes not reported separately for in-scope population receiving testosterone monotherapy.
Campos-Munoz L, Lopez-De Lara D, Rodriguez-Rojo ML, Conde-Taboada A, Lopez-Bran E. Transgender adolescents and acne: A cases series. Pediatr Dermatol. 2018;35(3):e155-e8.	Study design out-of-scope - case report (n=5 AFAB individuals). Study focus on use of medications to treat acne and the outcome of acne treatment.
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.	Outcomes not reported separately for in-scope population receiving testosterone monotherapy.

Study reference	Reason for exclusion
Carrillo N, McGurran M, Melton BL, Moeller KE. Comparison of inpatient psychiatric medication management in gender diverse youth with cisgender peers. Ment. 2023;13(4):169-75.	Insufficient information to confirm intervention in-scope as GAHT use on hospital admission included GnRH analogues and/or hormones. Insufficient information to confirm whether the n=21 participants taking GAHT were in-scope AFAB individuals.
Catlow C, Goffin S, Cunningham V, Abraham A, Grant C. The Health Needs and Management of Young People Accessing Paediatric Hauora Tahine (Transgender Health) Services in Te Tai Tokerau. J Paediatr Child Health. 2025;23:23.	Intervention out-of-scope as all participants had received previous treatment with GnRH anlogues.
Chelliah P, Lau M, Kuper LE. Changes in Gender Dysphoria, Interpersonal Minority Stress, and Mental Health Among Transgender Youth After One Year of Hormone Therapy. J Adolesc Health. 2024;74(6):1106-11.	Reports results for participants treated with masculinising medicines and patients treated with feminising medicines together. Outcomes not reported separately for in-scope population.
Chen D, Abrams M, Clark L, Ehrensaft D, Tishelman AC, Chan YM, et al. Psychosocial Characteristics of Transgender Youth Seeking Gender-Affirming Medical Treatment: Baseline Findings From the Trans Youth Care Study. J Adolesc Health. 2021;68(6):1104-11.	Out-of-scope as mental health demographics reported at baseline only, no follow-up data reported.
Chen D, Berona J, Chan YM, Ehrensaft D, Garofalo R, Hidalgo MA, et al. Psychosocial Functioning in Transgender Youth after 2 Years of Hormones. N Engl J Med. 2023;388(3):240-50.	Majority of outcome data not reported separately for in-scope population; subgroup analysis comparing gender incongruence and mental health in AFAB individuals initiating GAHT in early vs late puberty only provides baseline scores - no follow-up outcome data following treatment with GAHT.
Chew D, Anderson J, Williams K, May T, Pang K. Hormonal Treatment in Young People With Gender Dysphoria: A Systematic Review. Pediatrics. 2018;141(4):04.	Systematic review including in-scope and out-of-scope studies. No meta-analysis or pooled results. In-scope studies have been included separately in the evidence review.
Chiniara LN, Bonifacio HJ, Palmert MR. Characteristics of Adolescents Referred to a Gender Clinic: Are Youth Seen Now Different from Those in Initial Reports? Horm Res Paediatr. 2018;89(6):434-41.	Insufficient information to confirm whether the outcomes reported are for the in-scope population receiving testosterone monotherapy (>50% of AFAB youth were not receiving testosterone).
Chu L, Gold S, Harris C, Lawley L, Gupta P, Tangpricha V, et al. Incidence and Factors Associated With Acne in Transgender Adolescents on Testosterone: A Retrospective Cohort Study. Endocr Pract. 2023;29(5):353-5.	Larger, higher level evidence identified reporting on incidence of acne.
Crabtree L, Connelly KJ, Guerriero JT, Battison EAJ, Tiller-Ormord J, Sutherland SM, et al. A More Nuanced Story: Pediatric Gender-Affirming Healthcare is Associated With Satisfaction and Confidence. J Adolesc Health. 2024;75(5):772-9.	Out-of-scope as outcomes not reported separately for participants receiving testosterone monotherapy (58% had undergone surgery).

Study reference	Reason for exclusion
Cuellar-Flores I, Martin-Bejarano M, Garcia-Ron A, Villanueva S, Arias-Vivas E, Lopez-de Lara D. Androgen treatment effects on neurocognition in female-to-male transgender adolescents. Rev Neurol. 2024;78(3):83-9.	Non-English language paper (Spanish) - out-of-scope.
Dopp AR, Peipert A, Buss J, De Jesus-Romero R, Palmer K, Lorenzo-Luaces L. Interventions for Gender Dysphoria and Related Health Problems in Transgender and Gender-Expansive Youth: A Systematic Review of Benefits and Risks to Inform Practice, Policy, and Research. Rand health q. 2025;12(2):2.	Systematic review including in-scope and out-of-scope studies. No meta-analysis or pooled results. In-scope studies have been included separately in the evidence review.
Elkadi J, Chudleigh C, Maguire AM, Ambler GR, Scher S, Kozlowska K. Developmental Pathway Choices of Young People Presenting to a Gender Service with Gender Distress: A Prospective Follow-Up Study. Children (Basel). 2023;10(2):07.	Reports results for participants treated with masculinising medicines and patients treated with feminising medicines together; and insufficient information to confirm whether testosterone monotherapy used. Outcomes not reported separately for in-scope population.
Feigerlova E. Prevalence of detransition in persons seeking gender-affirming hormonal treatments: a systematic review. J Sex Med. 2025;22(2):356-68.	Systematic review including in-scope and out-of-scope studies. No meta-analysis or pooled results. In-scope studies have been included separately in the evidence review.
Glintborg D, Moller JK, Rubin KH, Lidegaard O, T'Sjoen G, Larsen MJO, et al. Gender-Affirming Treatment and Employment Rate in 3812 Danish Transgender Persons and 38 120 Controls. J Clin Endocrinol Metab. 2024;109(12):3076-86.	Population out-of-scope (median age >18 years). Outcome out-of-scope (employment levels).
Glintborg D, Moller JK, Rubin KH, Lidegaard O, T'Sjoen G, Larsen MJO, et al. Gender-affirming treatment and mental health diagnoses in Danish transgender persons: a nationwide register-based cohort study. Eur. 2023;189(3):336-45.	Population out-of-scope - median age 19 (range 15 to 24) years; outcomes not reported separately for in-scope individuals
Grannis C, Leibowitz SF, Gahn S, Nahata L, Morningstar M, Mattson WI, et al. Testosterone treatment, internalizing symptoms, and body image dissatisfaction in transgender boys. Psychoneuroendocrinology. 2021;132:105358.	To avoid reporting duplication of participant outcomes, the expansion paper (Grannis et al 2023) is included in the evidence review as it includes n=8 additional participants.
Green AE, DeChants JP, Price MN, Davis CK. Association of Gender-Affirming Hormone Therapy With Depression, Thoughts of Suicide, and Attempted Suicide Among Transgender and Nonbinary Youth. J Adolesc Health. 2022;70(4):643-9.	Population out-of-scope as majority nonbinary (63%). Outcomes not reported separately for in-scope population.

Study reference	Reason for exclusion
Gupta P, Patterson BC, Chu L, Gold S, Amos S, Yeung H, et al. Adherence to Gender Affirming Hormone Therapy in Transgender Adolescents and Adults: A Retrospective Cohort Study. J Clin Endocrinol Metab. 2023;108(11):e1236-e44.	Intervention out-of-scope as >50% of whole paediatric cohort had received previous GnRH analogues (unclear what proportion of the cohort were binary individuals AFAB).
Hisle-Gorman E, Schvey NA, Adirim TA, Rayne AK, Susi A, Roberts TA, et al. Mental Healthcare Utilization of Transgender Youth Before and After Affirming Treatment. J Sex Med. 2021;18(8):1444-54.	Outcomes not reported separately for in-scope population.
Hranilovich JA, Millington K. Headache prevalence in transgender and gender diverse youth: A single-center case-control study. Headache. 2023;63(4):517-22.	Insufficient information to confirm population in-scope as mean or median age not reported.
Insogna IG, Ginsburg E, Srouji S. Fertility preservation for adolescent transgender male patients: A case series. J Adolesc Health. 2020;66(6):750-3.	Study design out-of-scope - case report (n=4 AFAB individuals).
Jarin J, Pine-Twaddell E, Trotman G, Stevens J, Conard LA, Tefera E, et al. Cross-Sex Hormones and Metabolic Parameters in Adolescents With Gender Dysphoria. Pediatrics. 2017;139(5).	Higher level, comparator evidence identified reporting safety outcomes.
Jensen RK, Jensen JK, Simons LK, Chen D, Rosoklija I, Finlayson CA. Effect of Concurrent Gonadotropin-Releasing Hormone Agonist Treatment on Dose and Side Effects of Gender-Affirming Hormone Therapy in Adolescent Transgender Patients. Transgend Health. 2019;4(1):300-3.	Higher level, comparator evidence identified reporting safety outcomes.
Kain EJ, Fuqua JS, Eugster EA. A Retrospective Study of the Use of Gonadotropin-Releasing Hormone Analogs and Testosterone in Transgender Boys: Who, What, When, and for How Long? Transgend Health. 2024;9(4):357-60.	Retrospective cohort study reporting dose of testosterone at which menses stopped and dose at which menses continued, and additional masculinising physical change. Number (%) of participants experiencing outcomes not reported. Outcome out-of-scope.
Kaltiala R, Heino E, Tyolajarvi M, Suomalainen L. Adolescent development and psychosocial functioning after starting cross-sex hormones for gender dysphoria. Nord J Psychiatry. 2020;74(3):213-9.	Population out-of-scope as mean age >18 years. Outcomes not reported separately for in-scope population receiving testosterone monotherapy.
Karakilic Ozturan E, Ozturk AP, Bas F, Erdogdu AB, Kaptan S, Kardelen AI AD, et al. Endocrinological Approach to Adolescents with Gender Dysphoria: Experience of a Pediatric Endocrinology Department in a Tertiary Center in Turkey. J Clin Res Pediatr Endocrinol. 2023;15(3):276-84.	Intervention out-of-scope as >50% received GnRH analogues or testosterone plus surgery. Outcomes not reported separately for in-scope population and larger, higher level evidence identified reporting this outcome.

Study reference	Reason for exclusion
Karalexi MA, Georgakis MK, Dimitriou NG, Vichos T, Katsimpris A, Petridou ET, et al. Gender-affirming hormone treatment and cognitive function in transgender young adults: a systematic review and meta-analysis. Psychoneuroendocrinology. 2020;119:104721.	Systematic review including out-of-scope studies due to populations aged >18 years.
Khatchadourian K, Amed S, Metzger DL. Clinical management of youth with gender dysphoria in Vancouver. J Pediatr. 2014;164(4):906-11.	Insufficient information to confirm whether outcomes are reported for the in-scope population receiving testosterone monotherapy.
Knaus S, Steininger J, Klinger D, Riedl S. Body Mass Index Distributions and Obesity Prevalence in a Transgender Youth Cohort - A Retrospective Analysis. J Adolesc Health. 2024;75(1):127-32.	Intervention out-of-scope as all participants had received previous treatment with GnRH analogues.
Kuper LE, Stewart S, Preston S, Lau M, Lopez X. Body Dissatisfaction and Mental Health Outcomes of Youth on Gender-Affirming Hormone Therapy. Pediatrics. 2020;145(4):04.	Outcomes not reported separately for in-scope population receiving testosterone monotherapy.
Lee MK, Yih Y, Willis DR, Fogel JM, Fortenberry JD. The impact of gender affirming medical care during adolescence on adult health outcomes among transgender and gender diverse individuals in the United States: The role of state-level policy stigma. LGBT health. 2024;11(2):111-21.	Population out-of-scope (mean age >18 years).
Ludvigsson JF, Adolfsson J, Hoistad M, Rydelius PA, Kristrom B, Landen M. A systematic review of hormone treatment for children with gender dysphoria and recommendations for research. Acta Paediatr. 2023;112(11):2279-92.	Systematic review including in-scope and out-of-scope studies. No meta-analysis or pooled results. In-scope studies have been included separately in the evidence review.
Ma J, Ackley D, III, Reback CJ, Rusow JA, Skeen SJ, Miller-Perusse M, et al. Psychosocial correlates of gender-affirming hormone and medically necessary surgical intervention use among transgender and gender diverse youth and young adults. Psychology of Sexual Orientation and Gender Diversity. 2025:No Pagination Specified.	Population out-of-scope (mean age >18 years).
MacKinnon KR, Jeyabalan T, Strang JF, Delgado-Ron JA, Lam JS, Gould WA, et al. Discontinuation of gender-affirming medical treatments: Prevalence and associated features in a nonprobabilistic sample of transgender and gender-diverse adolescents and young adults in Canada and the United States. J Adolesc Health. 2024;75(4):569-77.	Population out-of-scope (mean age >18 years).

Study reference	Reason for exclusion
Martinez-Martin FJ, Kuzior A, Hernandez-Lazaro A, de Leon-Durango RJ, Rios-Gomez C, Santana-Ojeda B, et al. Incidence of hypertension in young transgender people after a 5-year follow-up: association with gender-affirming hormonal therapy. Hypertens Res. 2023;46(1):219-25.	Population out-of-scope (mean age >18 years).
Marwa A, Misra M, Lopez X. Determinants of Bone Mineral Density in Transgender Youth. Transgend Health. 2022;7(3):213-8.	Insufficient information to confirm whether outcomes are reported for the in-scope population receiving testosterone monotherapy.
Mattelin E, Strandell A, Bryman I. Fertility preservation and fertility treatment in transgender adolescents and adults in a Swedish region, 2013-2018. Hum. 2022;2022(2):hoac008.	Population out-of-scope as mean age >18 years.
McCallion S, Smith S, Kyle H, Shaikh MG, Wilkinson G, Kyriakou A. An appraisal of current service delivery and future models of care for young people with gender dysphoria. Eur J Pediatr. 2021;180(9):2969-76.	Outcomes not reported separately for in-scope population.
McFarlane T, Zajac JD, Cheung AS. Gender-affirming hormone therapy and the risk of sex hormone-dependent tumours in transgender individuals-A systematic review. Clin Endocrinol (Oxf). 2018;89(6):700-11.	Systematic review including in-scope and out-of-scope studies. No meta-analysis or pooled results. In-scope studies have been included separately in the evidence review.
Millington K, Barrera E, Daga A, Mann N, Olson-Kennedy J, Garofalo R, et al. The effect of gender-affirming hormone treatment on serum creatinine in transgender and gender-diverse youth: implications for estimating GFR. Pediatr Nephrol. 2022;37(9):2141-50.	Outcome out-of-scope – serum creatinine and implications for eGFR.
Millington K, Chan YM. Lipoprotein subtypes after testosterone therapy in transmasculine adolescents. J. 2021;15(6):840-4.	Median age of population >18 years. Population out-of-scope.
Miroshnychenko A, Ibrahim S, Roldan Y, Kulatunga-Moruzi C, Montante S, Couban R, et al. Gender affirming hormone therapy for individuals with gender dysphoria aged <26 years: a systematic review and meta-analysis. Arch Dis Child. 2025;110(6):437-45.	Systematic review including in-scope and out-of-scope studies. No meta-analysis or pooled results. In-scope studies have been included separately in the evidence review.
Nieder TO, Mayer TK, Hinz S, Fahrenkrug S, Herrmann L, Becker-Hebly I. Individual Treatment Progress Predicts Satisfaction With Transition-Related Care for Youth With Gender Dysphoria: A Prospective Clinical Cohort Study. J Sex Med. 2021;18(3):632-45.	Outcomes not reported separately for in-scope population.

Study reference	Reason for exclusion
Nokoff NJ, Scarbro SL, Moreau KL, Zeitler P, Nadeau KJ, Juarez-Colunga E, et al. Body Composition and Markers of Cardiometabolic Health in Transgender Youth Compared With Cisgender Youth. J Clin Endocrinol Metab. 2020;105(3):01.	Outcomes not reported separately for in-scope population. Larger studies identified reporting safety outcomes.
Nunes-Moreno M, Furniss A, Cortez S, Davis SM, Dowshen N, Kazak AE, et al. Mental Health Diagnoses and Suicidality Among Transgender Youth in Hospital Settings. LGBT health. 2025;12(1):20-8.	Population had received treatment with GnRH analogues or testosterone. Outcomes not reported separately for in-scope population
Nyquist CB, Torgersen L, David LW, Diseth TH, Gulbrandsen K, Waehre A. Treatment trajectories among children and adolescents referred to the Norwegian National Center for Gender Incongruence. Acta Paediatr. 2025;114(5):1006-14.	Population out-of-scope (mean age at GAHT initiation >18 years) and a proportion of individuals had received previous GnRH analogues or surgery. Outcomes not reported separately for in-scope population
Oda H, Kinoshita T. Efficacy of hormonal and mental treatments with MMPI in FtM individuals: Cross-sectional and longitudinal studies. BMC Psychiatry. 2017;17(1) (no pagination).	Mean age of population >18 years. Population out-of-scope.
Oliphant J, Barnett D, Veale J, Denny S, Farrant B. The wellbeing and health needs of a cohort of transgender young people accessing specialist medical gender-affirming healthcare in Auckland. N Z Med J. 2021;134(1541):33-44.	Population out-of-scope as mean age >18 years. Intervention out-of-scope as 49% of participants had received GnRH analogues, and a small proportion also underwent surgery. Outcomes not reported separately for in-scope population.
Olsavsky AL, Grannis C, Bricker J, Chelvakumar G, Indyk JA, Leibowitz SF, et al. Associations Among Gender-Affirming Hormonal Interventions, Social Support, and Transgender Adolescents' Mental Health. J Adolesc Health. 2023;72(6):860-8.	Reports results for participants treated with masculinising medicines and patients treated with feminising medicines together. Outcomes not reported separately for in-scope population.
Olson J, Schrager SM, Clark LF, Dunlap SL, Belzer M. Subcutaneous Testosterone: An Effective Delivery Mechanism for Masculinizing Young Transgender Men. LGBT health. 2014;1(3):165-7.	Mean age of population >18 years. Population out-of-scope.
Olson KR, Raber GF, Gallagher NM. Levels of Satisfaction and Regret With Gender-Affirming Medical Care in Adolescence. Jama, Pediatr. 2024;178(12):1354-61.	Intervention out-of-scope as all participants had received previous treatment with GnRH analogues.
Olson-Kennedy J, Okonta V, Clark LF, Belzer M. Physiologic Response to Gender-Affirming Hormones Among Transgender Youth. J Adolesc Health. 2018;62(4):397-401.	Population out-of-scope as mean age 18 years.

Study reference	Reason for exclusion
Olson-Kennedy J, Wang L, Wong CF, Chen D, Ehrensaft D, Hidalgo MA, et al. Emotional Health of Transgender Youth 24 Months After Initiating Gender-Affirming Hormone Therapy. J Adolesc Health. 2025;16:16.	Intervention out-of-scope as the majority of AFAB individuals (98%) had received GnRH analogues prior to receiving testosterone.
Pham AH, Eadeh HM, Garrison MM, Ahrens KR. A Longitudinal Study on Disordered Eating in Transgender and Nonbinary Adolescents. Acad Pediatr. 2023;23(6):1247-51.	Reports results for participants treated with masculinising medicines and patients treated with feminising medicines together. Outcomes not reported separately for in-scope population.
Reisner SL, Jadwin-Cakmak L, Sava L, Liu S, Harper GW. Situated Vulnerabilities, Sexual Risk, and Sexually Transmitted Infections' Diagnoses in a Sample of Transgender Youth in the United States. AIDS Patient Care STDS. 2019;33(3):120-30.	Population out-of-scope as mean age >18 years. Outcomes not reported separately for in-scope population receiving testosterone monotherapy.
Reisner SL, Vetters R, Leclerc M, Zaslow S, Wolfrum S, Shumer D, et al. Mental health of transgender youth in care at an adolescent urban community health center: a matched retrospective cohort study. J Adolesc Health. 2015;56(3):274-9.	Outcomes not reported separately for in-scope population taking testosterone monotherapy.
Roberts CM, Klein DA, Adirim TA, Schvey NA, Hisle-Gorman E. Continuation of Gender-affirming Hormones Among Transgender Adolescents and Adults. J Clin Endocrinol Metab. 2022;107(9):e3937-e43.	Population out-of-scope as majority of participants (60.9%) aged >18 years. Outcomes not reported separately for in-scope population <18 years.
Roy MK, Bothwell S, Kelsey MM, Ma NS, Moreau KL, Nadeau KJ, et al. Bone Density in Transgender Youth on Gender-Affirming Hormone Therapy. J. 2024;8(5):bvae045.	Larger, higher level evidence identified reporting safety outcomes.
Sanchez-Toscano E, Dominguez-Riscart J, Larran-Escandon L, Mateo-Gavira I, Aguilar-Diosdado M. Cardiovascular Risk Factors in Transgender People after Gender-Affirming Hormone Therapy. J. 2023;12(19) (no pagination).	Population and intervention out-of-scope. Median age of population 18 years and large proportion received GnRH analogues and/or surgery plus testosterone. Outcomes not reported separately for participants receiving testosterone monotherapy.
Segev-Becker A, Israeli G, Elkon-Tamir E, Perl L, Sekler O, Amir H, et al. Children and Adolescents with Gender Dysphoria in Israel: Increasing Referral and Fertility Preservation Rates. Endocr Pract. 2020;26(4):423-8.	Intervention out-of-scope as participants received GnRH analogues with or without hormones and surgery.
Sequeira GM, Kidd K, El Nokali NE, Rothenberger SD, Levine MD, Montano GT, et al. Early Effects of Testosterone Initiation on Body Mass Index in Transmasculine Adolescents. J Adolesc Health. 2019;65(6):818-20.	Retrospective chart review reporting BMI Z-score. Outcome out-of-scope.

Study reference	Reason for exclusion
Spanos C, Bretherton I, Zajac JD, Cheung AS. Effects of gender-affirming hormone therapy on insulin resistance and body composition in transgender individuals: A systematic review. World J Diabetes. 2020;11(3):66-77.	Systematic review including in-scope and out-of-scope studies. No meta-analysis or pooled results. In-scope studies have been included separately in the evidence review.
Strang JF, Chen D, Nelson E, Leibowitz SF, Nahata L, Anthony LG, et al. Transgender Youth Executive Functioning: Relationships with Anxiety Symptoms, Autism Spectrum Disorder, and Gender-Affirming Medical Treatment Status. Child Psychiatry Hum Dev. 2022;53(6):1252-65.	Reports results for participants treated with masculinising medicines and patients treated with feminising medicines together. Outcomes not reported separately for in-scope population.
Tack LJ, Craen M, Dhondt K, Vanden Bossche H, Laridaen J, Cools M. Consecutive lynestrenol and cross-sex hormone treatment in biological female adolescents with gender dysphoria: a retrospective analysis. Biol Sex Differ. 2016;7:14.	Intervention is out-of-scope - testosterone plus lynesterol.
Taillefer V, Kelley J, Marsolais S, Chiniara L, Chadi N. Expected vs. perceived effects of gender-affirming hormone therapy among transmasculine adolescents. J Pediatr Endocrinol Metab. 2023;36(11):1072-8.	Intervention out-of-scope as only 39.5% of individuals had not previously received GnRH analogues (42.1% had previously received GnRH analogues and it was unclear whether 18.4% of the population had previously received GnRH analogues as this was not specified).
Taylor J, Hall R, Langton T, Fraser L, Hewitt CE. Care pathways of children and adolescents referred to specialist gender services: a systematic review. Arch Dis Child. 2024;109(Suppl 2):s57-s64.	Systematic review including in-scope and out-of-scope studies. No meta-analysis or pooled results. In-scope studies have been included separately in the evidence review.
Thompson L, Sarovic D, Wilson P, Irwin L, Visnitchi D, Samfjord A, et al. A PRISMA systematic review of adolescent gender dysphoria literature: 3) treatment. PLOS Glob Public Health. 2023;3(8):e0001478.	Systematic review including in-scope and out-of-scope studies. No meta-analysis or pooled results. In-scope studies have been included separately in the evidence review.
Thoreson N, Grasso C, Potter J, King DS, Peebles JK, Dommasch ED. Incidence and Factors Associated With Androgenetic Alopecia Among Transgender and Gender-Diverse Patients Treated With Masculinizing Hormone Therapy. JAMA Dermatol. 2021;157(3):348-9.	Study type out-of-scope - letter.
Tollit MA, May T, Maloof T, Telfer MM, Chew D, Engel M, et al. The clinical profile of patients attending a large, Australian pediatric gender service: A 10-year review. Int J Transgend Health. 2023;24(1):59-69.	Insufficient information to be able to confirm the number of participants receiving testosterone monotherapy with no prior use of GnRH analogues.

Study reference	Reason for exclusion
Tordoff DM, Wanta JW, Collin A, Stepney C, Inwards-Breland DJ, Ahrens K. Mental Health Outcomes in Transgender and Nonbinary Youths Receiving Gender-Affirming Care. JAMA netw. 2022;5(2):e220978.	Insufficient information to confirm the proportion of AFAB binary participants receiving testosterone alone and outcomes not reported separately for in-scope population. Higher level evidence identified reporting mental health outcomes.
Turban JL, King D, Kobe J, Reisner SL, Keuroghlian AS. Access to gender-affirming hormones during adolescence and mental health outcomes among transgender adults. PLoS ONE. 2022;17(1):e0261039.	Reports results for participants treated with masculinising medicines and patients treated with feminising medicines together. Outcomes not reported separately for in-scope population.
Valentine A, Nokoff N, Bonny A, Chelvakumar G, Indyk J, Leibowitz S, et al. Cardiometabolic Parameters Among Transgender Adolescent Males on Testosterone Therapy and Body Mass Index-Matched Cisgender Females. Transgend Health. 2021;6(6):369-73.	Larger, higher level evidence identified reporting comparator data for cardiovascular outcomes.
van der Loos MA, Hellinga I, Vlot MC, Klink DT, den Heijer M, Wiepjes CM. Development of Hip Bone Geometry During Gender-Affirming Hormone Therapy in Transgender Adolescents Resembles That of the Experienced Gender When Pubertal Suspension Is Started in Early Puberty. J Bone Miner Res. 2021;36(5):931-41.	Intervention out-of-scope as all participants received GnRH analogues, followed by GAHT, and a small proportion also underwent surgery.
Van Donge N, Schvey NA, Roberts TA, Klein DA. Transgender Dependent Adolescents in the U.S. Military Health Care System: Demographics, Treatments Sought, and Health Care Service Utilization. Mil Med. 2019;184(5-6):e447-e54.	Reports results for participants treated with masculinising medicines and patients treated with feminising medicines together. Outcomes not reported separately for in-scope population.
Vehmas N, Holopainen E, Savolainen-Peltonen H. Metabolic and Anthropometric Changes and Adverse Effects in Finnish Adolescents Using Gender-Affirming Hormonal Treatment. Transgender Health. 2024.	Population out-of-scope (median age >18 years).
Abbreviations AFAB: assigned female at birth; BMI: body mass index; eGFR: estimated glomerular filtration rate; GAHT: gender-affirming hormone therapy; GnRH: gonadotropin-releasing hormone; n: number; QoL: quality of life.	

Appendix E Evidence table

For abbreviations see list after table

The language used in this table is that of the study authors and may not reflect current language used by NHS England or NHS Gender Identity Services.

Study details	Population	Interventions	Study outcomes	Appraisal and funding
Allen LR, Watson LB, Egan AM, Moser CN. Well-being and suicidality among transgender youth after gender-affirming hormones. Clin. 2019;7(3):302-11. **Study location** USA (single centre) **Study type** Retrospective cohort study[35] **Study aim** To assess the long-term effectiveness of puberty suppression medication and	TG youth **Inclusion criteria** TG youth (adolescents and young adults aged approximately 13 to 20 years) receiving care for gender dysphoria, with a final assessment at least 3 months after treatment initiation **Exclusion criteria** TG youth without a second follow-up at	**Interventions** Testosterone[38] **Comparators** No relevant comparator Use of concomitant treatments: NR Treatment duration (days) (whole population; AMAB and AFAB), mean (SD): 349 (193)	Follow-up: at least 3 months **Critical outcomes** **Impact on mental health** Suicidality measured using the ASQ at baseline and final assessment,[39] mean (SE) Baseline: 1.01 (0.23) At least 3 months: 0.29 (0.13); *p value not reported* **Impact on QoL** Wellbeing measured at baseline and final assessment,[40] mean (SE) At baseline: 64.95 (2.66) At least 3 months: 70.94 (2.35); *p value not reported*	This study was appraised using the JBI critical appraisal checklist for case series. 1. Yes 2. Yes 3. Yes 4. No 5. No 6. Yes 7. No 8. No

[35] Although this study was a retrospective cohort study, as the comparator group was out-of-scope for this evidence review (ie TG youth AMAB), it was treated as a case series.
[38] The authors reported that the endocrinologists in their clinic sometimes begin participants at hormone levels lower than the recommended protocol, and typically, patients' doses are gradually increased every three to six months so that the dosage levels recommended by suggested protocols are reached by the end of treatment.
[39] ASQ is a four-item measure used to identify patients who are at risk of attempting suicide. Questions include: In the past few weeks have you… "…wished you were dead?", "…felt that you or your family would be better off if you were dead?", "…been having thoughts about harming or killing yourself?", or "…done anything to hurt yourself or to end your life?" A response of "no" was scored as 0 and a response of "yes" was scored as 1; with an overall score for suicidality on a scale ranging from 0 to 4, with higher scores indicating greater levels of suicidal ideation.
[40] Wellbeing was measured using the Paediatric Quality of Life Inventory General Wellbeing Scale (PedsQL GWBS) which is a 5-point response scale, containing seven items, and measures "general well-being" and "general health". The general well-being subscale includes six items (eg "I feel happy" and "I think my health will be good in the future"). Participants are asked to consider each item over the past month and rate responses from 0 (never) to 4 (almost always). The general health subscale contains one item, "In general, how is your health?" ranging from 0 (Bad) to 4 (Excellent). All items are scored and linearly transformed to a 0 to 100 scale (initial score of 0 = 0, 1 = 25, 2 = 50, 3 = 75, and 4 = 100). Higher scores indicate perceptions of minimal problems, high wellbeing.

Study details	Population	Interventions	Study outcomes	Appraisal and funding
GAHT on psychological outcomes among TG youth **Study dates** Not stated	least 3 months after initiation of treatment **Total sample size** N=47 AFAB: n=33 (70.2%) **Baseline characteristics** *Baseline characteristics are for the whole population (AMAB and AFAB)* Age (years) at GAHT administration, mean (SD; range) 16.59[36] (1.19; 13.73 to 19.04)[37] Race/Ethnicity, n (%) White: 39 (83) Biracial or multiracial: 2 (4.3) Latinx or Hispanic: 3 (6.4) Black or African American: 1 (2.1) American Indian or Alaskan Native: 1 (2.1) Asian: 1 (2.1)			9. Yes 10. No **Other comments:** This was a small cohort study that was treated as a case series due to the lack of a relevant comparator group. The study was conducted in a single, multidisciplinary gender clinic in a children's hospital in the USA and included only participants with a final assessment that was at least three months after administration of treatment. The author described the services provided by the clinic and that TG youth were seen once a year by a multidisciplinary team, with interim visits to endocrinologists, nurses, and psychologists available for follow-up care. TG youth were included in the study after diagnostic evaluation by a mental health provider to identify youth with a history of gender nonconformity or gender dysphoria. The author did not describe the

[36] There was a discrepancy in the mean age reported in the main text (16.59) and table (16.50) and we have reported the figure presented in the main text.
[37] The authors reported that most participants (90%) were at or below the age of 18.01 years at initiation of GAHT.

Study details	Population	Interventions	Study outcomes	Appraisal and funding
				circumstances under which TG youth aged 15 years or younger received testosterone monotherapy for gender transition. The author acknowledged that they did not distinguish between binary and non-binary youth and it was unclear what proportion of participants may have been non-binary and therefore out-of-scope for this evidence review. Only limited details on clinical characteristics of included participants were reported. Eight participants were administered GnRH analogues prior to beginning GAHT but it was unclear what proportion were AFAB. The author identified duration of treatment as a confounding factor and included this in statistical analyses, but acknowledged that the study may have included other confounding factors that were not taken into consideration, such as the level of familial support (most participants in the study had some degree of parental support),

Study details	Population	Interventions	Study outcomes	Appraisal and funding
				differences in GAHT (eg dose), and whether participants were receiving other active treatments. Statistical analyses explored outliers, heterogeneity of variance, and missing data. No potential confounders were identified or controlled for in ANCOVA analyses. P values were not reported for sub-analyses (AFAB and GAH-only groups). No subgroup analyses were conducted. The author acknowledged that the comprehensive care provided by the experienced multidisciplinary team at their clinic, and the demographics of the study population (ie mainly white), means that the findings may not be generalisable to all TG youth. The generalisability of the findings to the UK is also limited. **Source of funding:** Not stated.
Baines HK, Connelly KJ. A prospective comparison study of subcutaneous and intramuscular testosterone injections in transgender	TGD adolescents AFAB **Inclusion criteria**	**Interventions** All participants initiated treatment with testosterone	Follow-up: 3, 6 and 9 months **Critical outcomes** **Impact on QoL**	This study was appraised using the JBI critical appraisal checklist for case series.

Study details	Population	Interventions	Study outcomes	Appraisal and funding
male adolescents. J Pediatr Endocrinol Metab. 2023;36(11):1028-36. **Study location** USA (single centre) **Study type** RCT[41] **Study aim** To compare the effectiveness of testosterone injection type on masculinising physical changes, QoL, and safety in TGD adolescents AFAB **Study dates** January 2018 to March 2019	Testosterone naïve TGD adolescents AFAB aged 14 to 18 years and commencing testosterone **Exclusion criteria** TGD adolescents AFAB with prior testosterone use **Total sample size** N=26 **Baseline characteristics** Age (years), median (IQR) 15.5 (15.0 to 17.0) Race, n (%) Caucasian: 21 (81) Hispanic or Latino: 4 (15) Unknown/NR: 1 (4) Weight (kg), median (IQR) 68.25 (58.5 to 82.2)	cypionate 200 mg/mL (participants transitioned to testosterone enanthate if skin irritation occurred) Testosterone method, n (%) Subcutaneous injection: starting dose (n=16) 14 mg every week; at 3 (n=14) to 6 (n=13) months 26 mg every week; at 6 to 9 months (n=3) 100 mg every 2 weeks Intramuscular injection: starting dose (n=10) 50 mg every 2 weeks; at 3 (n=9) to 6 (n=6) months 75 mg every 2 weeks; at 6 to 9 months (n=2) 40 mg every week **Comparators** No relevant comparator Use of concomitant treatments: five	QoL measured using PedsQL,[42] median (IQR) At baseline (n=24): 66.08 (60.87 to 80.98) At 3 months (n=20): 64.31 (54.97 to 76.63); *p value not reported* At 6 months (n=18): 61.41 (56.52 to 85.87); *difference in QoL from baseline to 6 months: p=0.71* At 9 months (n=5): 59.78 (42.39 to 66.30); *p value not reported* **Important outcomes** **Masculinising physical changes** Perceived physical changes using unvalidated masculinising effects questionnaire,[43] median (IQR) At 3 months (n=20):16.0 (12.5 to 20.0) At 6 months (n=18): 19.5 (17.0 to 23.0) *Difference in perceived physical changes from 3 to 6 months: p<0.001* At 9 months (n=5): 21.0 (20.0 to 22.0); *p value not reported* **Safety** Number of individuals reporting adverse effects, n (%) Skin irritation at initiation of treatment: 3 (11.5)[44] Elevated haemoglobin up to 9 months: 0 (0)	1. Yes 2. Unclear 3. Unclear 4. Yes 5. Yes 6. Yes 7. No 8. Yes 9. No 10. No **Other comments:** This was a small optional cross-over RCT that was treated as a case series due to the lack of a relevant comparator group. The study was conducted in a single centre in the USA, with patients recruited between 2018 and 2019. The authors did not describe the circumstances under which TGD adolescents AFAB aged 15 years or

[41] Although this was an RCT, it was treated as a case series as no relevant comparator group (as stated in the PICO document) was included in the study.
[42] Higher scores indicate better health-related quality of life.
[43] Masculinising effects questionnaire assessed perceived physical changes after starting testosterone, including skin oiliness/acne, facial hair, body hair, increased muscle mass/strength, menstrual cessation, and clitoral enlargement. Higher scores indicate greater perceived physical changes.
[44] Skin irritation in all three participants was following initiation of testosterone cypionate and all transitioned to testosterone enanthate.

Study details	Population	Interventions	Study outcomes	Appraisal and funding
	BMI (kg/m^2), median (IQR) 24.7 (20.6 to 30.1)	participants received GnRH analogues treatment (leuprolide-depot) during the study period Treatment duration: NR	Elevated haematocrit up to 9 months: 0 (0) Significant dyslipidaemia up to 9 months: 0 (0) Elevated ALT (≤60 U/L) at 6 months: 1 Elevated AST (≤36 U/L) at 3 months: 1	younger received testosterone monotherapy for gender transition or provide details on what monitoring was in place for TGD adolescents AFAB receiving testosterone monotherapy. TGD adolescents AFAB were included in the study but it was unclear what methods were used to identify individuals as TGD or whether any individuals identified as non-binary. Limited details were provided by the authors in relation to clinical information for the participants and the services provided by the clinic. Five participants opted to participate in the optional cross-over part of the study: n=2 changed from subcutaneous to intramuscular injections and n=3 change from intramuscular to subcutaneous injections. Although by the end of the cross-over period, n=4 participants opted for subcutaneous injections and n=1 participant changed back to intramuscular injections.

Study details	Population	Interventions	Study outcomes	Appraisal and funding
				Safety outcomes were measured using laboratory measures, QoL used a validated measure, but masculinising physical changes were measured using a non-validated measure which may impact on the validity of the findings. Two participants withdrew from the study due to skin irritation or insurance concerns and one participant was not included in the analysis due to delaying testosterone initiation. The authors reported that four participants were lost to follow-up, three of whom returned to the clinic to continue gender affirming care (but after the study period had ended) and the fourth participant had not returned to clinic. The authors acknowledged that the follow-up duration was only short-term, which means that this may not have been sufficient to capture the outcomes measured. No subgroup analysis was reported.

Study details	Population	Interventions	Study outcomes	Appraisal and funding
				The authors acknowledged that the starting dose of testosterone was lower than currently used in many paediatric clinics and the generalisability of the findings may be limited. In addition, the generalisability of the findings to the UK is limited. **Source of funding:** Paediatric Endocrine Society Rising Star Award, the OHSU Tartar Trust Fellowship, the OHSU Department of Paediatrics for biostatistics support, and OHSU Oregon Clinical & Translation Research Institute REDCap.
Grannis C, Mattson WI, Leibowitz SF, Nahata L, Chen D, Strang JF, et al. Expanding upon the relationship between gender-affirming hormone therapy, neural connectivity, mental health, and body image dissatisfaction. Psychoneuroendocrinology. 2023;156:106319.	TNB youth **Inclusion criteria** TNB youth with a diagnosis of gender dysphoria,[45] and aged between 9 and 21 years Participants must not have MRI contraindications (e.g.	**Interventions** Testosterone injections: n=21 **Comparators** No testosterone: [46] n=29 Use of concomitant treatments: NR	Follow-up: none **Critical outcomes** **Impact on gender incongruence** Gender dysphoria measured using the BID,[47] mean (SD) Testosterone: 92.29 (19.74) No testosterone: 103.14 (18.99)	This study was appraised using the JBI critical appraisal checklist for cross-sectional studies. 1. Yes 2. Yes 3. Yes 4. Yes

[45] The authors stated that gender dysphoria was diagnosed in participants after several months of assessments for mental health and gender dysphoria by licensed mental health care providers.
[46] Three individuals in the no testosterone group were receiving puberty blockers.
[47] Higher values indicate greater dissatisfaction with one's body image.

Study details	Population	Interventions	Study outcomes	Appraisal and funding
Study location USA (single centre) **Study type** Retrospective cross-sectional study **Study aim** To assess the effectiveness of GAHT on internalising symptoms, body image dissatisfaction, and activation patterns within the amygdala-prefrontal cortex circuit in TNB youth **Study dates** 2018 to 2022	braces, metal implants, etc) Participants were eligible for inclusion if they met *"diagnostic classification of gender dysphoria based on comprehensive mental health evaluation by a provider specialising in gender development"* **Exclusion criteria** Not stated **Total sample size** N=82 AFAB: n=50 (binary 45 [90%]; non-binary 1 [2%]; *"binary and non-binary"* 4 [8%]) **Baseline characteristics** *Baseline characteristics are for TNB youth AFAB* Age (years), mean (SD)	Treatment duration (months), mean (SD): 12.87 (9.94)	*Difference between groups (ANOVA): F-value 7.46 (df between groups: 1, df within groups: 47); p<0.01* **Impact on mental health** Symptoms of generalised anxiety, measured using SCARED,[48] mean (SD) Testosterone: 38.75 (17.41) No testosterone: 50.25 (14.12) *Difference between groups (ANOVA): F-value 7.76 (df between groups: 1, df within groups: 45); p<0.01* Symptoms of depression, measured using the CDI,[49] mean (SD) Testosterone: 13.38 (7.81) No testosterone: 18.34 (7.02) *Difference between groups (ANOVA): F-value 5.62 (df between groups: 1, df within groups: 47); p<0.05* Symptoms of social anxiety, measured using the LSAS,[50] mean (SD) Testosterone: 50.21 (22.34) No testosterone: 78.62 (31.41) *Difference between groups (ANOVA): F-value 14.8 (df between groups: 1, df within groups: 42); p<0.001*	5. No 6. Yes 7. Unclear 8. Yes **Other comments:** This was a small cross-sectional study conducted in a single clinic in the USA between 2018 and 2022. Limited details were provided on inclusion criteria and exclusion criteria were not stated. The authors did not describe the circumstances under which TNB youth aged 15 years or younger received testosterone monotherapy for gender transition or provide details on what monitoring was in place for TNB youth receiving testosterone monotherapy. Gender identify was self-reported and expressed as male, female, transgender, female to male transgender/FTM, male to female transgender/MTF, trans male/trans man, trans

[48] Higher values indicate greater generalised anxiety symptoms.
[49] Higher values indicate greater depression symptoms.
[50] Higher values indicate greater social anxiety.

Study details	Population	Interventions	Study outcomes	Appraisal and funding
	Testosterone: 17.04 (1.18) No testosterone: 15.24 (1.72) <u>Race, n (%)</u> Black or African American: Testosterone 3 (14.29); no testosterone 1 (3.45) Multiracial: Testosterone 3 (14.29); no testosterone 3 (10.34) Native American or American Indian: Testosterone 1 (4.76); no testosterone 1 (3.45) White: Testosterone 14 (66.67); no testosterone 22 (75.86) Prefer not to answer: Testosterone 0 (0); no testosterone 2 (6.90) <u>Ethnicity, n (%)</u> Hispanic/Latinx: Testosterone 4 (19.05); no testosterone 1 (3.44) Non-Hispanic/Latinx:		<u>Suicidality (defined as frequency of suicidal ideation and/or attempts in the past year), mean (SD)</u> Testosterone: 1.95 (0.92) No testosterone: 2.86 (1.46) *Difference between groups (ANOVA): F-value 3.05 (df between groups: 1, df within groups: 47); p=NS*	female/trans woman, trans*r/trans asterisk, genderqueer, gender nonconforming, gender variant, gender fluid, gender expansive, intersex, androgynous, non-binary, transsexual, cross-dresser, two-spirited, third gender, agender, not sure, other. A small proportion of individuals identified as non-binary or *"binary and non-binary"* (10%). Details on use or no use of testosterone were taken from electronic medical records. The authors stated that the reasons for not receiving testosterone included lack of parental consent, deferring GAHT in favour of fertility preservation, uncertainty of interest in pursuing GAHT, not interested in GAHT, interested but not yet started GAHT. The authors identified age as a confounding factor, which was included as a covariate in all statistical analyses. Internalising symptomatology and body image dissatisfaction were assessed using self-reported

Study details	Population	Interventions	Study outcomes	Appraisal and funding
	Testosterone 17 (80.95); no testosterone 28 (96.56)			scales, which may have resulted in over- or under-reporting of symptoms. The authors acknowledged the small sample size and that the findings should therefore be interpreted cautiously. The authors also acknowledged the limitations of using a cross-sectional study design, the omission of mental health related considerations such as psychiatric comorbidities, and the inclusion of youth in the 'no testosterone' group who had received puberty blockers. No subgroup analysis was reported. The authors acknowledged that the study sample did not capture youth whose mental health prevented them from being able to appropriately engage in decision-making for irreversible treatment, which means that the study sample may not be generalisable to all TNB youth seeking hormone treatment. Furthermore, the generalisability of the findings to the UK is also limited. **Source of funding:**

Study details	Population	Interventions	Study outcomes	Appraisal and funding
				National Centre for Advancing Translational Sciences/National Institutes of Health and the Abigail Wexner Research Institute.
Grimstad F, Kremen J, Shim J, Charlton BM, Boskey ER. Breakthrough Bleeding in Transgender and Gender Diverse Adolescents and Young Adults on Long-Term Testosterone. J Pediatr Adolesc Gynecol. 2021;34(5):706-16. **Study location** USA (single centre) **Study type** Retrospective case series **Study aim** To assess breakthrough bleeding patterns in TGD adolescents and young adults AFAB who had been on testosterone for longer than 1 year **Study dates**	TGD adolescents and young adults AFAB **Inclusion criteria** TGD adolescents and young adults assigned female or intersex at birth with functional uterus and ovaries present at the start of testosterone use **Exclusion criteria** TGD adolescents and young adults AFAB <1 year on testosterone; no uterine bleeding documented in medical records; Mayer-Rokintansky-Kuster-Hauser Syndrome;	**Interventions** Type of testosterone used, n (%) Injectable:[51] no breakthrough bleeding 170 (97.7); breakthrough bleeding 51 (87.9) Topical gel: no breakthrough bleeding 1 (0.6); breakthrough bleeding 0 (0) Injectable then subcutaneous pellets: no breakthrough bleeding 1 (0.6); breakthrough bleeding 3 (5.2) Topical gel then injectable: no breakthrough bleeding 1 (0.6); breakthrough bleeding 1 (1.7)	Follow-up: NR[52] **Important outcomes** **Masculinising physical changes** Number of participants reporting breakthrough bleeding,[53] n (%) No breakthrough bleeding: 174 (75%) Breakthrough bleeding: 58 (25%)[54]; *p value not reported*	This study was appraised using the JBI critical appraisal checklist for case series. 1. Yes 2. Unclear 3. Unclear 4. No 5. No 6. Yes 7. Yes 8. No 9. No 10. No **Other comments:** This was a retrospective case series conducted in a

[51] Includes intramuscular and subcutaneous testosterone cypionate and testosterone enanthate [reported as ethanate in the paper].
[52] The authors stated that to observe breakthrough bleeding effects after a medication change, they followed participants up three months after the change, but no further details were reported.
[53] Breakthrough bleeding was defined as any bleeding presumed to originate in the uterus while on testosterone.
[54] For participants with breakthrough bleeding, bleeding started a mean of 24.3 (SD 7.2) months after initiation of testosterone. Mean age at time of first breakthrough bleeding was 18.4 (SD 2.8) years. Eight (13.8%) of these participants had never become amenorrhoeic in the first year and continued to bleed beyond the first year. A total of 48 patients (82.8%) had more than one episode of bleeding after one year on testosterone.

Study details	Population	Interventions	Study outcomes	Appraisal and funding
2010 to 2020	hypogonadal hypogonadism **Total sample size** N=232 **Baseline characteristics** *Baseline characteristics are for participants with no breakthrough bleeding (n=174) or breakthrough bleeding (n=58)* Age (years) at initiation of testosterone, mean (SD) No breakthrough bleeding: 16.3 (1.8) Breakthrough bleeding: 16.3 (2.2) BMI (kg/m^2), mean (SD) No breakthrough bleeding: 27.2 (7.1) Breakthrough bleeding: 26.2 (6.9) Race/ethnicity, n (%) White (non-Hispanic): no breakthrough bleeding 130 (74.7);	Injectable then topical gel then subcutaneous pellets: no breakthrough bleeding 1 (0.6); breakthrough bleeding 0 (0) Injectable then topical gel: no breakthrough bleeding 0 (0); breakthrough bleeding 3 (5.2) **Comparators** No relevant comparator Use of concomitant treatments, n On menstrual suppression but discontinued during study: 74 On menstrual suppression throughout study: 43 Never on menstrual suppression or GnRH analogues: 106 *Menstrual suppression use, n (%)* No breakthrough bleeding 93 (53.4)		single centre in the USA, with patients recruited between 2010 and 2020. The authors did not clearly describe the services provided by the clinic but stated that all participants in the service are followed by both medical and mental health clinicians. Only participants with explicitly documented presence or absence of breakthrough bleeding in their chart review were included in the study. The authors stated that participants changing medications were followed for 3 months after the change, but no further details were reported. The authors stated that the age range for initiation of testosterone was 13 to 28 years, but the proportion of those aged >18 years was not reported. The authors did not describe the circumstances under which TGD adolescents AFAB aged 15 years or younger received testosterone monotherapy for gender transition. TGD adolescents AFAB were included in the study but it was unclear what

Study details	Population	Interventions	Study outcomes	Appraisal and funding
	breakthrough bleeding 42 (72.4) White (Hispanic): no breakthrough bleeding 6 (3.4); breakthrough bleeding 8 (13.8) Asian: no breakthrough bleeding 4 (2.3); breakthrough bleeding 1 (1.7) Black: no breakthrough bleeding 3 (1.7); breakthrough bleeding 3 (5.2) American Indian: no breakthrough bleeding 2 (1.1); breakthrough bleeding 0 (0) Additional race: no breakthrough bleeding 3 (1.7); breakthrough bleeding 0 (0) Declined to answer: no breakthrough bleeding 26 (14.9); breakthrough bleeding 4 (6.9) <u>Comorbid conditions, n (%)</u>	Breakthrough bleeding 33 (56.9) *Type of menstrual suppression, n (%)* Oral norethindrone: no breakthrough bleeding 65 (37.4); breakthrough bleeding 21 (36.2) Combined oral contraceptive pills: no breakthrough bleeding 5 (2.9); breakthrough bleeding 3 (5.2) Depo medroxyprogesterone acetate: no breakthrough bleeding 4 (2.3); breakthrough bleeding 3 (5.2) Etonogestrel implant: no breakthrough bleeding 2 (1.1); breakthrough bleeding 1 (1.7) Oral medroxyprogesterone: no breakthrough bleeding 1 (0.6); breakthrough bleeding 3 (5.2) Danazol: no breakthrough bleeding 1 (0.6); breakthrough bleeding 0 (0)		methods were used for participants to identify as TGD or whether any individuals identified as non-binary. Sixteen participants received GnRH analogues (n=8 discontinued during the study and n=8 continued taking throughout the study). Eighteen participants underwent hysterectomy aged >18 years. Eight (13.8%) participants had never become amenorrhoeic in the first year of testosterone use and continued to bleed beyond the first year. The authors reported that very few participants had breakthrough bleeding when only withdrawing from a concomitant method of menstrual suppression (n=2, 3.4%) or if a testosterone dose was missed (n=10, 17.2%); n=46 (79.3%) had no known cause for breakthrough bleeding. No statistical analysis was reported. No subgroup analysis was reported.

Study details	Population	Interventions	Study outcomes	Appraisal and funding
	Endometriosis: no breakthrough bleeding 1 (0.6); breakthrough bleeding 3 (5.2) Hypothyroidism: no breakthrough bleeding 2 (1.1); breakthrough bleeding 0 (0)	Levonorgestrel intrauterine device and oral norethindrone: NR Treatment duration (months), mean (SD) No breakthrough bleeding 28.5 (14.6) Breakthrough bleeding: 37.3 (16.9)		The authors acknowledged that little or no clinical evaluation was undertaken to determine the underlying cause(s) of breakthrough bleeding. They also acknowledged the inability to control which medications were used by participants, hence the variability. The authors also acknowledged that breakthrough bleeding was self-reported by participants and symptoms may therefore have varied. They also highlighted the lack of standardisation for documentation of breakthrough bleeding by clinicians. The generalisability of the findings to the UK is limited. **Source of funding:** One author was supported by a grant from the American Cancer Society.
Kramer R, Aarnio-Peterson CM, Conard LA, Lenz KR, Matthews A. Eating disorder symptoms among transgender and gender diverse youth. Clin. 2024;29(1):30-44. **Study location**	TGD youth **Inclusion criteria** TGD youth seeking gender-affirming treatment between 2015 and 2018 **Exclusion criteria**	**Interventions** Testosterone, n (%): 39 (21.5) **Comparators** No testosterone, n (%): 142 (78.5)	Follow-up: none **Critical outcomes** **Impact on mental health**	This study was appraised using the JBI critical appraisal checklist for cross-sectional studies. 1. Yes 2. No 3. No

<table>
<tr><th>Study details</th><th>Population</th><th>Interventions</th><th>Study outcomes</th><th>Appraisal and funding</th></tr>
<tr>
<td>USA (single centre)
Study type
Retrospective cross-sectional study
Study aim
To explore the potential associations between GAHT use and eating disorder symptoms among TGD youth
Study dates
January 2015 to September 2018</td>
<td>Not stated
Total sample size
N=251
Binary AFAB, n (%): 181 (72.1)
Baseline characteristics
Baseline characteristics are for TGD youth AFAB
Age (years), mean (SD)
16.77 (SD 2.74)
BMI, mean (SD)
26.87 (7.88)
EDE-Q score, mean (SD)
1.34 (1.17)
Ethnicity, n (%)
Non-Hispanic: 130 (96.3)
Hispanic: 5 (3.7)
Race, n (%)
White: 122 (89.7)
Black: 3 (2.2)
Asian: 3 (2.2)</td>
<td>Use of concomitant treatments: NR
Treatment duration: NR</td>
<td>Number of individuals reporting eating disorder behaviours, measured using the EDE-Q score,[55] n (%)
<table>
<tr><th rowspan="2">Eating disorder behaviour</th><th colspan="2">Taking GAHT</th><th>Not taking</th></tr>
<tr><th>Present</th><th>Not present</th><th>Present</th></tr>
<tr><td>Subjective binge episode</td><td>5 (13.5)</td><td>32 (86.5)</td><td>16 (18.0)</td></tr>
<tr><td>Objective binge episode</td><td>4 (10.8)</td><td>33 (89.2)</td><td>19 (21.1)</td></tr>
<tr><td>Self-induced vomiting</td><td>1 (2.7)</td><td>36 (97.3)</td><td>1 (1.1)</td></tr>
<tr><td>Laxative use</td><td>2 (5.1)</td><td>37 (94.9)</td><td>6 (6.3)</td></tr>
<tr><td>Compensatory exercise</td><td>7 (18.9)</td><td>30 (81.1)</td><td>16 (17.8)</td></tr>
</table>
The authors reported that there were no statistically significant interactions between GAHT use and eating disorder behaviours (all p>0.05)</td>
<td>4. Unclear
5. No
6. Yes
7. Yes
8. Yes
Other comments:
This was a small retrospective cross-sectional study conducted in a single centre in the USA between 2015 and 2018. Limited details were provided on inclusion criteria, and exclusion criteria were not reported.
The authors did not describe the circumstances under which TGD youth aged 15 years or younger received testosterone monotherapy for gender transition or provide details on what monitoring was in place for TGD youth receiving testosterone monotherapy.
TGD adolescents AFAB were included in the study but it was unclear what methods were used to identify individuals as TGD. The authors excluded three</td>
</tr>
</table>

[55] EDE-Q subscales include Restraint, Weight Concern, Shape Concern, and Eating Concern. Individual EDE-Q items reflect specific eating disorder behaviours, ie objective and subjective binge eating episodes, self-induced vomiting, laxative use, and compensatory exercise, and these items were used in the study analyses.

Study details	Population	Interventions	Study outcomes	Appraisal and funding
	Other: 8 (5.9)			youth identifying as non-binary from the study analyses due to the small sample size. Details on the use of gender-affirming hormones prior to the initial clinic visit were extracted via retrospective chart review. The authors acknowledged that the use of chart review prohibited standardised assessment of gender dysphoria. Eating disorder behaviours were assessed using the EDE-Q self-report scale, and the authors acknowledged that the use of EDE-Q items may not adequately represent unique symptoms experienced by TGD youth. The authors identified age and BMI as confounding factors based on research supporting that these variables are associated with EDE-Q scores, and these were included as covariates in all statistical analyses. Limited p values were presented for findings of interest (only significance). Only statistically significant ORs were presented.

Study details	Population	Interventions	Study outcomes	Appraisal and funding
				No subgroup analysis was reported. The generalisability of the findings to the UK is limited. **Source of funding:** None.
Laurenzano SE, Newfield RS, Lee E, Marinkovic M. Subcutaneous Testosterone Is Effective and Safe as Gender-Affirming Hormone Therapy in Transmasculine and Gender-Diverse Adolescents and Young Adults: A Single Center's 8-Year Experience. Transgend Health. 2021;6(6):343-52. **Study location** USA (single centre) **Study type** Retrospective case series **Study aim** To assess the effectiveness and safety of subcutaneous testosterone in achieving cessation of menses in TM/GD adolescents and young adults AFAB **Study dates** August 2012 to February 2020	TM/GD adolescents and young adults AFAB **Inclusion criteria** TM/GD adolescents and young adults AFAB who started subcutaneous testosterone at age 13 to 19 years and had received treatment for a minimum of 6 months **Exclusion criteria** Not stated **Total sample size** N=119 **Baseline characteristics** Age at presentation (years), mean (range) 16 (10.1 to 19.8)	**Interventions** Initial subcutaneous testosterone dose The authors stated that *"Nearly all subjects were started on 50–100 mg SC-T monthly divided into every 2 weeks (biweekly) doses; two subjects misunderstood instructions and started on 120–140 mg SC-T monthly"* Final follow-up dose of monthly subcutaneous testosterone, n (%) 100 to 200 mg: 94 (79) 240 to 320 mg: 21 (18) **Comparators** No relevant comparator	Follow-up: median 1.9 years (range 6 months to 5.5 years) **Important outcomes** **Detransition after receipt of masculinising medicines** Number of participants discontinuing treatment *"due to desire to end masculinizing therapy"*, n (%) The authors reported that *"two stopped because they were satisfied with effects of SC-T achieved and one stopped within 6 months of starting after reassessing their gender identity… due to concerns about body changes and potential impact on fertility"* **Safety** Change in laboratory values from baseline to final subcutaneous testosterone dose, mean (SD) *Total cholesterol, mg/dL:* Baseline: 158.0 (29.8) Final dose: 154.6 (31.1) *Difference from baseline to final dose: -3.5 (0.194); p=NS* *Haematocrit, %:* Baseline: 39.2 (2.6)	This study was appraised using the JBI critical appraisal checklist for case series. 1. Yes 2. Yes 3. Yes 4. No 5. No 6. Yes 7. No 8. Yes 9. Yes 10. Yes/No **Other comments:** This was a small retrospective case series conducted in a single centre in the USA between 2012 and 2020. It was unclear whether participants were

Study details	Population	Interventions	Study outcomes	Appraisal and funding
	Age at start of testosterone (years, mean (range) 16.5 (13 to 19.9) Gender identity, n (%) Transmale/male: 110 (92.5) Non-binary: 3 (2.5) Other:[56] 6 (5) Race, n (%) Caucasian: 79 (66.4) Asian: 7 (5.9) Native American/Alaska Native: 4 (3.4) African American: 2 (1.7) Native Hawaiian/Pacific Islander: 1 (0.8) Unknown/unavailable: 20 (16.8) Other:[57] 17 (14.3)	Use of concomitant treatments: NR Treatment duration: NR	Final dose: 44.1 (3.3) *Difference from baseline to final dose: 4.9 (SD not reported); p<0.001* *ALT, U/L:* Baseline: 20.8 (10.0) Final dose: 21.4 (12.9) *Difference from baseline to final dose: 0.7 (0.60); p=NS* *AST, U/L:* Baseline: 21.8 (8.0) Final dose: 23.0 (8.8) *Difference from baseline to final dose: 1.1 (0.29); p=NS* Number of adverse events, n (%) Mild injection site reactions: 14 (11.8)[58] Hypertension: 0 (0) Progression of acne: 77 (64.7) [advanced acne management, ie oral treatment and/or referral to dermatology, was reported in n=23 of these participants][59] Transaminitis: 0 (0) Dyslipidaemia: 1 (0.8) [described as worsening of dyslipidaemia]	recruited consecutively or completely. TM/GD AFAB participants were included in the study if they met diagnostic criteria for gender dysphoria in accordance with WPATH Standards of Care guidelines. A small proportion of individuals identified as non-binary (2.5%). The authors provided only limited details on clinical information for the participants. The authors reported that 12 (10.1%) participants had used puberty blockers prior to starting testosterone and eight (6.7%) participants had previously received progesterone only or the combined oral contraceptive. The authors stated that their centre typically initiates subcutaneous testosterone around the age of 14 years or older, following multidisciplinary assessment

[56] Other gender identities included gender queer (n=1), gender queer or fluid (n=1), gender fluid (n=1), agender (n=1), on the gender spectrum (n=1), no preferred gender identity or pronoun stated at first visit (n=1).
[57] Other self-reported races included: Mexican (n=1), Peruvian (n=1), Puerto Rican (n=1), and Hispanic not otherwise specified (n=14).
[58] Four of 14 participants changed treatment: n=3 changed from subcutaneous testosterone cypionate to subcutaneous testosterone enanthate and symptoms resolved; n=1 switched to testosterone gel because of pain and swelling with injections.
[59] Sixty one of 106 (57.5%) participants were documented to have acne at baseline, with the majority being mild to moderate.

Study details	Population	Interventions	Study outcomes	Appraisal and funding
	Ethnicity, n (%) Non-Hispanic/Latinx: 88 (73.9) Hispanic/Latinx: 29 (24.4) Unknown: 2 (1.7) BMI Z-score, mean (SD) 0.55 (1.22) BMI Z-score category, n (%) Obese: 24 (20.2) Overweight: 19 (16) Underweight: 3 (2.5)		Haematocrit >55%: 0 (0)	of readiness in youth with gender dysphoria and who have no contraindications. Six participants commenced subcutaneous testosterone younger than the age of 14 years, but no details were provided as to the circumstances. The authors did not provide details on what monitoring was in place for TM/GD AFAB individuals receiving testosterone monotherapy. Total cholesterol, haematocrit, ALT and AST were assessed across three dose ranges (<160 mg, 160 to 240 mg and >240 mg); with statistically significant trends for final haematocrit (p=0.024), final ALT (p=0.05), and difference between baseline and final AST (p=0.01). No subgroup analysis was reported. The authors stated that 31.1% of participants deviated from the dose or schedule of subcutaneous testosterone recommended by the provider: minimally different doses than prescribed or injections being occasionally delayed or missed due to interruption

Study details	Population	Interventions	Study outcomes	Appraisal and funding
				in supply, follow-up, or adherence. One participant was lost to long-term follow-up after discontinuing treatment at six months. The generalisability of the findings to the UK is limited. **Source of funding:** None.
Millington K, Lee JY, Olson-Kennedy J, Garofalo R, Rosenthal SM, Chan YM. Laboratory Changes During Gender-Affirming Hormone Therapy in Transgender Adolescents. Pediatrics. 2024;153(5):01. **Study location** US (four centres) **Study type** Retrospective cohort study[60] **Study aim** To assess changes in laboratory measures in TGD adolescents receiving GAHT **Study dates**	TGD adolescents **Inclusion criteria[62]** TGD adolescents aged 8 to 20 years who had not initiated GAHT **Exclusion criteria** Previous use of GAHT, presence of serious psychiatric symptoms (eg active hallucinations, thought disorder), visibly distraught (eg suicidal), intoxicated or under the influence of alcohol or other substances **Total sample size**	**Interventions** Testosterone formulation at baseline, n (%) Subcutaneous: 195 (97) Transdermal gel: 5 (3) Testosterone dose at baseline visit, median (IQR) Subcutaneous administration, mg per week: 25.0 (25.0 to 26.0) Transdermal administration, mg per day: 25.0 (20.25 to 25.0)	Follow-up: 6, 12 and 24 months **Safety** Change in laboratory values from baseline to follow-up, median (IQR) *Haemoglobin, mg/dL* Baseline (n=189): 13.2 (12.5 to 13.9) At 6 months (n=155): 14.2 (13.3 to 15.1);[63] *difference compared to baseline p<0.001* At 12 months (n=136): 14.7 (13.6 to 15.6); *difference compared to 6 months follow-up p<0.001* At 24 months (n=119): 15.0 (14.1 to 15.8); *difference compared to 12 months follow-up p=0.01* *Haematocrit, %:* Baseline (n=191): 39.9 (37.8 to 41.6)	This study was appraised using the JBI critical appraisal checklist for case series. 1. Yes 2. Yes 3. Yes 4. No 5. No 6. Yes 7. Yes 8. Yes 9. No 10. Yes

[60] Participants were recruited for the Trans Youth Care-United States (TYCUS) Study, which was a cohort study but was treated as a case series in this evidence review as there was no in-scope comparator group.
[62] Inclusion and exclusion criteria were extracted from the study protocol.
[63] There was a slight discrepancy in the IQR reported in the Table (13.3 to 15.1) and the figure reported in the main text (13.3 to 15.2) and we have reported the figure presented in the Table.

Study details	Population	Interventions	Study outcomes	Appraisal and funding
July 2016 to September 2018[61]	N=293 AFAB: n=200 **Baseline characteristics** *Baseline characteristics are for TGD adolescents AFAB* Age at presentation (years), median (IQR) 16.2 (15.1 to 17.6) Gender identity, n (%) Male: 81 (41) TGD male: 106 (53) Gender fluid: 1 (1) Gender queer: 1 (0.5)	Testosterone dose at 12 month visit, median (IQR) Subcutaneous administration, mg per week: 50.0 (40.0 to 50.0) Transdermal administration, mg per day: 40.5 (37.5 to 40.5) Testosterone dose at 24-month visit, median (IQR) Subcutaneous administration, mg per week: 50.0 (50.0 to 60.0) Transdermal administration, mg per	At 6 months (n=157): 43.8 (41.6 to 45.8); *difference compared to baseline p<0.001* At 12 months (n=136): 44.6 (41.6 to 47.0); *difference compared to 6 months follow-up p<0.001* At 24 months (n=119): 45.4 (42.9 to 47.6); d*ifference compared to 12 months follow-up p=0.03*[64] *HbA1c, %:*[65] Baseline (n=105): 5.2 (5.0 to 5.4) At 6 months (n=66): 5.1 (5.0 to 5.4); *difference compared to baseline p=NS* At 12 months (n=64): 5.1 (4.9 to 5.3); *difference compared to 6 months follow-up p=NS* At 24 months (n=59): 5.1 (4.9 to 5.3); d*ifference compared to 12 months follow-up p=NS*[66] *ALT, U/L:*[67] Baseline (n=77): 17 (11 to 25)	**Other comments:** This was a prospective cohort study that was treated as a case series and was conducted across four centres in the US between 2016 and 2018. The authors did not clearly describe the services provided by the centres. It was unclear whether participants were recruited consecutively or completely. TGD adolescents were included in the study if they met diagnostic criteria for gender dysphoria as determined by a clinician. A small proportion of TGD adolescents AFAB identified as non-binary (5%).

[61] Study dates were extracted from the study protocol.
[64] There were 13 participants (6.5%) DFAB, one of whom had recent tobacco use, who had haematocrit above the typical cisgender male range (haematocrit 40% to 50%) during treatment with testosterone, ranging from 50.1% to 53.1%. Apart from a decrease in the testosterone dose, there were no other clinical interventions (eg therapeutic phlebotomy) required to address the increased haematocrit.
[65] Five participants with pre-existing diabetes mellitus were excluded from the HbA1c analysis. There were 10 (5.1%) participants with baseline HbA1c measurements in the prediabetes range (5.7% to 6.4%). Of these, three participants did not have additional follow-up measurements, five had subsequently normal measurements, and two had persistently elevated HbA1c measurements of 5.7%.

[66] One participant had an increase in HbA1c from 5.4% to 5.7% over the 24-month treatment period. This participant also had an increase in BMI from 24.5 to 27.8 kg/m^2. Another participant had an increase in HbA1c from 5.6% to 6.1% at the 12-month follow-up visit, but HbA1c was in the normal range at the 24-month follow-up visit (4.3%). This participant also had a baseline BMI in the obese range (32.0 kg/m2). In total, there were 15 (7.7%) participants who had an elevated HbA1c measurement at any point during the study period.
[67] Four (2%) participants had elevations in liver enzymes >2 times the upper limit of normal during testosterone therapy (range 118 to 263 U/L). The authors stated that all four participants were receiving concurrent treatment with other medications that have been associated with abnormal liver function tests (ie isotretinoin, quetiapine, lamotrigine, bupropion,). Three of these four participants had subsequent ALT and AST measurements that returned to normal during the study visit. The fourth participant, who was taking isotretinoin and sertraline, continued to have elevated ALT at the 24-month follow-up visit.

Study details	Population	Interventions	Study outcomes	Appraisal and funding
	Non-binary: 10 (5) <u>Tanner stage at baseline, n (%)</u> 3: 1 (0.5) 4: 17 (9) 5: 168 (91)	day: 40.5 (20.25 to 60.75) **Comparators** No relevant comparator <u>Use of concomitant treatments, n (%)</u> Progesterone use: Norethindrone acetate: 31 (16) Medroxyprogesterone acetate: 6 (3) Etonogestrel implant: 1 (0.5) Combined oestrogen and progesterone oral contraceptive: 22 (11) Treatment duration: NR	At 6 months (n=58): 19.5 (14 to 28); *difference compared to baseline p=NS* At 12 months (n=57): 18 (13 to 26); *difference compared to 6 months follow-up p=NS* At 24 months (n=42): 19 (13 to 27); d*ifference compared to 12 months follow-up p=NS* *AST, U/L:* Baseline (n=77): 20 (17 to 25); *difference compared to baseline p=NS* At 6 months (n=58): 23 (19 to 29); *difference compared to baseline p=NS* At 12 months (n=57): 22 (18 to 27); d*ifference compared to 6 months follow-up p=NS* At 24 months (n=42): 22 (18 to 28); *difference compared to 12 months follow-up p=NS*	The authors stated that the minimum age for inclusion in the study was 8 years in order to ensure that potential participants who might be eligible for hormones based on their Tanner stage would not be excluded due to age alone. Seven youths under the age of 13 years at the time of recruitment were enrolled into the study, it was unclear how many were AFAB and details on the circumstances for prescribing testosterone monotherapy to children aged 15 years or younger were not stated. The authors did not provide details on what monitoring was in place for TGD adolescents receiving testosterone monotherapy. Not all participants were free of the outcomes of interest at the start of the study, eg a proportion of participants had elevated laboratory measures or diabetes at baseline. The authors reported that 31 (16%) participants had used GnRH analogues in late or post-puberty (Tanner stage 4 or 5) prior to receiving testosterone; the number of

Study details	Population	Interventions	Study outcomes	Appraisal and funding
				participants who had received GnRH analogues in early puberty (Tanner stage 2 or 3) was not reported. The authors included age and baseline BMI in multivariate models. Haemoglobin and haematocrit analyses were controlled for recent tobacco use. with diabetes mellitus were excluded from the HbA1c, ALT, and AST analyses. Participants taking bicalutamide were excluded from the ALT and AST analyses. The authors stated that missing data may have skewed the results, but the reasons for missing data were not always clear. A mixed model repeated measures was used to model the longitudinal association between the outcome variables. No subgroup analysis was reported. The authors acknowledged that due to the study being conducted in large, urban, tertiary centres with clinics specialising in the care of TGD adolescents, the study sample may not represent all TGD youth, particularly those from minority

Study details	Population	Interventions	Study outcomes	Appraisal and funding
				backgrounds, lower socio-economic status, or living in rural settings. The generalisability of the findings to the UK is limited. **Source of funding:** Eunice Kennedy Shriver National Institute of Child Health and Human Development, the Doris Duke Charitable Foundation, the Biostatistics Core at the Saban Research Institute, and the National Institutes of Health.
Moussaoui D, Elder CV, O'Connell MA, McLean A, Grover SR, Pang KC. Pelvic pain in transmasculine adolescents receiving testosterone therapy. Int J Transgend Health. 2024;25(1):10-8. **Study location** Australia (single centre) **Study type** Retrospective case series **Study aim**	TGD children and adolescents AFAB **Inclusion criteria** TGD children and adolescents AFAB aged up to age 18 years who commenced testosterone **Exclusion criteria** Individuals with follow-up of less than 6 months after	**Interventions** Topical testosterone: 17 (10.8%) 3-weekly intramuscular testosterone undecanoate: 63 (39.9%) 3- monthly intramuscular testosterone enanthate or mixed testosterone esters: 78 (49.4%)	Follow-up: median 22.1 (IQR median value 15.4) months **Safety** Number of individuals reporting pelvic pain, abdomino-pelvic pain or pain in the lower part of the abdomen [no validated measurement tool used], n (%)[70] Pelvic pain: 37 (23.4)[71] (pelvic pain description: cramps 17 [45.9); similar to period pain 8 [21.6]; sex-related pain 10 [27]; early morning/or waking up at night 5 [13.5]; suspicion of pelvic floor spasms 5 [13.5]; associated with breakthrough bleeding 11 [29.7]; associated with nausea	This study was appraised using the JBI critical appraisal checklist for case series. 1. Yes 2. Yes 3. Unclear 4. No 5. No 6. Yes 7. Yes

[70] Pelvic pain was defined as the timing of onset as reported by the individual, or as the date of first documentation in medial chart if no mention of timing of onset. The median interval between testosterone initiation and onset of pain was 1.6 months (range 0.3 to 6.4).
[71] Thirty six of 37 participants reporting pelvic pain had not received past puberty blockers. **P**ain intensity was reported for n**=**11 adolescents: n=1 mild (self-reported score of 1 to 3 out of 10, with a score of 10 being most severe); n=10 severe (self-reported score of 7 to 10 out of 10, with a score of 10 being most severe).

Study details	Population	Interventions	Study outcomes	Appraisal and funding
To assess the prevalence of pelvic pain, potential predicting factors, and treatment of pelvic pain among TGD adolescents receiving testosterone **Study dates** January 2007 to April 2020	initiation of testosterone **Total sample size** N=158 **Baseline characteristics** Age at testosterone initiation (years), median (range) 16.6 (13.9 to 18.6)	Dose at initiation, n (%) Low dose:[68] 74 (46.8) High dose:[69] 84 (53.2) **Comparators** No relevant comparator Use of concomitant treatments, n (%) Hormonal medication to suppress menstruation: 137 (86.7), including levonorgestrel intrauterine device: 10 (7.3); oral combined contraceptive pill: 2 (1.5); etonogestrel implant: 3 (2.2); norethisterone: 106 (77.4); medroxyprogesterone acetate injection: 13 (9.5); oral medroxyprogesterone acetate: 3 (2.2) Treatment duration: NR	and/vomiting 4 [10.8]; pain radiating into lower limbs 2 [5.4]) No pelvic pain: 121 (76.6); *p value not reported*	8. Yes 9. No 10. No **Other comments:** This was a small retrospective case series conducted in a primary care clinic in Australia, with patients assessed across a wide time period (2007 to 2020). The authors did not clearly describe services provided by the clinic. It was unclear whether participants were recruited consecutively or completely. TGD individuals were included in the study but it was unclear what methods were used for participants to identify as TGD or whether any individuals identified as non-binary. The authors did not describe the circumstances under which TGD children and adolescents AFAB aged 15 years or younger received testosterone monotherapy for gender transition or

[68] 'Low' dose of testosterone was defined as topical testosterone 12.5 to 25 mg, intramuscular testosterone enanthate or mixed testosterone esters 125 mg or intramuscular testosterone undecanoate 500 mg.
[69] 'High' dose of testosterone was defined as topical testosterone 32.5 to 50 mg, intramuscular testosterone enanthate or mixed testosterone esters 250 mg or intramuscular testosterone undecanoate 1,000 mg (doses equivalent to those used for standard maintenance dosing in adult men).

Study details	Population	Interventions	Study outcomes	Appraisal and funding
				provide details on what monitoring was in place for TGD children and adolescents AFAB receiving testosterone monotherapy. The authors reported that 15 (9.5%) participants had used puberty blockers in the past. The authors compared pelvic pain between participants taking or not taking menstrual suppression and provided descriptions of treatment for pelvic pain, but these details are not reported in this evidence review. No statistical measures were reported and no subgroup analyses were conducted. The authors acknowledged that due to poor documentation, it was unclear whether pelvic pain pre-existed prior to testosterone use and that the follow-up duration may not have been sufficient to capture the onset and development of pelvic pain. Due to pelvic pain being self-reported by participants, there may have been a risk of bias in terms of over- or under-reporting. The authors acknowledged that the retrospective nature of the

Study details	Population	Interventions	Study outcomes	Appraisal and funding
				study may have resulted in under-reporting of pelvic pain and highlighted the limitations in the generalisability of the findings due to limited documentation of details in medical records in relation to the nature and severity of pelvic pain. The generalisability of the findings to the UK is limited. **Source of funding:** Two authors received funding from the Fonds de Perfectionnement, Geneva University Hospitals, the Swiss National Science Foundation, and the Hugh Williamson Foundation, the Royal Children's Hospital Foundation and the NHMRC.
Mullins ES, Geer R, Metcalf M, Piccola J, Lane A, Conard LAE, et al. Thrombosis Risk in Transgender Adolescents Receiving Gender-Affirming Hormone Therapy. Pediatrics. 2021;147(4):04. **Study location**	TG adolescents and young adults **Inclusion criteria** TG adolescents and young adults who initiation GAHT prior to March 2019 and aged between 13 and 24 years	**Interventions** Testosterone (n=429): median dose 70 mg (IQR 60 to 80)[73] Testosterone formulation, n (%) Subcutaneous: 312 (72.7)	Follow-up: median 577 days (IQR 283 to 923) **Safety** Incidence of thrombosis during treatment, n (%) VTE or arterial thrombosis (including stroke) (n=429): 0 (0)	This study was appraised using the JBI critical appraisal checklist for case series. 1. Yes 2. Unclear 3. Unclear 4. Yes

[73] The authors stated that TG men had testosterone levels followed during the study period and doses were adjusted to maintain physiologic levels.

Study details	Population	Interventions	Study outcomes	Appraisal and funding
USA (single centre) **Study type** Retrospective case series **Study aim** To assess the incidence of thrombosis and thrombosis risk factors among TG adolescents and young adults **Study dates** July 2013 to March 2019	**Exclusion criteria** Aged <13 years at initiation of GAHT **Total sample size** N=611 AFAB: n=428[72] (n=12 [2.8%] non-binary) **Baseline characteristics** *Baseline characteristics are for the whole population (AMAB and AFAB)* Age at presentation (years), median (IQR) 17 (15 to 19) Race, n (%) White: 544 (89) African American: 50 (8.2) Asian American: 8 (1.3) Other: 12 (1.9) Not documented: 8 (1.3)	Intramuscular: 105 (24.4) Gel: 11 (2.8) Transdermal: 1 (0.7) **Comparators** No relevant comparator Use of concomitant treatments, n (%) Anticoagulation (rivaroxaban): 1 (0.2)[74] Treatment duration, median (IQR): 577 days (283 to 923)		5. Yes 6. Yes 7. Yes 8. Yes 9. No 10. Not applicable **Other comments:** This was a retrospective case series including 429 TG adolescents and young adults AFAB and was conducted in a single centre in the USA between 2013 and 2019. TG adolescents and young adults were included in the study but it was unclear what methods were used for participants to identify as TG. A small proportion of TG adolescents and young adults AFAB identified as non-binary (2.8%). The authors did not describe the circumstances under which TG adolescents aged 15 years or younger received testosterone monotherapy for gender

[72] There was an unexplained discrepancy between the number of TG men (reported as n=428 [70%]) and the number of TB men receiving testosterone (n=429 [70.2%]) and we have reported the number as n=429 for the purposes of reporting outcomes.
[74] One TG male had a thrombotic event of VTE prior to starting testosterone and was maintained on rivaroxaban during treatment with testosterone.

Study details	Population	Interventions	Study outcomes	Appraisal and funding
	Ethnicity, n (%) Hispanic: 14 (2.3) Non-Hispanic: 595 (97.4) Not documented: 2 (0.3) BMI, kg/m^2, n (%) 18.5: 40 (6.5) 18.5 to 25: 212 (34.7) 25 to 30 148 (24.2) 30: 211 (34.5) Tobacco use, n (%) Yes: 94 (15.4) No: 516 (84.5) Migraine with aura, n (%) Yes: 28 (4.6) No: 524 (85.8) Not documented: 59 (9.7) Family history of thrombosis, n (%) Yes: 49 (8.0 No: 388 (63.5) Not documented: 174 (28.5)			transition or provide details on what monitoring was in place for TG adolescents and young adults receiving testosterone monotherapy. The authors stated that 17 participants (unclear number of AFAB participants) with a personal or family history of thrombosis or thrombosis risk or personal history of thrombophilia were referred for evaluation with haematology prior to starting GAHT. Ten participants (the majority of whom were TG adolescents and young adults AFAB) had elevated haemoglobin (>17.7 g/dL), and one TG youth AFAB with the prothrombin gene mutation was started on thromboprophylaxis (rivaroxaban) before starting testosterone, but it was unclear whether this treatment continued during treatment with testosterone. The authors reported previous hormonal use, including norethindrone contraceptive pill (n=148 [24.2%], depo-medroxyprogesterone acetate (n=113 [18.5%], combined oral contraceptive pill (n=35 [5.7%], norethindrone acetate (n=15

Study details	Population	Interventions	Study outcomes	Appraisal and funding
	Family history of risk factors for thrombosis, n (%) Yes: 5 (0.8) No: 374 (61.2) Not documented: 232 (38.0) Other personal risk factors, n (%) Inflammatory bowel disease: 3 (0.5) Juvenile rheumatoid arthritis: 1 (0.2)			[2.5%], LNG-IUS (n=15 [2.5%], etonogestrel implant (n=2 [0.3%]. No subgroup analysis was reported. The authors acknowledged some of the limitations of the study, including the retrospective nature which made it difficult to determine from medical records whether participants had discontinued GAHT or were lost to follow-up (eg may have transitioned to adult care), and whether contraception or menstrual suppression was stopped (making it difficult to determine which TG adolescents and young adults AFAB also used concomitant oestrogen-containing medicines, which the authors suggested is associated with increased risk of thrombosis). In addition, the authors acknowledged that not all participants had reached target physiologic hormone levels during the study period due to them being in varying stages of gender transition, and continued titration of hormone dosage after the study end date to

Study details	Population	Interventions	Study outcomes	Appraisal and funding
				achieve physiologic levels would therefore not have been captured. The generalisability of the findings to the UK is limited. **Source of funding:** Cincinnati Children's Hospital Medical Centre and Cancer and Blood Diseases Institute.
Persky RW, Apple D, Dowshen N, Pine E, Whitehead J, Barrera E, et al. Pubertal Suppression in Early Puberty Followed by Testosterone Mildly Increases Final Height in Transmasculine Youth. J. 2024;8(6):bvae089. **Study location** USA (five centres) **Study type** Retrospective cohort study[75] **Study aim** To compare the effect of GnRH analogues use prior to testosterone vs testosterone	TM youth **Inclusion criteria** TM youth receiving testosterone with or without prior GnRH analogues and reached final adult height;[76] TM youth receiving testosterone alone had to be at least 15 years old at the start of treatment **Exclusion criteria** TM youth on growth-altering medications, such as systemic glucocorticoids, or with any history of a	**Interventions** Formulation of testosterone at initiation, n (%) Testosterone cypionate/enanthate intramuscular injection: 9 (14.5) Testosterone cypionate/enanthate subcutaneous injection: 53 (85.5) Testosterone gel (topical): 0 (0) Initial starting dose of injectable	Follow-up: NR **Important outcomes** **Masculinising physical changes** Final adult height (cm) from baseline to follow-up, mean (SD) Baseline: 164.1 (6.8) Follow-up: 164.7 (6.7)[77] *Difference between baseline and follow-up: NS* Final adult height Z-scores[78] from baseline to follow-up, mean (SD) Baseline: 0.21 (1.0) Follow-up: 0.21 (1.0) *Difference between baseline and follow-up: NS*	This study was appraised using the JBI critical appraisal checklist for case series. 1. Yes 2. Yes 3. Yes 4. Yes 5. Yes 6. Yes 7. No 8. No 9. Yes

[75] Although this was a cohort study, it was treated as a case series as there was no relevant in-scope comparator group.
[76] Final adult height was defined as a growth velocity ≤1.5 cm per year, three years post menarche, or a bone age ≥15 years.
[77] The authors stated that due to the variability in height measurements reported in the medical records, the average of all heights recorded after the final adult height was reached was calculated and this was used for data analysis purposes.
[78] Final adult height Z-scores were calculated based on the CDC 20 year old growth chart for girls.

Study details	Population	Interventions	Study outcomes	Appraisal and funding
without preceding GnRH analogues on final adult height, midparental target height and predicted adult height **Study dates** January 2008 to June 2020	growth-altering disorder, such as precocious puberty, growth hormone deficiency, or an advanced or delayed baseline bone age **Total sample size** N=94 (n=62 receiving testosterone only) **Baseline characteristics** *Baseline characteristics are for TM youth receiving testosterone only* Age at start of testosterone (years), median (IQR) 16.8 (15.8 to 17.8) Race, n (%) White: 49 (79) Other: 13 (21) [Black/African American: 3; Asian or Pacific Islander: 1; American Indian or Alaskan: 1; Other/Unknown: 8] Ethnicity, n (%) Hispanic: 5 (8)	testosterone, median (IQR) Testosterone cypionate/enanthate intramuscular injection (mg), median (IQR) (n=9): 25 (25 to 25); frequency (days), median (range): 7 (7 to 14) Testosterone cypionate/enanthate subcutaneous injection (mg), median (IQR) (n=53): 26 (25 to 50); frequency (days), median (range): 7 (7 to 14) **Comparators** No relevant comparator Use of concomitant treatments: the authors reported that participants on stimulant medications were included, but no further details were provided Treatment duration: NR		10. No **Other comments:** This was a small retrospective cohort study that was treated as a case series and was conducted across five centres in the USA from 2008 to 2020. The follow-up duration was not reported and it was unclear whether any participants were lost to follow-up. TM youth were included in the study if they met diagnostic criteria for ICD-10 F64.0, F64.9, and E34.9 or ICD-9 302.85. It was unclear whether any individuals identified as non-binary. The authors did not describe the circumstances under which TM youth aged 15 years or younger received testosterone monotherapy for gender transition or provide details on what monitoring was in place for TM youth receiving testosterone monotherapy. No statistical measures were reported. No subgroup analysis was reported. The authors acknowledged the variability in treatment regimens and data collection

Study details	Population	Interventions	Study outcomes	Appraisal and funding
	BMI (at initial visit; Z-score), mean (SD) 0.71 (1.1)			across the five centres, and that the demographics of the included population (ie mainly white and non-Hispanic) may have limited the generalisability of the findings to other populations. The generalisability of the findings to the UK is also limited. **Source of funding:** Eunice Kennedy Shriver National Institute of Child Health and Human Development.
Valentine A, Davis S, Furniss A, Dowshen N, Kazak AE, Lewis C, et al. Multicenter Analysis of Cardiometabolic-related Diagnoses in Transgender and Gender-Diverse Youth: A PEDSnet Study. J Clin Endocrinol Metab. 2022;107(10):e4004-e14. **Study location** USA (PEDSnet study; six centres) **Study type**	TGD youth **Inclusion criteria** TGD youth (aged >2 years at last visit) with a diagnosis of gender dysphoria or related diagnosis and at least one outpatient visit **Exclusion criteria** Not stated **Total sample size** N=4,172[79] AFAB: n=2,766	**Interventions** Testosterone: n=832 **Comparators** No testosterone: n=1,934[80] Use of concomitant treatments, n (%) Progestin norethindrone and medroxyprogesterone: 112 (4.1)	Follow-up: none **Safety**[81] Odds of cardiometabolic-related diagnoses between patients receiving testosterone vs patients not receiving testosterone *Overweight/obesity* OR (95% CI) 1.8 (1.5 to 2.1); p<0.0001 *Dyslipidaemia* OR (95% CI) 1.7 (1.3 to 2.3); p<0.01 *Liver dysfunction* OR (95% CI) 1.5 (1.1 to 1.9); p≤0.01	This study was appraised using the JBI critical appraisal checklist for cross-sectional studies. 1. Yes 2. No 3. Yes 4. No 5. Yes 6. Yes 7. Yes 8. Yes

[79] There was a discrepancy in the total sample size which was reported as n=4,172, but the number of AMAB and AFAB individuals totals n=4,173.
[80] The number of participants who were not receiving testosterone has been calculated as this figure was not reported in the paper.
[81] Analyses were adjusted for electronic health record sex, age at last visit, duration on PEDSnet, overweight/obesity, depression, and antipsychotic prescription.

Study details	Population	Interventions	Study outcomes	Appraisal and funding
Retrospective cross-sectional study **Study aim** To evaluate the potential effects of GAHT on cardiometabolic-related outcomes among TGD youth **Study dates** 2009 to November 2019	AMAB: n=1,407 **Baseline characteristics** *Baseline characteristics are reported for the whole population (AMAB and AFAB)* Age (years) at first visit, median (IQR) 10.0 (4.4 to 14.6) Age (years) at last visit, median (IQR) 16.7 (14.6 to 18.3) Race, n (%) White: 3,027 (72.5) Unknown: 401 (9.6) Other: 390 (9.3) Black: 257 (6.2) Asian: 98 (2.3) Ethnicity, n (%) Non-Hispanic: 3,538 (84.8) Hispanic: 354 (8.5) Unknown: 281 (6.7)	COCP: 199 (7.2) The authors mentioned the use of antipsychotics, but no further details were provided Treatment duration: NR	*Dysglycaemia* OR (95% CI) NR; p=NS *Hypertension* OR (95% CI) 1.6 (1.2 to 2.2); p<0.01 *PCOS* OR (95% CI) NR; p=NS *The authors reported that the differences in outcomes were not statistically significant in unadjusted models for any cardiometabolic-related diagnoses*	**Other comments:** This was a large cross-sectional study conducted across six centres in the USA, with patients recruited between 2009 and 2019. The median duration of participants in the PEDSnet study was 5.7 years (IQR 1.7 to 11.1) and the median number of outpatient visits was 10 (IQR 4 to 26). The inclusion criteria were limited as were details on the included individuals in terms of baseline demographics and clinical information. The authors mentioned depression and medications including antipsychotics but did not provide any further details. TGD adolescents AFAB were included in the study but it was unclear what methods were used to identify individuals as TGD or whether any individuals identified as non-binary. The authors did not describe the circumstances under which TGD adolescents AFAB aged 15 years or younger received testosterone monotherapy

Study details	Population	Interventions	Study outcomes	Appraisal and funding
				for gender transition or provide details on what monitoring was in place for TGD adolescents AFAB receiving testosterone monotherapy. The authors mentioned that they did not evaluate any gender-affirming surgical procedures such as masculinising chest surgery, and it was therefore unclear whether a proportion of out-of-scope participants may have been included in the study population. No subgroup analysis was reported. The authors acknowledged that they were unable to determine relationships between the timing of the outcomes, timing of gender dysphoria diagnosis (ie whether participants who received testosterone were experiencing worse gender dysphoria and worse minority stress compared to those not receiving testosterone), or initiation of GAHT (ie whether diagnoses preceded or followed the initial testosterone prescription).

Study details	Population	Interventions	Study outcomes	Appraisal and funding
				The generalisability of the findings to the UK is limited. **Source of funding:** National Institutes of Health/National Institute of Child Health and Human Development/National Heart, Lung, and Blood Institute, Doris Duke Foundation, National Institute of Diabetes and Digestive and Kidney Diseases, Paediatric Endocrine Society, and the Society for Adolescent Health and Medicine.

Abbreviations

AFAB: assigned female at birth; ALT: Alanine aminotransferase; AMAB: assigned male at birth; ANOVA: analysis of variance; ASQ: Ask Suicide-Screening Questions; AST: aspartate aminotransferase; BID: Body Image Dissatisfaction; BMI: body mass index; CDI: Children's Depression Inventory; CI: confidence interval; COCP: combined oral contraceptive pill; df: degrees of freedom; EDE-Q: Eating Disorder Examination Questionnaire; g/dL: grams per decilitre; GAHT: gender affirming hormones; GnRH: gonadotropin-releasing hormone; HbA1c: glycated haemoglobin; ICD: International Statistical Classification of Diseases; IQR: interquartile range; kg: kilogram; kg/m^2: kilogram per square metre; LNG-IUS: levonorgestrel intrauterine device; LSAS: Leibowitz Social Anxiety Scale; mg: milligrams; mg/dL: milligrams per decilitre; mg/mL: milligrams per millilitre; n: number; NA: not applicable; NR: not reported; NS: not significant; OHSU: Oregon Health & Science University; OR: odds ratio; PCOS: polycystic ovary syndrome; PEDSnet: Paediatric Learning Health System network; PedsQL GWBS: Paediatric Quality of Life General Wellbeing Scale; QoL: quality of life; RCT: randomised controlled trial; REDCap: Research Electronic Data Capture; SCARED: Screen for Child Anxiety Related Emotional Disorders; SC-T: subcutaneous testosterone; SD: standard deviation; SE: standard error; TG: transgender; TGD: transgender and gender diverse; TM: transmasculine; TM/GD: transmasculine and gender diverse; TNB: transgender and non-binary; UK: United Kingdom; USA: United States of America; WPATH: World Professional Association for Transgender Health.

Appendix F Quality appraisal checklists

JBI Critical Appraisal Checklist for Case Series

1. Were there clear criteria for inclusion in the case series?
2. Was the condition measured in a standard, reliable way for all participants included in the case series
3. Were valid methods used for the identification of the condition for all participants included in the case series?
4. Did the case series have consecutive inclusion of participants?
5. Did the case series have complete inclusion of participants?
6. Was there clear reporting of the demographics of the participants in the study?
7. Was there clear reporting of clinical information of the participants?
8. Were the outcomes or follow up results of cases clearly reported?
9. Was there clear reporting of the presenting site(s)/clinic(s) demographic information?
10. Was statistical analysis appropriate?

JBI Critical Appraisal Checklist for Analytical Cross-sectional Studies

1. Were the criteria for inclusion in the sample clearly defined?
2. Were the study subjects and the setting described in detail?
3. Was the exposure measured in a valid and reliable way?
4. Were objective, standard criteria used for measurement of the condition?
5. Were confounding factors identified?
6. Were strategies to deal with confounding factors stated?
7. Were the outcomes measured in a valid and reliable way?
8. Was appropriate statistical analysis used?

Appendix G GRADE profiles

The language used in this table is that of the study authors and may not reflect current language used by NHS England or NHS Gender Identity Services

For abbreviations and footnotes see end of tables

Table 1: Masculinising medicines compared to no masculinising medicines

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of patients		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Testosterone	No testosterone	Result		
Impact on gender incongruence (1 cross-sectional study)									
Gender dysphoria measured using the BID, mean (SD) at mean 12.87 (SD 9.94) months treatment duration [higher values indicate greater dissatisfaction with body image]									
1 cross-sectional study Grannis et al 2023	Very serious limitations[1]	Very serious indirectness[2]	Not applicable	Not calculable	21	29	Testosterone: 92.29 (19.74) No testosterone: 103.14 (18.99) *Difference between groups (ANOVA): F-value 7.46 (df between groups: 1, df within groups: 47); p<0.01*	Critical	Very low
Impact on mental health (2 cross-sectional studies)									
Symptoms of generalised anxiety, measured using SCARED, mean (SD) at mean 12.87 (SD 9.94) months treatment duration [higher values indicate greater generalised anxiety symptoms]									
1 cross-sectional study Grannis et al 2023	Very serious limitations[1]	Very serious indirectness[2]	Not applicable	Not calculable	21	29	Testosterone: 38.75 (17.41) No testosterone: 50.25 (14.12) *Difference between groups (ANOVA): F-value 7.76 (df between groups: 1, df within groups: 45); p<0.01*	Critical	Very low
Symptoms of depression, measured using the CDI, mean (SD) at mean 12.87 (SD 9.94) months treatment duration [higher values indicate greater depression symptoms]									
1 cross-sectional study	Very serious limitations[1]	Very serious indirectness[2]	Not applicable	Not calculable	21	29	Testosterone: 13.38 (7.81) No testosterone: 18.34 (7.02)	Critical	Very low

Grannis et al 2023							*Difference between groups (ANOVA): F-value 5.62 (df between groups: 1, df within groups: 47); p<0.05*		
Symptoms of social anxiety, measured using the LSAS, mean (SD) at mean 12.87 (SD 9.94) months treatment duration [higher values indicate greater social anxiety]									
1 cross-sectional study Grannis et al 2023	Very serious limitations[1]	Very serious indirectness[2]	Not applicable	Not calculable	21	29	Testosterone: 50.21 (22.34) No testosterone: 78.62 (31.41) *Difference between groups (ANOVA): F-value 14.8 (df between groups: 1, df within groups: 42); p<0.001*	Critical	Very low
Suicidality (defined as frequency of suicidal ideation and/or attempts in the past year), mean (SD) at mean 12.87 (SD 9.94) months treatment duration [higher values indicate greater frequency]									
1 cross-sectional study Grannis et al 2023	Very serious limitations[1]	Very serious indirectness[2]	Not applicable	Not calculable	21	29	Testosterone: 1.95 (0.92) No testosterone: 2.86 (1.46) *Difference between groups (ANOVA): F-value 3.05 (df between groups: 1, df within groups: 47); p=NS*	Critical	Very low
Number (%) of individuals reporting eating disorder behaviours (subjective binge episode), measured using the EDE-Q score,[a] at unknown treatment duration/follow-up									
1 cross-sectional study Kramer et al 2024	Very serious limitations[3]	Very serious indirectness[4]	Not applicable	Not calculable	39	142	Testosterone: Behaviour present: 5 (13.5) Behaviour not present: 32 (86.5) No testosterone: Behaviour present: 16 (18.0) Behaviour not present: 73 (82.0) *Difference between groups: p>0.05*	Critical	Very low
Number (%) of individuals reporting eating disorder behaviours (objective binge episode), measured using the EDE-Q score,[a] at unknown treatment duration/follow-up									
1 cross-sectional study Kramer et al 2024	Very serious limitations[3]	Very serious indirectness[4]	Not applicable	Not calculable	39	142	Testosterone: Behaviour present: 4 (10.8) Behaviour not present: 33 (89.2) No testosterone: Behaviour present: 19 (21.1) Behaviour not present: 71 (78.9)	Critical	Very low

							Difference between groups: p>0.05		
Number (%) of individuals reporting eating disorder behaviours (self-induced vomiting), measured using the EDE-Q score,[a] at unknown treatment duration/follow-up									
1 cross-sectional study Kramer et al 2024	Very serious limitations[3]	Very serious indirectness[4]	Not applicable	Not calculable	39	142	Testosterone: Behaviour present: 1 (2.7) Behaviour not present: 36 (97.3) No testosterone: Behaviour present: 1 (1.1) Behaviour not present: 89 (98.8) *Difference between groups: p>0.05*	Critical	Very low
Number (%) of individuals reporting eating disorder behaviours (laxative use), measured using the EDE-Q score,[a] at unknown treatment duration/follow-up									
1 cross-sectional study Kramer et al 2024	Very serious limitations[3]	Very serious indirectness[4]	Not applicable	Not calculable	39	142	Testosterone: Behaviour present: 2 (5.1) Behaviour not present: 37 (94.9) No testosterone: Behaviour present: 6 (6.3) Behaviour not present: 90 (93.8) *Difference between groups: p>0.05*	Critical	Very low
Number (%) of individuals reporting eating disorder behaviours (compensatory exercise), measured using the EDE-Q score,[a] at unknown treatment duration/follow-up									
1 cross-sectional study Kramer et al 2024	Very serious limitations[3]	Very serious indirectness[4]	Not applicable	Not calculable	39	142	Testosterone: Behaviour present: 7 (18.9) Behaviour not present: 30 (81.1) No testosterone: Behaviour present: 16 (17.8) Behaviour not present: 74 (82.2) *Difference between groups: p>0.05*	Critical	Very low
Safety (1 cross-sectional study)									
Odds of cardiometabolic-related diagnoses (overweight/obese), n (%) at unknown treatment duration/follow-up									
1 cross-sectional study Valentine et al 2022	Very serious limitations[1]	Serious indirectness[5]	Not applicable	Not calculable	NR	NR	Testosterone: 491 (52.3) No testosterone: 1,273 (39.4)	Important	Very low

							Difference between testosterone vs no testosterone groups: OR (95% CI) 1.8 (1.5 to 2.1); p<0.0001		
Odds of cardiometabolic-related diagnoses (dyslipidaemia), n (%) at unknown treatment duration/follow-up									
1 cross-sectional study Valentine et al 2022	Very serious limitations[1]	Serious indirectness[5]	Not applicable	Not calculable	NR	NR	Testosterone: 133 (14.2) No testosterone: 218 (6.7) *Difference between testosterone vs no testosterone groups: OR (95% CI) 1.7 (1.3 to 2.3); p<0.01*	Important	Very low
Odds of cardiometabolic-related diagnoses (liver dysfunction), n (%) at unknown treatment duration/follow-up									
1 cross-sectional study Valentine et al 2022	Very serious limitations[1]	Serious indirectness[5]	Not applicable	Not calculable	NR	NR	Testosterone: 134 (14.3) No testosterone: 326 (10.1) *Difference between testosterone vs no testosterone groups: OR (95% CI) 1.5 (1.1 to 1.9); p≤0.01*	Important	Very low
Odds of cardiometabolic-related diagnoses (dysglycaemia), n (%) at unknown treatment duration/follow-up									
1 cross-sectional study Valentine et al 2022	Very serious limitations[1]	Serious indirectness[5]	Not applicable	Not calculable	NR	NR	Testosterone: 18 (1.9) No testosterone: 72 (2.2) *Difference between testosterone vs no testosterone groups: OR (95% CI) NR; p=NS*	Important	Very low
Odds of cardiometabolic-related diagnoses (hypertension), n (%) at unknown treatment duration/follow-up									
1 cross-sectional study Valentine et al 2022	Very serious limitations[1]	Serious indirectness[5]	Not applicable	Not calculable	NR	NR	Testosterone: 105 (11.2) No testosterone: 268 (8.3) *Difference between testosterone vs no testosterone groups: OR (95% CI) 1.6 (1.2 to 2.2); p<0.01*	Important	Very low
Odds of cardiometabolic-related diagnoses (PCOS), n (%) at unknown treatment duration/follow-up									
1 cross-sectional study Valentine et al 2022	Very serious limitations[1]	Serious indirectness[5]	Not applicable	Not calculable	NR	NR	Testosterone: 9 (1.0) No testosterone: 33 (1.0) *Difference between testosterone vs no testosterone groups: OR (95% CI) NR; p=NS*	Important	Very low

Abbreviations
ANOVA: analysis of variance; BID: Body Image Dissatisfaction; CI: confidence interval; CDI: Children's Depression Inventory; df: degrees of freedom; EDE-Q: Eating Disorders Examination-Questionnaire; LSAS: Leibowitz Social Anxiety Scale; n: number; NR: not reported; NS: not significant; OR: odds ratio; PCOS: polycystic ovary syndrome; SCARED: Screen for Child Anxiety Related Emotional Disorders; SD: standard deviation.

1. Risk of bias: very serious limitations due unclear measurement of the exposure, and use of subjective outcome measures
2. Indirectness: very serious indirectness due to inclusion of proportion of participants identifying as non-binary (10%) and inclusion of participants in the no testosterone group who had received GnRH analogues
3. Risk of bias: very serious limitations due to limited reporting on study details and unclear measurement of the exposure
4. Indirectness: serious indirectness due to uncertainty around inclusion of participants receiving GnRH analogues
5. Indirectness: very serious indirectness due to uncertainty around the inclusion of participants identifying as non-binary and whether a proportion of participants had undergone surgery

a. EDE-Q subscales include Restraint, Weight Concern, Shape Concern, and Eating Concern. Individual EDE-Q items reflect specific eating disorder behaviours, ie objective and subjective binge eating episodes, self-induced vomiting, laxative use, and compensatory exercise, and these items were used in the study analyses

Table 2: Masculinising medicines (no comparator)

QUALITY					Summary of findings			IMPORTANCE	CERTAINTY
					No of patients		Effect		
Study	Risk of bias	Indirectness	Inconsistency	Imprecision	Testosterone	No comparator	Result		
Impact on mental health (1 retrospective case series)									
Suicidality measured using the ASQ,[a] mean (SE) to at least 3 months follow-up [higher scores indicate greater levels of suicidal ideation]									
1 retrospective case series Allen et al 2019	Very serious limitations[1]	Very serious indirectness[2]	Not applicable	Not calculable	33	NA	Baseline: 1.01 (0.23) At least 3 months: 0.29 (0.13); *p value not reported*	Critical	Very low
Impact on quality of life (1 retrospective case series, 1 case series)									
QoL measured using PedsQL, median (IQR) at 3 months follow-up [higher scores indicate benefit]									
1 prospective case series Baines et al 2023	Very serious limitations[3]	Very serious indirectness[4]	Not applicable	Not calculable	20	NA	Baseline (n=24): 66.08 (60.87 to 80.98) At 3 months (n=20): 64.31 (54.97 to 76.63); *p value not reported*	Critical	Very low
Wellbeing measured using PedsQL GWBS,[b] mean (SE) to at least 3 months follow-up [higher scores indicate benefit]									
1 retrospective case series Allen et al 2019	Very serious limitations[1]	Very serious indirectness[2]	Not applicable	Not calculable	33	NA	Baseline: 64.95 (2.66) At least 3 months: 70.94 (2.35); *p value not reported*	Critical	Very low
QoL measured using PedsQL, median (IQR) at 6 months follow-up [higher scores indicate benefit]									
1 prospective case series Baines et al 2023	Very serious limitations[5]	Very serious indirectness[4]	Not applicable	Not calculable	18	NA	Baseline (n=24): 66.08 (60.87 to 80.98) At 6 months (n=18): 61.41 (56.52 to 85.87); *difference in QoL from baseline to 6 months: p=0.71*	Critical	Very low
QoL measured using PedsQL, median (IQR) at 9 months follow-up [higher scores indicate benefit]									
1 prospective case series	Very serious limitations[3]	Very serious indirectness[4]	Not applicable	Not calculable	5	NA	Baseline (n=24): 66.08 (60.87 to 80.98)	Critical	Very low

Baines et al 2023							At 9 months (n=5): 59.78 (42.39 to 66.30); *p value not reported*		
Masculinising physical changes (1 prospective case series, 2 retrospective cohort studies)									
Perceived physical changes using unvalidated masculinising effects questionnaire,[c] median (IQR) at 3 months follow-up [higher scores indicate greater perceived physical changes]									
1 prospective case series Baines et al 2023	Very serious limitations[5]	Very serious indirectness[4]	Not applicable	Not calculable	20	NA	16.0 (12.5 to 20.0)	Important	Very low
Perceived physical changes using unvalidated masculinising effects questionnaire,[c] median (IQR) at 6 months follow-up [higher scores indicate greater perceived physical changes]									
1 prospective case series Baines et al 2023	Very serious limitations[5]	Very serious indirectness[4]	Not applicable	Not calculable	18	NA	19.5 (17.0 to 23.0) *Difference in perceived physical changes from 3 to 6 months: p<0.001*	Important	Very low
Perceived physical changes using unvalidated masculinising effects questionnaire,[c] median (IQR) at 9 months follow-up [higher scores indicate greater perceived physical changes]									
1 prospective case series Baines et al 2023	Very serious limitations[3]	Very serious indirectness[4]	Not applicable	Not calculable	5	NA	21.0 (20.0 to 22.0); *p value not reported*	Important	Very low
Number (%) of participants reporting breakthrough bleeding,[d] at mean 28.5 (SD 14.6) [no breakthrough bleeding] and 37.3 (SD 16.9) [breakthrough bleeding] months treatment duration									
1 retrospective case series Grimstad et al 2021	Very serious limitations[6]	Very serious indirectness[7]	Not applicable	Not calculable	232	NA	No breakthrough bleeding: 174 (75) Breakthrough bleeding: 58 (25);[e] *p value not reported*	Important	Very low
Final adult height (cm), mean (SD) at unknown treatment duration/follow-up [higher scores indicate benefit]									
1 retrospective case series Persky et al 2024	Serious limitations[8]	Very serious indirectness[9]	Not applicable	Not calculable	62	NA	Baseline: 164.1 (6.8) Follow-up: 164.7 (6.7)[f] *Difference between baseline and follow-up: NS*	Important	Very low

Final adult height Z-scores,[9] mean (SD) at unknown treatment duration/follow-up [higher scores indicate benefit]									
1 retrospective case series Persky et al 2024	Serious limitations[8]	Very serious indirectness[9]	Not applicable	Not calculable	62	NA	Baseline: 0.21 (1.0) Follow-up: 0.21 (1.0) *Difference between baseline and follow-up: NS*	Important	Very low
Detransition after receipt of masculinising medicines (1 retrospective case series)									
Number of participants discontinuing treatment *"due to desire to end masculinizing therapy"*, at median 1.9 years (range 6 months to 5.5 years) follow-up									
1 retrospective case series Laurenzano et al 2021	Very serious limitations[10]	Very serious indirectness[11]	No applicable	Not calculable	119	NA	The authors reported that *"two stopped because they were satisfied with effects of SC-T achieved and one stopped within 6 months of starting after reassessing their gender identity… due to concerns about body changes and potential impact on fertility"; p value not reported*	Important	Very low
Safety (1 prospective case series, 3 retrospective cohort studies)									
Number (%) of individuals reporting skin irritation, at initiation of treatment									
1 prospective case series Baines et al 2023	Very serious limitations[3]	Very serious indirectness[4]	No applicable	Not calculable	26	NA	3 (11.5)*; p value not reported*	Important	Very low
Incidence of VTE or arterial thrombosis (including stroke), n (%) at median 577 days (IQR 283 to 923) follow-up									
1 retrospective case series Mullins et al 2021	Serious limitations[12]	Very serious indirectness[13]	No applicable	Not calculable	429	NA	0 (0)	Important	Very low
Number (%) of individuals reporting elevated AST, at 3 months follow-up									
1 prospective case series	Very serious limitations[3]	Very serious indirectness[4]	No applicable	Not calculable	26	NA	1 (3.8) [≤36 U/L]*; p value not reported*	Important	Very low

Baines et al 2023									
Change in haemoglobin (mg/dL), median (IQR) at 6 months follow-up [direction of benefit unclear]									
1 retrospective case series Millington et al 2024	Very serious limitations[14]	Very serious indirectness[15]	No applicable	Not calculable	155	NA	Baseline (n=189): 13.2 (12.5 to 13.9) At 6 months (n=155): 14.2 (13.3 to 15.1) *Difference compared to baseline p<0.001*	Important	Very low
Change in haematocrit (%) median (IQR) at 6 months follow-up [direction of benefit unclear]									
1 retrospective case series Millington et al 2024	Very serious limitations[14]	Very serious indirectness[15]	No applicable	Not calculable	157	NA	Baseline (n=191): 39.9 (37.8 to 41.6) At 6 months (n=157): 43.8 (41.6 to 45.8) *Difference compared to baseline p<0.001*	Important	Very low
Change in HbA1c (%) median (IQR) at 6 months follow-up[h] [direction of benefit unclear]									
1 retrospective case series Millington et al 2024	Very serious limitations[14]	Very serious indirectness[15]	No applicable	Not calculable	66	NA	Baseline (n=105): 5.2 (5.0 to 5.4)[i] At 6 months (n=66): 5.1 (5.0 to 5.4) *Difference compared to baseline p=NS*	Important	Very low
Change in ALT (U/L) median (IQR) at 6 months follow-up[j] [direction of benefit unclear]									
1 retrospective case series Millington et al 2024	Very serious limitations[14]	Very serious indirectness[15]	No applicable	Not calculable	58	NA	Baseline (n=77): 17 (11 to 25) At 6 months (n=58): 19.5 (14 to 28) *Difference compared to baseline p=NS*	Important	Very low
Number (%) of individuals reporting elevated ALT, at 3 months follow-up									
1 prospective case series Baines et al 2023	Very serious limitations[3]	Very serious indirectness[4]	No applicable	Not calculable	26	NA	1 (3.8) [≤60 U/L]*; p value not reported*	Important	Very low
Change in AST (U/L) median (IQR) at 6 months follow-up [direction of benefit unclear]									
1 retrospective case series	Very serious limitations[14]	Very serious indirectness[15]	No applicable	Not calculable	58	NA	Baseline (n=77): 20 (17 to 25) At 6 months (n=58): 23 (19 to 29)	Important	Very low

Millington et al 2024							*Difference compared to baseline p=NS*		
Number (%) of individuals reporting elevated haemoglobin at 9 months follow-up									
1 prospective case series Baines et al 2023	Very serious limitations[3]	Very serious indirectness[4]	No applicable	Not calculable	26	NA	0 (0)	Important	Very low
Number (%) of individuals reporting elevated haematocrit at 9 months follow-up									
1 prospective case series Baines et al 2023	Very serious limitations[3]	Very serious indirectness[4]	No applicable	Not calculable	26	NA	0 (0)	Important	Very low
Number (%) of individuals reporting significant dyslipidaemia at 9 months follow-up									
1 prospective case series Baines et al 2023	Very serious limitations[3]	Very serious indirectness[4]	No applicable	Not calculable	26	NA	0 (0)	Important	Very low
Change in haemoglobin (mg/dL), median (IQR) at 12 months follow-up [direction of benefit unclear]									
1 retrospective case series Millington et al 2024	Very serious limitations[14]	Very serious indirectness[15]	No applicable	Not calculable	136	NA	At 6 months (n=155): 14.2 (13.3 to 15.1) At 12 months (n=136): 14.7 (13.6 to 15.6) *Difference compared to 6 months follow-up p<0.001*	Important	Very low
Change in haematocrit (%) median (IQR) at 12 months follow-up [direction of benefit unclear]									
1 retrospective case series Millington et al 2024	Very serious limitations[14]	Very serious indirectness[15]	No applicable	Not calculable	136	NA	At 6 months (n=157): 43.8 (41.6 to 45.8) At 12 months (n=136): 44.6 (41.6 to 47.0) *Difference compared to 6 months follow-up p<0.001*	Important	Very low
Change in HbA1c (%) median (IQR) at 12 months follow-up[h] [direction of benefit unclear]									

1 retrospective case series Millington et al 2024	Very serious limitations[14]	Very serious indirectness[15]	No applicable	Not calculable	66	NA	At 6 months (n=66): 5.1 (5.0 to 5.4)[i] At 12 months (n=64): 5.1 (4.9 to 5.3) *Difference compared to 6 months follow-up p=NS*	Important	Very low
Change in ALT (U/L) median (IQR) at 12 months follow-up [direction of benefit unclear]									
1 retrospective case series Millington et al 2024	Very serious limitations[14]	Very serious indirectness[15]	No applicable	Not calculable	57	NA	At 6 months (n=58): 19.5 (14 to 28) At 12 months (n=57): 18 (13 to 26) *Difference compared to baseline p=NS*	Important	Very low
Change in AST (U/L) median (IQR) at 12 months follow-up [direction of benefit unclear]									
1 retrospective case series Millington et al 2024	Very serious limitations[14]	Very serious indirectness[15]	No applicable	Not calculable	57	NA	At 6 months (n=58): 23 (19 to 29) At 12 months (n=57): 22 (18 to 27) *Difference compared to 6 months follow-up p=NS*	Important	Very low
Number (%) of individuals reporting pelvic pain, abdomino-pelvic pain or pain in the lower part of the abdomen [no validated measurement tool used],[k] at median 22.1 (IQR median value 15.4) months follow-up									
1 retrospective case series Moussaoui et al 2024	Very serious limitations[16]	Very serious indirectness[17]	No applicable	Not calculable	158	NA	Pelvic pain: 37 (23.4)[l] (pelvic pain description: cramps 17 [45.9]; similar to period pain 8 [21.6]; sex-related pain 10 [27]; early morning/or waking up at night 5 [13.5]; suspicion of pelvic floor spasms 5 [13.5]; associated with breakthrough bleeding 11 [29.7]; associated with nausea and/vomiting 4 [10.8]; pain radiating into lower limbs 2 [5.4]) No pelvic pain: 121 (76.6); *p value not reported*	Important	Very low
Change in total cholesterol (mg/dL), mean (SD) at median 1.9 years (range 6 months to 5.5 years) follow-up [benefit indicated by lower total cholesterol]									
1 retrospective case series Laurenzano et al 2021	Very serious limitations[18]	Very serious indirectness[11]	No applicable	Not calculable	119	NA	Baseline: 158.0 (29.8) Final dose: 154.6 (31.1) *Difference from baseline to final dose: -3.5 (0.194); p=NS*	Important	Very low
Change in haematocrit (%), mean (SD) at median 1.9 years (range 6 months to 5.5 years) follow-up [direction of benefit unclear]									

1 retrospective case series Laurenzano et al 2021	Very serious limitations[18]	Very serious indirectness[11]	No applicable	Not calculable	119	NA	Baseline: 39.2 (2.6) Final dose: 44.1 (3.3) *Difference from baseline to final dose: 4.9 (SD not reported); p<0.001*	Important	Very low
Change in ALT (U/L), mean (SD) at median 1.9 years (range 6 months to 5.5 years) follow-up [direction of benefit unclear]									
1 retrospective case series Laurenzano et al 2021	Very serious limitations[18]	Very serious indirectness[11]	No applicable	Not calculable	119	NA	Baseline: 20.8 (10.0) Final dose: 21.4 (12.9) *Difference from baseline to final dose: 0.7 (0.60); p=NS*	Important	Very low
Change in AST (U/L), mean (SD) at median 1.9 years (range 6 months to 5.5 years) follow-up [direction of benefit unclear]									
1 retrospective case series Laurenzano et al 2021	Very serious limitations[18]	Very serious indirectness[11]	No applicable	Not calculable	119	NA	Baseline: 21.8 (8.0) Final dose: 23.0 (8.8) *Difference from baseline to final dose: 1.1 (0.29); p=NS*	Important	Very low
Number (%) of individuals reporting mild injection site reactions, at median 1.9 years (range 6 months to 5.5 years) follow-up									
1 retrospective case series Laurenzano et al 2021	Very serious limitations[10]	Very serious indirectness[11]	No applicable	Not calculable	119	NA	14 (11.8); *p value not reported*	Important	Very low
Number (%) of individuals reporting hypertension, at median 1.9 years (range 6 months to 5.5 years) follow-up									
1 retrospective case series Laurenzano et al 2021	Very serious limitations[18]	Very serious indirectness[11]	No applicable	Not calculable	119	NA	0 (0)	Important	Very low
Number (%) of individuals reporting progression of acne, at median 1.9 years (range 6 months to 5.5 years) follow-up									
1 retrospective case series Laurenzano et al 2021	Very serious limitations[10]	Very serious indirectness[11]	No applicable	Not calculable	119	NA	77 (64.7) [advanced acne management, ie oral treatment and/or referral to dermatology, was reported in n=23 of these participants]; *p value not reported*	Important	Very low
Number (%) of individuals reporting transaminitis, at median 1.9 years (range 6 months to 5.5 years) follow-up									

1 retrospective case series Laurenzano et al 2021	Very serious limitations[18]	Very serious indirectness[11]	No applicable	Not calculable	119	NA	0 (0)	Important	Very low
Number (%) of individuals reporting dyslipidaemia, at median 1.9 years (range 6 months to 5.5 years) follow-up									
1 retrospective case series Laurenzano et al 2021	Very serious limitations[10]	Very serious indirectness[11]	No applicable	Not calculable	119	NA	1 (0.8) [described as worsening of dyslipidaemia]; *p value not reported*	Important	Very low
Number (%) of individuals reporting haematocrit >55%, at median 1.9 years (range 6 months to 5.5 years) follow-up									
1 retrospective case series Laurenzano et al 2021	Very serious limitations[18]	Very serious indirectness[11]	No applicable	Not calculable	119	NA	0 (0)	Important	Very low
Change in haemoglobin (mg/dL), median (IQR) at 24 months follow-up [direction of benefit unclear]									
1 retrospective case series Millington et al 2024	Very serious limitations[14]	Very serious indirectness[15]	No applicable	Not calculable	119	NA	At 12 months (n=136): 14.7 (13.6 to 15.6) At 24 months (n=119): 15.0 (14.1 to 15.8) *Difference compared to 12 months follow-up p=0.01*	Important	Very low
Change in haematocrit (%) median (IQR) at 24 months follow-up [direction of benefit unclear]									
1 retrospective case series Millington et al 2024	Very serious limitations[14]	Very serious indirectness[15]	No applicable	Not calculable	119	NA	At 12 months (n=136): 44.6 (41.6 to 47.0) At 24 months (n=119): 45.4 (42.9 to 47.6)[m] *Difference compared to 12 months follow-up p=0.03*	Important	Very low
Change in HbA1c (%) median (IQR) at 24 months follow-up[h] [direction of benefit unclear]									
1 retrospective case series	Very serious limitations[14]	Very serious indirectness[15]	No applicable	Not calculable	59	NA	At 12 months (n=64): 5.1 (4.9 to 5.3)[i] At 24 months (n=59): 5.1 (4.9 to 5.3)	Important	Very low

Millington et al 2024							*Difference compared to 12 months follow-up p=NS*[n]		
Change in ALT (U/L) median (IQR) at 24 months follow-up [direction of benefit unclear]									
1 retrospective case series Millington et al 2024	Very serious limitations[14]	Very serious indirectness[15]	No applicable	Not calculable	42	NA	At 12 months (n=57): 18 (13 to 26) At 24 months (n=42): 19 (13 to 27)[j] *Difference compared to baseline p=NS*	Important	Very low
Change in AST (U/L) median (IQR) at 24 months follow-up [direction of benefit unclear]									
1 retrospective case series Millington et al 2024	Very serious limitations[14]	Very serious indirectness[15]	No applicable	Not calculable	42	NA	At 12 months (n=57): 18 to 27) At 24 months (n=42): 22 (18 to 28) *Difference compared to 12 months follow-up p=NS*	Important	Very low

Abbreviations

ALT: alanine transaminase; ASQ: Ask Suicide-Screening Questions; AST: aspartate aminotransferase; cm: centimetre; HbA1c: glycated haemoglobin; IQR: interquartile range; mg/dL: milligrams per decilitre; n: number; NA: not applicable; NR: not reported; NS: not significant; Peds QL: Paediatric Quality of Life; QoL: quality of life; SC-T: subcutaneous testosterone; SD: standard deviation; SE: standard error; U/L: units per litre.

1. Risk of bias: very serious limitations due to lack of clinical information for participants, lack of clarity about whether inclusion was consecutive and complete, and lack of statistical measures
2. Indirectness: very serious indirectness due to lack of a comparator group, inclusion of 10% of participants who started treatment with masculinising medicines at >18 years of age, unspecified number of AFAB participants who received GnRH analogues, and an unspecified proportion who identified as non-binary
3. Risk of bias: very serious limitations due to lack of clinical information for participants, lack of clarity about whether inclusion was consecutive and complete, and lack of statistical measures
4. Indirectness: very serious indirectness due to lack of a comparator group, unspecified proportion of participants who were receiving GnRH analogues, and uncertainty around the inclusion of participants who identified as non-binary
5. Risk of bias: very serious limitations due to lack of clinical information for participants, and lack of clarity about whether inclusion was consecutive and complete
6. Risk of bias: very serious limitations due to uncertainty around identification and measurement of gender incongruence, non-consecutive and incomplete inclusion of participants, and lack of statistical measures
7. Indirectness: very serious indirectness due to lack of a comparator group, an unspecified proportion of participants who started treatment with masculinising medicines at >18 years of age, 6.9% of participants who received GnRH analogues and 7.8% who had undergone hysterectomy, and uncertainty around the inclusion of participants who identified as non-binary
8. Risk of bias: serious limitations due to lack of clinical information for participants, and lack of statistical measures
9. Indirectness: very serious indirectness due to lack of a comparator group and uncertainty around the inclusion of participants who identified as non-binary
10. Risk of bias: very serious limitations due to lack of information on clinical information for participants, lack of clarity about whether inclusion was consecutive and complete, and lack of statistical measures
11. Indirectness: very serious indirectness due to lack of a comparator group, an unspecified proportion of participants who started treatment with masculinising medicines at >18 years of age, 10.1% of participants who received GnRH analogues, and a proportion who identified as non-binary (2.5%)
12. Risk of bias: serious limitations due to very serious limitations due to uncertainty around identification and measurement of gender incongruence
13. Indirectness: very serious indirectness due to lack of a comparator group, uncertainty around potential use of GnRH analogues, and a proportion of participants who identified as non-binary (2.8%)
14. Risk of bias: very serious limitations due to lack of clinical information for participants, and lack of clarity about whether inclusion was consecutive and complete

15. Indirectness: very serious indirectness due to lack of a comparator group, inclusion of 16% of participants who had used GnRH analogues in late or post-puberty and unspecified proportion who had used GnRH analogues in early puberty, and inclusion of non-binary participants (5%)
16. Risk of bias: very serious limitations due to uncertainty around identification and measurement of gender incongruence, lack of information on clinical information for participants, lack of clarity about whether inclusion was consecutive and complete, and lack of statistical measures
17. Indirectness: very serious indirectness due to lack of a comparator group, an unspecified proportion of participants who started treatment with masculinising medicines at >18 years of age, 9.5% of participants who received GnRH analogues, and uncertainty around inclusion of participants who identified as non-binary
18. Risk of bias: very serious limitations due to lack of information on clinical information for participants, and lack of clarity about whether inclusion was consecutive and complete

a. ASQ is a four-item measure used to identify patients who are at risk of attempting suicide. Questions include: In the past few weeks have you… "…wished you were dead?", "…felt that you or your family would be better off if you were dead?", "…been having thoughts about harming or killing yourself?", or "…done anything to hurt yourself or to end your life?" A response of "no" was scored as 0 and a response of "yes" was scored as 1; with an overall score for suicidality on a scale ranging from 0 to 4
b. Wellbeing was measured using the Paediatric Quality of Life Inventory General Wellbeing Scale (PedsQL GWBS) which is a 5-point response scale, containing seven items, and measures "general well-being" and "general health". The general well-being subscale includes six items (eg "I feel happy" and "I think my health will be good in the future"). Participants are asked to consider each item over the past month and rate responses from 0 (never) to 4 (almost always). The general health subscale contains one item, "In general, how is your health?" ranging from 0 (Bad) to 4 (Excellent). All items are scored and linearly transformed to a 0 to 100 scale (initial score of 0 = 0, 1 = 25, 2 = 50, 3 = 75, and 4 = 100)
c. Masculinising effects questionnaire assessed perceived physical changes after starting testosterone, including skin oiliness/acne, facial hair, body hair, increased muscle mass/strength, menstrual cessation, and clitoral enlargement
The authors stated that due to the variability in height measurements reported in the medical records, the average of all heights recorded after the final adult height was reached was calculated and this was used for data analysis purposes
d. Breakthrough bleeding was defined as any bleeding presumed to originate in the uterus while on testosterone
e. For participants with breakthrough bleeding, bleeding started at a mean of 24.3 (SD 7.2) months after initiation of testosterone. Mean age at time of first breakthrough bleeding was 18.4 (SD 2.8) years. Eight (13.8%) of these participants had never become amenorrhoeic in the first year and continued to bleed beyond the first year. A total of 48 patients (82.8%) had more than one episode of bleeding after one year on testosterone
f. The authors stated that due to the variability in height measurements reported in the medical records, the average of all heights recorded after the final adult height was reached was calculated and this was used for data analysis purposes
g. Final adult height Z-scores were calculated based on the CDC 20 year old growth chart for girls
h. Five participants with pre-existing diabetes mellitus were excluded from the HbA1c analysis
i. There were 10 (5.1%) participants with baseline HbA1c measurements in the prediabetes range (5.7% to 6.4%). Of these, three participants did not have additional follow-up measurements, five had subsequently normal measurements, and two had persistently elevated HbA1c measurements of 5.7%
j. Four (2%) participants had elevations in liver enzymes >2 times the upper limit of normal during testosterone therapy (range 118 to 263 U/L). The authors stated that all four participants were receiving concurrent treatment with other medications that have been associated with abnormal liver function tests (ie isotretinoin, quetiapine, lamotrigine, bupropion,). Three of these four participants had subsequent ALT and AST measurements that returned to normal during the study visit. The fourth participant, who was taking isotretinoin and sertraline, continued to have elevated ALT at the 24-month follow-up visit
k. Pelvic pain was defined as the timing of onset as reported by the individual, or as the date of first documentation in medical chart if no mention of timing of onset. The median interval between testosterone initiation and onset of pain was 1.6 months (range 0.3 to 6.4)
l. Thirty six of 37 participants reporting pelvic pain had not received past puberty blockers. Pain intensity was reported for n=11 adolescents: n=1 mild (self-reported score of 1 to 3 out of 10, with a score of 10 being most severe); n=10 severe (self-reported score of 7 to 10 out of 10, with a score of 10 being most severe)
m. There were 13 participants (6.5%) DFAB, one of whom had recent tobacco use, who had haematocrit above the typical cisgender male range (haematocrit 40% to 50%) during treatment with testosterone, ranging from 50.1% to 53.1%. Apart from a decrease in the testosterone dose, there were no other clinical interventions (eg therapeutic phlebotomy) required to address the increased haematocrit
n. One participant had an increase in HbA1c from 5.4% to 5.7% over the 24-month treatment period. This participant also had an increase in BMI from 24.5 to 27.8 kg/m2. Another participant had an increase in HbA1c from 5.6% to 6.1% at the 12-month follow-up visit, but HbA1c was in the normal range at the 24-month follow-up visit (4.3%). This participant also had a baseline BMI in the obese range (32.0 kg/m2). In total, there were 15 (7.7%) participants who had an elevated HbA1c measurement at any point during the study period

Glossary

Term	Definition[82]
Adverse event	Any undesirable event experienced by a person while they are having a drug or any other treatment or intervention, regardless of whether the event is suspected to be related to or caused by the drug, treatment, or intervention.
Baseline	The set of measurements at the beginning of a study (after any initial 'run-in' period with no intervention), with which subsequent results are compared.
Case series	Reports of several patients with a given condition, usually covering the course of the condition and the response to treatment. There is no comparison (control) group of patients.
Cisgender	Used to describe a person whose personal identity and gender identity is the same as their birth registered sex.
Clinical importance	A benefit from treatment that relates to an important outcome such as length of life and is large enough to be important to patients and health professionals.
Comparator	The standard (for example, another intervention or usual care) against which an intervention is compared in a study. The comparator can be no intervention (for example, best supportive care).
Cost-effectiveness analysis	An analysis that assesses the cost of achieving a benefit by different means. The benefits are expressed in non-monetary terms related to health, such as life years gained (that is, the number of years by which life is extended as a result of the intervention). Options are often compared on the cost incurred to achieve 1 outcome (for example, cost per life year gained).
Cross-sectional study	A 'snapshot' observation of a set of people at one time. This type of study (sometimes called a cross-sectional survey) contrasts with a longitudinal study, which follows a set of people over a period of time.
Detransition/detransitioners	The process of discontinuing or reversing a gender transition, often in connection with a change in how the individual identifies or conceptualises their sex or gender since initiating transition (MacKinnon et al 2023).
Diagnostic and Statistical Manual of Mental Disorders Fifth Edition (DSM-5)	The standard classification of mental disorders used by mental health professionals in the UK and internationally, published by the American Psychiatric Association (2013). The Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, Text Revision (DSM-5-TR) was released in 2022.
Endocrine treatment	Sometimes referred to hormone treatment/therapy. 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).

[82] These definitions are taken from the NICE glossary https://www.nice.org.uk/glossary and the glossary from the Cass Review [ARCHIVED CONTENT] Final Report – Cass Review

Term	Definition[82]
Feminising and masculinising hormones (also known as cross-sex hormones, and gender affirming hormones)	Sex hormones given as part of a medical transition for gender dysphoric individuals (testosterone for transgender males and oestrogen for transgender females).
Gender dysphoria	Diagnostic term used by health professionals and found in DSM-5. 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 criteria.
Gender fluid	An experience of gender that is not fixed, but changes between two or more identities.
Gender identity	Diagnostic term used by health professionals, found in the WHO International Classification of Diseases ICD-11 (see below). Gender incongruence is characterised by "a marked and persistent incongruence between an individual's experienced gender and the assigned sex".
Gender incongruence	Diagnostic term used by health professionals, found in the WHO International Classification of Diseases ICD-11 (see below). Gender incongruence is characterised by "a marked and persistent incongruence between an individual's experienced gender and the assigned sex".
Gonadotropin releasing hormone analogues (also known as hormone blockers and puberty blockers)	Taking these hormones stops the progress of puberty. The GnRH analogues (puberty blockers) act by competing with the body's natural gonadotrophin releasing hormone. This competition blocks the release of two gonadotrophin hormones important in puberty called Follicular Stimulating Hormone (FSH) and Luteinising Hormone (LH) from the pituitary gland.
GRADE (Grading of recommendations assessment, development and evaluation)	A systematic and explicit approach to grading the quality of evidence and the strength of recommendations developed by the GRADE working group.
International Classification of Diseases 11th Revision	The International Classification of Diseases (ICD) is a globally used medical classification of anything that is relevant to health care and is used clinically for medical diagnosis. (https://icd.who.int/en). It is developed and annually updated by the World Health Organization (WHO) and is the mandatory global data standard for recording health information. It is currently in its 11th revision (ICD-11).
Masculinising hormones (also known as cross-sex hormones, and gender affirming hormones)	Sex hormones given as part of a medical transition for gender dysphoric individuals (testosterone for transgender males).
Minimal clinically important difference	The smallest change in a treatment outcome that people with the condition would identify as important (either beneficial or harmful), and that would lead a person or their clinician to consider a change in treatment.

Term	Definition[82]
Non-binary	A gender identity that does not fit into the traditional gender binary of male and female (Twist & de Graaf, 2018).
PICO (population, intervention, comparison and outcome) framework	A structured approach for developing review questions that divides each question into 4 components: the population (the population being studied); the interventions (what is being done); the comparators (other main treatment options); and the outcomes (measures of how effective the interventions have been).
Retrospective study	A research study that focuses on the past and present. The study examines past exposure to suspected risk factors for the disease or condition. Unlike prospective studies, it does not cover events that occur after the study group is selected.
Psychosocial	Describes the psychological and social factors that encompass broader wellbeing.
Standard deviation	A measure of the spread, scatter or variability of a set of measurements. Usually used with the mean (average) to describe numerical data.
Statistical significance	A statistically significant result is one that is assessed as being due to a true effect rather than random chance.
Subgroup analysis	A way to find out from a study if a treatment is more effective in one group of people (for example, who are a particular age or have particular symptoms) than another. It uses evidence from a defined subgroup within the whole analysis set.
Tanner Stage	Tanner Staging, also known as Sexual Maturity Rating, is a classification of puberty by stage of development. This ranges from Stage 1, before physical signs of puberty appear, to Stage 5 at full maturity. The name originates from Professor JM Tanner, a child development expert, was the first to identify the visible stages of puberty.
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.

References

Included studies

- Allen LR, Watson LB, Egan AM, Moser CN. Well-being and suicidality among transgender youth after gender-affirming hormones. Clin. 2019;7(3):302-11.
- Baines HK, Connelly KJ. A prospective comparison study of subcutaneous and intramuscular testosterone injections in transgender male adolescents. J Pediatr Endocrinol Metab. 2023;36(11):1028-36.
- Grannis C, Mattson WI, Leibowitz SF, Nahata L, Chen D, Strang JF, et al. Expanding upon the relationship between gender-affirming hormone therapy, neural connectivity, mental health, and body image dissatisfaction. Psychoneuroendocrinology. 2023;156:106319.
- Grimstad F, Kremen J, Shim J, Charlton BM, Boskey ER. Breakthrough Bleeding in Transgender and Gender Diverse Adolescents and Young Adults on Long-Term Testosterone. J Pediatr Adolesc Gynecol. 2021;34(5):706-16.
- Kramer R, Aarnio-Peterson CM, Conard LA, Lenz KR, Matthews A. Eating disorder symptoms among transgender and gender diverse youth. Clin. 2024;29(1):30-44.
- Laurenzano SE, Newfield RS, Lee E, Marinkovic M. Subcutaneous Testosterone Is Effective and Safe as Gender-Affirming Hormone Therapy in Transmasculine and Gender-Diverse Adolescents and Young Adults: A Single Center's 8-Year Experience. Transgend Health. 2021;6(6):343-52.
- Millington K, Lee JY, Olson-Kennedy J, Garofalo R, Rosenthal SM, Chan YM. Laboratory Changes During Gender-Affirming Hormone Therapy in Transgender Adolescents. Pediatrics. 2024;153(5):01.
- Moussaoui D, Elder CV, O'Connell MA, McLean A, Grover SR, Pang KC. Pelvic pain in transmasculine adolescents receiving testosterone therapy. Int J Transgend Health. 2024;25(1):10-8.
- Mullins ES, Geer R, Metcalf M, Piccola J, Lane A, Conard LAE, et al. Thrombosis Risk in Transgender Adolescents Receiving Gender-Affirming Hormone Therapy. Pediatrics. 2021;147(4):04.
- Persky RW, Apple D, Dowshen N, Pine E, Whitehead J, Barrera E, et al. Pubertal Suppression in Early Puberty Followed by Testosterone Mildly Increases Final Height in Transmasculine Youth. J. 2024;8(6):bvae089.
- Valentine A, Davis S, Furniss A, Dowshen N, Kazak AE, Lewis C, et al. Multicenter Analysis of Cardiometabolic-related Diagnoses in Transgender and Gender-Diverse Youth: A PEDSnet Study. J Clin Endocrinol Metab. 2022;107(10):e4004-e14.

Other references

- American Psychiatric Association, DSM-5 Task Force. (2013). Diagnostic and statistical manual of mental disorders: DSM-5™ (5th ed.). American Psychiatric Publishing, Inc.. https://doi.org/10.1176/appi.books.9780890425596
- American Psychiatric Association (2022). Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, Text Revision (DSM-5-TR). https://doi/book/10.1176/appi.books.9780890425787
- American Psychiatric Association (2022). Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, Text Revision (DSM-5-TR). https://doi/book/10.1176/appi.books.9780890425787

NHS England
Wellington House
133-155 Waterloo Road
London
SE1 8UG