OPEN ACCESS

How does hormone transition in transgender women change body composition, muscle strength and haemoglobin? Systematic review with a focus on the implications for sport participation

Joanna Harper,[1] Emma O'Donnell,[1] Behzad Sorouri Khorashad,[2] Hilary McDermott,[1] Gemma L Witcomb [1]

[1]School of Sport, Exercise and Health Sciences, Loughborough University, Loughborough, UK
[2]Department of Women's and Children's Health, Karolinska Institute, Stockholm, Sweden

Correspondence to
Dr Gemma L Witcomb, School of Sport, Exercise and Health Sciences, Loughborough University, Loughborough, UK; G.L.Witcomb@lboro.ac.uk

Accepted 4 February 2021
Published Online First 1 March 2021

ABSTRACT

Objectives We systemically reviewed the literature to assess how long-term testosterone suppressing gender-affirming hormone therapy influenced lean body mass (LBM), muscular area, muscular strength and haemoglobin (Hgb)/haematocrit (HCT).

Design Systematic review.

Data sources Four databases (BioMed Central, PubMed, Scopus and Web of Science) were searched in April 2020 for papers from 1999 to 2020.

Eligibility criteria for selecting studies Eligible studies were those that measured at least one of the variables of interest, included transwomen and were written in English.

Results Twenty-four studies were identified and reviewed. Transwomen experienced significant decreases in all parameters measured, with different time courses noted. After 4 months of hormone therapy, transwomen have Hgb/HCT levels equivalent to those of cisgender women. After 12 months of hormone therapy, significant decreases in measures of strength, LBM and muscle area are observed. The effects of longer duration therapy (36 months) in eliciting further decrements in these measures are unclear due to paucity of data. Notwithstanding, values for strength, LBM and muscle area in transwomen remain above those of cisgender women, even after 36 months of hormone therapy.

Conclusion In transwomen, hormone therapy rapidly reduces Hgb to levels seen in cisgender women. In contrast, hormone therapy decreases strength, LBM and muscle area, yet values remain above that observed in cisgender women, even after 36 months. These findings suggest that strength may be well preserved in transwomen during the first 3 years of hormone therapy.

INTRODUCTION

Currently the world of sport, from grassroots level to elite, is facing the challenge of how to include transgender people in sporting competitions. Regulations governing the participation of athletes from outside the sex/gender binary have existed since the 1940s.[1–4] Presently, World Athletics requires that transgender athletes[5] and athletes with differences of sexual development[6] have testosterone levels ≤5 nmol/L in order to be eligible for the female category. There has been heavy criticism of this, and previous, testosterone-based regulations.[7–9] Although no openly transgender athlete has competed in the Olympics to date, the increasing visibility of gender-diverse people in society[10] means that the sports administrators and legislators must create rules to accommodate athletes from outside the sex/gender binary.[11]

There are many quantifiable performance-related differences between male and female athletes. In contrast, the performance-related differences between transwomen who have received gender affirming hormone treatment (GAHT) and cisgender women are less clear. GAHT for transwomen consists of an antiandrogen agent plus the introduction of exogenous oestrogen,[12] with the goal of altering the hormonal milieu and, as a result, feminisation of the body.[13] To date, there have been no prospective studies investigating the changes in athletic performance in transgender athletes after hormonal transition. In non-athletic transgender populations, studies are commonly focused on clinical outcomes, such as bone health.[14] However, studies in non-athletic transwomen undergoing GAHT also report changes in lean body mass (LBM),[15] muscle cross-sectional area (CSA),[16] muscular strength[17] and haemoglobin (Hgb)[18] and/or haematocrit (HCT).[19] These parameters are of relevance to athletic performance.

In endurance sports, Hgb is of importance. Hgb is a protein carried by the red blood cells that is responsible for transporting oxygen from the lungs to peripheral tissues.[20] Low Hgb, or low HCT, the volume of red blood cells compared with total blood volume, can lead to a diminished supply of oxygen to the tissues, and therefore have a direct effect on endurance performance. Typical values for Hgb differ between males and females, with 'normal' values ranging between 131–179 g/L for men and 117–155 g/L for women.[21] HCT values are also higher in males (42%–52%) than females (37%–47%).[22] Testosterone exerts erythrogenic effects that results in increases in both HCT and Hgb.[23] Since GAHT significantly lowers testosterone levels in transgender women,[24] it is possible that they may experience reductions in HCT and Hgb, which would be anticipated to negatively affect endurance performance.

In sports demanding speed and power, muscular strength and the ability to generate high rates of force are recognised as key determinant in athletic success.[25] In cisgender males, increases in testosterone due to puberty promote muscular strength

To cite: Harper J, O'Donnell E, Sorouri Khorashad B, *et al*. *Br J Sports Med* 2021;**55**:865–872.

in association with increased muscle CSA, and increased lean muscle mass.[26] It has been hypothesised that muscle retains a long-term memory allowing it to perform tasks that it has undertaken many times previously and myonuclei retention is thought to play an important role in such muscle memory.[27] Myonuclei number is increased with training and with use of anabolic steroids.[28] However, detraining does not diminish the myonuclei number,[27] and it has been hypothesised that cessation of steroids may also not lead to reductions in myonuclei number.[28] Hence, it is possible that strength advantages gained when training in a high-testosterone environment may not be fully reversed by testosterone suppression.

Understanding both the physiological effects of GAHT on athletic performance, and the time course of these effects, is of importance to decision-makers and those undertaking policy reviews. While it is known that testosterone levels are markedly reduced in transgender women taking testosterone suppressing GAHT,[29] the effects of this hormonal change on physiology, and the time course in which these changes occur, are less clear. Individual studies provide crucial, primary research on the topic, but a systematic review is warranted to provide a robust summary of the available evidence. Because bone mineral density studies have already been subject to systematic review,[30 31] this review focuses on physiological changes induced by GAHT in transwomen that affect athletic performance; specifically, LBM, CSA, strength and Hgb/HCT.

Aim

The aim of this systematic review was to: (1) summarise the current state of knowledge as it relates to the changes, and the time course of these changes, in physiological parameters associated with athletic performance in non-athletic transwomen resulting from GAHT (suppression of testosterone and supplementation with oestrogen), and (2) consider the potential implications for the participation of transwomen in elite sport.

MATERIALS AND METHODS

Search strategy and selection criteria

This systematic review was conducted in line with Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.[32] Two electronic searches of four online databases (BioMed Central, PubMed, Scopus and Web of Science) were completed 15 months apart. The first was performed by BSK in January 2019 and the second by JH in April 2020. The two sets of search results were compared by GLW. The second search identified novel data from three additional studies using the same cohorts as three studies identified in the first search. The more recent search also identified three additional recent papers. Reference lists were also searched for additional citations pertinent to the review. The searches combined terms related to transwomen, GAHT, muscle and blood parameters (online supplemental table 1).

Study selection, quality assessment, and data extraction

Each study was initially categorised based on its design (eg, cohort, case–control) and examined for quality in line with the Effective Public Health Practise Project (EPHPP) tool.[33] This is a generic tool used to evaluate a variety of intervention study designs and is suitable for use in systematic reviews,[34] having content and construct validity.[35] Based on the EPHPP, six domains are evaluated: (1) selection bias; (2) study design; (3) confounders; (4) blinding; (5) data collection method; and (6) withdrawals/dropouts. Each domain is rated as strong (3 points), moderate (2 points) or weak (1 point), and domain scores are averaged to provide the overall mean rating. Based on the overall mean rating, studies are rated as weak (1.00–1.50), moderate (1.51–2.50) or strong (2.51–3.00).

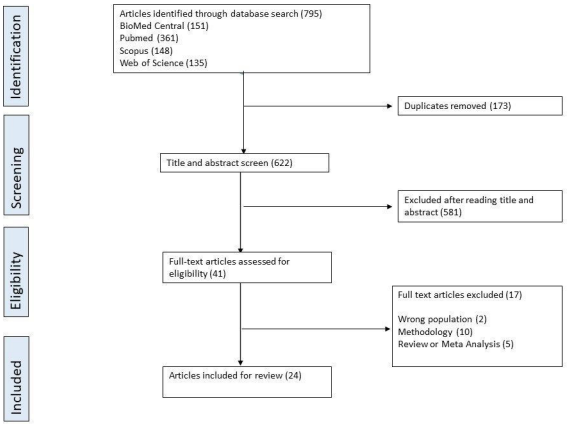

Figure 1 PRISMA flow chart illustrating search strategy. PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

For longitudinal studies, data were extracted to examine changes in LBM, CSA, strength and Hgb/HCT in transwomen taking GAHT. In cross-sectional studies, data in transwomen were compared with data from both cisgender men and cisgender women. The study authors were contacted if there were any questions regarding the presented data. In this regard, authors of the nine studies carried out by the European Network for the Investigation of Gender Incongruence (ENIGI) were contacted regarding potential overlapping participants[15 17 19 36–41] and another author was contacted to clarify graphical data content.[16]

RESULTS

Search results

Figure 1 shows the search strategy following PRISMA guidelines. From an initial yield of 795 articles, 24 studies[15–19 36–54] were included in this review. The following information was extracted from each study: name of the first author, country, year of publication, number of transfemale participants, number of cisgender male and female participants (where applicable), duration of any follow-up, type of medical treatment, method of measurement, evaluation time, and results.

Quality assessment

Based on the mean EPHPP scores, all studies were categorised as moderate in quality. The individual scores are listed in the online supplemental table 2.

Study characteristics

A summary of the study characteristics is reported in table 1. The sample sizes of the studies varied from 12 to 249. Three large studies from the ENIGI group published in 2018 and 2019[15 17 19] contained much novel data, but also included many participants from previous studies making it impossible to accurately state the number of unique participants.

Study designs

Thirteen studies[15 17 19 36–40 42 43 46–48] utilised a follow-up study design comparing participants' measurements before initiating hormone transition (baseline) to several months after hormone transition. Two studies[41 51] used both follow-up and cross-sectional designs with cisgender controls. Six studies[18 45 50 52–54]

Table 1 Characteristics of reviewed studies

Author (year)	Study type	Country	Study quality rating	Participants (N)				Age (years)	Timing (months post GAHT)	T (nmol/L)	Measures			
				TW	CM	CW	HNTW	Mean±SD med (min–max)		Baseline post GAHT	LBM	CSA	MS	Hgb or HCT
Elbers *et al* (1999)[42]	Follow-up	Netherlands	Mod	20	–	–	–	26±6	Baseline 12	22 1.0	N	Y	N	N
Gooren and Bunck (2004)[43]	Follow-up	Netherlands	Mod	19	–	–	–	NR	Baseline 12 36	21.5 1.0 0.9	N	Y	N	Y
Mueller *et al* (2011)[44]	Prospective	Germany	Mod	84	–	–	–	36.3±11.3	Baseline 12 24	13.6 0.6 0.7	Y	N	N	N
Wierckx *et al* (2014)[45]	Follow-up	Norway and Belgium	Mod	53	–	–	–	31.7±14.8 19.3±2.4	Baseline 12	18.4 0.4	Y	N	N	Y
Gava *et al* (2016)[38]	Follow-up	Italy	Mod	40	–	–	–	32.9±9.4 29.4±10.2	Baseline 12	19.2 0.7	Y	N	N	N
Auer *et al* (2016)[46]	Follow-up	Belgium	Mod	20	–	–	–	NR	Baseline 12	20.5 2.0	N	N	Y	Y
Auer *et al* (2018)[40]	Follow-up	Belgium	Mod	45	–	–	–	34.8±1.4	Baseline 12	17.5 1.9	Y	N	N	N
Jarin *et al* (2017)[39]	Follow-up	USA	Mod	13	–	–	–	18 (14–25)	Baseline 6	13.6 6.9	N	N	N	Y
Defreyne *et al* (2018)[19]	Follow-up	Netherlands and Belgium	Mod	239	–	–	–	28.5 (16–65)	Baseline 3 6 24	17.4 0.7 0.6 0.6	N	N	N	Y
Vita *et al* (2018)[48]	Follow-up	Italy	Mod	21	–	–	–	25.2±7.0	Baseline 30	20.5 1.1	N	N	N	Y
Klaver *et al* (2018)[15]	Follow-up	Netherlands and Belgium	Mod	179	–	–	–	29.0 (18–66)	Baseline 12		Y	N	N	N
Olson-Kennedy *et al* (2018)[49]	Prospective	USA	Mod	23	–	–	–	18 (12–23)	Baseline 24	14.8 5.9	N	N	N	Y
Tack *et al* (2018)[36]	Follow-up	Belgium	Mod	21	–	–	–	16.3±1.2	Baseline 5–31	15.2 8.8	Y	Y	Y	N
Tack *et al* (2017)[47]	Follow-up	Belgium	Mod	21	–	–	–	16.3±1.2	Baseline 12–31	15.8 7.8	N	N	N	Y
Scharff *et al* (2019)[17]	Follow-up	Netherlands and Belgium	Mod	249	–	–	–	28 (23–40)	Baseline 12	18.3 0.8	N	N	Y	N
Wiik (2020)[16]	Prospective	Sweden	Mod	11	–	–	–	27±4	Baseline 4 12	18.0 0.5 0.5	N	Y	Y	Y
Van Caenegem *et al* (2014)[45]	Follow-up and cross-sectional	Belgium	Mod	49	49	–	–	33±12 30 (17–67) 33±12	Baseline 12 24 TW Baseline vs CM	19.0 0.5 0.5	Y	Y	Y	N
Haraldsen *et al* (2007)[51]	Follow-up and cross-sectional	Norway	Mod	12	77	–	–	29.3±7.8 33.9±9.3	Baseline 12 TW Baseline vs CM	16.8 6.8	Y	Y	N	N

Continued

Table 1 Continued

Author (year)	Study type	Country	Study quality rating	Participants (N)				Age (years)	Timing (months post GAHT)	T (nmol/L)	Measures			
				TW	CM	CW	HNTW	Mean±SD med (min–max)		Baseline post GAHT	LBM	CSA	MS	Hgb or HCT
SoRelle *et al* (2019)[52]	Cross-sectional	USA	Mod	133	–	–	87	33±12 31±12	TW>6 m vs HNTW	1.9 12.7	N	N	N	Y
Greene *et al* (2019)[18]	Cross-sectional	USA	Mod	93	–	–	–	35.1 (18–69)	TW>12 m vs CW ranges	1.4	N	N	N	Y
Roberts *et al* (2014)[53]	Cross-sectional	USA	Mod	55	20	20	–	46 (27–67) 58 (21–84) 56 (23–88)	TW>6 m vs CM TW>6 m vs CW		N	N	N	Y
Lapauw *et al* (2008)[54]	Cross-sectional	Belgium	Mod	23	20	–	–	41±7 40±7	TW>48 m vs CM	1.1 20.1	Y	Y	Y	Y
Jain *et al* (2019)[50]	Cross-sectional	USA	Mod	277	–	–	102	31±7.1 31±7.1	TW vs HNTW		N	N	N	Y
Sharula (2012)[37]	Cross-sectional	Japan	Mod	129	–	–	22	33.9±10.0 31.5±9.9	TW vs HNTW	2.5 20.5	N	N	N	Y

CM, cismen; CSA, cross-sectional area; CW, ciswomen; HCT, haematocrit; Hgb, haemoglobin; HNTW, hormone-naive transwomen; LBM, lean body mass; TW, transwomen.

used an exclusively cross-sectional design; three comparing transwomen on GAHT with cisgender controls[18 53 54] and three comparing transwomen on GAHT with hormone-naive transwomen.[45 50 52] Three studies[16 44 49] used a prospective method gathering data over 12–24 months. Aside from these three studies, data were extracted from medical charts (nine of which were from the same research group,[15 17 19 36–41]) posing a risk of selective data reporting and publication bias.

Medical treatments

Medical treatments for endocrine transition were varied, in line with the individualised approach advised by the WPATH Standards of Care.[55] Fourteen studies[15 17 19 36–43 46 48 54] used cyproterone acetate (50–100 mg daily) as an antiandrogen. In six studies[16 38 40 44 46 49] a form of gonadotropin-releasing hormone agonist was administered either to supress puberty or androgens. In four studies[18 49 50 52] spironolactone was used as an antiandrogen. Seventeen studies[15 17–19 36–39 41 44 45 47–50 52 53] used 2–4 mg/day of oral oestradiol valerate. Eleven studies[15–17 19 39 42 43 45 46 48 49] used transdermal 17-beta-oestradiol releasing 100 mcg/day. Four studies[16 18 47 49] used an injection of oestradiol valerate (10 mg/ampoule, every 1–4 months). Two studies[45 54] used 0.625–2.5 mg/day of conjugated equine oestrogen. Four studies,[42 43 51 54] all undertaken prior to 2010, used 25–50 mcg/day of ethinyl oestradiol. Ethinyl oestradiol was not used in any study after 2010, primarily due to increased risk of thrombogenesis.[56]

Based on the variability in drug regimens used, there is substantial heterogeneity in the hormone levels achieved. Although the transwomen in most of the studies achieved testosterone levels within the reference range for cisgender women, there were five studies[38 40 47 49 51] in which the transfemales had post-GAHT testosterone values greater than 5 nmol/L. Four of the five studies[38 40 47 49] were carried out on adolescent transfemales; two of the five studies[38 51] did not involve the use of an antiandrogen agent; one study[40] did not involve the use of any form of oestrogen. The high post-GAHT testosterone is a possible confounder, and potential physiological differences between adolescent and adult participants may also confound results.

Muscle mass and body fat changes

Table 2 summarises the studies reporting muscle mass and body fat. Eight studies[15 36 39–41 44 46 51] used a follow-up design to assess changes in LBM; seven studies assessed after 12 months,[15 36 39 41 44 46 51] and one[40] study reviewed patients who had been under treatment for 5–31 months. Seven of these studies,[15 36 39–41 44 51] including the large (n=179) ENIGI study,[15] and two studies[40 51] with high post-GAHT testosterone (~8 nmol/L), showed that total LBM was decreased by 3.0%–5.4% following hormone transition ($p<0.05$). The one study that failed to demonstrate significant changes in LBM[46] was not an outlier in any obvious way. The large ENIGI study[15] was the only study in which the limits of agreement would indicate a change in LBM at the 95% CI. All studies reported an increase in total body fat mass in transwomen after hormone transition. Three cross-sectional studies[41 51 54] compared transwomen with cisgender men. Two studies included hormone-naive transwomen.[41 51] These studies reported 6.4% and 8.0% lower LBM than in cisgender men and reductions of 4% in LBM in the transwomen with 12 months of GAHT. The third cross-sectional study compared transwomen who had undergone at least 48 months of GAHT with cisgender men[54] and reported 17% lower LBM in transwomen than in cisgender men.

Table 2 Changes in total LBM in kilograms

Longitudinal studies

Author (year)	Participants (N) TW	Baseline mean±SD (95% CI)	12 Months mean±SD (95% CI)	12–31 months mean±SD	% Change	P	T (nmol/L) Base-post GAHT
Mueller *et al* (2011)[11]	84	59.6 (54.6–64.6)	57.2 (54.0–64.1)		−4.0	<0.005	13.6–0.6
Wierckx *et al* (2014)[45]	40 (oral oestrogen) 12 (transdermal oestrogen)	56.0±7.5 62.6±9.3	53±8 59.7±8.1		−5.4 −4.6	<0.001 <0.05	18.0–0.4 19.7–0.5
Gava *et al* (2016)[38]	20 (cyproterone acetate) 20 (leuprolide acetate)	51.7±8.3 50.2±7.0	49.9±7.8 49.8±6.7		−3.5 −0.8	>0.05 >0.05	16.3–0.7 22.2–0.7
Auer *et al* (2018)[40]	45	59.5±8.7 (56.9–62.0)	57.5±12 (53.9–60.2)		−3.4	<0.001	17.5–1.9
Klaver *et al* (2018)[15]	179	57.2±8.3	55.5 (54.9–56.1)		−3.0	<0.001	
Tack *et al* (2018)[36]	21	47.0±6.4		44.8±6.3	−4.7	<0.01	15.2–8.8
Haraldsen *et al* (2007)[51]	12	54.4±6.2	52.2		−4.0	<0.001	16.8–8.6
Van Caenegem *et al* (2015)[41]	49	57.4±8.7	55.1±8.7		*−4.0*	<0.001	19.0–0.5

Cross-sectional studies

Author (year)	Participants (N) TW	CM	TW baseline mean±SD	TW 48 months mean±SD	CM mean±SD	% Difference	P	T (nmol/L) TW
Lapauw *et al* (2008)[54]	23	46		51.2±8.4	61.8±7.9	−17.2	<0.001	1.1
Haraldsen *et al* (2007)[51]	12	77	54.4±6.2		59.1±5.7	−8.0	<0.05	16.8
Van Caenegem *et al* (2015)[41]	49	49	57.4±8.7		61.3±6.8	−6.4	<0.05	19.0

Data are from dual energy X-ray absorptiometry scans.
CM, cismen; LBM, lean body mass; TW, transwomen.

CSA changes

Four follow-up studies[16 40–42] investigated the CSA either in the quadriceps, forearm or calf regions using MRI[16 42] or peripheral quantitative computed tomography (pQCT).[40 41] Of note, two of the studies measured the total CSA of the individual MRI[42] or pQCT[41] image while two studies measured the isolated muscle.[16 40] A decrease in CSA of 1.5%–11.7% was reported over periods ranging from 12 to 36 months. One of these studies[40] examined adolescent participants who only reached a final testosterone level of 8.8 nmol/L and exhibited forearm and calf CSA decreases of 4.1% and 8.9%, respectively. There were two studies[41 42] that assessed muscle CSA at both 12 months and at either 24 or 36 months. The first study[42] reported a 9.5% decrease in quadriceps CSA compared with baseline after 12 months and an 11.7% decrease in quadriceps CSA compared with baseline after 36 months. The second study[41] reported a 1.5% decrease in tibia CSA compared with baseline after 12 months and a 3.8% decrease compared with baseline after 24 months. The same study reported that compared with baseline, forearm CSA was decreased by 8.6% after 12 months, yet at 24 months was 4.4% lower than baseline, indicating that forearm CSA was 4.2% larger at 24 months than at 12 months. There was only one study[42] in which the limits of agreement indicated a change at the 95% CI. Two cross-sectional studies[41 54] compared transwomen with cisgender men. One study reported 9% smaller CSA in hormone-naive transwomen[41] than in cisgender men, with the transwomen undergoing a further 4% decrease in CSA with 24 months of GAHT. The transwomen in the second study had all undergone at least 48 months of GAHT[54] and had 24% smaller CSA than cisgender men. See table 3.

Muscular strength changes

Table 4 summarises the studies reporting muscular strength. Five longitudinal studies[16 17 37 40 41] investigated the muscular strength of transwomen. Four of the studies[17 37 40 41] measured hand grip strength in participants on the ENIGI study. The largest of the three (n=249) ENIGI studies[17] and one other study[41] found significant (p<0.001) reductions (4.3% and 7.1%, respectively) after 12 months on GAHT. Two ENIGI studies[37 40] found no significant strength differences, although one of these studies[40] was carried out on adolescents who failed to reach typical female testosterone levels (8.8 nmol/L after GAHT). The large ENIGI study[17] was the only study in which the limits of agreement would indicate a change in strength at the 95% CI. The fifth longitudinal study to assess strength measured upper leg strength using knee flexion and extension and found no significant difference after 12 months.[16] Two studies[41 54] used a cross-sectional design to compare the strength of transwomen to cisgender men. One study found 14% lower hand grip strength in hormone-naive transwomen than in cisgender men (p<0.001)[41] and a further 7% reduction in hand grip strength of the transwomen after 12 months of GAHT. The other study[54] found 24% lower hand grip and quadriceps strength in transwomen than in cisgender men after 48 months or more on GAHT (p<0.001).

Hgb and HCT changes

Nine studies[16 19 36–38 43 47–49] reported the levels of Hgb or HCT in transwomen before and after GAHT, from a minimum of three to a maximum of 36 months post hormone therapy. Eight of these studies,[16 19 36–38 43 48 49] including the large (n=239) ENIGI study,[19] found that hormone therapy led to a significant (4.6%–14.0%) decrease in Hgb/HCT (p<0.01), while one study found no significant difference after 6 months.[47] The mean age of participants in the latter study was 18 years and the range was 14–25 years. The participants also failed to reach typical female testosterone levels (after 6 months mean testosterone=6.9 nmol/L), while in six[16 19 36 37 43 48] of the eight other studies mean testosterone after GAHT was less than 2.0 nmol/L. The large ENIGI study[19] was the only study in which the limits of agreement would indicate a change in Hgb/HCT at the 95%

Table 3 Changes in muscle CSA

Longitudinal studies

Author (year)	Participants (N) TW	CSA region (units)	Baseline CSA mean±SD (95% CI)	Follow-up CSA mean±SD (95% CI)	Number of months of GAHT	% Change	P	T (nmol/L) Base-post GAHT
Elbers *et al* (1999)[42]	20	Thigh (cm^2)	307±47	278±37 (269–287) 271±39	12 36	–9.5 –11.7	<0.001 <0.001	22.0–1.0 22.0–0.9
Wiik (2020)[16]	11	Quadriceps (mm^2)	6193±679	5931±671 (5680– 6190)	12	–4.2	<0.05	18.0–0.5
Tack *et al* (2018)[36]	21	Forearm (mm^2) Calf (mm^2)	3275±541 4204±1282	3142±574 3828±478	12–31 12–31	–4.1 –8.9	<0.05 >0.05	15.2–8.8
Van Caenegem *et al* (2015)[41]	49	Forearm (mm^2) Tibia (mm^2)	3999±746 7742±1361	3664±783 3825±867 7623±1479 7448±1390	12 24 12 24	–8.6 –4.4 –1.5 –3.8	<0.001 <0.001 <0.01 <0.01	19.0–0.5 19.0–0.5

Cross-sectional studies

Author (year)	Participants (N) TW	CM	CSA region (units)	TW mean±SD	CM mean±SD	Number of months of GAHT	% Difference	P	T (nmol/L) TW
Lapauw *et al* (2008)[54]	23	46	Forearm (mm^2) Tibia (mm^2)	3500±700 6600±1300	4600±700 8700±1100	48 48	–23.9 –24.1	<0.001 <0.001	1.1
Van Caenegem *et al* (2015)[41]	49	49	Forearm (mm^2) Tibia (mm^2)	3999±746 7742±1361	4512±579 8233±1498	Baseline Baseline	–11.4 –6.0	<0.001 <0.01	19.0

Data are from MRI or pQCT.
CM, cismen; CSA, cross-sectional area; TW, transwomen.

CI. Three cross-sectional studies[18 53 54] compared HCT in transwomen post GAHT with cisgender controls (table 5). Two studies found that transwomen on GAHT for 6 or 48 months had lower (10%) HCT than cisgender men[53 54] ($p<0.005$), while two studies found no difference between transwomen after 6 and 12 months of GAHT and cisgender women.[18 53] Three cross-sectional studies[45 50 52] found significant differences[45 50] ($p<0.05$) or large effect sizes[52] (Cohen's $d=1.0$) in HCT between transwomen after 6 months of GAHT and hormone-naive transwomen, and HCT decreases of 7.4%–10.9%. See table 5.

DISCUSSION

We summarise changes induced by GAHT in non-athletic transwomen in four characteristics strongly associated with athletic performance: LBM, muscle CSA, muscular strength, and Hgb/ HCT levels. Overall, the findings demonstrate a reduction in these parameters over time. However, the time course of these reductions was not consistent across the parameters assessed.

In keeping with the muscular anabolic effects of testosterone[57] and the mixed effects of oestrogens,[58] studies using dual energy X-ray absorptiometry report decreased LBM (0.8%–5.4%) in association

Table 4 Changes in strength measures

Longitudinal studies

Author (year)	Participants (N) TW	Strength measure (units)	Baseline mean±SD (95% CI)	12 months mean±SD (95% CI)	21–31 months Mean±SD	% Change	P	T (nmol/L) Base-post GAHT
Van Caenegem *et al* (2015)[41]	49	Hand grip (kg)	42±9	39±9		–7.1	<0.001	19.0–0.5
Auer *et al* (2016)[46]	20	Hand grip (kg)	41.7±7.8	41.9±7		0.5	>0.05	17.5–1.9
Tack *et al* (2018)[36]	21	Hand grip (kg)	33.8±8.1		34.3±5.6	1.5	>0.05	15.2–8.8
Scharff (2019)	249	Hand grip (kg)	41.8±8.9	40.0±8.9 (39.2– 40.8)		–4.3	<0.001	18.3–0.8
Wiik (2020)[16]	11	Knee extension (N-m) Knee flexion (N-m)	239.7±44.0 99.5±16.8	242.6±41.5 (230–252) 101.5±15.5 (92–109)		1.2 2.0	>0.05 >0.05	18.0–0.5

Cross-sectional studies

Author (year)	Participants (N) TW	CM	Strength measure (units)	TW baseline mean±SD	TW 48 months mean±SD	CM mean±SD	% Difference	P	T (nmol/L) TW
Van Caenegem *et al* (2015)[41]	49	49	Hand grip (kg)	42±9		49±6	–14.3	<0.001	19.0
Lapauw *et al* (2008)[54]	23	46	Hand grip (kg) Knee extension (N-m)		41±8 150±49	53±8 200±44	–22.6 –25	<0.001 <0.001	1.1

CM, cismen; TW, transwomen.

Table 5 Changes in HCT and Hgb levels

Longitudinal studies

Author (year)	Participants (N) TW	Measure (units)	Baseline mean±SD (95% CI)	Follow-up mean±SD (95% CI)	Number of months	% Change	P	T (nmol/L) Base-post GAHT
Wierckx (2014)	40 (oral oestrogen) 12 (transdermal oestrogen)	HCT (%)	45±2.5 45.5±1.7	42±5.7 42.2±2.3	12 12	−7.0 −4.6	<0.01 <0.001	18.0–0.4 19.7–0.5
Auer *et al* (2016)[46]	20	HCT (%)	45.2±2.7	42.7±1.8	12	−5.5	<0.01	17.5–1.9
Jarin *et al* (2017)[39]	13	HCT (%)	43.8	42.3	6	−3.4	>0.05	13.6–6.9
Vita *et al* (2018)[48]	21	HCT (%)	44.8±2.9	40.1±2.6	6–30	−10.5	<0.001	20.5–1.1
Defreyne *et al* (2018)[19]	239	HCT (%)	45.0±2.5 (44.9–45.5)	41.0±3.1 (40.9– 41.7) 41.1±3.2 (40.5– 41.2) 40.7±3.2 (40.0– 40.8)	3 6 24	−8.9 −8.7 −9.6	<0.001 <0.001 <0.001	17.4–0.7 17.4–0.6 17.4–0.6
Tack *et al* (2017)[47]	21	HCT (%)	43.8±1.9	39.9±2.2	12–31	−8.9	<0.001	15.2–8.8
Gooren and Bunck (2004)[43]	19	Hgb (mmol/L)	9.3±0.7	8.0±0.7 8.1±0.6	12 36	−14.0 −12.9	<0.001 <0.001	21.5–1.0 21.5–0.9
Olson-Kennedy *et al* (2018)[49]	23	Hgb (g/dL)	153±11	140±12	12	−8.3	<0.001	14.8–5.9
Wiik (2020)[16]	9 10	Hgb (g/L)	148.3±10.1 150.3±9.1	132.7±9.1 133.3±9.0	4 12	−10.5 −11.7	<0.001 <0.001	18.0–0.5 18.0–0.5

Cross-sectional studies

Author (year)	Participants (N) TW	CM	CW	HNTW	Measure (units)	TW mean±SD or (range)	Control mean±SD or (range)	Number of months	% Difference	P	T (nmol/L) TW
Lapauw *et al* (2008)[54]	23	46			HCT (%)	41.2±2.3	45.3±2.3	>48	−9.1	<0.001	1.1
SoRelle *et al* (2019)[52]	105			73	HCT (%)	(35.9– 48.7)	(39.0– 50.6)	>6	–	d=1.0	1.9
Greene *et al* (2019)[18]	93				HCT (%)	(35–47)	(35.5– 46) CW	>12	–	>0.05	1.4
Roberts *et al* (2014)[53]	55	20	20		HCT (%)	(34.6–43.7)	(38.4– 45.7) CM (34.4– 41.9) CW	>6	– –	<0.01 >0.05	
Jain (2019)	182 (oestrogen) 95 (oestrogen +progesterone)			92	HCT (%)	42.5 40.9	45.9±2.0	>3	−7.4 −10.9	<0.05 <0.05	
Sharula (2012)[37]	129			22	HCT (%)	40.2±3.1	44.4±2.4	>3	−9.5	<0.001	2.5

CM, cismen; CW, ciswomen; HCT, haematocrit; Hgb, haemoglobin; HNTW, hormone-naive transwomen; TW, transwomen.

with GAHT. Twelve months of GAHT also decreased muscle CSA (1.5%–9.7%). However, a further 12 or 24 months of GAHT did not always elicit further decreases in muscle CSA. Strength loss with 12 months of GAHT also ranged from non-significant to 7%. Taking these strength parameter data collectively, and in consideration of cisgender women demonstrating 31% lower LBM,[59] 36%[60] lower hand-grip strength and 35%[61] lower knee extension strength than cisgender men, the small decrease in strength in transwomen after 12–36 months of GAHT suggests that transwomen likely retain a strength advantage over cisgender women. Whether longer duration of GAHT would yield further decrements in strength in transgender women is unknown.

In contrast to strength-related data, blood cell findings revealed a different time course of change. After 3–4 months on GAHT, the HCT[19] or Hgb[16] levels of transwomen matched those of cisgender women, with levels remaining stable within the 'normal' female range for studies lasting up to 36 months. Given the rapid fall in Hgb/HCT to 'normal' female levels with GAHT, it is possible that transfemale athletes experience impaired endurance performance in part due to reduced oxygen transport from the lungs to the working muscles.[62] This postulate is consistent with findings reported in one of the few studies conducted in athletic transwomen.[63] In this study, the race times of eight transfemale distance runners were compared at baseline and after one or more years of GAHT. After adjusting performance for age, the eight runners were not more competitive in the female category (after GAHT) than they had been in the male category (before GAHT). Given this, and that the changes in Hgb/HCT follow a different time course than strength changes, sport-specific regulations for transwomen in endurance ver strength sports may be needed.

Of interest, compared with cisgender men, hormone-naive transwomen demonstrate 6.4%–8.0% lower LBM,[41 51] 6.0%–11.4% lower muscle CSA and ~10%–14% lower handgrip strength.[17 41 60] This disparity is noteworthy given that hormone-naive transwomen and cisgender men have similar testosterone levels.[16 17 19 42] Explanations for this strength difference are unclear but may include transwomen actively refraining from building muscle and/or engaging in disordered eating[64] or simply not being athletically inclined, perhaps influenced by feelings of an unwelcome presence in sporting arenas.[65] Taken together, hormone-naive transwomen may not, on average, have the same athletic attributes as cisgender men. The need to move beyond simple comparisons of cisgender men and women to assess the sporting capabilities of transwomen is imperative.

This systematic review identified studies that assessed the changes in LBM, CSA, muscular strength and Hgb/HCT in non-athletic transgender women following GAHT. However, several limitations are noted. Although the data we present are meaningful, the effects of GAHT on these parameters, or indeed athletic performance in transgender people who engage in training and competition, remain unknown. The levels of physical activity of the transwomen compared with cisgender women in the studies were not reported. Other limitations include the studies being written in English only,

and the research being conducted in Western countries, contributing to geographical bias. Furthermore, as with much research with transgender individuals, there is a sparse data risk[66] because of small sample sizes and short study durations, indicative of the relatively small population, difficulties with recruitment and high drop-out rates over time. Indeed, the overlap of participants in the ENIGI studies and the heterogenous methodology in the other studies precluded the possibility of meaningful meta-analysis. However, overall, the results across different study groups and methods (ie, longitudinal vs follow-up studies) are largely consistent, suggesting that the risk of selective reporting and publication bias are low and the data in the reviewed studies are reliable. This review only focused on binary transgender individuals; those who medically transition from their birth assigned gender to the opposite gender and did not consider non-binary individuals. Not only are there even more limited data on non-binary individuals, but also, for many, their affirmed gender expression does not require GAHT, thus there are no hormone-induced changes to observe which would be relevant to this review. That is not to say that non-binary inclusivity in sport is not an important issue, only that the central tenets are not focused on physiology.

As previously stated, a major limitation in this area of research is the absence of studies in transgender athletes. However, a very recent study reported changes in fitness levels of 29 transmen and 46 transwomen in the United States Air Force, from before and after 30 months of GAHT.[67] Enlisted Air Force members are required to engage in regular physical activity and to complete annual assessments of number of sit-ups and push-ups in 1 min, and 1.5 mile race time. Although not athletes per se, enlisted members could at least be considered exercise trained. The study reported that after 2 years on GAHT there were no significant differences between ciswomen and transwomen in the number of push-ups or sit-ups performed in 1 min. However, transwomen ran significantly faster during the 1.5 mile fitness test than ciswomen. These observations in trained transgender individuals are consistent with the findings of the current review in untrained transgender individuals, whereby 30 months of GAHT may be sufficient to attenuate some, but not all, influencing factors associated with muscular endurance and performance.

Overall, this review reports decreases in muscle strength, LBM and muscle CSA in response to 12–36 months, and decreases in Hbg after 3–4 months, of GAHT in transwomen. These findings may help to shape future studies with transgender athletes and provide data for valuable and rigorous research going forward. Sporting bodies wish to be inclusive to all athletes, and there is a critical desire and need for more research to be able to develop evidence-based policies around this topic. Given that the range of physical parameters important for success varies considerably between sports, and that the physiological effects of GAHT vary in their time course (eg, muscle vs blood), future research should be sport specific as well as athlete centric. Although a level playing field in sport is illusory, it is important that opportunities for women to engage in meaningful competition within the female category exist.[68] Whether transgender and cisgender women can engage in meaningful sport, even after GAHT, is a highly debated question. However, before this question can be answered with any certainty, the intricacies and complexity of factors that feed into the development of high-performance athletes warrant further investigation of attributes beyond those assessed herein.

Contributors GLW devised the study. BSK completed an initial search in 2019 with GLW and HMD. JH completed a second search in 2020 with GLW and EOD. All authors contributed to the manuscript.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

What is already known

- There is much debate on whether (and when) transwomen should be permitted to compete in the female category in sport.

What are the new findings

- Longitudinal and cross-sectional studies identify that hormone therapy in transwomen decreases muscle cross-sectional area, lean body mass, strength and haemoglobin levels, with noted differences in the time course of change.
- Haemoglobin levels decrease to those seen in cisgender women after 4 months of hormone therapy. In contrast, despite significant decreases in muscle cross-sectional area, lean body mass and strength after 12–36 months of hormone therapy, values remain higher than that in cisgender women.
- It is possible that transwomen competing in sports may retain strength advantages over cisgender women, even after 3 years of hormone therapy.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

ORCID iD
Gemma L Witcomb http://orcid.org/0000-0003-4668-2441

REFERENCES

1 Ferguson-Smith MA, Ferris EA. Gender verification in sport: the need for change? *Br J Sports Med* 1991;25:17–20.
2 Simpson JL, Ljungqvist A, de la Chapelle A, *et al*. Gender verification in competitive sports. *Sports Med* 1993;16:305–15.
3 Ljungqvist A. Sex segregation and sport. *Br J Sports Med* 2018;52:3.
4 International Olympic Committee. IOC consensus meeting on sex reassignment and hyperandrogenism, 2015. Available: https://stillmed.olympic.org/Documents/Commissions_PDFfiles/Medical_commission/201511_ioc_consensus_meeting_on_sex_reassignment_and_hyperandrogenism-en.pdf
5 World Athletics. Eligibility regulations transgender regulation, 2019. Available: https://www.worldathletics.org/about-iaaf/documents/book-of-rules
6 World Athletics. Eligibility regulations for the female classification, 2019. Available: https://www.worldathletics.org/about-iaaf/documents/book-of-rules
7 Karkazis K, Carpenter M. Impossible "Choices": The Inherent Harms of Regulating Women's Testosterone in Sport. *J Bioeth Inq* 2018;15:579–87.
8 Hutcheon D. Hyperandrogenism in athletics: a review of Chand V. IAAF. law in sport, 2015. Available: https://www.lawinsport.com/topics/item/hyperandrogenism-in-athletics-a-review-of-chand-v iaaf
9 Karkazis K, Jordan-Young R, Davis G, *et al*. Out of bounds? A critique of the new policies on hyperandrogenism in elite female athletes. *Am J Bioeth* 2012;12:3–16.
10 Floyd MJ, Martin O, Eckloff KJ. A qualitative study of transgender individuals' experiences of healthcare including radiology. *Radiography* 2020;26:e38–44.

11 Buzuvis E. Challenging gender in Single-Sex spaces: lessons from a feminist Softball League. law and contemporary problems. *Sex in Sport* 2017;80.
12 Seal LJ. A review of the physical and metabolic effects of cross-sex hormonal therapy in the treatment of gender dysphoria. *Ann Clin Biochem* 2016;53:10–20.
13 T'Sjoen G, Arcelus J, Gooren L, *et al*. Endocrinology of transgender medicine. *Endocr Rev* 2019;40:97–117.
14 Wiepjes CM, de Jongh RT, de Blok CJ, *et al*. Bone safety during the first ten years of Gender-Affirming hormonal treatment in Transwomen and Transmen. *J Bone Miner Res* 2019;34:447–54.
15 Klaver M, de Blok CJM, Wiepjes CM, *et al*. Changes in regional body fat, lean body mass and body shape in trans persons using cross-sex hormonal therapy: results from a multicenter prospective study. *Eur J Endocrinol* 2018;178:163–71.
16 Wiik A, Lundberg TR, Rullman E, *et al*. Muscle strength, size, and composition following 12 months of Gender-affirming treatment in transgender individuals. *J Clin Endocrinol Metab* 2020;105:e805–13.
17 Scharff M, Wiepjes CM, Klaver M, *et al*. Change in grip strength in trans people and its association with lean body mass and bone density. *Endocr Connect* 2019;8:1020–8.
18 Greene DN, McPherson GW, Rongitsch J, *et al*. Hematology reference intervals for transgender adults on stable hormone therapy. *Clin Chim Acta* 2019;492:84–90.
19 Defreyne J, Vantomme B, Van Caenegem E, *et al*. Prospective evaluation of hematocrit in gender-affirming hormone treatment: results from European network for the investigation of gender incongruence. *Andrology* 2018;6:446–54.
20 Weed RI, Reed CF, Berg G. Is hemoglobin an essential structural component of human erythrocyte membranes? *J Clin Invest* 1963;42:581–8.
21 Otto JM, Montgomery HE, Richards T. Haemoglobin concentration and mass as determinants of exercise performance and of surgical outcome. *Extrem Physiol Med* 2013;2:33.
22 Cohen E, Kramer M, Shochat T, *et al*. Relationship between hematocrit levels and intraocular pressure in men and women: a population-based cross-sectional study. *Medicine* 2017;96:e8290.
23 Coviello AD, Kaplan B, Lakshman KM, *et al*. Effects of graded doses of testosterone on erythropoiesis in healthy young and older men. *J Clin Endocrinol Metab* 2008;93:914–9.
24 Weinand JD, Safer JD. Hormone therapy in transgender adults is safe with provider supervision; a review of hormone therapy sequelae for transgender individuals. *J Clin Transl Endocrinol* 2015;2:55–60.
25 Cronin J, Lawton T, Harris N, *et al*. A brief review of handgrip strength and sport performance. *J Strength Cond Res* 2017;31:3187–217.
26 Handelsman DJ. Sex differences in athletic performance emerge coinciding with the onset of male puberty. *Clin Endocrinol* 2017;87:68–72.
27 Bruusgaard JC, Johansen IB, Egner IM, *et al*. Myonuclei acquired by overload exercise precede hypertrophy and are not lost on detraining. *Proc Natl Acad Sci U S A* 2010;107:15111–6.
28 Kadi F, Eriksson A, Holmner S, *et al*. Effects of anabolic steroids on the muscle cells of strength-trained athletes. *Med Sci Sports Exerc* 1999;31:1528–34.
29 Hembree WC, Cohen-Kettenis PT, Gooren L, *et al*. Endocrine treatment of Gender-Dysphoric/Gender-Incongruent persons: an endocrine Society clinical practice guideline. *J Clin Endocrinol Metab* 2017;102:3869–903.
30 Delgado-Ruiz R, Swanson P, Romanos G. Systematic review of the long-term effects of transgender hormone therapy on bone markers and bone mineral density and their potential effects in implant therapy. *JCM* 2019;8:784.
31 Fighera TM, Ziegelmann PK, Rasia da Silva T, *et al*. Bone mass effects of Cross-Sex hormone therapy in transgender people: updated systematic review and meta-analysis. *J Endocr Soc* 2019;3:943–64.
32 Liberati A, Altman DG, Tetzlaff J, *et al*. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS Med* 2009;6:e1000100.
33 Armijo-Olivo S, Stiles CR, Hagen NA, *et al*. Assessment of study quality for systematic reviews: a comparison of the Cochrane collaboration risk of bias tool and the effective public health practice project quality assessment tool: methodological research. *J Eval Clin Pract* 2012;18:12–18.
34 Storebø OJ, Pedersen N, Ramstad E, *et al*. Methylphenidate for attention deficit hyperactivity disorder (ADHD) in children and adolescents - assessment of adverse events in non-randomised studies. *Cochrane Database Syst Rev* 2018;5:CD012069.
35 Thomas BH, Ciliska D, Dobbins M, *et al*. A process for systematically reviewing the literature: providing the research evidence for public health nursing interventions. *Worldviews Evid Based Nurs* 2004;1:176–84.
36 Wierckx K, Van Caenegem E, Schreiner T, *et al*. Cross-sex hormone therapy in trans persons is safe and effective at short-time follow-up: results from the European network for the investigation of gender incongruence. *J Sex Med* 2014;11:1999–2011.
37 Auer MK, Cecil A, Roepke Y, *et al*. 12-Months metabolic changes among gender dysphoric individuals under cross-sex hormone treatment: a targeted metabolomics study. *Sci Rep* 2016;6:37005.
38 Tack LJW, Heyse R, Craen M, *et al*. Consecutive cyproterone acetate and estradiol treatment in Late-Pubertal transgender female adolescents. *J Sex Med* 2017;14:747–57.
39 Auer MK, Ebert T, Pietzner M, *et al*. Effects of sex hormone treatment on the metabolic syndrome in transgender individuals: focus on metabolic cytokines. *J Clin Endocrinol Metab* 2018;103:790–802.
40 Tack LJW, Craen M, Lapauw B, *et al*. Proandrogenic and antiandrogenic progestins in transgender youth: differential effects on body composition and bone metabolism. *J Clin Endocrinol Metab* 2018;103:2147–56.
41 Van Caenegem E, Wierckx K, Taes Y, *et al*. Preservation of volumetric bone density and geometry in trans women during cross-sex hormonal therapy: a prospective observational study. *Osteoporos Int* 2015;26:35–47.
42 Elbers JM, Asscheman H, Seidell JC, *et al*. Effects of sex steroid hormones on regional fat depots as assessed by magnetic resonance imaging in transsexuals. *Am J Physiol* 1999;276:E317–25.
43 Gooren LJG, Bunck MCM. Transsexuals and competitive sports. *Eur J Endocrinol* 2004;151:425–9.
44 Mueller A, Zollver H, Kronawitter D, *et al*. Body composition and bone mineral density in male-to-female transsexuals during cross-sex hormone therapy using gonadotrophin-releasing hormone agonist. *Exp Clin Endocrinol Diabetes* 2011;119:95–100.
45 Sharula, Chekir C, Emi Y, *et al*. Altered arterial stiffness in male-to-female transsexuals undergoing hormonal treatment. *J Obstet Gynaecol Res* 2012;38:932–40.
46 Gava G, Cerpolini S, Martelli V, *et al*. Cyproterone acetate vs leuprolide acetate in combination with transdermal oestradiol in transwomen: a comparison of safety and effectiveness. *Clin Endocrinol* 2016;85:239–46.
47 Jarin J, Pine-Twaddell E, Trotman G, *et al*. Cross-Sex hormones and metabolic parameters in adolescents with gender dysphoria. *Pediatrics* 2017;139:e20163173.
48 Vita R, Settineri S, Liotta M, *et al*. Changes in hormonal and metabolic parameters in transgender subjects on cross-sex hormone therapy: a cohort study. *Maturitas* 2018;107:92–6.
49 Olson-Kennedy J, Okonta V, Clark LF, *et al*. Physiologic response to Gender-Affirming hormones among transgender youth. *J Adolesc Health* 2018;62:397–401.
50 Jain J, Kwan D, Forcier M. Medroxyprogesterone acetate in Gender-Affirming therapy for Transwomen: results from a retrospective study. *J Clin Endocrinol Metab* 2019;104:5148–56.
51 Haraldsen IR, Haug E, Falch J, *et al*. Cross-sex pattern of bone mineral density in early onset gender identity disorder. *Horm Behav* 2007;52:334–43.
52 SoRelle JA, Jiao R, Gao E, *et al*. Impact of hormone therapy on laboratory values in transgender patients. *Clin Chem* 2019;65:170–9.
53 Roberts TK, Kraft CS, French D, *et al*. Interpreting laboratory results in transgender patients on hormone therapy. *Am J Med* 2014;127:159–62.
54 Lapauw B, Taes Y, Simoens S, *et al*. Body composition, volumetric and areal bone parameters in male-to-female transsexual persons. *Bone* 2008;43:1016–21.
55 Coleman E, Bockting W, Botzer M, *et al*. Standards of care for the health of transsexual, transgender, and Gender-Nonconforming people, version 7. *Int J Transgend* 2012;13:165–232.
56 Dragoman MV, Tepper NK, Fu R, *et al*. A systematic review and meta-analysis of venous thrombosis risk among users of combined oral contraception. *Int J Gynaecol Obstet* 2018;141:287–94.
57 Handelsman DJ. Performance Enhancing Hormone Doping in Sport. In: Feingold KR, Anawalt B, Boyce A, *et al*, eds. *Endotext*. South Dartmouth (MA): MDText.com, Inc, 2000.
58 Rosa-Caldwell ME, Greene NP. Muscle metabolism and atrophy: let's talk about sex. *Biol Sex Differ* 2019;10:43.
59 Lee DH, Keum N, Hu FB, *et al*. Development and validation of anthropometric prediction equations for lean body mass, fat mass and percent fat in adults using the National health and nutrition examination survey (NHANES) 1999-2006. *Br J Nutr* 2017;118:858–66.
60 Bohannon RW, Wang Y-C, Yen S-C, *et al*. Handgrip strength: a comparison of values obtained from the NHANES and NIH Toolbox studies. *Am J Occup Ther* 2019;73:7302205080p1–9.
61 Neder JA, Nery LE, Shinzato GT, *et al*. Reference values for concentric knee isokinetic strength and power in nonathletic men and women from 20 to 80 years old. *J Orthop Sports Phys Ther* 1999;29:116–26.
62 Mairbäurl H. Red blood cells in sports: effects of exercise and training on oxygen supply by red blood cells. *Front Physiol* 2013;4:332.
63 Harper J. Race times for transgender athletes. *Journal of Sporting Cultures and Identities* 2015;6:1–9.
64 Witcomb GL, Bouman WP, Brewin N, *et al*. Body image dissatisfaction and eating-related psychopathology in trans individuals: a matched control study. *Eur Eat Disord Rev* 2015;23:287–93.
65 Jones BA, Arcelus J, Bouman WP, *et al*. Barriers and facilitators of physical activity and sport participation among young transgender adults who are medically transitioning. *Int J Transgend* 2017;18:227–38.
66 Greenland S, Mansournia MA, Altman DG. Sparse data bias: a problem hiding in plain sight. *BMJ* 2016;352:i1981.
67 Roberts TA, Smalley J, Ahrendt D. Effect of gender affirming hormones on athletic performance in transwomen and transmen: implications for sporting organisations and legislators. *Br J Sports Med* 2020:bjsports-2020-102329.
68 Harper J, Martinez-Patino M-J, Pigozzi F, *et al*. Implications of a third gender for elite sports. *Curr Sports Med Rep* 2018;17:42–4.

Supplementary Table 1. Selection criteria and search terms using the PICOS framework.

PICOS	Selection Criteria	Search Term
Population	Hormone-naïve transfemales at Tanner stage 4 or later	#1 transgender OR transsexual OR trans people OR transwomen #1A transgender OR transsexual OR transwomen
Intervention	Gender Affirming Hormone Therapy	#2 cross-sex hormone therapy OR hormonal treatment OR testosterone OR oestradiol #2A hormone OR testosterone OR oestradiol #2B hormone OR testosterone OR oestrogen #2C hormone OR therapy OR testosterone OR oestrogen
Comparison		
Outcome	Changes in LBM, CSA, strength, Hgb or HCT	#3 grip strength OR muscle OR lean body mass OR haematocrit OR haemoglobin OR metabolic changes #3A strength OR muscle OR lean OR haematocrit OR haemoglobin OR metabolic #3B haematocrit OR haemoglobin OR muscle OR strength #3C haematocrit OR haemoglobin OR muscle OR strength OR lean
Study design		

Note: Pubmed searched as #1 AND #2 AND #3. Web of Science searched as #1 AND #2A AND #3A. Scopus searched as #1A AND #2B AND #3B. BioMed Central as #1A AND #2C AND #3C. The searches were limited to studies published in English.

Supplementary Table 2. EPHPP Scores

	Selection bias	Study design	Confounders	Blinding	Data collection method	Withdrawals/ dropouts	Average score
Elbers (1999)	2	2	2	1	3	cc	2
Gooren (2004)	2	2	2	1	3	cc	2
Mueller (2011)	2	2	2	1	3	3	2.2
Wierckx (2014)	2	2	2	1	3	cc	2
Gava (2016)	2	2	2	1	3	cc	2
Auer (2016)	2	2	2	1	3	cc	2
Auer (2018)	2	2	2	1	3	cc	2
Jarin (2017)	2	2	2	1	3	cc	2
Defrayne (2018)	2	2	2	1	3	cc	2
Vita (2018)	2	2	2	1	3	cc	2
Klaver (2018)	2	2	2	1	3	cc	2
Olson-Kennedy (2018)	2	2	2	1	3	1	1.8
Tack (2018)	2	2	2	1	3	cc	2
Tack (2017)	2	2	2	1	3	cc	2
Scharff (2019)	2	2	2	1	3	cc	2
Wiik (2019)	2	2	2	1	3	3	2.2
Van Caenegem (2014)	2	2	2	1	3	cc	2
Haraldsen (2007)	2	2	2	1	3	cc	2
SoRelle (2019)	2	2	2	1	3	cc	2
Greene (2019)	2	2	2	1	3	cc	2
Roberts (2014)	2	2	2	1	3	cc	2
Lapauw (2008)	2	2	2	1	3	cc	2
Jain (2019)	2	2	2	1	3	cc	2
Sharula (2012)	2	2	2	1	3	cc	2

Notes on EPHPP scores: The studies recruited transgender participants from gender clinics and were assessed as moderate for selection bias. All studies were assessed as moderate for study design as they used either retrospective or prospective cohort studies, in which measurements were conducted before and after hormone transition to assess possible changes. The studies were assessed as moderate for controlling confounding factors. Since none of the studies were blinded, all were considered weak on this variable of assessment, while all studies used valid and reliable medical records and as such were considered strong in terms of quality of data collection. Withdrawal and drop-outs were not applicable in the case-control studies but were appropriately described in two cohort studies[31,57] with low dropout levels and these were assessed as strong. A third cohort study[66] had a dropout level of greater than 40% and was assessed as weak. Based on the mean scores, all studies were categorized as moderate in quality (average scores between 1.8 and 2.2).