Effects of moderately increased testosterone concentration on physical performance in young women: a double blind, randomised, placebo controlled study

Angelica Lindén Hirschberg ,[1,2] Jona Elings Knutsson,[3,4] Torbjörn Helge,[5] Manne Godhe,[5] Maria Ekblom,[5] Stephane Bermon,[6,7] Björn Ekblom[8]

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/bjsports-2018-100525).

[1]Department of Women's and Children's Health, Karolinska Institutet, Stockholm, Sweden
[2]Department of Gynaecology and Reproductive Medicine, Karolinska University Hospital, Stockholm, Sweden
[3]Women's and Children's Health, Karolinska Institutet, Stockholm, Sweden
[4]Theme Children's and Women's Health, Karolinska University Hospital, Stockholm, Sweden
[5]Swedish School of Sports and Health Sciences, Stockholm, Sweden
[6]Monaco Institute of Sports Medicine, Monaco, Monaco
[7]LAMHESS, Universite Cote d'Azur, Nice, France
[8]Astrand Laboratory of Work Physiology, The Swedish School of Sport and Health Sciences, Stockholm, Sweden

Correspondence to
Professor Angelica Lindén Hirschberg, Department of Women's and Children's Health, Karolinska Institutet, Stockholm 171 77, Sweden; angelica.linden-hirschberg@sll.se

Accepted 7 July 2019

To cite: Hirschberg AL, Elings Knutsson J, Helge T, *et al*. *Br J Sports Med* Epub ahead of print: [*please include* Day Month Year]. doi:10.1136/bjsports-2018-100525

ABSTRACT

Objective To investigate the effects of a moderate increase in serum testosterone on physical performance in young, physically active, healthy women.

Methods A double blind, randomised, placebo controlled trial was conducted between May 2017 and June 2018 (ClinicalTrials.gov ID: NCT03210558). 48 healthy, physically active women aged 18–35 years were randomised to 10 weeks of treatment with 10 mg of testosterone cream daily or placebo (1:1). All participants completed the study. The primary outcome measure was aerobic performance measured by running time to exhaustion (TTE). Secondary outcomes were anaerobic performance (Wingate test) and muscle strength (squat jump (SJ), counter movement jump (CMJ) and knee extension peak torque). Hormone levels were analysed and body composition assessed by dual energy X-ray absorptiometry.

Results Serum levels of testosterone increased from 0.9 (0.4) nmol/L to 4.3 (2.8) nmol/L in the testosterone supplemented group. TTE increased significantly by 21.17 s (8.5%) in the testosterone group compared with the placebo group (mean difference 15.5 s; P=0.045). Wingate average power, which increased by 15.2 W in the testosterone group compared with 3.2 W in the placebo group, was not significantly different between the groups (P=0.084). There were no significant changes in CMJ, SJ and knee extension. Mean change from baseline in total lean mass was 923 g for the testosterone group and 135 g for the placebo group (P=0.040). Mean change in lean mass in the lower limbs was 398 g and 91 g, respectively (P=0.041).

Conclusion The study supports a causal effect of testosterone in the increase in aerobic running time as well as lean mass in young, physically active women.

INTRODUCTION

Biological sex is one of the most decisive factors for physical performance in humans. In sports that rely on strength and endurance, male athletes have, in general, an advantage of 10–15% in comparison with female athletes, which most likely is explained by men having, on average, more than 15 times higher circulating concentrations of testosterone than women.[1 2] Androgens are considered beneficial for athletic performance because of their potent anabolic effects on muscle mass and bone tissue, by stimulating erythropoiesis and by promoting competitive behaviour.[1 3]

Some rare inborn conditions in women (ie, differences in sex development (DSD)) may cause a greatly increased production of testosterone, into the male range, produced by functioning male gonads.[1] Women with these conditions and normal androgen receptor sensitivity will develop a higher proportion of lean body mass and enhanced physical performance. The prevalence of 46XY, DSD among elite female athletes is estimated to be about 140 times greater than in the general population.[4] There is controversy internationally on whether it is fair to allow these hyperandrogenic individuals with high testosterone levels to compete against women with normal female androgen levels.[5–9] Indeed, some scientists and human rights advocates consider that an individual who has been assigned female gender at birth, raised as a girl and who is socially accepted as a woman, should be allowed to compete in the women's category, irrespective of her levels of testosterone and androgen receptor sensitivity.[5 7]

The International Association of Athletics Federations (IAAF) implemented eligibility regulations before the Olympic Games in London 2012 that permitted hyperandrogenic athletes to participate in the female category provided they first lowered their endogenous testosterone levels. The rationale was to facilitate participation under conditions that preserved fair and meaningful competition in the female classification. However, the policies have been criticised and challenged.[5–9] Indeed, some authors expressed concern about testosterone being the main driver of the sex difference in sport performances.[5 7] In 2015, the Court of Arbitration of Sport suspended the IAAF regulations pending receipt of further scientific evidence on the role of testosterone for athletic performance in female athletes. Later, after another trial, the Court of Arbitration of Sport delivered an award supporting the validity of the DSD regulations. The decision was sent for appeal to the Swiss Federal Tribunal, which dismissed the athlete's case on 31 July 2019.

There is surprisingly little information about the effect of testosterone on physical performance in women.[10] Few studies of female athletes have demonstrated associations between endogenous androgens levels, muscle mass and muscle strength.[11–14] Furthermore, mild hyperandrogenism, such as polycystic ovary syndrome, is overrepresented among elite female athletes.[15–19] This

is considered indirect evidence of the performance enhancing effect of testosterone in female athletes.

The overall aim of the present study was to investigate the causal effects of 10 weeks of exposure to moderately increased testosterone concentration on physical performance and body composition in young, healthy and physically active women in a double blind, randomised, placebo controlled trial. We hypothesised that increased levels of testosterone will enhance physical performance and increase muscle mass in these women.

MATERIAL AND METHODS

Study design and study population

This was a randomised, double blind, placebo controlled, parallel study conducted at the Karolinska University Hospital, Stockholm, Sweden, between May 2017 and June 2018. Healthy women were recruited by advertisement, mainly from the Swedish School of Sports and Health Sciences. Women fulfilling all of the inclusion criteria, aged 18–35 years, body mass index 19–25 kg/m^2, non-smoking, a moderate to high self-reported level of recreational physical activity, not taking hormonal contraception and willing to use highly efficient non-hormonal contraception during the study period were included in the study. In order not to infringe antidoping rules, the women had to agree not to participate in any sports competition event during the study period and for 1 month after termination of the study. Exclusion criteria were the presence of cardiovascular, liver, biliary or renal disease; hyperlipidaemia; uncontrolled high blood pressure; endocrinological disorder; oligomenorrhoea or amenorrhoea; pregnancy; history of thromboembolic disorder; any malignancy; and intake of hormonal contraception in the 2 months prior to the study.

The study was registered at ClinicalTrials.gov ID: NCT03210558. We were alerted to minor discrepancies between our original trial registration and our trial protocol during peer review of the manuscript and we amended trial registration in August 2019 in keeping with research best practice. The study was approved by the regional ethics committee in Stockholm (2016/1485-32, amendment 2017/779–32) and was carried out in accordance with Good Clinical Practice and the World Medical Association Declaration of Helsinki—ethical principles for medical research involving human subjects. All women gave written informed consent.

Participants were randomly assigned to treatment with placebo cream (1 mL) or testosterone cream 10 mg (1 mL) (Andro-Feme 1) applied every evening to the upper outer thigh for 10 weeks. Randomisation was provided by the pharmacy at Karolinska University Hospital. A balanced block randomisation (1:1) was carried out with four participants in each block of fixed size. The study investigators, research coordinators and the participating women were blinded to the treatment allocation.

Baseline data collection was performed in the early follicular phase of the menstrual cycle (cycle days 1–7) before the start of treatment, and final data collection at the end of 10 weeks of treatment. On both occasions, fasting blood samples were collected, body composition was determined, a gynaecological examination was undertaken, body hair growth was evaluated using the Ferriman–Gallwey score [20] and physical performance tests were carried out.

Outcome measures

The primary outcome measure was running time to exhaustion (TTE) during an individually standardised work load on a treadmill. Secondary outcomes measures were average and peak power output (Wingate test), muscle strength and body composition.

Physical performance tests

Physical performance tests were performed at the Swedish School of Sports and Health Sciences. Prior to baseline data collection, participants visited the test lab for familiarisation tests and to determine performance values. Before the test days, participants were instructed to have a standardised rest day and not to have a heavy meal, caffeinated drinks, cigarettes or snuff for a minimum of 2 hours before the tests. The tests were carried out in an identical manner at baseline and at the end of treatment.

Jump tests

Squat jumps (SJs) and countermovement jumps (CMJs) were measured on a force plate (Kistler-model 9281EA, Kistler Nordic AB, Sweden). After a short warm up, three maximal SJs with a start position at 90° knee flexion with hands placed on the hips during the entire jump and three maximal dynamic CMJs with voluntary knee flexion were performed. Vertical forces during SJs and CMJs were recorded and used to calculate jump height in metres (Qualisys Track Manager, V.2.12, Qualisys AB, Gothenburg, Sweden). Reproducibility for the SJ and CMJ has been reported as high (Cronbach's α of 0.97 and 0.98, respectively).[21]

Knee extension torque

Participants were positioned in a seated position in an isokinetic dynamometer (Isomed 2000) with their shoulders and hips securely strapped and with the centre of their knee aligned with the movement axis of the dynamometer and their right lower leg attached to the lever arm of the dynamometer at a knee angle of 60°. Knee extensor torque was low pass filtered at 50 Hz and sampled at 5000 Hz (Spike2, V.7.09, CED, Cambridge, UK). Participants performed three maximal voluntary knee extensions separated by 1 min of rest, and the highest torque from the trials was registered as knee extension peak torque.

Running test on treadmill

All running was performed on a treadmill (Rodby Electronics, Sweden) with gas composition of expired air of oxygen uptake (oxygen consumption, VO_2), volume of carbon dioxide (VCO_2), respiratory exchange ratio and ventilation measured continuously using an online ergo spirometry system (Oxycon Pro, Erich Jaeger GmbH, Hoechberg, Germany). The system was calibrated before each test according to the manufacturer's instructions: ambient conditions, volume sensor and gas calibration (15.00%±0.01% O2 and 6.00%±0.01% CO2; Air Liquid, Kungsängen, Sweden). This system has been reported to achieve high validity and stability.[22] The maximal test protocol, at baseline and at the end of the treatment period, started at either 10 or 12 km/h with 1° incline (depending on pretest results), with 1 km/h speed increases every 1 min. After 4 min, workload increased by increasing the incline by 1° every 1 min until exhaustion. Our primary outcome measure (TTE) was registered on a separate timer, from the start of the running protocol to the point of exhaustion when the test was stopped. Maximal running protocol, peak values of VO_2 and heart rate was calculated from 30 s recordings.

Wingate test

Exactly 10 min after the maximal running test ended, the 30 s Wingate test[23] was performed on a cycle ergometer (Monark Peak bike 894e, Vansbro, Sweden). The initial 5 min was

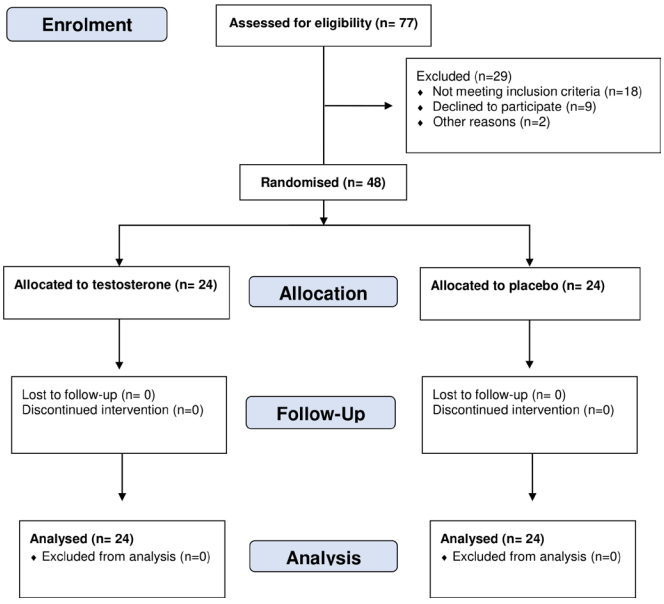

Figure 1 Flow diagram of the trial. All participants completed the study and were included in the final analysis.

recovery cycling without any load at 50–70 rpm. Brake load, at baseline and at the end of the treatment period, was set as 8.5% of body mass.[24] At a given signal, the subject accelerated to a cadence of 120 rpm, at which point the weight loaded basket automatically dropped on to the ergometer wheel and the test time started. Participants' peak power (peak W) (data in online supplementary table 1) and mean power (average W) were analysed (Monark anaerobic test software, V.3.3.0.0, Monark AB, Vansbro). High reproducibility has been reported for this test in both men and women.[23 25]

In a pilot study, eight women performed TTE and the Wingate test on two different days. No significant differences were observed between the two tests: TTE=288.0±67.6 s versus 287.8±77.7 s; Wingate average W=471.3±94.1 W versus 469.4±87.2 W.

Body composition

Body composition (total lean mass, lower limb lean mass, fat mass, and total and spine bone mineral densities (BMD)) was determined by dual energy X-ray absorptiometry (Lunar Prodigy Advance, GE Healthcare, Madison, Wisconsin, USA) at baseline and after treatment, at the Radiology Department, Karolinska University Hospital, as previously described.[13] The accuracy of the whole body BMD is calculated as <0.01 SD g/cm^2.[26]

Endocrine analyses

Serum levels of testosterone were determined by liquid chromatography–tandem mass spectrometry (LC-MS/MS) at the Endoceutics Laboratory, Quebec, Canada.[27] Follicle stimulating hormone, luteinising hormone and sex hormone binding globulin were determined by electrochemiluminescence immunoassay at the Department of Clinical Chemistry, Karolinska University Hospital.[13] Free androgen index was calculated (testosterone (nmol/L) divided by sex hormone binding globulin (nmol/L) × 100).

Table 1 Baseline characteristics of the women in the two treatment groups

	Testosterone (n=24)	Placebo (n=24)
Age (years)	28.4 (3.2)	28.4 (4.3)
Weight (kg)	67.1 (7.2)	65.1 (7.1)
Height (cm)	170 (5)	168 (6)
Body mass index (kg/m^2)	23.3 (1.8)	23.0 (1.9)
Systolic blood pressure (mm Hg)	117.6 (9.1)	117.2 (8.4)
Diastolic blood pressure (mm Hg)	72.9 (6.7)	74.4 (6.2)

Values are mean (SD).

Haemoglobin measurement

Serum haemoglobin levels were measured by a standard spectrophotometric method at the Department of Clinical Chemistry, Karolinska University Hospital.

Patient involvement

Patients were not involved in the study as the purpose was to study the physiological effects of testosterone in healthy women. The results will be disseminated to study participants by letter.

Statistical analyses

Continuous data are presented as mean (SD) or median (IQR), according to distribution. All analyses were based on an intention to treat population, defined as all randomised subjects. The data were analysed using a mixed model, with subjects as a random factor, and treatment (testosterone and placebo), time (baseline and exit) and treatment × time as fixed factors. Treatment was the between group factor and time was the within group factor. Differences between groups were evaluated by the interaction treatment × time, and differences within groups by the factor time, by holding the other factor (treatment) fixed. To adjust for baseline differences, analysis of covariance (ANCOVA) was used, with treatment as a fixed factor and baseline as a covariate in the model. Prior to these analyses, some of the variables were log transformed to compensate for their positively skewed distributions. Also, the assumption with regard to equal variances between groups was explored, and where evidence for differences were present, the ANCOVA was modelled using the robust variance model. Correlations between variables were calculated using Spearman's rank correlation. A P value <0.05 was considered statistically significant. There was no adjustment for multiplicity as the primary outcome variable was 1. Secondary endpoints were considered as exploratory.

Based on power analysis, a sample size of 23 subjects in each group provided 80% power to detect a difference in means of 21 s for the primary outcome TTE, assuming a common SD of 25 s (effect size=0.84).

RESULTS

Baseline characteristics

Seventy-seven women were assessed for eligibility at a screening visit, and 48 fulfilled the criteria for participation. The 48 women were scheduled for a baseline visit and randomly assigned to either testosterone treatment or placebo (figure 1). Baseline characteristics were similar between the two treatments groups (table 1). No participant was lost to follow-up.

Table 2 Endocrine values, body composition and physical performance results at baseline and after 10 weeks of treatment in the two treatment groups

	Testosterone			Placebo			Testosterone–placebo change	P value
	Baseline (n=24)	After treatment (n=24)	P value	Baseline (n=24)	After treatment (n=24)	P value	Least square means (SE)	I Adjusted for baseline II Adjusted for baseline and RV*
Endocrine variables								
FSH (IU/L)	6.0 (2.9)	4.1 (1.8)	0.003	6.1 (1.7)	5.2 (1.8)	0.132	−1.91 (0.37); −0.82 (0.37)	0.031
LH (IU/L)	6.0 (3.0)	6.5 (4.8)	0.534	6.9 (2.9)	7.9 (3.4)	0.257	0.29 (0.81); 1.23 (0.81)	0.417
Testosterone (nmol/L)	0.9 (0.4)	4.3 (2.8)	<0.001	1.0 (0.4)	1.1 (0.4)	0.343	2.54; 0.12†	<0.001;<0.001
Free androgen index	1.2 (0.8–1.7)	6.4 (3.6–12.1)	<0.001	1.1 (0.9–1.5)	1.2 (1.0–1.4)	0.887	1.73 (0.14); −0.04 (0.13)	<0.001;<0.001
Body composition								
Weight (kg)	67.1 (7.2)	67.2 (7.3)	0.351	65.1 (7.1)	65.4 (7.1)	0.397	0.37 (0.36); 0.27 (0.35)	0.856
Total lean mass (g)	47 034 (4 894)	47 773 (4 909)	0.002	45 418 (5 383)	45 582 (5 307)	0.527	923 (264); 135 (258)	0.040
Lower limbs lean mass (g)	15 937 (1 588)	16 276 (1 608)	<0.001	15 340 (1 947)	15 440 (1 960)	0.326	398 (103); 91 (101)	0.041
Body fat (%)	26.5 (7.4)	25.5 (7.4)	0.030	27.2 (5.4)	27.2 (6.0)	0.962	−1.01 (0.45); 0.03 (0.44)	0.107
BMD total (g/cm^2)	1.22 (0.08)	1.21 (0.08)	0.300	1.18 (0.08)	1.19 (0.08)	**0.035**	0.004 (0.004); 0.004 (0.004)	0.498
BMD spine (g/cm^2)	1.06 (0.11)	1.05 (0.10)	0.182	1.01 (0.12)	1.03 (0.12)	0.063	−0.01 (0.006); 0.01 (0.006)	0.052
Physical performance								
Endurance TTE (s)	250 (49)	271 (57)	<0.001	258 (43)	264 (45)	0.264	21.17 (5.14); 5.95 (5.26)	0.045
VO_2 max (L/min)	3.02 (0.38)	3.08 (0.40)	0.013	2.89 (0.37)	2.90 (0.36)	0.283	0.058 (0.22); 0.023 (0.023)	0.284
Wingate 30 s (average W)	470 (67)	485 (70)	0.005	458 (83)	462 (73)	0.445	15.16 (4.78); 3.21 (4.78)	0.084
SJ (m)	0.21 (0.04)	0.21 (0.04)	0.376	0.20 (0.04)	0.20 (0.04)	0.845	−0.003 (0.005); −0.0006 (0.005)	0.663
CMJ (m)	0.25 (0.04)	0.26 (0.05)	0.054	0.25 (0.03)	0.25 (0.04)	0.771	0.001 (0.005); −0.002 (0.005)	0.703
Knee extension peak torque (Nm)	219 (31)	224 (40)	0.267	214 (42)	218 (47)	0.433	5.35 (4.83); 3.78 (4.83)	0.82

Values are mean (SD) or median (IQR).
*RV, robust variance (ie, where there was evidence for that variance within a variable was not equal in the two treatment groups).
†Geometric least square means based on back transformed (natural log) values where it is inappropriate to transform the corresponding standard errors.
BMD, bone mineral density; CMJ, countermovement jump; FSH, follicle stimulating hormone; LH, luteinising hormone; SHBG, sex hormone binding globulin; SJ, squat jump; TTE, time to exhaustion; VO_2, oxygen consumption.

Hormone levels

Serum levels of hormones before and during treatment in the two treatment groups are shown in table 2. Testosterone increased in the active treatment group to a mean level of 4.3 nmol/L, 4.8 times higher than the baseline level (figure 2, table 2). Free androgen index increased, and follicle stimulating hormone levels decreased in the active treatment group with no corresponding changes in the placebo group.

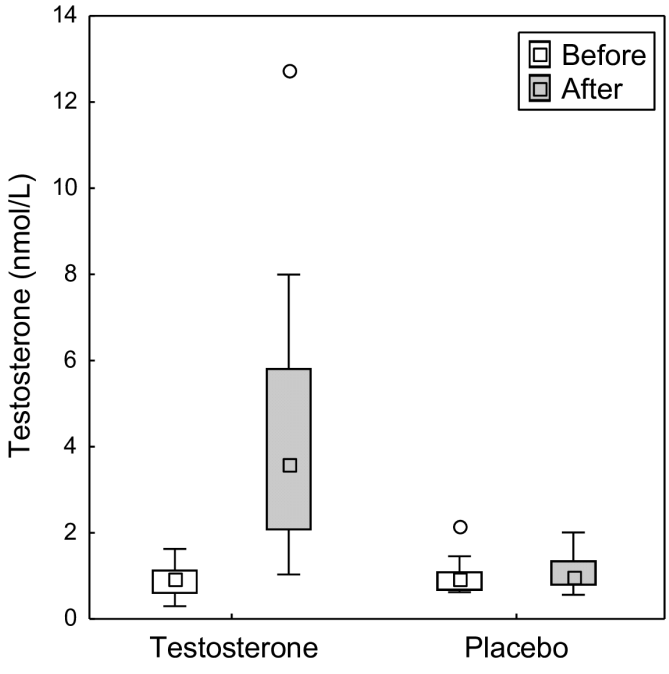

Figure 2 Serum levels of total testosterone before and after 10 weeks of treatment with testosterone or placebo. Values are median (IQR).

Haemoglobin levels

There was no significant change in haemoglobin concentration in either of the treatment groups (data not shown).

Body composition

Body weight showed similar changes in the two groups. In the testosterone group, total lean mass and lower limbs lean mass increased and were significantly different from the placebo group (P=0.040 and P=0.041, respectively). Body fat percentage decreased significantly in the testosterone group (P=0.031) but was not different from the placebo group (P=0.107) (table 2). There was a slight decrease in spinal BMD in the testosterone group and a slight increase in the placebo group, but no significant group difference (P=0.052) (table 2).

Physical performance outcomes

Running TTE increased by 21.1 s (8.5%) in the active treatment group (figure 3A, table 2). The between group treatment effect was significant (mean difference 15.2 s, P=0.045). Maximal oxygen uptake as well as average power obtained in the Wingate test increased in the testosterone group (figure 3B, table 2) but neither was statistically different compared with placebo. Changes in both SJ and CMJ performances were not statistically different after this intervention. Furthermore, there was no significant difference with regard to change in isokinetic knee extension peak torque (figure 3D, table 2).

Correlations

When the data in both groups were pooled, changes in serum testosterone correlated positively with changes in running

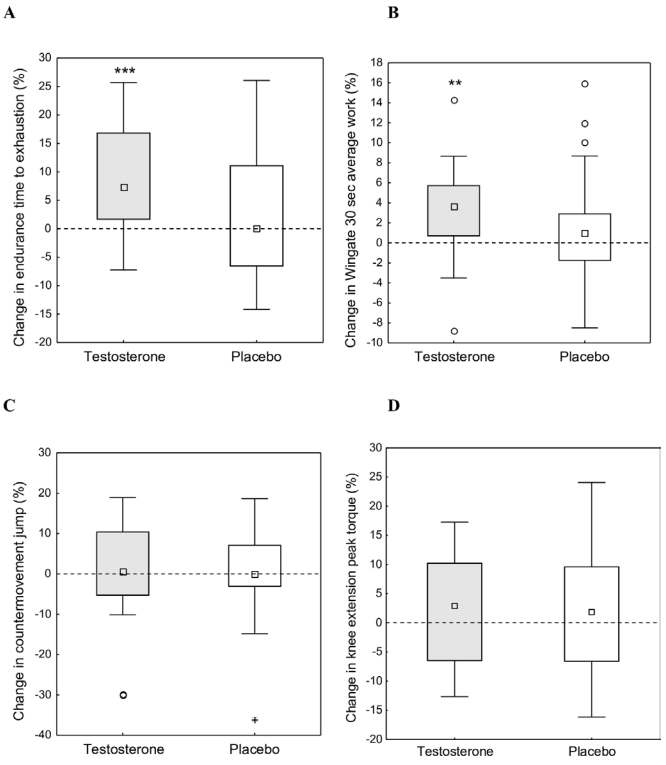

Figure 3 Percentage change in (A) endurance time to exhaustion, (B) Wingate 30 s average work, (C) countermovement jump and (D) knee extension peak torque in the testosterone and placebo groups. Values are median (IQR). **p <0.05, ***p<0.001.

TTE (r=0.31, P=0.034). Changes in serum testosterone also correlated positively with changes in total lean mass (r=0.34, P=0.020). There were significant positive correlations between changes in total lean mass and lower limb lean mass and Wingate average power (r=0.40, P=0.006 and r=0.41, P=0.004, respectively).

Safety and adverse events

Adverse events were documented in 16 and 7 subjects in the testosterone and placebo groups, respectively. In the testosterone group, 58% of the women reported increased acne compared with 25% in the placebo group. Five and four women showed increased body hair growth in the testosterone and placebo groups, respectively. One woman in each group had an increased Ferriman–Gallwey score, indicating hirsutism. Furthermore, menstrual disturbances (ie, prolonged cycles) were reported by four individuals in the testosterone group. All of these adverse events returned to normal 4 months after study completion.

DISCUSSION

This study is the first randomised, placebo controlled trial investigating the effect of testosterone on physical performance in young, physically active women. Ten weeks of exposure to a moderate increase in circulating testosterone caused a significant treatment response in our primary outcome measure, running TTE, by a mean difference of 15.5 s in comparison with placebo. Secondary outcomes, including power output (Wingate), increased in the testosterone group but not significantly compared with placebo (mean difference 12 average W) whereas there was no difference in the treatment response for jump tests and muscle strength. Body weight was unchanged but total lean mass and lean lower limb mass were significantly increased by testosterone compared with placebo.

Aerobic and anaerobic performance

There was a 6.2% between group increase in TTE but no between group difference in VO_2 max. However, when using the equation of Leger and Mercier[28] to convert VO_2 max to maximal aerobic speed (MAS), we calculated that the testosterone group improved its MAS from 3.57 m/s to 3.64 m/s. Bellenger *et al*[29] showed that 1500 m and 2000 m time trials could be run at a speed equal to 1.0627 × MAS and 1 × MAS, respectively. In our case, the observed improvement in MAS would represent performance improvements of 7.7 s and 10.8 s in 1500 m and 2000 m time trials, respectively.

Our results of improved running time to exhaustion extend the cross sectional studies[12 14] reporting that higher circulating testosterone was associated with enhanced performance in female middle distance running events primarily relying on the aerobic energy pathway. Given that haemoglobin concentration remained unchanged in the testosterone supplemented group, we can speculate that other central (cardiac) and/or peripheral (capillary density, oxidative capacity) mechanisms may account for the observed increase in VO_2 max. Our results support a causal aerobic performance enhancing effect of moderate testosterone exposure in young women.

Both average and peak anaerobic power, as measured by the Wingate test, increased in the testosterone group but was not statistically different between groups. Therefore, the present study lends no support for an improvement in anaerobic performance by testosterone.

Body composition

Men increase their muscle mass and strength in a dose dependent manner with increasing doses of testosterone.[30–32] However, there are few data on the effects of exogenous testosterone on physical performance in women. This limitation is likely due to ethical concerns about the risks of adverse effects when administrating increasing doses of testosterone in women. In a previous study in postmenopausal women, a dose dependent increase in lean mass and muscle performance by testosterone was observed.[10] The highest dose (25 mg testosterone enanthate weekly) produced supraphysiological levels of circulating testosterone (mean 7.3 nmol/L) and caused an increase in muscle mass of 4.4% and muscle performance of 12–26% after 24 weeks of treatment.[1 10] In comparison, we obtained a significant change in total and lower limb muscle mass of 2.0% and 2.5%, respectively, with a dose resulting in a moderate increase in testosterone level (mean 4.3 nmol/L). However, we found no significant improvement in isokinetic muscle strength and jump test performance after treatment. This may be explained by the relatively short treatment period (ie, 10 weeks). Furthermore, body fat percentage decreased in the testosterone group resulting in a leaner body composition.

Hyperandrogenism in female athletes

The physiological effect of testosterone is the same whether the source of testosterone is exogenous or endogenous.[1] Our results are therefore of great importance for the ongoing discussion of whether it is fair to allow athletes with naturally producing high testosterone to compete in the female category without reducing their hormonal concentration to the female range. The IAAF requires that athletes with naturally produced high testosterone and normal androgen receptor sensitivity reduce

their testosterone level to <5 nmol/L to be eligible to compete in the female category at the international level in middle distance races (400 m to 1 mile).[9 33]

The decision has been criticised by some scientists and human rights advocates who believe that social and/or legal gender only should determine eligibility, and by others who request more research on the relation between testosterone levels and athletic performance in women.[5 7] As determined by tandem mass spectrometry, the normal female range of circulating testosterone is 0.1–1.8 nmol/L and does not overlap with the normal male range (7.7–29.4 nmol/L).[1 2] The dose we used resulted in an increase in mean serum testosterone of 4.3 nmol/L, which is clearly below the male range, and yet increased some indices of aerobic and anaerobic performances.

Strengths and limitations

The present study has several strengths, including the double blind, randomised, placebo controlled design. The dropout rate was zero and adherence to treatment was judged to be good. All procedures and physical performance tests were standardised according to time of day, food intake and physical activity. Serum levels of testosterone were measured by the gold standard method. Limitations of the study include short treatment duration and a dose of testosterone resulting in moderately elevated levels that were not in the male range. For secondary outcome variables, the power to detect group differences may have been insufficient. Furthermore, the magnitude of the observed findings in our physically active women could not be directly generalised to an elite female population.

CONCLUSION

Our study supports a causal effect of testosterone on physical performance, as measured by running time to exhaustion, in young healthy women. Thus the ergogenic effect of short term moderately increased testosterone concentration seemed to apply for aerobic performance only. Testosterone also promoted a leaner body composition with an increase in muscle mass although body weight was unchanged.

What are the new findings?

- This study is the first randomised, placebo controlled trial investigating a possible causal effect of testosterone on physical performance in young women.
- 10 weeks of exposure to a moderate increase in circulating testosterone caused a significant increase in aerobic running time in comparison with placebo.
- Secondary outcomes, including anaerobic capacity (Wingate), jumping and muscle strength, did not change significantly between groups but lean body mass increased significantly following testosterone exposure.

Acknowledgements The authors thank research assistants Berit Legerstam and Lotta Blomberg at the Women's Health Research Unit, Karolinska University Hospital, for their logistical support. They also thank the statistician Elisabeth Berg, Karolinska Institutet, for ongoing advice regarding study design and statistical analysis, as well as Jan Kowalski for independent statistical advice. The authors are grateful for helpful support from Professor Emeritus Arne Ljungqvist (International Olympic Committee) throughout the study and for suggestions and insights on testosterone treatment from Professor Susan Davis (Monash University, Australia).

Contributors All authors listed met the conditions required for full authorship. Concept and study design: ALH, SB and BE. Acquisition of the data, data analysis and interpretation of the data: all authors. Drafting of the manuscript: ALH. Critical revision and approval of the manuscript for important intellectual content: all authors.

Funding The study was financed by grants from the International Athletics Foundation, the Swedish Research Council (2017-02051), the Swedish Research Council for Sport Science, the Karolinska Institutet and the Swedish Military Forceś Research Authorities (#AF9220914). Lawley Pharmaceuticals provided the treatment of testosterone and placebo.

Competing interests ALH is a medical and scientific consultant for the Swedish Olympic Committee and a member of the International Association of Athletics Federation (IAAF) and the International Olympic Committee (IOC) working groups on hyperandrogenic female athletes and transgender athletes. SB is a medical and scientific consultant for the IAAF and a member of the IAAF and IOC working groups on hyperandrogenic female athletes and transgender athletes. The authors have no other involvement with any entity having a financial interest in the material discussed in the manuscript.

Patient consent for publication Not required.

Ethics approval The study was approved by the regional ethics committee in Stockholm (2016/1485-32, amendment 2017/779-32). All women gave written informed consent.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement The trial is registered at ClinicalTrials.gov ID: NCT03210558. De-identified data from the trial are available on reasonable request from the corresponding author.

ORCID iD

Angelica Lindén Hirschberg http://orcid.org/0000-0001-6481-6277

REFERENCES

1 Handelsman DJ, Hirschberg AL, Bermon S. Circulating testosterone as the hormonal basis of sex differences in athletic performance. *Endocr Rev* 2018;39:803–29.
2 Clark RV, Wald JA, Swerdloff RS, *et al*. Large divergence in testosterone concentrations between men and women: frame of reference for elite athletes in sex-specific competition in sports, a narrative review. *Clin Endocrinol* 2019;90:15–22.
3 Notelovitz M. Androgen effects on bone and muscle. *Fertil Steril* 2002;77(Suppl 4):34–41.
4 Bermon S, Garnier PY, Hirschberg AL, *et al*. Serum androgen levels in elite female athletes. *J Clin Endocrinol Metab* 2014;99:4328–35.
5 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.
6 Bermon S, Ritzén M, Hirschberg AL, *et al*. Are the new policies on hyperandrogenism in elite female athletes really out of bounds? Response to "out of bounds? A critique of the new policies on hyperandrogenism in elite female athletes". *Am J Bioeth* 2013;13:63–5.
7 Sönksen P, Ferguson-Smith MA, Bavington LD, *et al*. Medical and ethical concerns regarding women with hyperandrogenism and elite sport. *J Clin Endocrinol Metab* 2015;100:825–7.
8 Harper J, Lima G, Kolliari-Turner A, *et al*. The fluidity of gender and implications for the biology of inclusion for transgender and intersex athletes. *Curr Sports Med Rep* 2018;17:467–72.
9 Hirschberg AL. Hyperandrogenism in female athletes. *J Clin Endocrinol Metab* 2019;104:503–5.
10 Huang G, Basaria S, Travison TG, *et al*. Testosterone dose-response relationships in hysterectomized women with or without oophorectomy: effects on sexual function, body composition, muscle performance and physical function in a randomized trial. *Menopause* 2014;21:612–23.
11 Cardinale M, Stone MH. Is testosterone influencing explosive performance? *J Strength Cond Res* 2006;20:103–7.
12 Bermon S, Garnier P-Y. Serum androgen levels and their relation to performance in track and field: mass spectrometry results from 2127 observations in male and female elite athletes. *Br J Sports Med* 2017;51:1309–14.
13 Eklund E, Berglund B, Labrie F, *et al*. Serum androgen profile and physical performance in women Olympic athletes. *Br J Sports Med* 2017;51:1301–8.
14 Bermon S, Hirschberg AL, Kowalski J, *et al*. Serum androgen levels are positively correlated with athletic performance and competition results in elite female athletes. *Br J Sports Med* 2018;52:1531–2.
15 Rickenlund A, Carlström K, Ekblom B, *et al*. Hyperandrogenicity is an alternative mechanism underlying oligomenorrhea or amenorrhea in female athletes and may improve physical performance. *Fertil Steril* 2003;79:947–55.
16 Rickenlund A, Thorén M, Carlström K, *et al*. Diurnal profiles of testosterone and pituitary hormones suggest different mechanisms for menstrual disturbances in endurance athletes. *J Clin Endocrinol Metab* 2004;89:702–7.
17 Hagmar M, Berglund B, Brismar K, *et al*. Hyperandrogenism may explain reproductive dysfunction in Olympic athletes. *Med Sci Sports Exerc* 2009;41:1241–8.
18 Coste O, Paris F, Galtier F, *et al*. Polycystic ovary-like syndrome in adolescent competitive swimmers. *Fertil Steril* 2011;96:1037–42.

19 Javed A, Kashyap R, Lteif AN. Hyperandrogenism in female athletes with functional hypothalamic amenorrhea: a distinct phenotype. *Int J Womens Health* 2015;7:103–11.
20 Yildiz BO, Bolour S, Woods K, *et al*. Visually scoring hirsutism. *Hum Reprod Update* 2010;16:51–64.
21 Markovic G, Dizdar D, Jukic I, *et al*. Reliability and factorial validity of squat and countermovement jump tests. *J Strength Cond Res* 2004;18:551–5.
22 Foss Ø, Hallén J. Validity and stability of a computerized metabolic system with mixing chamber. *Int J Sports Med* 2005;26:569–75.
23 Bar-Or O, Dotan R, Inbar O. 30-sec all-out ergometric test-its reliability and validity for anaerobic capacity. *Isr J Med Sci* 1977;13:326–7.
24 Bradley AL. The Wingate test: effect of load on the power outputs of female athletes and nonathletes. *J Strength Cond Res* 1992;6:193–9.
25 Weinstein Y, Bediz C, Dotan R, *et al*. Reliability of peak-lactate, heart rate, and plasma volume following the Wingate test. *Med Sci Sports Exerc* 1998;30:1456–60.
26 Brismar TB, Ringertz J. Effect of bone density of the head on total body DEXA measurements in 100 healthy Swedish women. *Acta Radiol* 1996;37:101–6.
27 Ke Y, Bertin J, Gonthier R, *et al*. A sensitive, simple and robust LC–MS/MS method for the simultaneous quantification of seven androgen- and estrogen-related steroids in postmenopausal serum. *J Steroid Biochem Mol Biol* 2014;144:523–34.
28 Léger L, Mercier D. Gross energy cost of horizontal treadmill and track running. *Sports Med* 1984;1:270–7.
29 Bellenger CR, Fuller JT, Nelson MJ, *et al*. Predicting maximal aerobic speed through set distance time-trials. *Eur J Appl Physiol* 2015;115:2593–8.
30 Bhasin S, Woodhouse L, Casaburi R, *et al*. Testosterone dose-response relationships in healthy young men. *Am J Physiol Endocrinol Metab* 2001;281:E1172–E1181.
31 Crewther BT, Cook C, Cardinale M, *et al*. Two emerging concepts for elite athletes: the short-term effects of testosterone and cortisol on the neuromuscular system and the dose-response training role of these endogenous hormones. *Sports Med* 2011;41:103–23.
32 Finkelstein JS, Lee H, Burnett-Bowie S-AM, *et al*. Gonadal steroids and body composition, strength, and sexual function in men. *N Engl J Med Overseas Ed* 2013;369:1011–22.
33 IAAF eligibility regulations for the female classification. Available: https://www.iaaf.org/about-iaaf/documents/health-science