Concerns regarding the research design, participant selection, and data interpretation of 'Body composition, exercise-related performance parameters and associated health factors of transgender women, cisgender women and cisgender men volleyball players'

For correspondence:
c.kirk@shu.ac.uk

Christopher Kirk[1], Kerry McGawley[2], Grant Abt[3], Tommy R Lundberg[4], Cathy Devine[3], Gregory A Brown[5], Eva Carneiro[6]

[1] Sheffield Hallam University, UNITED KINGDOM, [2] Mid Sweden University, SWEDEN, [3] Independent researcher, UNITED KINGDOM, [4] Karolinska Institutet, SWEDEN, [5] University of Nebraska at Kearney, USA, [6]The Sports Medical Group, UNITED KINGDOM

Please cite as: Kirk, C., et al., (2025). Concerns regarding the research design, participant selection, and data interpretation of 'Body composition, exercise-related performance parameters and associated health factors of transgender women, cisgender women and cisgender men volleyball players'. *SportRχiv*.

This pre-print represents an open letter to the Editor of the British Journal of Sports Medicine (BJSM) in lieu of this journal offering a 'letters to the editor' option as part of its publishing practice. A shorter version of this work has been posted on the 'rapid response' section of the BJSM website attached to the article being discussed. This expanded version is posted here to provide a more detailed discussion of the study's flaws as highlighted by the authors, and to ensure a permanent DOI for citation purposes.

The BJSM recently published an article by Alvares et al. (1) which claimed that transgender women (males that self identify as women, TW) volleyball players had similar physiological and performance characteristics to females (cisgender women, CW, in the article), but were statistically different to non-transgender males (cisgender men, CM, in the article). The results and conclusions of this study are at best misleading, and most likely invalid due to vital flaws in

All authors have read and approved this version of the manuscript. This article was last modified on May, 16 2025.

Authors X @*ChrisKirk_ASP* can be reached on Twitter.

the methods, sample, and resulting inferences. These issues raise concerns regarding how this work progressed through peer review into a journal that is considered by many to be an outlet for rigorous research.

The authors state their aim was to compare 'matched' groups of TW, female and non-transgender male volleyball players. This 'matching', however, resulted in groups that cannot be appropriately compared for the study's stated aims. The female group's weekly training duration was an impressive 13.9 (11.3-16.4) hours·week^{-1}. Such an amount of training would be in keeping with high-level team-sport athletes, likely placing this group in the 'highly trained/national level' tier (tier 3) or above of McKay et al.'s classification framework (2). The TW group in comparison only reported training 4.1 (3.4-4.8) hours·week^{-1}. This would make them active but recreational (tier 2) participants *at best*. Any comparison of athletic or physical performance between these groups is therefore meaningless given: a) the performance adaptations the females will have experienced due to their extensive training; b) the genetic characteristics of these females that likely contributed to their selection as high-level volleyball players. With neither of these being the case for the TW group, it is no surprise that females displayed large to very large advantages over TW in absolute handgrip strength (d = .90), absolute countermovement jump (d =1.42) and absolute squat jump (d = .90). The TW group do outperform the females in both absolute (d=0.49) and relative (d=0.96) $\dot{V}O_2$max to a moderate-large extent. Each of these comparisons being 'non-significant' is used to infer that the TW do not differ to females in terms of performance. A more correct interpretation would be that females who train at a high level may outperform recreationally trained males in some variables, but males will still retain moderate-large advantages in others, even in the absence of equivalent training.

Further examination of the group characteristics reveal more matching errors. The TW group were on average 5 years older, 5 cm shorter, and 8 kg lighter than the female group. Such group descriptives not only deviate from well known sex-based differences in stature and mass (3), but also do not align with data showing male volleyball players being both taller and of greater mass than equivalent standard female players (4). When compared to the non-transgender males the TW were roughly comparable in age, but 18 kg lighter and 16 cm shorter (this latter variable being unaffected by gender hormone therapy (5)). Accordingly, the authors report between-group effect sizes for these variables of η^2p = .23-.33, effects that could only be described as extreme, demonstrating that the groups are in no way 'matched'. The authors are, therefore, comparing short, light males (the TW group in this instance) to tall, athletic females who are 5 years younger. Such inappropriate comparisons between non-matched groups have been highlighted previously as a critical problem in transgender athlete studies (6).

The authors state that 12 TW were excluded as participants due to "inappropriate use of hormones (GAHT)" without providing any details about what was considered "inappropriate" or

why this would preclude their participation. Similarly opaque exclusion criteria were applied to females and non-transgender males based on a claimed attempt to match the groups by age and BMI. The resulting large differences reported in both age and BMI shows that these attempts were unsuccessful. We therefore find it surprising that a group of male volleyball players with such short stature was found and then placed into an experimental group purely by chance of the stated exclusion criteria. The authors do (appropriately) state that this study was conducted in partnership with and at the request of a TW volleyball team. Whilst this declaration is welcome and encouraged, the veracity and trustworthiness of the results must be questioned in light of the ineffective matching processes and group allocations. Given the team's status as a TW team, there is a potential incentive for these particular participants to underperform during laboratory tests, or for the team organisation itself to suggest certain participants, whilst excluding others.

We note there was no *a priori* sample size estimation or rationale presented for the smallest effect size of interest for each outcome, but a post-hoc power calculation was performed, which is conceptually flawed (7). Given that the study is clearly underpowered for certain comparisons (even to detect some large differences) and with considerable uncertainty in population estimates, the claims of there being no differences between the TW and females in this study, therefore, cannot be supported. To suggest that groups 'are similar' or 'the same' based on $p > 0.05$ is a common misconception that is highlighted as being responsible for misinterpretation of data (8, 9). As such, claiming that the performance of the females and TW did not differ, or were equivalent, when based on $p > 0.05$ is an invalid statement. With such large effects and low statistical power being reported, it is clear that any non-statistically significant differences between the TW and females results from small sample sizes. The two groups being drawn from distinct performance populations and then compared using inappropriate statistical methods are both fatal shortcomings of this work, shortcomings that violate the CHAMP guidelines (10) that BJSM specifically request all submissions adhere to (https://bjsm.bmj.com/pages/authors). Accordingly, it would be expected that a sufficiently rigorous review process would have resulted in, if not a rejection of this paper, an appropriate revision to address these problems.

In addition to these fundamental methodological weaknesses, the authors make statements regarding the current status of transgender participation in sport that are common, but nevertheless, unsubstantiated. Firstly, the authors make the claim that "the inclusion of transgender athletes has promoted significant dialogue and controversy". This is a commonly used misdirection in this debate. The very real controversy in question is not about 'inclusion' of transgender athletes, as no serious person has suggested that transgender athletes should be excluded from sport. In actuality the debate regards whether *males* should be allowed to compete against *females*. Whilst international sports federations (IFs) have been accused of 'excluding' transgender athletes (11), the reality is that no organisation has installed any policies that do this. Some IFs have reinforced their eligibility rules to maintain the 'women's' category for

females only (https://worldathletics.org/news/press-releases/world-athletics-launches-new-stakeholder-consultation-on-female-eligibility), meaning TW may only compete against other males. Others have changed the 'men's' category to being 'open' for both sexes to compete in regardless of gender identity, whilst maintaining a separate 'women's' category for females only (https://d2cx26qpfwuhvu.cloudfront.net/worldrowing/wp-content/uploads/2025/04/01043112/2025-World-Rowing-Rules-of-Racing-Overall-Classic-rowing-Related-byelaws_March2025.pdf) . Regardless of the specific policy, all sports still allow all people of both sexes to compete against *others of their own sex* with no-one being excluded. Continuing to frame the debate as one of 'inclusion/exclusion' is, therefore, disingenuous and feeds into much of the inaccurate and unhelpful commentary around the topic.

Secondly, the authors claim that "…differentiation between cisgender women (CW) and cisgender men (CM) is largely attributed to testosterone.". This represents another common but misleading discussion point that attempts to reduce the clear differences between males and females to a simple matter of testosterone levels. This notion is simply untrue. Examination of the available data makes it clear that male advantage in sport is provided *by the very nature of being male*, with greater levels of testosterone, particularly historic exposure during male puberty, being just one of many sub-factors related to this status, as detailed by Lundberg et al. (12).

Each of the flaws discussed here mirror those of a previous article published recently (13) that reached similarly unsupportable conclusions. These conclusions were challenged independently by several members of the scientific community at the time of publication (https://bjsm.bmj.com/content/58/11/586.responses#concerns-regarding-respiratory-data-interpretation-athlete-definition-and-group-matching-in-strength-power-and-aerobic-capacity-of-transgender-athletes-a-cross-sectional-study). Seeing such poor scientific standards repeated again within the pages of the same journal further raises concerns regarding the robustness of the editorial and peer review processes in this area of research. Authors and reviewers carry the ethical responsibility to present the truth and preserve standards in publication and presentation of evidence particularly in the present era of misinformation.

In conclusion, it is imperative that research on transgender athletes (and all athletes) is conducted in a rigorous manner, consistent with the basic principles of science to ensure that decisions are based on accurate and trustworthy information. Instead, this study has fundamental and clear flaws in research design resulting in unsupportable conclusions and as such should not have been published in its current form.

The academic debate on this topic is having a direct, rapid and meaningful influence on global sports (https://worldathletics.org/news/press-releases/world-athletics-launches-new-stakeholder-consultation-on-female-eligibility) and governmental policy

(https://www.ohchr.org/en/documents/thematic-reports/a79325-report-special-rapporteur-violence-against-women-and-girls-its). Due to such potential impact these studies are having on people's lives and careers, *no research would be preferable to bad research*.

Contributions

Substantial contributions to conception and design: CK, KM, GA, TL, CD, GB, EC
Drafting the article or revising it critically for important intellectual content: CK, KM, GA, TL, CD, GB, EC
Final approval of the version to be published: CK, KM, GA, TL, CD, GB, EC

REFERENCES

[1]. L.A. Alvares, M.V.L dos Santos Quaresma, F.P Nakamoto,... R. Ferreira. "Body composition, exercise-related performance parameters and associated health factors of transgender women, cisgender women and cisgender men volleyball players". British Journal of Sports Medicine (2025) doi:10.1136/bjsports-2024-108601

[2]. A.,McKay, T. Stellingwerff. E. Smith, D. Martin, I. Mujika, V. Goosey-Tolfrey, ... Burke, L. (2021). "Defining training and performance caliber: a participant classification framework". International Journal of Sports Physiology and Performance 17(20). (2021). pp 317–331.

[3]. M.J. Joyner, S.K. Hunter, J.W. Senefeld. Evidence on sex differences in sports performance. (2025) Journal of Applied Physiology 138(1). pp 274-281. doi: 10.1152/japplphysiol.00615.2024

[4]. T. Sattler, V. Hadzic, E. Dervisevic, G. Markovic. "Vertical jump performance of professional male and female volleyball players: effects of playing position and competition level". (2015) Journal of Strength & Conditioning Research 29(6). pp 1486-93.

[5]. L.S Boogers, C.M. Wiepjes, D.T. Klink, I. Hellinga I, A van Trotsenburg, M. den Heijer, S. Hannema. "Transgender girls grow tall: adult height is unaffected by GnRH analogue and estradiol treatment". (2022). Journal of Clinical Endocrinology and Metabolism 107(9) doi: 10.1210/clinem/dgac349

[6]. T. Lundberg,, M. O'Connor, C. Kirk, N. Pollock, G. Brown, (2024), "Comment on: "A unique pseudo-eligibility analysis of longitudinal laboratory performance data from a transgender female competitive cyclist". (2024) Translational Exercise Biomedicine 1(3-4) https://doi.org/10.1515/teb-2024-0026

[7]. Y. Zhang, R. Hedo, A. Rivera, R. Rull, S. Richardson, X.M. Tu. "Post hoc power analysis: is it an informative and meaningful analysis?". (2019. General Psychiatry, 32(4), e100069. https://doi.org/10.1136/gpsych-2019-100069

[8]. S. Greenland, S.J. Senn, K.J. Rothman, J.B. Carlin, C. Poole, S.N. Goodman, D.G. Altman (2016). "Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations." (2016) European Journal of Epidemiology, 31(4), pp 337–350. https://doi.org/10.1007/s10654-016-0149-3

[9]. D. Lakens. (2017). "Equivalence Tests: A practical primer for t Tests, Correlations, and Meta-Analyses". (2017). Social Psychology and Personality Science 8(4), pp. 355–362. https://doi.org/10.1177/1948550617697177

[10]. M.A. Mansournia, G.S. Collins, R. Nielsen,, M. Nazemipour, N. Jewell, D. Altman, M. Campbell (2021). "A CHecklist for statistical Assessment of Medical Papers (the CHAMP statement): explanation and elaboration." (2021). British Journal of Sports Medicine, 55(18), pp.1009–1017. https://doi.org/10.1136/bjsports-2020-103652

[11]. S. Harris, S. Jedlicka, R. Pielke, H. Ryan, (2023), "The politics of exclusion: Analyzing U.S. state responses to interscholastic transgender athletes." (2023) International Journal of Sport Policy and Politics, 15:4, https://doi.org/10.1080/19406940.2023.2242878

[12]. T. Lundberg, R. Tucker, K. McGawley, A. Williams, G. Millet, O. Sandbakk, G. Howatson, G. Brown, L. Carlson, S. Chantler, M. Chen, S. Heffernan, N. Heron, C. Kirk, M. Murphy, N. Pollock, J. Pringle, A. Richardson, J. Santos-Concejero, G. Stebbings, A. Christiansen, S. Phillips, C. Devine, C. Jones, J. Pike, E. Hilton, "The International Olympic Committee framework on fairness, inclusion and nondiscrimination on the basis of gender identity and sex variations does not protect fairness for female athletes," (2024), Scandinavian Journal of Medicine and Science in Sports, https://doi.org/10.1111/sms.14581

[13]. B. Hamilton, A. Brown, S. Montagner-Moraes, C. Comeras-Chueca, P. Bush, F. Guppy, Y. Pitsiladis. "Strength, power and aerobic capacity of transgender athletes: a cross-sectional study". (2024) British Journal of Sports Medicine 58(11).