INTERNATIONAL JOURNAL OF
–EXERCISE SCIENCE–

Inclusive

*Brief Report*

# An Example Analysis for a Gender-inclusive Approach in Sport and Exercise Science Research using Fitbit Outcomes from the *All of Us* Research Program Dataset

James W. Navalta‡1, Dustin W. Davis†1, Jafra D. Thomas‡2, Whitley J. Stone‡3

[1]Department of Kinesiology and Nutrition Sciences, University of Nevada, Las Vegas, Las Vegas, NV, USA; [2]Department of Kinesiology and Public Health, California Polytechnic State University, San Luis Obispo, CA, USA; [3]School of Kinesiology, Recreation, and Sport, Western Kentucky University, Bowling Green, KY, USA

†Denotes early-career investigator, ‡Denotes established investigator

## Abstract

***International Journal of Exercise Science 18(1): 1133-1141, 2025.*** https://doi.org/10.70252/HRNQ4501 Prevalence studies with wearable devices are used to understand disparities in health-related physical activity behaviors and whether interventions are efficacious. However, studies have been limited to a binary definition of sex. This example analysis aimed to demonstrate how researchers can investigate differences in data beyond the sex-gender binary. Using a cross-sectional analysis of the All of Us Research Program dataset, participants' self-identified gender was categorized into Cisgender Female ($n$ = 10,401), Additional Options ($n$ = 27), Non-binary ($n$ = 84), Transgender ($n$ = 17), and Cisgender Male ($n$ = 4,470). Fitbit data on active calories, steps, sedentary minutes, and very active minutes were analyzed following a valid statistical decision framework found in the companion editorial to this paper. Data were checked for normality using the Shapiro-Wilk test, and because data were not normally distributed, homogeneity was evaluated using the Brown-Forsyth test. The omnibus test for significant group differences was determined using the Kruskal-Wallis test, with significance accepted at $p < 0.05$. Effect sizes (ES) for omnibus test results were calculated using Epsilon squared. Results provide evidence for differences in physical activity metrics among gender groups ($p < 0.001$; active calories ES = 0.069, steps ES = 0.005, and very active minutes ES = 0.026). Cisgender males had higher active calories, steps, and very active minutes than cisgender females (40% more) and non-binary individuals (45% more). No differences were observed among other gender groups studied. These findings highlight that activity patterns vary beyond traditional binary classifications, emphasizing the need for gender-inclusive research in sport and exercise science. Specifically, the disparities observed underscore the importance of nuanced interpretations and tailored recommendations for diverse populations, addressing systemic gaps in supporting gender-diverse individuals in health and exercise behaviors.

Keywords: Wearables, diversity and inclusion, representation and underrepresentation, equity, wellness interventions, community-based participatory research

## Introduction

Research indicates a gender difference may be present with respect to how people interact with wearable devices.[1] This contrast may be due to certain clothing having a lower capacity to carry smartphones,[2] or to differences in socialization that affect the ability to multitask when using devices.[3] Devices used specifically for tracking fitness are similar to the previous examples, suggesting gender differences may occur in how people interact with fitness trackers or their proficiency to use them when multitasking.

To date, qualitative studies indicate that gender is an important factor when considering data derived from fitness trackers. A duo-ethnographic study provided evidence that gendered design features reinforce the dominant socio-cultural understandings of gender and femininity using the Jawbone UP3 (a now obsolete tracker).[4] It is unclear how gendered design may impact the usage of fitness trackers.

There is evidence that fitness tracking app usage is different among sexes. Sex was a determinant in single-predictor models for sharing results or using the Runtastic application live tracking feature.[5] Male participants in the study shared results and used live tracking with greater frequency than female participants.[5] Another study on the use of wearables during popular world marathons found that women tend to utilize non-specific lighter devices (such as the Garmin Vivo), while men employ specialized trail and multisport devices (such as the Garmin Forerunner).[6] While the previous studies referenced gender in their titles, only female and male participants were reported, and it is unclear how sex data were obtained. Based on the available literature, we suggest an investigation is needed to explore fitness wearables and gender using transparent and inclusive study designs.

As identified previously, the need for inclusive gender options in wearable design has been noted. It is unknown whether quantitative data from wearables differ among genders beyond the binary definition of sex (female, male). Thus, a purpose of the present study was to determine whether observable gender differences exist using Fitbit data obtained from the *All of Us* dataset. As an application of a recently published statistical decision framework for gender-inclusive research in sport and exercise science,[7] we utilized the *All of Us* dataset to conduct an example investigation into ways genders may differ among physical activity metrics measured through wearable devices. Given the quantitative and qualitative research findings previously summarized, the hypothesis of the present study of this example analysis was that differences among genders would be present for the measures of active calories, steps, sedentary minutes, and very active minutes when more than a binary definition of sex was employed. This present brief report serves as an example study to accompany the IJES editorial "A Step-by-step Statistical Decision Framework for a Gender-inclusive Approach in Sport and Exercise Science Research".[7] That editorial contains a detailed framework for an approach to conducting statistical testing for sex or gender differences which was followed employed for the present study.

## Methods

The research conducted for this example analysis was carried out fully in accordance to the ethical standards of the *International Journal of Exercise Science*.[8] As mentioned previously, the framework for conducting statistical testing for sex or gender differences is presented in the accompanying editorial.[7]

*Participants*

Participants in the *All of Us* Research Program provided self-identified responses for their gender, and responses were sorted into five mutually exclusive categories: Cisgender Female, Additional Options, Non-binary, Transgender, and Cisgender Male. People in the transgender category (that is, individuals who selected the transgender response option) included transgender women and transgender men. Furthermore, it was presumed that people who self-identified as simply female or male were cisgender, given that all participants were given the option to specify their gender as transgender or non-binary. It is acknowledged that overlapping gender identities exist (for example, someone who identifies as transgender could also identify as non-binary), however for the purposes of this investigation the singular identity noted in the All of Us survey was considered mutually exclusive. Cases were excluded only if they did not identify any gender on the survey or did not select the "prefer not to answer" option.

*Protocol*

We conducted a cross-sectional analysis using the controlled tier v7 data repository of the *All of Us* Research Program.[9] Data were collected for U.S. residents aged 18 years or older who enrolled from May 31, 2017, to January 1, 2022. Data sources included health surveys administered in English or Spanish and Fitbit device measurements. The *All of Us* Research Program Science Committee provided approval, and because the present study was considered non-human participants research (i.e., a secondary data analysis); the requirement for informed consent was not applicable (i.e., informed consent was already obtained when participants enrolled into the All of Us Research Program).

The recently released (April 2024) Fitbit data from the *All of Us* v7 data repository were used to obtain the following outcomes: active calories, steps, sedentary minutes, and very active minutes per day. Active calories represent energy expended through exercise and physical activity beyond the basal metabolic rate. Fitbit devices track this outcome through a combination of heart rate, movement data, and user data entered into the personal profile (e.g., sex, age, height, and body mass).[10] Steps represent the number of registered steps taken throughout the day. Fitbit devices use an accelerometer to detect movement and translate movement data into steps. Sedentary minutes represent the total time being inactive, usually periods of prolonged sitting or lying down. Fitbit devices use the accelerometer to identify periods of inactivity. Very active minutes represent time spent engaged in vigorous physical activity with elevated heart rate. Fitbit devices use heart rate data and movement data to track and count this activity as very active minutes.

*Statistical Analysis*

For teaching purposes, a more detailed description of the decision framework for statistical testing is found in the Discussion of the accompanying editorial.[7] Specific to the current example analysis, and aligning with the statistical decision framework for gender-inclusive analysis by Navalta et al,[7] data were checked for normality using the Shapiro-Wilk test. Because data were not normally distributed, data were evaluated for homogeneity using the Brown-Forsythe test which aligns with the statistical testing framework.[7] Following the statistical framework,[7] the omnibus test for significant group differences was determined using the Kruskal-Wallis test. Significance was accepted at the $p < 0.05$ level. Effect sizes for omnibus test results were calculated using Epsilon squared ($\varepsilon^2$) with interpretation being small effect < 0.01, medium effect = 0.06, and large effect > 0.14.[11] Post hoc pairwise comparisons were determined using Dunn's Test which mitigates measurement error when test assumptions are violated, including when groups are unequal.[12] As required by the *All of Us* Research Program at the time the analysis was conducted, all data analyses were performed using Python within the Jupyter Notebook online environment.

**Results**

Fitbit data from the *All of Us* data set revealed significant differences among multiple genders considered in the sample. Significant differences in physical activity metrics were observed across gender groups (Kruskal-Wallis $p < 0.001$, medium effect size $\varepsilon^2 = 0.0684$). Cisgender males had a higher number of active calories, steps, and very active minutes than people who identified as cisgender female (40% more in males vs females; $p < 0.001$) and people who identified as non-binary (45% more than non-binary individuals; $p = 0.002$) (Table 1). Furthermore, people who identified as cisgender male had a higher number of active calories than people who identified as Additional Options ($p = 0.012$) (Table 1). These differences suggest systemic disparities in physical activity engagement, likely influenced by intersecting social, environmental, and psychological determinants.

There were no differences for any outcome among people who identified as Additional Options, Non-Binary, and Transgender (Table 1). Sedentary minutes did not differ significantly among gender groups ($p = 0.0738$), indicating that despite differences in active calories and steps, time spent in sedentary behavior remained similar across participants (Table 1). The effect sizes for the omnibus test and pairwise comparisons for each outcome were small and medium. This suggests that factors influencing engagement in high-intensity activity may differ from those affecting overall inactivity levels, warranting further investigation into behavioral and environmental contributors.

Table 1 shows that while cisgender males had the highest levels of physical activity (mean steps = 8,429), non-binary individuals and those in the Additional Options category recorded lower levels (6,583 and 6,546 steps, respectively). The differences between these groups suggest that gender identity may interact with physical activity behaviors in ways that extend beyond traditional sex-based classifications.

**Table 1.** Fitbit outcomes from the *All of Us* data set among multiple genders.

| | Sample *n* (%) | Active Calories | Steps | Sedentary Minutes | Very Active Minutes |
|---|---|---|---|---|---|
| Cisgender Female | 10,401 (69.3) | 875 (517)* | 7,634 (5,088)* | 885 (314) | 19 (28)* |
| Additional Options | 27 (0.2) | 820 (498)* | 6,546 (3,813) | 833 (306) | 17 (19) |
| Non-binary | 84 (0.6) | 887 (554)* | 6,583 (5,063)* | 822 (277) | 17 (25)* |
| Transgender | 17 (0.1) | 869 (554) | 6,727 (4,546) | 1,037 (303) | 15 (19) |
| Cisgender Male | 4,470 (29.8) | 1,224 (665) | 8,429 (5,372) | 872 (297) | 30 (38) |
| Kruskal-Wallis *p*-Value | | < 0.001 | < 0.001 | 0.0738 | < 0.001 |
| Effect Size | | 0.0684 (medium) | 0.0048 (small) | 0.0006 (small) | 0.0264 (small) |

Groups are compared using means, with standard deviation shown in parenthesis.* = different from cisgender male.

## Discussion

The authors' aim in conducting this example analysis of gender inclusive research was two-fold: (1) to provide a sport and exercise science example for conducting and reporting gender-inclusive research that others may employ in their study design and analysis, and (2) illustrate presented principles using an original study implemented expressly for the accompanying editorial (i.e., simultaneous with writing this Brief Report, the authors addressed a knowledge gap in exercise and sport science, by investigating whether Fitbit outcomes differ among genders).[7] The outlined approach encourages an a priori rationale for testing for sex or gender differences, identifying the number of independent groups being evaluated, and using appropriate statistical analyses. Fitbit outcomes from the *All of Us* Research Program dataset were utilized in the current example, which may be used by sport and exercise scientists in the future.

The findings from the present study indicate that gender-inclusive approaches require a more nuanced interpretation of outcome data, as activity patterns appear to vary beyond traditional binary classifications. Specifically, cisgender male participants exhibited higher activity levels than non-binary participants and those categorized as Additional Options. These results align with previous research indicating that gender minorities may face additional barriers to engaging in structured exercise.[13] Such disparities likely reflect intersecting social, psychological, and institutional influences, including gendered expectations of athleticism, vulnerability to discrimination in exercise spaces, and limited access to safe, inclusive exercise environments.[13] For example, binary-only restrooms and locker rooms, sex-based exercise programming, and insufficient staff training in gender inclusivity create systemic barriers to participation. These forms of structural exclusion may discourage or restrict engagement among gender-diverse individuals. Judge et al. highlight the importance of structural diversity, equity, and inclusion strategies within kinesiology and allied health programs.[13] Although their work does not explicitly focus on sex and gender diversity, their emphasis on institutional barriers, such as insufficient targeted outreach to diverse communities and cultural competency training, underscores the need for systemic reforms that would also benefit gender-diverse populations in exercise settings.[13]

It is evident that sex- and gender-inclusive investigations are desperately needed, particularly so that sex and gender minorities, their exercise professionals, and their healthcare providers have appropriate data upon which to base recommendations. For example, a recent systematic review focused on physical activity levels in Lesbian, Gay, Bisexual, Transgender, and Queer or Questioning (LGBTQ) individuals, but none of the eligible peer-reviewed studies ($K = 19$) included sex and gender minorities as categorized in the current investigation (i.e., the focus was on sexual orientation instead of sex or gender).[14] Our findings complement previous research indicating that individuals who identify as transgender or gender nonconforming report lower levels of physical activity. For example, Bishop et al found that 23.9% of gender expansive individuals reported being physically active zero times over the last week compared to cisgender peers (10.2%).[15] A survey of Spanish youth reported that transgender individuals are less likely to practice sports than cisgender counterparts.[16] Despite significant differences in active calorie expenditure and step count, sedentary time did not significantly differ among gender groups in the current investigation ($p = 0.0738$). This suggests that while some gender identities engage in lower-intensity movement, they may not necessarily exhibit greater sedentary behaviors. One possible explanation is that factors influencing engagement in moderate or vigorous activity (e.g., exercise motivation, social norms) differ from factors affecting total sedentary time, which may be influenced by occupational, academic, or home environments.[15] These findings are underscored by the report that transgender youth often feel unsafe in physical education classes, which is a further barrier to physical activity.[17] Future research should explore whether these findings hold across different age groups, ethnicity, country of origin or socioeconomic backgrounds. To facilitate inclusion, research teams would benefit from staff training, diverse representation on the team (including members of gender expansive groups), using gender-neutral language and respecting the use of pronouns, as well as outward displays of allyship (such as a diversity and inclusion statement on the laboratory or project website, and inclusive signage within the facility).

Investigations included in the aforementioned systematic review did not capture potential disparities on the basis of gender beyond the presumed binary.[14] We need to acknowledge that the current study, while including a greater number of genders, may be subject to some misclassification bias as the dataset restricts respondents to a single gender expression. This limitation diminishes understanding of factors that influence frequency and quality of physical activity pursuits by people with diverse genders. It is understood that physical activity is salutary for physical, psychological, spiritual and social health,[18] but insights characterizing physical activity opportunities for diverse genders remain obscure, which is problematic. These patterns reflect broader systemic gaps in institutional readiness to support gender-diverse populations in kinesiology and allied health fields.[13] In the context of the present study for the accompanying editorial,[7] our results indicate that sex and gender minorities within the *All of Us* research project may have activity patterns (step count, very active minutes and the resultant active calories) that are more closely aligned with cisgender females than with cisgender males, which may constitute a health behavior disparity. The observed differences in activity levels across gender identities highlight the need for more tailored exercise recommendations. Sport and exercise professionals should:

1. **Develop gender-inclusive exercise testing protocols**—current guidelines primarily focus on binary sex classifications, which may not reflect the physiological and behavioral characteristics of gender-diverse individuals.[19]

2. **Improve accessibility of fitness spaces**—non-binary and transgender individuals may avoid exercise settings due to discrimination or lack of gender-affirming facilities (e.g., locker rooms, gender-neutral changing areas).

3. **Consider psychosocial factors in exercise programming**—social support, identity affirmation, and inclusivity in team-based activities may play a critical role in encouraging participation among gender-diverse individuals.

Future research should explore whether customized exercise prescriptions improve physical activity engagement and long-term adherence in gender minorities. Further investigation is required to determine the long-term consequences of the present study's findings for cisgender females and gender minorities; investigations are also needed that examine the social-cultural and institutional factors that perpetuate disparities in physical activity behaviors among diverse genders and ways to effectively mitigate them.

**Acknowledgement**

The authors wish to acknowledge and thank Dr. Trevor Dean Ruiz (Assistant Professor, Department of Statistics, California Polytechnic State University, San Luis Obispo) for providing a review, particularly with regards to statistical approach and interpretation.

University of Nevada, Las Vegas (UNLV) is situated on the traditional homelands of Indigenous groups, including the Nuwu or Nuwuvi, Southern Paiute People, descendants of the Tudinu, or Desert People.

Western Kentucky University (WKU) honors and acknowledges the Indigenous peoples' land on which this University was built. All land in the state of Kentucky was once Indigenous territory, which is why it is our duty to acknowledge that WKU exists on Native land. The particular region of Kentucky wherein WKU sits was home to both the Shawnee (Shawandasse Tula) and Cherokee East (ᏣᎳᎫᏪᏘᏱ Tsalaguwetiyi) tribes.

The California Polytechnic State University in San Luis Obispo, California (Cal Poly) sits on the traditional lands of the Yak titʸu titʸu yak tiłhini Northern Chumash Tribe of San Luis Obispo County and Region. The Yak titʸu titʸu yak tiłhini have a documented presence for over 10,000 years. The Tiłhini Peoples have stewarded their ancestral and unceded homelands which

include all of the cities, communities, federal and state open spaces within the San Luis Obispo County region. These homelands extend East into the Carrizo Plains toward Kern County, South to the Santa Maria River, North to Ragged Point, and West beyond the ocean's shoreline in an unbroken chain of lineage, kinship, and culture.

## References

1. Zhang Y, Rau P-LP. Playing with multiple wearable devices: Exploring the influence of display, motion and gender. *Comp Hum Behav*. 2015;50:148–158. https://doi.org/10.1016/j.chb.2015.04.004

2. Elgan M. Will women dominate the wearable computing market? *Cult of Android*. 2013.

3. Stoet G, O'Connor DB, Conner M, Laws KR. Are women better than men at multi-tasking? *BMC Psychol*. 2013;1:1–10. https://doi.org/10.1186/2050-7283-1-18

4. Cifor M, Garcia P. Gendered by design: A duoethnographic study of personal fitness tracking systems. *ACM Trans Soc Comput*. 2020;2(4):1–22. https://doi.org/10.1145/3364685

5. Klenk S, Reifegerste D, Renatus R. Gender differences in gratifications from fitness app use and implications for health interventions. *Mob Media Commun*. 2017;5(2):178–193. https://doi.org/10.1177/2050157917691557

6. Lluch J, Abad F, Calduch-Losa Á, Rebollo M, Juan M-C. Gender differences and trends in the use of wearables in marathons. *Sustain Technol Entrepr*. 2024;3(3):100063. https://doi.org/10.1016/j.stae.2023.100063

7. Navalta JW, Davis DW, Thomas DJ, Stone JW. Editorial: A step-by-step statistical decision framework for a gender-inclusive approach in sport and exercise science Research. *Int J Exerc Sci*. 2025;18(1):1010–1029. https://doi.org/10.70252/ITKQ9186

8. Navalta JW, Stone WJ, Lyons TS. Ethical issues relating to scientific discovery in exercise science. *Int J Exerc Sci*. 2019;12(1):1-8. https://doi.org/10.70252/EYCD6235

9. All of Us Research Program Investigators. The "All of Us" research program. *N Engl J Med*. 2019;381(7):668–676. https://doi.org/10.1056/NEJMsr1809937

10. Fitbit Help Center. How does my Fitbit device calculate my daily activity? Fitbit. Accessed February 7, 2025, 2025. https://support.google.com/fitbit/answer/14237111?hl=en#zippy=%2Chow-does-my-fitbit-device-calculate-calories-burned

11. Kirk RE. Practical significance: A concept whose time has come. *Educ Psychol Meas*. 1996;56(5):746–759. https://doi.org/10.1177/0013164496056005002

12. Dolgun A, Demirhan H. Performance of nonparametric multiple comparison tests under heteroscedasticity, dependency, and skewed error distribution. *Commun Stat Simul Comput*. 2017;46(7):5166–5183. https://doi.org/10.1080/03610918.2016.1146761

13. Judge L, Livergood K, Smith A, Razon S. Advancing diversity, equity, and inclusion in kinesiology departments and allied health professions: Barriers, facilitators, and effective strategies. *J Equity Soc Justice Educ*. 2024;3:1–18. https://doi.org/10.62889/2024/jkas1127

14. Alshuwaiyer G. Differences in obesity pattern, physical activity level, and dietary habits among LGBTQ Individuals: A systematic review. *Int J Health Well Soc*. 2021;11(2):209. https://doi.org/10.18848/2156-8960/CGP/v11i02/209-239

15. Bishop A, Overcash F, McGuire J, Reicks M. Diet and physical activity behaviors among adolescent transgender students: school survey results. *J Adolesc Health*. 2020;66(4):484–490. https://doi.org/10.1016/j.jadohealth.2019.10.026

16. Aparicio-García ME, Díaz-Ramiro EM, Rubio-Valdehita S, López-Núñez MI, García-Nieto I. Health and well-being of cisgender, transgender and non-binary young people. *Int J Environ Res Public Health*. 2018;15(10):2133. https://doi.org/10.3390/ijerph15102133

17. Kosciw JG, Clark CM, Truong NL, Zongrone AD. *The 2019 National School Climate Survey: The Experiences of Lesbian, Gay, Bisexual, Transgender, and Queer Youth in Our Nation's Schools. A Report from GLSEN*. ERIC; 2020.

18. Thomas JD, Cardinal BJ. How credible is online physical activity advice? The accuracy of free adult educational materials. *Trans J Am Coll Sports Med*. 2020;5(9):82–91. https://doi.org/10.1249/TJX.0000000000000122

19. Heidari S, Babor TF, Castro DP, Tort S, Curno M. Sex and Gender Equity in Research: rationale for the SAGER guidelines and recommended use. *Res Integr Peer Rev*. 2016;1(1). https://doi.org/10.1186/s41073-016-0007-6

Corresponding author: James Navalta; james.navalta@unlv.edu