OPEN ACCESS

Centering the female athlete voice in a sports science research agenda: a modified Delphi survey with Team USA athletes

Julie McCleery ,[1] Ellie Diamond ,[2] Rose Kelly ,[3] Longxi Li,[1] Kathryn E Ackerman ,[3,4] William M Adams ,[5,6] Emily Kraus [2]

► Additional supplemental material is published online only. To view, please visit the journal online (https://doi.org/10.1136/bjsports-2023-107886).

[1]Center for Leadership in Athletics, University of Washington College of Education, Seattle, Washington, USA
[2]Female Athlete Science and Translational Research Program, Stanford University Department of Orthopaedic Surgery, Redwood City, California, USA
[3]Wu Tsai Female Athlete Program, Division of Sports Medicine, Boston Children's Hospital, Boston, Massachusetts, USA
[4]Harvard Medical School, Boston, Massachusetts, USA
[5]Department of Sports Medicine, United States Olympic & Paralympic Committee, Colorado Springs, Colorado, USA
[6]United States Coalition for the Prevention of Illness and Injury in Sport, Colorado Springs, Colorado, USA

Correspondence to
Dr Emily Kraus, Female Athlete Science and Translational Research Program, Stanford University Department of Orthopaedic Surgery, Redwood City, CA 94063, USA; ekraus@stanford.edu

Accepted 6 June 2024

To cite: McCleery J, Diamond E, Kelly R, *et al*. *Br J Sports Med* Epub ahead of print: [*please include* Day Month Year]. doi:10.1136/bjsports-2023-107886

ABSTRACT

Objectives To co-construct a sports medicine and exercise science research and translational agenda with Team USA elite female athletes serving as the experts on their health, performance and well-being.

Methods 40 Team USA female athletes across sports disciplines participated in an online, anonymous, modified Delphi survey by ranking topics on a Likert scale (1='strongly disagree' and 5='strongly agree') and providing qualitative justification regarding whether they believed having more information and research on each topic would support their athletic performance, health and well-being. After each Delphi round, quantitative rankings of topics and qualitative justifications were analysed, informing revisions to the list of topics for review in the subsequent round. Researchers provided athletes with a detailed report of findings and revisions following each round.

Results The final list contained 14 ranked topics. The top five were menstrual cycle symptoms (4.58±0.74), recovery (4.58±0.59), birth control (4.55±0.89), mental health (4.50±0.55) and fueling and the menstrual cycle (4.43±0.74). New topics originating from athletes included recovery, menstrual cycle symptoms, fueling and the menstrual cycle, mental health and sports performance, team dynamics, and institutionalised sexism.

Conclusion This is the first study to co-construct a research and translational agenda with Team USA elite female athletes. The list of sports science research topics developed by focusing on elite female athletes' voices lays the foundation for future research and provides valuable insight into the specific needs of female athletes.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Despite decades of increased female participation in sports, there remains a gap in sports medicine and exercise science research studying female athletes. Additionally, female athletes are rarely regarded as experts nor are their voices centred when determining research agendas intended to elevate female athlete performance and well-being.

WHAT THIS STUDY ADDS

⇒ This is the first study to co-construct a research and translational agenda with elite female athletes. Three rounds of a modified Delphi process with US Olympic and Paralympic Committee-affiliated female athletes produced a research agenda consisting of 14 diverse topics, including menstrual cycle symptoms, recovery, effects of birth control, institutionalised sexism and male-dominated systems, and mental health.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ By focusing on athletes' perspectives, this research contributes valuable and novel insights to inform future research questions and studies, policy changes and practice improvements that address a consensus of athletes' stated needs and knowledge interests, thereby directing better health, performance and well-being support for female athletes.

BACKGROUND

Despite the increased participation of women and girls in sport in the USA, including anticipated gender parity in the 2024 Olympic and Paralympic Games,[1] the inclusion of women and girls in sports medicine and exercise science (SMES) research remains poor. Of articles published in six leading SMES journals from 2014 to 2020, only 6% exclusively studied women.[2] This lack of research hinders the development of evidence-based, effective strategies to support the health, well-being and performance of female athletes to reach their full potential. Furthermore, research has identified sex-specific differences in injury rates and time lost from sports.[3–13] For example, previous literature has shown that female athletes experience higher rates of bone stress injuries[3 4] and anterior cruciate ligament (ACL) ruptures,[5–10] longer duration of symptoms after concussion[11 12] and a greater risk of disordered eating or eating disorders [13] than their male counterparts.

Recently, in response to these long-standing inequities, SMES researchers have been investigating sports ecosystems with an interdisciplinary, gendered lens. Parsons and colleagues[14] discussed the 'curious absence of gender as an influencer in the dialogue surrounding ACL injuries' and Thorpe *et al*[15] reviewed the interplay of gender, sport and health, and opportunities for multidisciplinary and intersectional advancements. Furthermore, female athletes, including those from USA Wheelchair Basketball,[16] the National Women's Soccer League,[17 18] USA Gymnastics[19] and USA Track and Field,[20–22] have raised such issues, highlighting

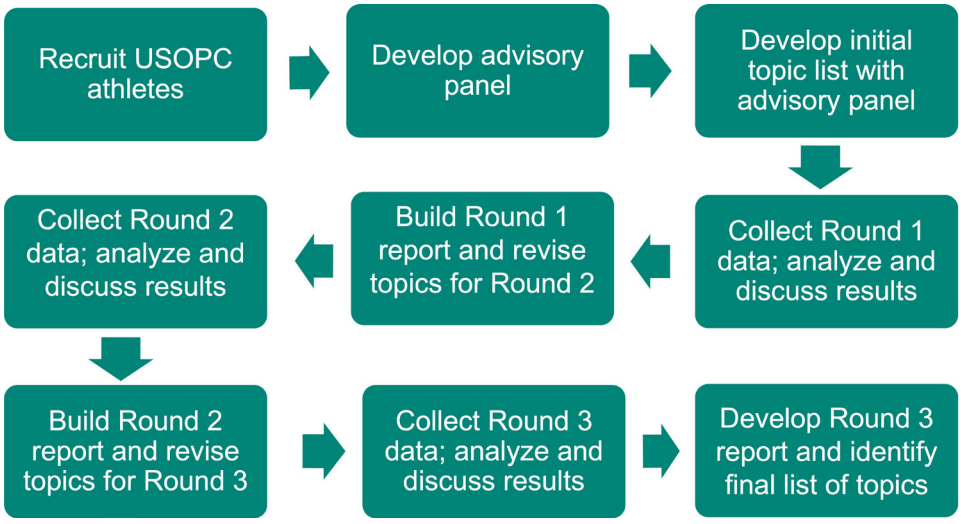

Figure 1 Delphi survey methodology. USOPC, US Olympic and Paralympic Committee.

questions and concerns ranging from coach and medical misconduct to the omission of pregnancy clauses in professional contracts. Female athletes' voices have reverberated across the sports world. Yet, even with the growing body of allegations, impactful, lasting change has remained elusive.

Objective

Female athletes are rarely centred and regarded as experts when determining topics of paramount importance to their performance and well-being. Therefore, the overall aim of this study was to centre the voices of Team USA elite female athletes and co-construct an SMES research and translational agenda. The findings could then guide athlete-informed research and translational work for scientists, medical providers, elite sports teams, governing bodies and educators.

METHODS

Study design: Delphi method

A three-round, internet-based, modified Delphi survey was conducted from September 2022 to January 2023 to build consensus on a research and translation agenda for Team USA elite female athletes. The Delphi method collects and distills knowledge from an expert group using rounds of questionnaires that include write-in responses to generate consensus about a specific topic.[23] The Delphi Method's key features—anonymity, iteration, controlled feedback, and statistical aggregation of responses—make it a strong method for developing a research agenda and translational practices informed by elite female athletes (figure 1).

Participants and recruitment

The primary criterion for Delphi panel participation is expertise on the study topic.[24–26] Experts were defined as elite athletes competing in the women's division of their sports, >18 years old, who had competed as part of Team USA at a sanctioned international competition within the previous 5 years with the intention to continue competing. We, the research team, worked with sports representatives from the Athletes' Advisory Council (AAC) of the US Olympic and Paralympic Committee (USOPC) to recruit participants. We asked each representative to contact their sports athletes via email and invite them to contact us to express their interest in participating. Assuming 15%–20% attrition and targeting a sample of 40–45 participants, we invited a purposive sample of 50 to participate, representing various team/individual sports, summer/winter sports, Olympic/Paralympic sports and races/ethnicities. The study protocol was approved by the Institutional Review Boards at Boston Children's Hospital (approval #IRBP00041163), Stanford University (approval #65626) and the University of Washington (approval #STUDY00015242). All participants provided informed consent.

Advisory panel

We also convened an advisory panel of experts with significant research and application experience involving female athletes.[27] The six-person panel included physicians, dietitians, sports psychologists, mental health providers, physical therapists, coaches, physiologists, SMES researchers and former athletes. To enhance the trustworthiness of qualitative analysis and scientific validity and serve as a check on researcher bias, this group provided insights and feedback during initial Delphi topic development and during analyses and modifications between Delphi rounds.[24 28 29]

Equity, diversity and inclusion statement

We prioritised recruiting a diverse group of participants from different racial and ethnic backgrounds, team and individual sports, winter and summer sports and Olympic and Paralympic sports. Recruitment efforts aimed to enroll a purposive sample of athletes, regarded as experts, representing the above categories such that they reflected, as closely as feasible, the currently available data on the composition of the US Olympic and Paralympic teams. The study team worked with the USOPC AAC to recruit specific populations in order to develop a sample reflective of the Team USA delegation. To support inclusion and limit attrition, a $500 e-gift card was provided to participants who completed all survey rounds. Furthermore, the study team sent multiple, individualised email reminders as well as deadline extensions to increase the opportunity to involve as many enrolled participants as possible. The lead research team was 100% white, able-bodied and female. The advisory panel was 90% female and included experts in qualitative and quantitative research; training in athletic training, nutrition, psychology, sports medicine, physical medicine and rehabilitation and endocrinology and experience as athletes and support staff in Olympic and Paralympic sport.

Definitions

Our definition of 'female athlete' as it pertains to this study's inclusion criteria and how it is referenced throughout this paper is anyone competing in the women's division of their sport. Additionally, every time the term 'female physiology' or 'female specific' is used throughout this paper or in reference to the study, we are referring to how hormonal fluctuations, anatomy or the environment specific to the female athlete impacts the body systems.

Data collection and measures: the modified Delphi method

Using a biopsychological framework, we and the advisory panel developed an initial list of 17 research topics based on existing literature and expert opinion (see online supplemental file A). We created all three Delphi surveys using the research electronic data capture (REDCap),[30 31] a secure, web-based database hosted by Boston Children's Hospital. Before taking the first round of the Delphi survey, athletes used REDCap to consent to participate and completed a questionnaire to collect data on demographics and sports history.

For all three Delphi surveys, athletes were asked to indicate, using a Likert scale, the extent to which they agreed or disagreed with the prompt, 'Having more research and knowledge about this topic is important to my athletic performance, health, and overall well-being.' After rating the topic (1 indicating strongly disagree and 5 indicating strongly agree), athletes provided written justification for their ratings. In Rounds One and Two, athletes provided justifications for their ratings, suggested additional topics, reported their primary source of information about female physiology, health, performance and well-being and reported if they believed the information was sufficient and accurate.

Incomplete surveys were excluded from analysis and participants who did not complete a round were not involved in subsequent rounds.

Data analysis

The quantitative ratings, qualitative justifications and new topic suggestions were analysed concurrently to inform one another, build topic consensus, increase mean scores, reduce variability and achieve rating stability.[24 28 32 33] The final list of topics each had participant ratings >4 and an SD ≤1. After each round, descriptive statistics, including mean, SD, mode and median, were calculated to assess athletes' perceptions of topics. To further explore variance between participant groups, mean comparisons were performed across seven pre-established binary groups (team/individual sport, summer/winter sport, Olympic/Paralympic sport, athlete of colour/non-athlete of colour, ever pregnant/never pregnant, ever had mental health diagnosis/never had mental health diagnosis and ever injured/never injured). Prior to inferential analysis, data were assessed for basic assumptions, adherence and outliers (z-scores exceeding±3.29), although there were none. The non-parametric independent group Mann-Whitney U test was employed to compare means across groups.[34] An alpha level of 0.05 was used, and all statistical analyses were conducted using IBM SPSS Statistics software (V.27).[35]

For the qualitative analysis of each round, two members of the research team independently coded the justifications using thematic analysis[26] and then compared results and analytic memos. Researchers also coded and thematically grouped participants' suggested new topics.[26] These thematic analyses were then used to inform (1) revisions (or exclusions) of lower-scoring topics; (2) revisions of topics for which the mean comparison analysis revealed significant between-group differences and (3) new topics for consideration in subsequent rounds. All qualitative data were coded using Nvivo.[36]

To support the iterative, consensus-building process and enhance trustworthiness and credibility foundational to the Delphi method, we compiled the quantitative and qualitative findings into a comprehensive report with guidance and feedback from the advisory panel on topic modification to ensure accuracy and serve as a check on researcher bias.[26] The report, which included statistical analyses and illustrative athlete quotes to explain why topics were revised, added or removed, was then shared with participants to read prior to the next round.[24 28 29] Athletes then had the opportunity to provide comments on the report, which were also thematically analysed. See online supplemental file F.

RESULTS

58 athletes expressed interest in participating in the study and from this group, 50 participants were invited to the initial survey. Eight interested athletes whose sports were already represented in the sample were placed on a waitlist; no one was invited from the waitlist to participate in the survey due to the high response rate in Round One, and these waitlist participants were informed when the waitlist closed. 45 participants began the study (completing the non-Delphi survey); 43 completed Round One, and 40 completed Rounds Two and Three (overall attrition of ~7% (3/43)). Among the three participants who were lost to attrition, two were Paralympians. Of the 40 participants who completed all three rounds, most were white (80%), Olympians (82.5%) competing in individual (60%), summer sports (73%) and representing 24 unique sports (see table 1). Participants' average age was 29±4 years, with an average of 9±5 years of international experience. The majority had never been pregnant (85%), had experienced an injury (82.5%) and had never received a mental health diagnosis (52.5%).

Round One

In Round One, participants provided ratings and ~23 000 words of justifications for 17 initial topics. Four of the Round One topics had a mean score of ≥4.4/5. The highest rated was *mental health tools* (4.53±0.77). Five topics were rated below 4, and three of them were subsequently removed: *acute and systemic illness* (3.42±1.03), *performance testing* (3.47±1.03) and *environmental variables* (3.72±0.93). Athlete comments on removed items often reflected either a lack of relevance to their sports context or sufficient information about a topic. See figure 2.

Two lower-rated topics were rewritten. *Low energy availability and relative energy deficiency in sport* (REDs) (3.86±0.94) was rewritten based on participant feedback suggesting terminology confusion. For example, athletes commented that they had never heard of REDs or low energy availability, so we changed the topic to include a definition of low energy availability. *Equipment needs* (3.79±1.08), although rated 14th of the 17 topics, was also rewritten because many new topic suggestions focused on equipment needs. Two higher-ranking topics were also rewritten based on feedback. For example, *birth control* ranked fourth with a mean score of 4.40±0.85, but new topic suggestions illustrated a specific interest in the short-term and long-term impacts of birth control. Six original topics were unchanged for Round Two (see online supplemental file A).

Table 1 Descriptive statistics of participant characteristics

Characteristics	n	%
Race		
White	32	80.00
Black or African American	5	12.50
Asian	1	2.50
Multiracial	1	2.50
Prefer not to respond	1	2.50
Ethnicity		
Mexican, Mexican American and Chicano/a/x	2	5.00
Another Hispanic, Latina/o/x or Spanish origin	1	2.50
Not of Hispanic, Latina/o/x or Spanish origin	36	90.00
Prefer not to say	1	2.50
Type of sport		
Individual	24	60.00
Team	15	37.50
Prefer not to respond	1	2.50
Season		
Summer	29	72.50
Winter	10	25.00
Prefer not to respond	1	2.50
Type of athlete		
Olympian	33	82.50
Paralympian	6	15.00
Prefer not to respond	1	2.50
Pregnancy		
Never have been pregnant	34	85.00
Have been pregnant (given birth/miscarriage)	6	15.00
Injury		
Ever injured	30	75.00
Never injured	10	25.00
Mental health		
Never received mental health diagnosis	21	52.50
Ever received mental health diagnosis	19	47.50

Note: N=40 female athletes; race, type of sport and athlete, season, pregnancy, injury and mental health are reported in group size (n) and percentage (%). Two and three subjects were removed from Delphi Rounds one and two, respectively, due to incomplete data.

There were significant differences between groups for nine topics. For example, *periods and menstrual cycle* (4.42±1.10) had a significant mean difference between individual and team sport athletes (4.72±0.74 vs 3.94±1.39, respectively, p=0.024) and athletes who never received a mental health diagnosis and athletes who had (4.17±1.24 vs. 4.74±0.81, respectively, p=0.035). Two of the nine topics were removed because of low ratings and justifications suggesting the topics were not a priority need, and two topics were rewritten. The topics with persistent mean differences between groups across the rounds are discussed in detail in the Round Three results (table 3).

Participants suggested 49 new topics, which were coded into approximately 25 themes and analysed in conjunction with existing topics to identify six novel topics for Round Two. 12 athletes mentioned the connection between nutrition and the menstrual cycle. Nine athletes explicitly expressed interest in understanding the impact of male-dominated spaces on women's physical and mental health and performance. Six athletes raised questions about how their interpersonal relationships with teammates and coaches might relate to their well-being. All six new participant-suggested topics were retained throughout the rest of the rounds.

Round Two

In Round Two, participants generated ~36 000 words of comments. The aggregate topic ratings improved slightly from Round One to Two with an increase in mean score and a decrease in SD (see online supplemental files B and C for a detailed topic list and analytical overview from Rounds One and Two). The newly added topic, *fueling and the menstrual cycle*, achieved the highest rating (4.65±0.57); the rewritten topic, *birth control*, achieved the second highest score (4.63±0.80). One more topic was removed for the next round: *equipment needs* (3.88±1.14). Despite being rewritten after a low rating in Round One, this topic ranked 19th of 19 topics (see figure 2 for illustrative athlete comments).

This round included 33 athlete suggestions. However, instead of adding any new topics, we used the feedback to rewrite several existing topics as well as combine topics. For example, suggestions from both Rounds One and Two included the importance of sleep, which was added to the *recovery* (4.40±0.66) topic. The mean comparison results revealed significant group differences for six topics. Some topics had significant differences for multiple groups. For example, the topic of *injury management* (4.15±0.91) had a significant mean difference between team and individual sport athletes (3.60±0.91 vs. 4.50±0.78, respectively, p=0.003) and between Paralympic and Olympic athletes (3.00±1.27 vs. 4.36±0.70, respectively, p=0.014). Of the six topics with differences, one was removed, three were combined with other topics, and one was rewritten.

Round Three

In Round Three, participants rated 14 topics and provided ~28 000 words of justification. From Round Two to Three, the average mean score increased and the average SD decreased. The highest-rated topics were *recovery* and *menstrual cycle symptoms* (4.58±0.59 and 4.58±0.74, respectively). All topics were rated over 4; the lowest-rated topic was *strength training* (4.05±0.84). All topics were subsequently reviewed and revised, and none were eliminated. See table 2 for a list of final topics.

The mean comparison results revealed significant differences between groups for seven topics. Only three topics retained significant differences between groups across multiple rounds (see online supplemental file D). In Round One, *pregnancy and postpartum* (4.12±1.20) had a significant mean difference between athletes who were ever versus never pregnant (5±0 vs 3.94±1.24, respectively, p=0.012) and again in Round Three (5±0 vs 3.97±1.03, respectively, p=0.007). Athletes who ranked *pregnancy and postpartum* low primarily reported it was due to their not intending to become pregnant or return to sport postpregnancy. However, some participants noted that while the topic may not impact them directly, they understood the value for their teammates and other athletes (see figure 2).

In Rounds Two and Three, *institutionalised sexism* (Round Three: 4.13±0.98) had a significant difference between Olympic and Paralympic athletes (Round Two: 4.18±0.92 vs 5±0, respectively, p=0.035; Round Three: 3.97±1.02 vs 5.±0, respectively, p=0.012). In Rounds One and Three, summer sport athletes rated *mental health* (Round Three: 4.50±0.55) significantly higher than winter sport athletes (Round One: 4.68±0.65 vs 4.09±0.94, respectively, p=0.039; Round Three: 4.66±vs 4.10±0.57, respectively, p=0.009; see table 3).

In the final round, participants provided comments about the Delphi process. Overall, athletes shared gratitude for their voices and female-specific needs guiding future research. Many shared that they are often treated like lesser versions of men and how

Figure 2 Illustrative quotes from athletes justifying topic selection.

this study could help change that by 'focusing on our experience instead of catering to the existing structures.' Athletes expressed appreciation for and interest in learning from other athletes, stating that reading the findings after each round broadened their understanding of female athlete experiences and needs beyond their own. One said, this study 'reinforces the merits of a collaborative approach because of how many valid inputs people had that differed vastly from mine. I read the report alongside my own answers and saw obvious gaps or points I missed'. They also expressed feeling less alone in their concerns. Many were intrigued by the group mean differences: 'I loved reading the two reports we received and seeing the response commonalities and variations across identified groups. I finished both reports with more curiosity and a desire to explore further topics in greater detail'. Though most feedback on the process was positive and illuminated the benefits of the participant experience, negative feedback included comments about redundancy and length.

DISCUSSION

The aim of this study was to co-construct an SMES research and translational agenda with Team USA elite female athletes. Centering athlete voice in SMES research is a relatively new practice but calls for including athletes directly in research and policy are increasing.[37–39] As female athlete maltreatment is more widely acknowledged and protecting the health and well-being of athletes becomes an increasing priority,[16–22 40–48] it is important to understand female athletes' specific needs from athletes themselves.[38] A broad range of consensus statements[49–59] aim to orient the clinical and athlete support communities around the interventions, research directions and practical tools necessary for athlete health and well-being. However, the athletes' stated needs and interests relative to those interventions and research are typically lacking. As such, this study positioned athletes as experts on their bodies, experiences and needs to advance SMES research and translational agenda tailored to them.

The participants' novel recommendations led to the creation of 6 of the 14 final topics. The original list of SMES topics came from a biopsychological framework and did not include topics related to the social environment's impact on athlete health. However, athletes expanded the scope of interest by raising topics related to their training and competitive environments, as well as their interpersonal relationships. The inclusion of these topics resonated with the majority of participants (see figure 2 for an illustrative comment). This aligns with research about the effects of elite athletes' larger social environment on their mental and physical well-being[58] and with recent calls to investigate how the social environment may contribute to ACL injuries among elite footballers.[60 61]

Each topic from the final list (see table 2) warrants further investigation to understand ways in which researchers and practitioners can enhance knowledge and translation. According to the advisory panel, the topics vary as to whether they lack foundational research or whether there is sufficient research but a lack of translation of the well-substantiated research into practice. Furthermore, the consensus-driven nature of a Delphi doesn't account for individual variations or highly sports-specific needs. Thus, some excluded topics may be salient to specific female athlete populations. For example, *equipment needs* ranked 14/17 in Round One (3.79±1.08) and 19/19 in Round Two (3.88±1.14), with significant mean differences between summer and winter athletes and team and individual athletes. But those who rated it highly did so because the equipment in their sport is designed for men's bodies (see online supplemental file E for all removed topics).

Table 2 Topics and agreement ratings from Round 3

Topic name	Topic description	Agreement rating* (mean±SD)
Menstrual cycle symptoms	How the symptoms before, during and after the bleeding phase of a menstrual cycle impact training and performance.	4.58±0.74
Recovery	The impact of female physiology, including menstrual cycle, on the recovery process and the effectiveness of various recovery modalities and approaches, including passive recovery and sleep.	4.58±0.59
Birth control	The short-term and long-term effects of specific types of birth control options, including different oral hormonal contraceptive pills, intrauterine devices, the patch, etc on health, performance and well-being.	4.55±0.89
Mental health	Tools, strategies and interventions that support and strengthen mental skills, training and performance for female athletes with or without potential mental health diagnoses (eg, anxiety, depression, eating disorder, obsessive compulsive disorder and post-traumatic stress disorder).	4.50±0.55
Fueling and the menstrual cycle	How best to fuel at different stages of the menstrual cycle to enhance performance, health and well-being.	4.43±0.74
Fueling and hydration	Sport-specific fueling, supplementation and hydration strategies (timing, macro/micronutrients) while travelling and during different stages of training (in competition, out of competition) and competition (precompetition, during and postcompetition).	4.43±0.67
Physiology and performance	Strategies for optimising training and using sports performance testing that incorporates female physiology to attain peak physical performance over the course of a season and athletic career.	4.33±0.72
Mental health and sports performance	Understanding similarities and differences between sports performance support and mental health support and how, when and where to seek care for each.	4.30±0.84
Team dynamics	The impact of team dynamics and interpersonal relationships between and among athletes, coaches and support staff on performance, physical and mental health and well-being.	4.25±0.86
Low energy availability and REDs	The symptoms and potential health and performance consequences of low energy availability, which is when the body does not have enough calorie intake to account for exercise and the caloric needs of daily life. (This can be inadvertent from an eating disorder or disordered eating, aka REDs or female athlete triad.)	4.15±0.82
Institutionalised sexism	The impacts of institutionalised sexism and/or male-dominated coaching and support staff on female athlete mental health and performance.	4.13±0.98
Pregnancy and postpartum	Training, performance, physical health and mental health considerations during pregnancy and postpartum return to sport.	4.13±1.00
Injury management	Effective strategies for female athletes to decrease injury and reinjury risk and safely return to play postinjury (including those related to fueling, physical and mental training and rehabilitation).	4.13±0.75
Strength training	Female-specific and individualised strength training considerations to optimise performance and reduce injury risk in female athletes.	4.05±0.84

*Agreement rating on a Likert scale, with 1 indicating strongly disagree and 5 indicating strongly agree.
REDs, relative energy deficiency in sport.

Limitations

Methodologically, the Delphi creates the chance for non-response error due to its time commitment,[32 62] affecting the response rate. In this study, 40 of the initial 43 participants completed all three rounds. To minimise attrition, we offered financial compensation, multiple methods of contact and assistance from the Athlete Advisory Council. Two of the three study participants lost to attrition were Paralympians, decreasing paralympic representation by 25% (8–6 total athletes). Additionally, the recruitment methods led to some selection bias. Participants were predominantly white (80%), which does not reflect the racial makeup of the US elite female athlete population. Approximately 40% of USOPC athletes are athletes of colour.[63] Furthermore, even though the inclusion criteria allowed for transgender and non-binary athletes, none identified as such, highlighting another gap in our sample. While the sample size is appropriate for a Delphi study, the small sample limits transferability and generalisability beyond this population. Because of the importance of athlete context and experience in a study of this nature, there may be limited applicability to athletes outside of the USA. Additionally, the online-only interaction potentially reinforces difficulties in understanding participants' or the groups' responses and

Table 3 Agreement rating differences between groups that persisted through Delphi rounds

	Round One		Round Two		Round Three	
Topic	Mean±SD	Sig.	Mean±SD	Sig.	Mean±SD	Sig.
Institutionalised sexism						
Olympic athletes	NA		4.18±0.92	0.035*	3.97±1.02	0.012*
Paralympic athletes	NA		5.00±0.00		5.00±0.00	
Mental health tools						
Summer sport athletes	4.68±0.65	0.039*	4.59±0.73	0.332	4.66±0.48	0.009**
Winter sport athletes	4.09±0.94		4.30±0.83		4.10±0.57	
Pregnancy and postpartum						
Ever been pregnant	5.00±0.00	0.012*	5.00±0.00	0.558	5.00±0.00	0.007**
Never been pregnant	3.94±1.24		4.18±1.14		3.97±1.03	

*$p < 0.05$ and **$p < 0.01$.

collaborative exchange.[26] Researchers using Delphi must carefully craft each iteration of the survey to scaffold for maximum participant understanding, and as with any empirical method, researchers must account for how their biases and perspectives may influence the question formation and analysis of data.[26 33 62] The advisory panel and the reports to participants served as safeguards against researcher bias and provided guidance on each iteration of the topics. Lastly, another limitation of the current study relates to the statistical methodology employed in analysing multiple comparisons across subgroups, which inherently carries the risk of inflating the Type I error rate. We opted for methods that balance Type I error control with the ability to uncover meaningful insights through the integration of qualitative methods. Although this approach does carry a potential for Type I error, we believe it is justified given the study's scope. Nonetheless, the findings should be interpreted with an understanding of these statistical considerations.

Future directions

This study's findings provide direction for efforts to improve the health and well-being of Team USA female athletes on a variety of fronts, including research, funding, translation, coaching, education and clinical care. The final list of topics could help shape priorities for the collaborative design, funding and execution of research in the burgeoning space of female athlete SMES research internationally. Noting that the salience of topics shifts based on the individual needs of athletes and their contexts, practitioners on the front lines with athletes, coaches, trainers, dietitians and physicians, can use the list as a starting point to educate themselves on topics of interest and guide interactions with athletes in their care. Researchers in other countries may also use this methodological approach as a springboard to inform the incorporation of female athlete voices into their SMES. This research offers an opportunity to consider the ways athletes can continue to shape and inform future research questions and study design.

CONCLUSION

The insights provided by Team USA athletes through the modified Delphi survey underscore the importance of athlete-centred perspectives in SMES research. By incorporating their qualitative feedback, future studies, policy initiatives and education can better align with athletes' needs. The prioritisation of menstrual cycle-related topics and the recognition of team dynamics and sexism issues in sports underscore the potential for progress when athletes are at the forefront of health and performance discussions. Ultimately, this research illustrates how athlete voice can drive meaningful progress in women's health and sports performance research, translation and policy development.

X Kathryn E Ackerman @drkateackerman, William M Adams @william_m_adams and Emily Kraus @emilykrausmd

Acknowledgements We first thank the forty Team USA athletes who participated in this Delphi survey and contributed to the first female athlete research and translational agenda. We would also like to thank the expert panel for developing the initial list of sports science topics: KEA, WMA, Cheri Blauwet, Alicia Glass, Lindsay Golich, EK, JMC, Kim O'Brien, Aretha Thurmond and Gale Wetzler.

Contributors JC and EK proposed the idea of a Delphi consensus study using athlete voice. RK and ED helped with IRB protocol development. WMA helped in USOPC communications and research study approval through the USOPC. JC and ED performed the qualitative analysis using Nvivo. RK and ED managed athlete communications and survey creation and distribution. JC, EK, ED and RK reviewed and revised topics between each survey round. WMA and KEA reviewed manuscript drafts and provided feedback and revisions. LL performed the quantitative analysis using SPSS. All authors critically revised the manuscript for important intellectual content and approved the final version of the manuscript. EK is the guarantor for the manuscript and accepts full responsibility for the work and/or the conduct of the study, had access to the data and controlled the decision to publish.

Funding We thank the Voice in Sport Foundation and the Wu Tsai Human Performance Alliance for their support in funding this study.

Competing interests KEA is a Deputy Editor of the BJSM and an Associate Editor of the BJSM IPHP.

Patient and public involvement Patients and/or the public were involved in the design, or conduct, or reporting, or dissemination plans of this research. Refer to the Methods section for further details.

Patient consent for publication Not applicable.

Ethics approval This study involves human participants and was approved by the Boston Children's Hospital Institutional Review Board (IRB): #IRBP00041163, the Stanford University IRB (approval #65626) and the University of Washington IRB (approval #STUDY00015242). Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request. All data relevant to the study are included in the article or uploaded as supplementary information. None of the data presented in the article or in the supplemental material includes patient-identifiable information. Additional de-identified participant data is available upon reasonable request from EK.

ORCID iDs
Julie McCleery http://orcid.org/0000-0003-0737-7612
Ellie Diamond http://orcid.org/0009-0004-0273-7687
Rose Kelly http://orcid.org/0009-0004-9983-4255
Kathryn E Ackerman http://orcid.org/0000-0003-2626-7785
William M Adams http://orcid.org/0000-0001-7372-8455
Emily Kraus http://orcid.org/0000-0001-5544-2939

REFERENCES

1 #Genderequalolympics: celebrating full gender parity on the field of play at Paris 2024 [press release]. International Olympic Committee; 2024.
2 Cowley ES, Olenick AA, McNulty KL, *et al*. "Invisible sportswomen": the sex data gap in sport and exercise science research. *Women Sport Phys Act J* 2021;29:146–51.
3 Ohta-Fukushima M, Mutoh Y, Takasugi S, *et al*. Characteristics of stress fractures in young athletes under 20 years. *J Sports Med Phys Fitness* 2002;42:198–206.
4 Wentz L, Liu PY, Haymes E, *et al*. Females have a greater incidence of stress fractures than males in both military and athletic populations: a systemic review. *Mil Med* 2011;176:420–30.
5 Arendt E, Dick R. Knee injury patterns among men and women in collegiate basketball and Soccer. NCAA data and review of literature. *Am J Sports Med* 1995;23:694–701.
6 Arendt EA, Agel J, Dick R. Anterior cruciate ligament injury patterns among collegiate men and women. *J Athl Train* 1999;34:86–92.
7 Gwinn DE, Wilckens JH, McDevitt ER, *et al*. The relative incidence of anterior cruciate ligament injury in men and women at the United States naval academy. *Am J Sports Med* 2000;28:98–102.
8 Agel J, Arendt EA, Bershadsky B. Anterior cruciate ligament injury in national collegiate athletic association basketball and soccer: a 13-year review. *Am J Sports Med* 2005;33:524–30.
9 Hootman JM, Dick R, Agel J. Epidemiology of collegiate injuries for 15 sports: summary and recommendations for injury prevention initiatives. *J Athl Train* 2007;42:311–9.
10 Mall NA, Chalmers PN, Moric M, *et al*. Incidence and trends of anterior cruciate ligament reconstruction in the United States. *Am J Sports Med* 2014;42:2363–70.
11 Lincoln AE, Caswell SV, Almquist JL, *et al*. Trends in concussion incidence in high school sports: a prospective 11-year study. *Am J Sports Med* 2011;39:958–63.

12 Covassin T, Moran R, Elbin RJ. Sex differences in reported concussion injury rates and time loss from participation: an update of the National collegiate athletic association injury surveillance program from 2004-2005 through 2008-2009. *J Athl Train* 2016;51:189–94.
13 Sundgot-Borgen J, Torstveit MK. Prevalence of eating disorders in elite athletes is higher than in the general population. *Clin J Sport Med* 2004;14:25–32.
14 Parsons JL, Coen SE, Bekker S. Anterior cruciate ligament injury: towards a gendered environmental approach. *Br J Sports Med* 2021;55:984–90.
15 Thorpe H, Bekker S, Fullagar S, *et al*. Advancing feminist innovation in sport studies: a transdisciplinary dialogue on gender, health and wellbeing. *Front Sports Act Living* 2022;4:1060851.
16 Azzi A. *U.S. wheelchair basketball coach resigns after players allege emotional misconduct*. NBC Sports On Her Turf, 2021.
17 Azzi A. *2021 NWSL timeline: five male coaches ousted due to misconduct, abuse allegations*. On Her Turf, 2021.
18 Yates SQ. Report of the independent investigation to the U.S. Soccer federation concerning allegations of abusive behavior and sexual misconduct in women's professional soccer. King & Spalding; 2022.
19 Macur J. *Biles and her teammates rip the F.B.I. for botching nassar abuse case*. The New York Times, 2021.
20 Montaño A, Cantor M, Jensen T, *et al*. *Nike told me to dream crazy, until I wanted a baby*. The New York Times, 2019.
21 Felix A, Crouse L, Jensen T, *et al*. *Allyson felix: my own nike pregnancy story*. The New York Times, 2019.
22 Cain M. *I was the fastest girl in America, until I joined nike*. The New York Times, 2019.
23 Yousuf MI. Using experts' opinions through Delphi technique. *Pract Assess Res Eval* 2007;12.
24 Hasson F, Keeney S, McKenna H. Research guidelines for the Delphi survey technique. *J Adv Nurs* 2000;32:1008–15.
25 Costa CA. The status and future of sport management: a Delphi study. *J Sport Manag* 2005;19:117–42.
26 Fogo B. Core practices for teaching history: the results of a Delphi panel survey. *Theory Res Soc Educ* 2014;42:151–96.
27 Trego LL, Middleton MCA, Centi S, *et al*. The military women's health Delphi study: a research agenda for the next decade. *Mil Med* 2023;188:32–40.
28 Hasson F, Keeney S. Enhancing rigour in the Delphi technique research. *Technol Forecast Soc Change* 2011;78:1695–704.
29 Brady SR. Utilizing and adapting the Delphi method for use in qualitative research. *Int J Qual Methods* 2015;14:160940691562138.
30 Harris PA, Taylor R, Thielke R, *et al*. Research electronic data capture (REDCap). A metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009;42:377–81.
31 Harris PA, Taylor R, Minor BL, *et al*. The REDCap consortium: building an international community of software partners. *J Biomed Inform* 2019;95:103208.
32 Dia Seyaki AK. Qualitative Delphi method: a four round process with a worked example. *Qual Rep* 2017;22:2755–63.
33 Clayton MJ. Delphi: a technique to harness expert opinion for critical decision-making tasks in education. *Educ Psychol* 1997;17:373–86.
34 Field A. *Discovering statistics using IBM SPSS*. Sage, 2013.
35 Corporation I. *IBM SPSS statistics for windows*. 27.0 Ed. Armonk, NY: IBM Corp, 2020.
36 Jackson K, Bazeley P. *Qualitative data analysis with NVivo*. 3rd Ed. Sage Publications, 2019.
37 Weissensteiner JR. The importance of listening: engaging and incorporating the athlete's voice in theory and practice. *Br J Sports Med* 2015;49:839–40.
38 Chamari K. The crucial role of elite athletes and expert coaches with academic profiles in developing sound sport science. *Int J Sports Physiol Perform* 2019;14:413.
39 Thornton JS, Richards D. Learning from 'lived expertise': engaging athletes and patients in sport and exercise medicine research and policy. *Br J Sports Med* 2023;57:189–90.
40 Stirling AE, Bridges EJ, Cruz EL, *et al*. Canadian academy of sport and exercise medicine position paper: abuse, harassment, and bullying in sport. *Clin J Sport Med* 2011;21:385–91.
41 Mountjoy M, Brackenridge C, Arrington M, *et al*. International olympic committee consensus statement: harassment and abuse (non-accidental violence) in sport. *Br J Sports Med* 2016;50:1019–29.
42 Wojtys EM. Pediatric tragedy. *Sports Health* 2017;9:400–1.
43 Mountjoy M. 'Only by speaking out can we create lasting change': what can we learn from the Dr Larry Nassar tragedy? *Br J Sports Med* 2019;53:57–60.
44 McCradden MD, Cusimano MD. Voices of survivors: 'you will not destroy our light' *Br J Sports Med* 2019;53:1384.
45 Ackerman KE, Stellingwerff T, Elliott-Sale KJ, *et al*. REDS (relative energy deficiency in sport): time for a revolution in sports culture and systems to improve athlete health and performance. *Br J Sports Med* 2020;54:369–70.
46 Mountjoy ML, Verhagen E. Bethechange': the responsibility of sports medicine in protecting athletes from Harassment and abuse in sport. *BMJ Open Sport Exerc Med* 2022;8:e001303.
47 Tuakli-Wosornu YA, Amick M, Guiora AN, *et al*. Athlete abuse hurts everyone: vicarious and secondary traumatic stress in sport. *Br J Sports Med* 2022;56:119–20.
48 Stirling AE, Taylor AR, Mountjoy ML, *et al*. Canadian academy of sport and exercise medicine position paper: the clinician's role in addressing and preventing maltreatment in Sport-10-year anniversary. *Clin J Sport Med* 2023;33:103–9.
49 Bø K, Artal R, Barakat R, *et al*. Exercise and pregnancy in recreational and elite athletes: 2016 evidence summary from the IOC expert group meeting, Lausanne. Part 1-exercise in women planning pregnancy and those who are pregnant. *Br J Sports Med* 2016;50:571–89.
50 Bø K, Artal R, Barakat R, *et al*. Exercise and pregnancy in recreational and elite athletes: 2016 evidence summary from the IOC expert group meeting, Lausanne. Part 2-the effect of exercise on the fetus, labour and birth. *Br J Sports Med* 2016;50:1297–305.
51 Bø K, Artal R, Barakat R, *et al*. Exercise and pregnancy in recreational and elite athletes: 2016/17 evidence summary from the IOC expert group meeting, Lausanne. Part 3-exercise in the postpartum period. *Br J Sports Med* 2017;51:1516–25.
52 Bø K, Artal R, Barakat R, *et al*. Exercise and pregnancy in recreational and elite athletes: 2016/17 evidence summary from the IOC expert group meeting, Lausanne. Part 4-recommendations for future research. *Br J Sports Med* 2017;51:1724–6.
53 Hainline B, Drezner J, Baggish A, *et al*. Interassociation consensus statement on cardiovascular care of college student-athletes. *Br J Sports Med* 2017;51:74–85.
54 Drezner JA, O'Connor FG, Harmon KG, *et al*. AMSSM position statement on cardiovascular preparticipation screening in athletes: current evidence, knowledge gaps, recommendations, and future directions. *Clin J Sport Med* 2016;26:347–61.
55 McCrory P, Meeuwisse W, Dvorak J, *et al*. Consensus statement on concussion in sport-the 5(Th) international conference on concussion in sport held in Berlin, October 2016. *Br J Sports Med* 2017;51:838–47.
56 Bø K, Artal R, Barakat R, *et al*. Exercise and pregnancy in recreational and elite athletes: 2016/2017 evidence summary from the IOC expert group meeting, Lausanne. Part 5. recommendations for health professionals and active women. *Br J Sports Med* 2018;52:1080–5.
57 Mountjoy M, Sundgot-Borgen JK, Burke LM, *et al*. IOC consensus statement on relative energy deficiency in sport (RED-S): 2018 update. *Br J Sports Med* 2018;52:687–97.
58 Reardon CL, Hainline B, Aron CM, *et al*. Mental health in elite athletes: International Olympic committee consensus statement (2019). *Br J Sports Med* 2019;53:667–99.
59 Rivara FP, Tennyson R, Mills B, *et al*. Consensus statement on sports-related concussions in youth sports using a modified Delphi approach. *JAMA Pediatr* 2020;174:79.
60 Nicholson J. Leah Williamson is latest Footballer to fall victim to a sport designed for men. Football365; 2023.
61 Harpur C. The women's world cup is near: how are clubs managing fear of injuries? The Athletic; 2023.
62 Murry JW, Hammons JO. Delphi: a versatile methodology for conducting qualitative research. *Rev High Educ* 1995;18:423–36.
63 Committee USOaP. United States Olympic and Paralympic committee. 2023.