

Please cite the Published Version

Smith, Ella S, McKay, Alannah KA, Ackerman, Kathryn E, Harris, Rachel, Elliott-Sale, Kirsty J, Stellingwerff, Trent and Burke, Louise M (2022) Methodology review: a protocol to audit the representation of female athletes in sports science and sports medicine research. International Journal of Sport Nutrition and Exercise Metabolism, 32 (2). pp. 114-127. ISSN 1526-484X

DOI: https://doi.org/10.1123/ijsnem.2021-0257

Publisher: Human Kinetics

Version: Accepted Version

Downloaded from: https://e-space.mmu.ac.uk/632011/

Additional Information: Accepted author manuscript version reprinted, by permission, from International Journal of Sport Nutrition and Exercise Metabolism, 2022, 32(2), 114-127, https://doi.org/10.1123/ijsnem.2021-0257 © Human Kinetics, Inc.

Enquiries:

If you have questions about this document, contact openresearch@mmu.ac.uk. Please include the URL of the record in e-space. If you believe that your, or a third party's rights have been compromised through this document please see our Take Down policy (available from https://www.mmu.ac.uk/library/using-the-library/policies-and-guidelines) Methodology Review: A Protocol to Audit the Representation of Female Athletes in Sports
 Science and Sports Medicine Research

3

4 Abstract

Female-specific research on sports science and sports medicine (SSSM) fails to mirror the 5 6 increase in participation and popularity of women's sport. Females have historically been 7 excluded from SSSM research, particularly because their physiological intricacy necessitates 8 more complex study designs, longer research times and additional costs. Consequently, most 9 SSSM practices are based on research with men, despite potential problems in translation to females due to sexual dimorphism in biological and phenotypical parameters as well as 10 11 differences in event characteristics (e.g., race distances/durations). Recognition that erroneous extrapolations may hamper the efforts of females to maximize their athletic potential has 12 created an impetus to acknowledge and readdress the sex disparity in SSSM research. To direct 13 the priorities for future research, it is prudent to first develop a comprehensive understanding 14 of the gaps in current knowledge by systematically 'auditing' the literature. By conducting 15 16 audits of the literature to highlight underdeveloped topics or identify potential problems with 17 the quality of research, this information can then be used to expediently direct new research activities. This paper therefore presents a standardized audit methodology to establish the 18 19 representation of female athletes in sub-disciplines of existing SSSM research, including a template for reporting the results of key metrics. This standardized audit process will enable 20 comparisons over time and between research sub-disciplines. This working guide provides an 21 22 important step towards achieving sex equity across SSSM research, with the eventual goal of 23 providing evidence-based recommendations specific to the female athlete.

Keywords: women, physical activity, performance, menstrual status, oral contraceptive, metaanalysis

- 26
- 27

28 Introduction

The representation of women in high-performance sport has increased in recent decades. 29 Indeed, the Tokyo 2020 Olympic Games was the first to achieve near parity in medal 30 opportunities for women's and men's events, with female representation rate at 49% of total 31 competitors, an increase from 45% in Rio 2016 and 38% in Sydney 2000 (Houghton et al., 32 2017; International Olympic Committee, 2021). Meanwhile, the Paralympic Games lags 33 behind with fewer medal opportunities for women, and a representation of 42% of athletes in 34 Tokyo 2020 (Tokyo2020, 2021). Across the globe, professional female teams are becoming 35 36 firmly established, demands for equal prize money normalized, and increases in female participation rates across all levels are driving the increase in the popularity of women's sport 37 (Claus, 2020; Douglas, 2018; Oxley, 2021; Townes, 2019). These trends justify the support for 38 female-specific sports science and sports medicine (SSSM) research to investigate specific 39 needs of this athletic population and their events. Numerous anatomical and biological 40 differences exist between the sexes, which in turn can influence performance, fundamental 41 biomechanics, and physiological responses to exercise (Devries, 2016; Green et al., 2016; 42 McNulty et al., 2020). Women's sports may also differ from those of their male counterparts, 43 according to the event demands or characteristics of the typical playing styles, including shorter 44 45 distances and lighter equipment for women (Kovalchik & Reid, 2017; Sanders et al., 2019). It is, therefore, problematic to apply conclusions drawn from male athletes directly to women 46 without considering any influence of these sexual dimorphisms and event-specific demands. 47

Unfortunately, the conspicuous imbalance of female specific SSSM research is well known. 48 For example, Costello and colleagues (2014) surveyed nearly 1,400 studies involving more 49 than six million participants published in three principal SSSM journals across a two year 50 period. Overall, women contributed just 39% of the total accumulated participant pool, with 4 51 to 13% of studies (depending on the journal) exclusively investigating female participants, 52 compared to 18 to 34% focusing entirely on men (Costello et al., 2014). A smaller audit of 53 54 three journals over a five-month period reporting SSSM studies found a similar proportion (42%) of women across the total accumulated participant count, with just 4% of studies 55 56 investigating only women, while 27% involved male-only participant cohorts (Brookshire, 2016). Across SSSM topics, the bias against female participants was most striking in studies 57 investigating strategies to enhance athletic performance; only 3% of all participants were 58 women (Brookshire, 2016). Female specific SSSM research shows little evidence of mirroring 59 the recent increase in the popularity of women's sport. Cowley et al. (2021) examined 5,261 60 studies across six SSSM journals including 12.5 million participants and reported similar 61 findings to Costello et al. (2014) seven years earlier. Women accounted for 34% of overall 62 participants, with 6% of studies focussing exclusively on women compared to 31% studying 63 men in isolation (Cowley et al., 2021). Interestingly, both Cowley et al. (2021) and Costello et 64 al. (2014) reported that on average 63% of studies included *both* men and women. Thus, given 65 the substantially greater number of male participants in both audits (61-66% of total participant 66 pool), this suggests an influence of volunteer (or self-selection) bias on participation rates, 67 whereby women are perhaps less willing or available to volunteer for certain investigations 68 despite being eligible to participate (Nuzzo, 2021). 69

Traditionally, the "typical 70 kg man" reference participant has been considered a sufficient
proxy for women across SSSM research (Marts & Keitt, 2004; Miller, 2005). Involving female
participants necessitates additional methodological considerations, including controlling for

sex-differences in concentrations of the reproductive hormones, as well as intra-female 73 fluctuations in oestrogen and progesterone due to menstrual cycle (MC) phase, the use of 74 hormonal contraceptives (HC) or impaired menstrual function (Elliott-Sale et al., 2021). 75 Ultimately, this results in more expensive, labour-intensive, and time-consuming study designs 76 which has traditionally been viewed as an inconvenience (Bruinvels et al., 2017). Moreover, 77 female athletes are more likely to experience nutritional issues such as iron deficiency or low 78 79 energy availability (Areta et al., 2021; Coad & Conlon, 2011; Logue et al., 2020), adding to the screening burden or the risk of interference in study outcomes. The availability and 80 81 recruitment of female athletes can also be challenging due to the tendency for smaller team sizes and the disproportionately low number of professional female athletes with opportunities 82 to participate in research projects (Emmonds et al., 2019). Collectively, these factors have 83 84 created a tendency to exclude women from SSSM research, with findings in men extrapolated to recommendations for female athletes with minimal consideration of sexual dimorphisms. 85 While understandable in terms of the burden of cost, convenience, and complexity of research, 86 this bias is unacceptable and has likely hampered the opportunities for women to maximise 87 their athletic potential. 88

The urgency of the need to address the sex disparity across SSSM research is receiving 89 substantial publicity within scientific literature (Elliott-Sale et al., 2021; Hutchins et al., 2021; 90 91 Martínez-Rosales et al., 2021) and popular/social media (Yu, 2018, 2021). It is therefore 92 important to efficiently direct the current enthusiasm for research on the female athlete into high-quality and meaningful outputs. Already, meta-analyses of selected SSSM topics have 93 been undertaken to determine how female athletes respond to a particular intervention or 94 95 stressor (Delextrat et al., 2018; Gomez-Bruton et al., 2021; Saunders et al., 2021). However, such technically well-conducted meta-analyses are challenged by substantial limitations in the 96 quantity and quality of the original literature. Moving forward, a comprehensive understanding 97

of the quantity and quality of research pertaining to female athletes is needed before time and 98 resources are invested into further original research or undertaking meta-analyses of topics 99 where women are substantially underrepresented or the female-specific methodological 100 considerations do not align with current recommendations (Elliott-Sale et al., 2021). This can 101 be achieved through 'auditing' the literature in specific sub-disciplines of SSSM research to 102 determine how well female athletes are currently addressed. In the context of this paper, we 103 104 define an 'audit' as a systematic analysis of the literature in specific sub-disciplines of SSSM to help to create a gap analysis of areas in which there is little information/ representation of 105 106 women alongside research areas with the greatest scope for development or impact. Conducting such audits across SSSM topics may produce information to help direct future research 107 activities in a systematic and expedient manner. Examples of such audits are beginning to 108 109 appear (Brookshire, 2016; Costello et al., 2014; Cowley et al., 2021; Hutchins et al., 2021). However, implementing a more standardized audit protocol will optimise the outputs and 110 interpretation of such audits, ensuring that comparisons can be made across time and across 111 SSSM topics. In particular, it is important to assess key quantitative and qualitative aspects are 112 assessed, such as the performance/fitness level of female athletes who are absent or represented 113 in the literature, or the quality of methods used to tackle the challenges of menstrual function. 114 Our proposed audit protocol utilizes new frameworks that have been proposed as best-practice 115 to assess menstrual status (Elliott-Sale et al., 2021) and athletic calibre (McKay et al., (in 116 117 press)). Therefore, this paper proposes a standardized protocol to facilitate the efficient and effective conduct of audits of the representation of female athletes in SSSM research, detailing 118 both our proposed methods of conducting the audit (guided by crowd sourced ideas with 119 experts in the field) and a standardized process of reporting the results. Our aim is to encourage 120 a collaborative and harmonised effort that can expedite a systematic correction to the current 121 sex-based biases in SSSM research. 122

123

- 124
- 125
- 126

127 Methods

Proposed standardized protocol for auditing the representation of women in sports science and sports medicine research

130 This standardized protocol and spreadsheet tool (provided in the supplementary material) are designed to facilitate a uniform and coordinated audit of the representation of female 131 132 participants within a topic or sub-discipline of SSSM research (e.g., nutritional pre-event strategies, use of performance supplements, recovery strategies, altitude training). The topic of 133 134 a specific audit can be defined and justified by the authors. In addition to standardizing the quantification of female participation, the audit methodology assesses key qualitative features 135 of the study design. These include the characterisation of the calibre of the athletic populations 136 137 that have been included in the available research, alongside arguably the most crucial qualitative aspect of research on female athletes: characterisation of menstrual status or HC 138 use. Collation of this qualitative information facilitates a comparison of SSSM research 139 140 conducted on men compared to women, as well as between SSSM sub-disciplines and across time. The concept of the proposed auditing process is illustrated in Figure 1. 141

142

[Figure 1].

The spreadsheet tool (supplementary material) is provided to aid researchers in conducting
such audits through providing a template to extract metrics A-F. The tool contains three tabs:
"flowchart" (a copy of Figure 2, to be populated upon the completion of preliminary study

screening as explained below), "study screening" (basic study information is first inputted into the "general information" category, whereby information to complete metrics A-F are then populated in the subsequent sections as denoted by sub-headings), and "MC control" (where the information necessary to grade studies on the basis of menstrual status classification and control according to our tiering system is inputted). Example data is also provided in the spreadsheet to guide researchers as to how we propose each tab is completed.

152 Data selection/search strategy

153 An electronic literature search is first conducted to identify relevant papers using the proposed search terms: "(athlete OR sport OR healthy) AND (**discipline-specific terms**) AND 154 (exercise OR performance OR endurance OR aerobic OR strength OR power OR anaerobic 155 156 OR speed OR skill OR tactics) NOT animal NOT rodent)". It is recommended that the search 157 is exclusive to original research papers among human participants, with no date restriction. Following the literature search, papers are then initially screened for separate review papers 158 and to remove duplicates or papers meeting any of the following exclusion criteria: (a) 159 populations with lifestyle diseases such as obesity or hypertension, (b) outcomes irrelevant to 160 areas of interest (i.e., not related to performance, health or indirect associations with 161 performance/health), (c) the sub-discipline/area of interest is not investigated as the 162 independent variable or the primary outcome of interest, (d) failure to explicitly state the sex 163 164 of participants, even if this could be inferred from data (e.g., anthropometric information), or not specifying the male to female participant ratio in mixed sex studies. Inferring participant 165 sex from the use of pronouns within a report may be acceptable but requires explanation in the 166 167 methodology of the audit outcomes, due to potential differences between gender and sex. Audits may include additional search/exclusion criteria specific to the sub-discipline of SSSM, 168 alongside the population of interest (e.g., age, sport, pregnancy). We advise that papers whose 169 full text cannot be obtained are also excluded. Following the initial search, we suggest that 170

review papers are screened for additional relevant papers that were not detected in the primarysearch.

173 The use of specialist software which support systematic reviews, such as Covidence (Covidence, 2021) or Rayyan (Ouzzani et al., 2016) is encouraged to aid the initial screening 174 process. The selected papers are then exported from the chosen software (including title, date, 175 176 journal, authors and abstract) to a spreadsheet program such as Microsoft Excel (template provided in the supplementary material) for more detailed analysis. Single papers that report 177 multiple separate studies should be identified, and each discrete study analysed individually. 178 Hereafter, the term 'paper' refers to the entire single publication, whereas 'study' is used to 179 describe separate investigations within a single paper or across several papers. After 180 preliminary screening, the following metrics are extracted: (A) population (B) athletic calibre, 181 (C) menstrual status, (D) research theme, (E) study impact, and (F) sample size. A Microsoft 182 Excel template has been provided in the supplementary material to support this process as 183 explained above. The template provides a table to record extracted metrics ("study screening" 184 tab), alongside a flowchart to report the results. The flowchart pictured in Figure 2 provides an 185 illustration of the requisite information that we recommend is extracted in sections A-F to 186 ensure consistent reporting across audits. Additional discipline-specific sections may be added 187 to the flowchart as appropriate. It is suggested that both the absolute and relative (%) values, 188 189 where applicable, are reported as denoted on the flowchart (Figure 2). If for any reason it is not possible to complete the flowchart in full or a section is intentionally excluded, a clear 190 explanation is needed in the study text. However, we recommend that incomplete or omitted 191 sections are avoided if possible. 192

193

[Figure 2]

194 **A. Population**

It is valuable to describe the potential of the study to contribute to knowledge of sex-based 195 differences in SSSM or to directly inform evidence-based guidelines for female athletes. 196 Therefore, we recommend that studies be separated into five population categories (a) males 197 only, (b) females only, (c) mixed-sex cohort, (d) male versus female sub-analysis, and (e) male 198 versus female design features. Male versus female design features (e) describes studies that 199 have been purposely designed to investigate differences in the intervention response between 200 201 sexes and include a statement in the aims as well as clear features in the study design. Meanwhile, male versus female sub-analysis (d) describes studies in which sex-based 202 203 comparisons were completed within the statistical procedures, although this was not a primary aim of the study or the study was not specifically controlled or designed for best-practice 204 comparisons. In addition to quantifying the representation of female participants within the 205 206 SSSM discipline, this approach will differentiate research that excludes female participants (category A, males only), simply includes women without methodological consideration of 207 sexual dimorphisms (category C, mixed-sex cohort), evaluates the response of women to an 208 intervention or topic (category B, females only), or specifically targets sexual dimorphisms 209 within SSSM disciplines (category D, male versus female sub-analysis or category E, male 210 211 versus female design features).

B. Athletic Calibre

A general principle of research is that the results of a study apply to appropriately defined participant populations and scenarios that are similar to those included in its design (Schünemann et al., 2021). Therefore, it is important to provide a clear and standardized description of participants of SSSM research in terms of their calibre of athletic ability and/or level of competition. This information can determine the ecological validity of the application of the findings to high-performance female athletes and their direct inclusion in evidence-based recommendations/guidelines. A recently developed framework for classifying athletic calibre

(McKay et al., (in press)) allows participants to be ranked on a six-tiered classification system, 220 ranging from Sedentary/Healthy participants (Tier 0) up to World Class athletes (Tier 5). 221 Classification is made from information around performance indicators and training status that 222 is easily accessed and commonly reported in most papers. Importantly, classification is 223 determined, wherever possible, according to objective quantitative data (i.e., personal best 224 performances or world rankings) rather than subjective statements such as 'elite' or 'trained' 225 226 (McKay et al., (in press)). Studies that provide insufficient information to be robustly classified into a single tier are graded as 'unclassified'. For studies in which a range of athletic calibres 227 228 are included, a classification noting the majority of participants is stated. If such a determination is not possible (i.e., participant numbers in each tier are not reported), we 229 recommend that the mean tier is taken. However, studies comparing distinct athletic calibres 230 through purposeful methodological design should be recognised for their superior study design, 231 and therefore we suggest that each tier is reported individually. For example, a paper comparing 232 Tier 3 versus Tier 4 athletes would contribute to the census of each Tier, with a symbol in the 233 flowchart to note that multiple cohorts were counted in a study and thus there is a difference 234 between the total number of athlete cohorts and the number of studies. Importantly, this applies 235 only to a priori and not retrospective comparison between athletic calibres. Two authors should 236 independently classify studies, with discrepancies resolved through discussion, adjudication by 237 a third author, or, if possible, contact validation with the author(s) of the original study. Table 238 1 provides examples of how the Participant Classification System would be implemented. An 239 additional benefit of conducting audits in this manner is that it may provide the first systematic 240 quantification of the calibre of male participants in SSSM research, as well as allowing the 241 comparison to female athletes. 242

244

245

246 C. Menstrual Status

Because the ovarian hormones oestrogen and progesterone influence multiple biological 247 systems (Ansdell et al., 2020; Devries, 2016; Green et al., 2016; Laurent et al., 2014; 248 Wohlgemuth et al., 2021), they have the potential to cause downstream effects on the nature 249 and heterogeneity of the findings of various study topics. This is of relevance to high-250 performance athletes (specifically) whose fluctuations in ovarian hormone concentrations, and 251 252 any resulting influence on study outcomes, may mask marginal (i.e., <1-2%) changes in performance that are of significance within real-world elite sporting scenarios. There are 253 myriad of fluctuations in ovarian hormone concentrations across a MC acutely over a given 254 255 month and chronically over an athlete's training blocks (menstrual cycle status over months to 256 years) and over a lifespan (e.g., pre-puberty versus menopause) (Elliott-Sale et al., 2021). This highlights the importance of robust characterisation of overall menstrual status and the specific 257 phase of the cycle in female participants and/or the alignment of experimental design with 258 specific hormonal milieu. Table 2 provides a detailed guide to our proposed tiered system to 259 grade the standard of methodological control, underpinned by a new framework promoted as 260 best-practice for this theme (Elliott-Sale et al., 2021). Because the extent to which ovarian 261 262 hormones influence particular areas of SSSM and athletic performance is currently unclear, 263 failure to control for hormonal profiles within female participants in a research project introduces uncertainty in the application of the project's results to the larger female population 264 and sacrifices an opportunity to investigate the extent of sexual dimorphisms. 265

Our proposed tiering system provides detailed information regarding the consideration of participant menstrual status in study design, underpinned by best-practice guidelines (Elliott-Sale et al., 2021). Our tiered ranking system assesses studies with female participants on the basis of both (1) the classification of female participants according to menstrual status, and (2)
the standardization of methodological control relating to ovarian hormonal profiles. Table 3
details examples of how this process is undertaken:

Step 1. Study populations are initially characterised according to four categories of 272 menstrual status: (i) MC (i.e., including eumenorrheic and naturally menstruating 273 274 women); (ii) HC; (iii) menstrual irregularities; and (iv) mixed menstrual status (i.e., including women from more than one of the aforementioned categories, with each 275 group being distinguishable). If there is insufficient information to provide a robust 276 classification of participants, or a mixed female cohort in which individual menstrual 277 status cannot be discerned, it is reported as an 'unclassified' cohort. Studies that state 278 that HC use was not excluded, but do not provide further information as to the number 279 of HC users in the cohort and/or identify results from this group distinguishably, are 280 also graded as 'unclassified'. In the case of studies that have excluded HC users, it 281 282 cannot be assumed that participants were naturally menstruating, unless this is stated and described, and hence these cohorts are also graded as 'unclassified'. Explicit 283 information about the menstrual cycle should describe participants as either 284 eumenorrheic or naturally menstruating based on five criteria; MC length, number of 285 consecutive menses per year, evidence of LH surge, hormonal profiles and absence of 286 HC use (Elliott-Sale et al., 2021). Participants achieving all five criteria are described 287 as eumenorrheic, whereas those fulfilling fewer than five are described as naturally 288 menstruating (Elliott-Sale et al., 2021). 289

290

291 Step 2. Following the classification of participants, a further assessment of the standard 292 of methodological control regarding ovarian hormonal profiles of each category is 293 completed. This assessment places the methodological control into one of four tiers

(Gold, Silver, Bronze or Ungraded). It is recommended that mixed participant studies 294 assess each participant classification separately. For example, if a study includes 295 naturally menstruating women alongside those using HC, these are assessed as two 296 separate groups, with 0.5 assigned to each participant classification. Separating the 297 participant classifications in mixed studies allows the overall standard of studies to be 298 identified, while considering the methodological adjustments for ovarian hormone 299 300 concentration specific to the menstrual status of each female population in the cohort. Studies using 'unclassified' participants (as determined in step one) are not assessed for 301 302 methodological control. Judgement should be used for those cases which are not explicitly covered by the categories (i.e., implicit judgement). 303

304

[Table 2 and 3]

305 **D. Research Theme**

We recommend that the distribution of research in female athletes across different themes of 306 SSSM is evaluated to assess how this distribution compares to research conducted in men, 307 whereby male-oriented research provides control/normative values. Even within a given SSSM 308 309 topic (e.g., caffeine supplementation or altitude exposure), we recommend that studies are classified into three key research themes as follows: (1) performance focus/outcome; (2) 310 clinically established health focus/outcome or (3) indirect or emerging associations with 311 performance or health, including underlying mechanisms. Table 4 outlines these three 312 categories, with a non-exclusive list of examples to illustrate the definitions, while Table 5 313 provides an illustration of how themes might be decided for a selected range of studies. It is 314 315 noted that specific allocation of these themes will likely change according to the topic or subdiscipline of SSSM being assessed and may involve some subjectivity. Direct measures of 316 performance or health have the most relevance or practical application to high-performance 317

athlete populations (Schünemann et al., 2021). Meanwhile, indirect associations with 318 performance or health are one-step removed but may provide important insight regarding 319 potential performance/health outcomes if interpreted carefully. With this in mind, in scenarios 320 where the findings of an investigation could be classified into multiple categories, we suggest 321 that the study is assigned according to the following priority scale: performance or health, and 322 then indirect associations with performance or health. Two authors should independently 323 classify studies, with discrepancies resolved through discussion or involving an additional 324 author where necessary. 325

326

[Tables 4 and 5]

327 E. Study Impact

The trajectory of interest in a particular SSSM research topic can be summarised by plotting 328 the frequency of publications over time. Whether this trajectory is mirrored for studies 329 involving female athletes is also of interest. Consideration of the impact factor (IF) of the 330 journals that publish SSSM outputs involving male and/or female participants may illuminate 331 the actual, or perceived interest in, and value of, the work (Scully & Lodge, 2005; Waltman & 332 Traag, 2020). For example, substantial differences in the IF of the journals that support 333 different types of studies may highlight a sex-bias in the perception or reality of study quality 334 or, possibly, publication bias. Therefore, it is of interest to extract the IF of the journal in which 335 the study is published; we recommend that this is the most recent four-year IF, regardless of 336 publication date, to standardize changes in journal IF over time. Moreover, for the purpose of 337 regulating IF calculation method, we strongly recommend that IF is extracted from the 338 International Scientific Indexing (ISI) IF, or an index reporting this calculation method, rather 339 than the IF directly reported by the journal. It is noted that when a single paper contains several 340 discrete studies, we suggest that the IF of the journal is counted against each study included in 341

the audit. Such information may be able to highlight disparities to journal editors or provide an incentive for researchers to conduct more challenging study designs if there is an apparent reward in terms of publication in higher IF journals. When interpreting the results of the audit regarding IF, the limitations of this metric, including the propensity for manipulation to a journal's advantage, should be considered (Kaldas et al., 2020). Although this assessment is not essential to achieve the primary aims regarding directing new/targeted research and is therefore not a compulsory section of an audit, it is strongly recommended.

A more recent system to quantify study impact involves Altmetrics (alternative metrics). 349 Altmetrics involve bibliometrics that measure the attention, dissemination, influence, and 350 impact of scientific papers beyond their traditional academic outputs. The concept was first 351 promoted by Priem et al. (2010) and is now run by the "Altmetric.com" company, which 352 collects data on the online exposure to, and engagement with, scientific publications within 353 news, social media, article pageviews and downloads, article repositories, expert 354 commentaries, and public policy documents (Altmetric, 2016). Although there are limitations 355 to the gathering and interpretation of this metric, the advantages of including it within an audit 356 are that it provides a standardized characteristic of the wider community interest in a scientific 357 output and is accumulated more quickly than scientific citations (Sutton, 2014; Thelwall, 358 2020). Therefore, there is less disadvantage to including it for recently published papers that 359 360 have had less time to create traditional academic impact (Sutton, 2014; Thelwall, 2020). It is suggested that Altmetrics and journal IF are considered in combination when interpreting 361 findings, rather than in isolation, to build a more holistic view of study impact. We suggest that 362 the Altmetric Score is collected for all papers published from 2012 (the date of the initiation of 363 the Altmetric company); this can be derived using the company bookmarklet 364 (https://www.altmetric.com/products/free-tools/bookmarklet/). Again, if multiple studies are 365 included in the same paper, we advise that these are counted separately in the audit tally. 366

367

368 F. Sample Size

A difference in sample size between papers focussed on men, women and/or both sexes, may 369 suggest differences in the ease of recruitment or the requirements for study power (dependent 370 on the research topic). Furthermore, changes in the male to female participant ratio in mixed 371 372 sex studies over time or across research themes (metric D) may help to track whether there is greater interest in female study involvement by participants or researchers, or greater awareness 373 of opportunities to increase sample sizes to enable *a priori* comparisons between sexes. We 374 therefore recommend that sample size is extracted, separating between male and female 375 participants where applicable. 376

377 Summary

This audit protocol proposes a standardized approach to conducting literature audits on the 378 representation of female athletes in topics and sub-disciplines of SSSM (Figure 1). By 379 identifying strengths and gaps in the quantity and quality of the existing literature, sports 380 practitioners and academics can uncover underdeveloped topics or potential problems with the 381 382 quality of research in relation to female athletes. The standardization of the process of 383 conducting and reporting the results of such audits will facilitate an understanding of how well female athletes are addressed by currently available research, and how these might differ over 384 385 time or between sub-disciplines of interest. This will help to direct the priority for new research in a systematic way. In addition, these guidelines will contribute to best practices for future 386 studies by highlighting the key elements of appropriate scientific design in female athletes. Of 387 388 critical importance to the first auditing process is the assessment of how well studies classify participant menstrual status, alongside the implementation of appropriate methodological 389 controls for ovarian hormonal profiles (Elliott-Sale et al., 2021). An additional bonus of 390

conducting the proposed audits is that information on the SSSM literature involving male 391 athletes will also be collected, therefore capturing aspects that have not been previously 392 described in a systematic process in either men or women (e.g., the calibre of athletes, sample 393 sizes used to investigate various research themes). Although further analysis of the outcomes 394 of the studies included in individual audits, alongside specific aspects relating to the quality of 395 research methodology, is needed. The broader audits that we propose (undertaken using a 396 397 standardized format) may provide a greater understanding regarding the state of the general SSSM literature, before resources are expended in futile analysis of research areas where 398 399 females are not well represented, or sex-specific methodological considerations do not align with current recommendations (Elliott-Sale et al., 2021). The completion of audits across a 400 range of SSSM topics may therefore help to identify and support topics that merit a meta-401 402 analysis involving further and detailed screening of the research outcomes. Furthermore, audits may identify resources that could be developed to guide new research, areas in which specific 403 investigation of female athletes or comparisons between the responses of male and female 404 athletes should be undertaken as a priority, or the presence and absence of athletes from a 405 specific demographic of athletic calibre/achievement. Ultimately, this working guide provides 406 a crucial step towards accelerating sex parity in SSSM research, with the eventual goal of 407 delivering evidence-based recommendations specific to the female athlete. The potential 408 benefits of greater inclusion of women in SSSM research and more targeted research on female 409 410 athletes include better athletic performance, safer athlete practices, greater parity in access to resources and opportunities, and an increase in general SSSM knowledge and expertise. 411

412

413

414

415

416 Funding sources

- 417 We acknowledge support of this work by the Wu Tsai Human Performance Alliance and the
- 418 Joe and Clara Tsai Foundation.

419

420 **Potential conflicts of interest**

421 The authors declare no conflicts of interest.

422

423 Author Contributions

424 LMB, ESS and AKA formulated the concept. ESS wrote the manuscript with input from AKA,

425 KEA, RH, KJE-S, TS and LMB. All authors revised and approved the final version of the

426 manuscript.

References

- Ali, A., O'Donnell, J., Foskett, A., & Rutherfurd-Markwick, K. (2016). The influence of caffeine ingestion on strength and power performance in female team-sport players. J Int Soc Sports Nutr, 13(1), 46. doi:10.1186/s12970-016-0157-4
- Ali, A., O'Donnell, J. M., Starck, C., & Rutherfurd-Markwick, K. J. (2015). The Effect of Caffeine Ingestion during Evening Exercise on Subsequent Sleep Quality in Females. *Int J Sports Med*, 36(6), 433-439. doi:10.1055/s-0034-1398580
- Altmetric. (2016). What are Altmetrics? Retrieved from <u>https://www.altmetric.com/about-altmetrics/what-are-altmetrics/</u>
- Ansdell, P., Thomas, K., Hicks, K. M., Hunter, S. K., Howatson, G., & Goodall, S. (2020). Physiological sex differences affect the integrative response to exercise: acute and chronic implications. *Experimental physiology*, 105(12), 2007-2021.
- Areta, J. L., Taylor, H. L., & Koehler, K. (2021). Low energy availability: history, definition and evidence of its endocrine, metabolic and physiological effects in prospective studies in females and males. *Eur J Appl Physiol*, 121(1), 1-21. doi:10.1007/s00421-020-04516-0
- Boegman, S., Stellingwerff, T., Shaw, G., Clarke, N., Graham, K., Cross, R., & Siegler, J. C. (2020). The Impact of Individualizing Sodium Bicarbonate Supplementation Strategies on World-Class Rowing Performance. *Front Nutr*, 7, 138. doi:10.3389/fnut.2020.00138
- Brisola, G. M. P., de Souza Malta, E., Santiago, P. R. P., Vieira, L. H. P., & Zagatto, A. M. (2018). beta-Alanine Supplementation's Improvement of High-Intensity Game Activities in Water Polo. *Int J Sports Physiol Perform*, 13(9), 1208-1214. doi:10.1123/jjspp.2017-0636
- Brookshire, B. (2016). Women in sports are often underrepresented in science. Science News, 847-851.
- Bruinvels, G., Burden, R., McGregor, A., Ackerman, K., Dooley, M., Richards, T., & Pedlar, C. (2017). Sport, exercise and the menstrual cycle: where is the research?
- Bull, F. C., Al-Ansari, S. S., Biddle, S., Borodulin, K., Buman, M. P., Cardon, G., ... Willumsen, J. F. (2020). World Health Organization 2020 guidelines on physical activity and sedentary behaviour. *Br J Sports Med*, 54(24), 1451-1462. doi:10.1136/bjsports-2020-102955
- Burke, L. M., Hall, R., Heikura, I. A., Ross, M. L., Tee, N., Kent, G. L., ... Jones, A. M. (2021). Neither beetroot juice supplementation nor increased carbohydrate oxidation enhance economy of prolonged exercise in elite race walkers. *Nutrients*, 13(8), 2767.
- Carr, A. J., Slater, G. J., Gore, C. J., Dawson, B., & Burke, L. M. (2011). Effect of sodium bicarbonate on [HCO3-], pH, and gastrointestinal symptoms. *Int J Sport Nutr Exerc Metab*, 21(3), 189-194. doi:10.1123/ijsnem.21.3.189
- Claus, L. (2020). Why gender equal pay is so hard to achieve in sport. Retrieved from <u>https://www.mgmt.ucl.ac.uk/news/gender-equality-sports-%E2%80%93-why-gender-equal-pay-so-hard-achieve-sport</u>
- Coad, J., & Conlon, C. (2011). Iron deficiency in women: assessment, causes and consequences. *Curr Opin Clin Nutr Metab Care*, *14*(6), 625-634. doi:10.1097/MCO.0b013e32834be6fd
- Costello, J. T., Bieuzen, F., & Bleakley, C. M. (2014). Where are all the female participants in Sports and Exercise Medicine research? *European journal of sport science*, *14*(8), 847-851.
- Covidence. (2021). Covidence systematic review software, Veritas Health Innovation. Melbourne, Australia. Retrieved from <u>www.covidence.org</u>
- Cowley, E. S., Olenick, A. A., McNulty, K. L., & Ross, E. Z. (2021). "Invisible Sportswomen": The Sex Data Gap in Sport and Exercise Science Research. *Women in Sport and Physical Activity Journal*, *1*(aop), 1-6.
- Delextrat, A., Mackessy, S., Arceo-Rendon, L., Scanlan, A., Ramsbottom, R., & Calleja-Gonzalez, J. (2018). Effects of Three-Day Serial Sodium Bicarbonate Loading on Performance and Physiological Parameters During a Simulated Basketball Test in Female University Players. Int J Sport Nutr Exerc Metab, 28(5), 547-552. doi:10.1123/ijsnem.2017-0353
- Devries, M. C. (2016). Sex-based differences in endurance exercise muscle metabolism: impact on exercise and nutritional strategies to optimize health and performance in women. *Exp Physiol*, *101*(2), 243-249. doi:10.1113/EP085369
- Douglas, L. (2018). The rise of women's sports. Retrieved from <u>https://www.nielsen.com/wp-content/uploads/sites/3/2019/04/the-rise-of-womens-sports.pdf</u>

- Driller, M. W., Gregory, J. R., Williams, A. D., & Fell, J. W. (2012). The effects of serial and acute NaHCO3 loading in well-trained cyclists. J Strength Cond Res, 26(10), 2791-2797. doi:10.1519/JSC.0b013e318241e18a
- Elliott-Sale, K. J., Minahan, C. L., de Jonge, X. A. J., Ackerman, K. E., Sipilä, S., Constantini, N. W., . . . Hackney, A. C. (2021). Methodological considerations for studies in sport and exercise science with women as participants: a working guide for standards of practice for research on women. *Sports medicine*, 1-19.
- Emmonds, S., Heyward, O., & Jones, B. (2019). The challenge of applying and undertaking research in female sport. *Sports medicine-open*, *5*(1), 1-4.
- Flueck, J. L., Mettler, S., & Perret, C. (2014). Influence of caffeine and sodium citrate ingestion on 1,500-m exercise performance in elite wheelchair athletes: a pilot study. *Int J Sport Nutr Exerc Metab*, 24(3), 296-304. doi:10.1123/ijsnem.2013-0127
- Fowler, R., Jeffries, O., Tallent, J., Theis, N., Heffernan, S. M., McNarry, M. A., . . . Waldron, M. (2021). No thermoregulatory or ergogenic effect of dietary nitrate among physically inactive males, exercising above gas exchange threshold in hot and dry conditions. *European journal of sport science*, 21(3), 370-378.
- Gomez-Bruton, A., Marin-Puyalto, J., Muñiz-Pardos, B., Matute-Llorente, A., Del Coso, J., Gomez-Cabello, A., . . . Lozano-Berges, G. (2021). Does Acute Caffeine Supplementation Improve Physical Performance in Female Team-Sport Athletes? Evidence from a Systematic Review and Meta-Analysis. *Nutrients, 13*(10), 3663.
- Goulet, E. D., Rousseau, S. F., Lamboley, C. R., Plante, G. E., & Dionne, I. J. (2008). Pre-exercise hyperhydration delays dehydration and improves endurance capacity during 2 h of cycling in a temperate climate. *J Physiol Anthropol*, 27(5), 263-271. doi:10.2114/jpa2.27.263
- Green, D. J., Hopkins, N. D., Jones, H., Thijssen, D. H., Eijsvogels, T. M., & Yeap, B. B. (2016). Sex differences in vascular endothelial function and health in humans: impacts of exercise. *Exp Physiol*, *101*(2), 230-242. doi:10.1113/EP085367
- Heikura, I. A., Burke, L. M., Hawley, J. A., Ross, M. L., Garvican-Lewis, L., Sharma, A. P., . . . Ackerman, K. E. (2019). A Short-Term Ketogenic Diet Impairs Markers of Bone Health in Response to Exercise. *Front Endocrinol (Lausanne)*, 10, 880. doi:10.3389/fendo.2019.00880
- Hitchins, S., Martin, D. T., Burke, L., Yates, K., Fallon, K., Hahn, A., & Dobson, G. P. (1999). Glycerol hyperhydration improves cycle time trial performance in hot humid conditions. *Eur J Appl Physiol Occup Physiol*, 80(5), 494-501. doi:10.1007/s004210050623
- Houghton, E. J., Pieper, L., & Smith, M. (2017). Women in the 2016 Olympic and Paralympic Games: an analysis of participation, leadership, and media coverage.
- Hutchins, K. P., Borg, D. N., Bach, A. J. E., Bon, J. J., Minett, G. M., & Stewart, I. B. (2021). Female (under) representation in exercise thermoregulation research. *Sports Med Open*, 7(1), 43. doi:10.1186/s40798-021-00334-6
- International Olympic Committee. (2021). Retrieved from https://olympics.com/en/
- ISI Impact Factor. ISI Imapct Factor. Retrieved from http://isi-impactfactor.com/
- Kaldas, M., Michael, S., Hanna, J., & Yousef, G. M. (2020). Journal impact factor: a bumpy ride in an open space. J Investig Med, 68(1), 83-87. doi:10.1136/jim-2019-001009
- Kovalchik, S. A., & Reid, M. (2017). Comparing Matchplay Characteristics and Physical Demands of Junior and Professional Tennis Athletes in the Era of Big Data. J Sports Sci Med, 16(4), 489-497. Retrieved from <u>https://www.ncbi.nlm.nih.gov/pubmed/29238248</u>
- Lara, B., Gutierrez Hellin, J., Ruiz-Moreno, C., Romero-Moraleda, B., & Del Coso, J. (2020). Acute caffeine intake increases performance in the 15-s Wingate test during the menstrual cycle. *Br J Clin Pharmacol*, 86(4), 745-752. doi:10.1111/bcp.14175
- Laurent, C. M., Vervaecke, L. S., Kutz, M. R., & Green, J. M. (2014). Sex-specific responses to selfpaced, high-intensity interval training with variable recovery periods. J Strength Cond Res, 28(4), 920-927. doi:10.1519/JSC.0b013e3182a1f574
- Lewiecki, E. M., Watts, N. B., McClung, M. R., Petak, S. M., Bachrach, L. K., Shepherd, J. A., . . . Densitometry, I. S. f. C. (2004). Official positions of the international society for clinical densitometry. *The Journal of Clinical Endocrinology & Metabolism*, 89(8), 3651-3655.
- Logue, D. M., Madigan, S. M., Melin, A., Delahunt, E., Heinen, M., Donnell, S. M., & Corish, C. A. (2020). Low Energy Availability in Athletes 2020: An Updated Narrative Review of

Prevalence, Risk, Within-Day Energy Balance, Knowledge, and Impact on Sports Performance. *Nutrients*, *12*(3), 835. doi:10.3390/nu12030835

- Martínez-Rosales, E., Hernández-Martínez, A., Sola-Rodríguez, S., Esteban-Cornejo, I., & Soriano-Maldonado, A. (2021). Representation of women in sport sciences research, publications, and editorial leadership positions: are we moving forward? *Journal of science and medicine in sport*.
- Marts, S. A., & Keitt, S. (2004). Foreword: a historical overview of advocacy for research in sex-based biology. In *Advances in Molecular and Cell Biology* (Vol. 34, pp. v-xiii): Elsevier.
- McKay, A., Stellingwerff, T., Smith, E., Martin, D., Mujika, I., Goosey-Tolfrey, V., . . . Burke, L. ((in press)). Defining training and performance calibre: a participant classification framework. *International journal of sports physiology and performance*.
- McNulty, K. L., Elliott-Sale, K. J., Dolan, E., Swinton, P. A., Ansdell, P., Goodall, S., ... Hicks, K. M. (2020). The Effects of Menstrual Cycle Phase on Exercise Performance in Eumenorrheic Women: A Systematic Review and Meta-Analysis. *Sports Med*, 50(10), 1813-1827. doi:10.1007/s40279-020-01319-3
- Mero, A. A., Keskinen, K. L., Malvela, M. T., & Sallinen, J. M. (2004). Combined creatine and sodium bicarbonate supplementation enhances interval swimming. J Strength Cond Res, 18(2), 306-310. doi:10.1519/R-12912.1
- Miller, V. M. (2005). Sex-based physiology prior to political correctness. Am J Physiol Endocrinol Metab, 289(3), E359-360. doi:10.1152/classicessays.00035.2005
- Nattiv, A., Loucks, A. B., Manore, M. M., Sanborn, C. F., Sundgot-Borgen, J., Warren, M. P., & American College of Sports, M. (2007). American College of Sports Medicine position stand. The female athlete triad. *Med Sci Sports Exerc*, 39(10), 1867-1882. doi:10.1249/mss.0b013e318149f111
- Nuzzo, J. (2021). Volunteer bias and female participation in exercise and sports science research. *Quest*, 73(1), 82-101.
- Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan—a web and mobile app for systematic reviews. *Systematic reviews*.
- Oxley, S. (2021). Prize money in sport: what can football learn from cricket? Retrieved from https://www.bbc.co.uk/sport/56240686
- Priem, J., Taraborelli, D., Groth, P., & Neylon, C. (2010). Altmetrics: a manifesto. Retrieved from <u>http://altmetrics.org/manifesto/</u>
- Sanders, D., van Erp, T., & de Koning, J. J. (2019). Intensity and load characteristics of professional road cycling: differences between men's and women's races. *International journal of sports physiology and performance*, 14(3), 296-302.
- Saunders, B., Oliveira, L. F., Dolan, E., Durkalec-Michalski, K., McNaughton, L., Artioli, G. G., & Swinton, P. A. (2021). Sodium bicarbonate supplementation and the female athlete: A brief commentary with small scale systematic review and meta-analysis. *Eur J Sport Sci*, 1-10. doi:10.1080/17461391.2021.1880649
- Schünemann, H. J., Vist, G. E., Higgins, J. P., Santesso, N., Deeks, J. J., Glasziou, P., ... Group, C. G. M. (2021). Chapter 15: Interpreting results and drawing conclusions. *Cochrane handbook for systematic reviews of interventions*, 6.2 (updated February 2021), 403-431.
- Scully, C., & Lodge, H. (2005). Impact factors and their significance; overrated or misused? *British Dental Journal*, 198(7), 391-393.
- Skinner, T. L., Jenkins, D. G., Leveritt, M. D., McGorm, A., Bolam, K. A., Coombes, J. S., & Taaffe, D. R. (2014). Factors influencing serum caffeine concentrations following caffeine ingestion. *J Sci Med Sport*, 17(5), 516-520. doi:10.1016/j.jsams.2013.07.006
- Smith-Ryan, A. E., Fukuda, D. H., Stout, J. R., & Kendall, K. L. (2012). High-velocity intermittent running: effects of beta-alanine supplementation. J Strength Cond Res, 26(10), 2798-2805. doi:10.1519/JSC.0b013e318267922b
- Sutton, S. (2014). Altmetrics: What good are they to academic libraries. *Kansas Library Association College and University Libraries Section Proceedings*, 4(2), 1-7.
- Thelwall, M. (2020). The pros and cons of the use of altmetrics in research assessment.

- Tokyo2020. (2021). Tokyo 2020 sets the record for most athletes and women at a Paralympic Games. Retrieved from <u>https://www.paralympic.org/news/tokyo-2020-sets-record-most-athletes-and-women-paralympic-games</u>
- Torstveit, M., & Sundgot-Borgen, J. (2005). Low bone mineral density is two to three times more prevalent in non-athletic premenopausal women than in elite athletes: a comprehensive controlled study. *British journal of sports medicine*, *39*(5), 282-287.
- Townes, C. (2019). A decade of fighting for equal pay in sports. Retrieved from <u>https://www.forbes.com/sites/ceceliatownes/2020/12/31/a-decade-of-fighting-for-equal-</u> pay/?sh=6557a29a1376
- Van Nieuwenhoven, M. A., Brummer, R. M., & Brouns, F. (2000). Gastrointestinal function during exercise: comparison of water, sports drink, and sports drink with caffeine. J Appl Physiol (1985), 89(3), 1079-1085. doi:10.1152/jappl.2000.89.3.1079
- Waltman, L., & Traag, V. A. (2020). Use of the journal impact factor for assessing individual articles: Statistically flawed or not? *F1000Research*, 9.
- Whitfield, J., Burke, L. M., McKay, A. K. A., Heikura, I. A., Hall, R., Fensham, N., & Sharma, A. P. (2021). Acute Ketogenic Diet and Ketone Ester Supplementation Impairs Race Walk Performance. *Med Sci Sports Exerc*, 53(4), 776-784. doi:10.1249/MSS.00000000002517
- Wohlgemuth, K. J., Arieta, L. R., Brewer, G. J., Hoselton, A. L., Gould, L. M., & Smith-Ryan, A. E. (2021). Sex differences and considerations for female specific nutritional strategies: a narrative review. J Int Soc Sports Nutr, 18(1), 27. doi:10.1186/s12970-021-00422-8
- Yu, C. (2018). Where are the women in sports science research? Retrieved from <u>https://www.outsideonline.com/health/training-performance/where-are-women-sports-</u><u>science-research/</u>
- Yu, C. (2021). The gender gap. Retrieved from <u>https://www.physiology.org/publications/news/the-physiologist-magazine/2021/july/the-gender-gap?SSO=Y</u>
- Ziegenfuss, T. N., Rogers, M., Lowery, L., Mullins, N., Mendel, R., Antonio, J., & Lemon, P. (2002). Effect of creatine loading on anaerobic performance and skeletal muscle volume in NCAA Division I athletes. *Nutrition*, 18(5), 397-402. doi:10.1016/s0899-9007(01)00802-4

Table 1. Examples across a range of studies demonstrating how the Participant Classification Framework (McKay et al., (in press)) is specifically applied retrospectively for the purpose of an audit.

Study	Description of participant characteristics	Reported participant characteristics	Premise for classification	Classification
Influence of caffeine and sodium citrate ingestion on 1,500-m exercise performance in elite wheelchair athletes: a pilot study (Flueck et al., 2014)	 Elite wheelchair-racing athletes, including several Paralympic Games, World and European Championship medallists. Competed in the category T53/54 and were national team members 	None provided	Participants were Paralympic/World medallists.	Tier 5 World Class
The Impact of Individualizing Sodium Bicarbonate Supplementation Strategies on World- Class Rowing Performance (Boegman et al., 2020)	 23 elite male rowers. Recruited across two research centres (Canadian Sport Institute Pacific and the New South Wales Institute of Sport, Australia) 13 Olympic/World-Champs team members as well as one rowing ergometer world record holder 	 2,000-m ergometer time trial personal bests ranged from 5 min 39 s (open- weight) to 6 min 14 s (lightweight) 	The majority (13 out of 23) of participants competed at international (Olympic/World championship level). The 2,000m ergometer times are between 1 - 5% of world record.	Tier 4 Elite/ International Level
Combined creatine and sodium bicarbonate supplementation enhances interval swimming (Mero et al., 2004)	 Competitive national level male and female swimmers All subjects had a minimum of 4 years of experience in competitive swimming training 	• 100 m freestyle personal best: 57.9 \pm 1.5 s (males), 67.1 \pm 1.7 s (females)	Participants compete at the national level, with personal best times within 20-26% of world leading times in the year of publication.	Tier 3 Highly Trained/ National Level

The effects of serial and acute NaHCO3 loading in well-trained cyclists (Driller et al., 2012)	 8 well-trained male cyclists All the cyclists were competing at the state level, with some competing at national level (n = 5) 	•	VO ₂ max = 66.8 ± 8.4 ml.kg ⁻¹ .min ⁻¹	Participants identify with a specific sport. The majority (5 out of 8) of participants compete at the national level which justifies a Tier 3 classification.	Tier 3 Highly Trained/ National Level
The Effect of Caffeine Ingestion during Evening Exercise on Subsequent Sleep Quality in Females (Ali et al., 2015)	 Participants from a range of team sports (soccer, hockey, and netball), at various competitive levels (recreational to international) Trained 2–6 times per week 	•	None provided	Participants all identify with a specific sport. Cohort includes individuals from multiple tiers: Regional (Tier 2), National (Tier 3) and International (Tier 4). As the number of participants in each tier is not reported the "median" classification is taken.	Tier 3 Highly Trained/ National Level
Glycerol hyperhydration improves cycle time trial performance in hot humid conditions (Hitchins et al., 1999)	 Trained male cyclists were chosen on the basis of their success in recent, regional cycling race Subjects had at least 3 years of competition experience 	•	$VO_2max = 65.0 \pm 3.9 ml.kg^{-1}.min^{-1}$ Wattmax = 376 ± 24 W.	Participants meet physical activity guidelines (Bull et al., 2020), identify with a sport and are competing regionally. No further information to justify 'sub-elite' terminology.	Tier 2 Trained/ Developmental Level
High-velocity intermittent running: effects of beta-alanine supplementation (Smith- Ryan et al., 2012)	• All participants were moderately trained, engaging in 3–7 days per week of aerobic, resistance, or recreational activities	•	Recreationally active (1–5 h/ week).	Participants meet physical activity guidelines (Bull et al., 2020). No information to suggest participants are competitive or identify with a specific sport to justify a Tier 2 classification.	Tier 1 Recreationally Active

No thermoregulatory or ergogenic effect of dietary nitrate among physically inactive males, exercising above gas exchange threshold in hot and dry conditions (Fowler et al., 2021)	• Healthy, none of the participants trained for endurance exercise on a regular basis and were deemed to be physically inactive based on exercising < 30 min of moderate exercise per week	 VO₂max = 41.1 ± 3.6 ml.kg⁻¹.min⁻¹ 	Participants do not meet physical activity guidelines (Bull et al., 2020).	Tier 0 Sedentary
Factors influencing serum caffeine concentrations following caffeine ingestion (Skinner et al., 2014)	 Trained male cyclists/ triathletes and active males "Trained" group cycled competitively for >1 season; and consistently trained at high volume and intensity for >6 months "Active" group completed >150 min physical activity per week but not currently nor previously involved in regular, high volume and/or intensity endurance training 	 "Trained" group had a VO₂max >60 ml.kg⁻¹.min⁻¹ 	"Trained" group meet physical activity guidelines (Bull et al., 2020), identify with a sport and are competing which justifies a Tier 2 classification. "Active" group also meet physical activity guidelines (Bull et al., 2020), but there is no information to suggest participants are competitive or identify with a specific sport and are therefore classified as Tier 1.	Comparison – Tier 1 vs Tier 2 (n = 1 study are assigned to <i>both</i> Tiers 1 and 2)

Table 2. Tiered ranking system to assess studies with female participants; studies are assessed on the basis of participant classification and methodological control.

Tion	Monstruel avala studios	Hormonal contraceptive studies		Menstrual
Tier	Wenstrual cycle studies	Oral contraceptive pill	Other	irregularities studies
Gold	Participants are eumenorrheic:	OCP use \geq 3 months prior to	HC use \geq 3 months prior to	Condition diagnosed
	• have MC lengths ≥ 21 days	recruitment (i.e., <u>length</u> of usage),	recruitment (i.e., <u>length</u> of	by medical
	and \leq 35 days resulting in 9	with the type (e.g., mono, bi or	usage), with the <u>type</u> (e.g.,	professional as part of
	or more consecutive periods	triphasic; combined or	implants, injections,	the study
	per year	progesterone-only and	intrauterine devices/coils that	
	• evidence of LH surge	formulation (name and	are hormone releasing and	Length of condition
	 correct hormonal profile 	concentration of exogenous	NOT copper-based, vaginal	stated
	[from blood sample analysis]	hormones) stated	rings, contraceptive	
	• no HC use 3 months prior to		transdermal patches) and	
	recruitment	Stipulate and consider OCP-	tormulation (e.g., combined	
		taking (i.e., active OCP) days and	or progesterone-only; names	
	MC characteristics are tracked for	OCP-free (i.e., inactive/placebo	and concentration of	
	\geq 2 months prior to testing	OCP) days	exogenous hormones) stated	
	Outcome measures are repeated in	One brand/type of OCP per group	One type of HC per group of	
	a second cycle	of participants	participants	
Silver	Participants are naturally	Two of three stated: OCP length	Two of three stated: HC	Condition diagnosed
	menstruating with ovulatory	of usage, type, formulation	length of usage, type,	by medical
	cycles:		formulation	professional not as
	• they experience menstruation,	Do/do not stipulate and consider		part of the study –
	with MC lengths ≥ 21 days	OCP-taking (i.e., active OCP)	One or more than one type of	self-reported or via
	and \leq 35 days	days and OCP-free (i.e.,	HC per group of participants	medical records
	• confirmed ovulation (LH) but	inactive/placebo OCP) days		T 1 0 11.1
	without correct hormonal			Length of condition
	profile	One or more than one brand/type		stated/not stated
		of OCP per group of participants		

	 prior HC use not stated or less than 3 months prior to recruitment MC characteristics are tracked for 1 month prior to testing Outcome measures not repeated in a second cycle 			
Bronze	 Participants are naturally menstruating: they experience menstruation, with MC lengths ≥ 21 days and ≤ 35 days without confirmed ovulation and correct hormonal profile prior HC use not stated or < 3 months prior to recruitment No tracking of MC characteristics prior to testing 	One of three stated: OCP length of usage, type, formulation Do not stipulate and consider OCP-taking (i.e., active OCP) days and OCP-free (i.e., inactive/placebo OCP) days More than one brand/type of OCP per group of participants	One of three stated: HC length of usage, type, formulation More than one type of HC per group of participants	Self-reported condition without medical diagnosis OR not specified if/how the condition was diagnosed. Length of condition not stated
Ungraded	Insufficient detail to award a gold, silver, or bronze	Insufficient detail to award a gold, silver, or bronze	Insufficient detail to award a gold, silver, or bronze	Insufficient detail to award a gold, silver, or bronze

Abbreviations: menstrual cycle (MC), luteinising hormone (LH), hormonal contraception (HC), oral contraceptive pill (OCP).

Study	Information provided relating to menstrual status	Female participant classification	Standard of methodological control
The influence of caffeine ingestion on strength and power performance in female team-sport players (Ali et al., 2016)	 Monophasic OCP Same hormonal composition (30 µg ethinyl oestradiol and 150 µg levonorgestrel) Used HC for >3 months prior to recruitment All testing was performed during days 5–8 and 18–22 of one pill-cycle 	OCP	Gold
Acute caffeine intake increases performance in the 15-s Wingate test during the menstrual cycle (Lara et al., 2020)	 The MC length ranged from 24-31 days No information regarding number of consecutive periods per year Urinary measurements of LH Hormonal profile was not retrospectively confirmed by blood sampling No HC use 1 month prior to testing Blood sampling was not used to retrospectively confirm MC phase Cycles were tracked for 4 months prior to testing Outcomes were not repeated in a second cycle 	MC Score 2/5 = naturally menstruating	Silver
Pre-exercise hyperhydration delays dehydration and improves endurance capacity during 2 h of cycling in a temperate climate (Goulet et al., 2008)	Self-reported amenorrheaLength of condition not stated	Menstrual irregularities	Bronze
Effect of creatine loading on anaerobic performance and skeletal muscle volume in NCAA Division I athletes (Ziegenfuss et al., 2002)	 No information on MC length No information regarding number of consecutive periods per year 	MC Score 0/5 = naturally menstruating	Ungraded

Table 3. Examples demonstrating how different studies would be assessed using our classification system.

	 No measurements of LH concentration Hormonal profile was not retrospectively confirmed by blood sampling HC users were excluded, but no time frame for minimum abstinence from HC prior to testing
	 Blood sampling was not used to retrospectively confirm MC phase Cycles were not tracked at all prior to testing Outcomes were not repeated in a second cycle
Effects of Three-Day Serial Sodium Bicarbonate Loading on Performance and Physiological Parameters During a Simulated Basketball Test in Female University Players (Delextrat et al., 2018)	No information provided regarding the menstrual Unclassified N/A status of participants.
Abbreviations: menstrual cycle (MC), lutein	ising hormone (LH), hormonal contraception (HC), oral contraceptive pill (OCP).

Table 4. Definitions and examples across the three themes (performance, health or indirect associations with performance or health) of SSSM research.

Themes	Definition	Examples
Performance	Studies measuring a performance outcome following an intervention or in association with a topic of interest.	Time trial, time to exhaustion, one-repetition maximum for whole body movements that are directly relevant to sporting actions (i.e., squats, deadlifts, bench press, power cleans), repeated sprints, Wingate tests
	Note: although this might not be the primary outcome, a direct measurement of performance is a high priority theme of sports and exercise science/medicine. Ideally, the measured performance outcome is validated within the sport of application.	Competition outcomes, tests of skill or tactics, movement patterns within a competition, sport-specific tests (e.g., serving accuracy, intermittent exercise protocols designed to replicate physiological, tactical and/or skill demands of team sports)
Health	Studies measuring outcomes related to health status or condition.	Illness (type, prevalence, recovery)
		Injury (type, prevalence, success of rehabilitation or recurrence)
	Note: Studies considered to have a health focus should attempt to measure a validated clinical or functional outcome.	Side effects/safety outcomes (functional outcomes) associated with an intervention or practice
		Disease or medical/clinical condition arising from, or associated with, sports involvement (e.g., asthma, diabetes, disordered eating)
		Doping outcomes (e.g., where effects on urine or blood markers used to detect anti- doping rule violations are measured)

Indirect associations with performance or health	Studies measuring a physiological/ psychological adaptation or response that may subsequently transfer to athletic performance or health, but performance or health was not directly measured.	Exercise economy, oxygen cost of exercise (e.g., VO ₂ max), cardiovascular and thermoregulatory responses to exercise, cardiac measures that underpin exercise (i.e., cardiac output/stroke volume/heart rate) Body composition, resting metabolic rate, gastrointestinal symptoms/discomfort
	Note: Studies that measure changes in body systems that suggest alterations or associations (correlations), but are not	strength, one-repetition maximum for isolated joint movements where the rest of the body is stationary and has minimal direct transfer to sports performance) Cognitive performance, visual attention/pupil dilation, neurological pathways
	necessarily validated, to health status but preclude a clinical diagnosis or functional manifestation of a health issue may belong in this classification.	Muscle damage, inflammation, oxidative stress, cortisol, mitochondrial respiration/biogenesis, pharmacokinetics Measures of fluid balance, such as diuresis or fluid retention
		Changes in markers of organ or system metabolism that suggest perturbations or acute/sub-clinical alterations rather than chronic or functional changes (e.g., changes in blood biochemistry and markers of immunology, hormonal, and metabolic profiles)

Study	Key relevant information provided	Theme	Explanation
Acute Ketogenic Diet and Ketone Ester Supplementation Impairs Race Walk Performance (Whitfield et al., 2021)	Elite race walkers undertook a four-stage exercise economy test and real-life 10,000 m race before and after a 5-day isoenergetic high-carbohydrate or low-carbohydrate, high-fat diet.	Performance	A 10,000 m race is a performance measure.
B-Alanine Supplementation's Improvement of High-Intensity Game Activities in Water Polo (Brisola et al., 2018)	The participants performed a simulated water polo game before and after the supplementation period.	Performance	A simulated match scenario is a performance measure.
Low bone mineral density is two to three times more prevalent in non-athletic premenopausal women than in elite athletes: a comprehensive controlled study (Torstveit & Sundgot-Borgen, 2005)	The study included a questionnaire (part I), measurement of bone mineral density (part II), and a clinical interview (part III). Bone mineral density was measured with dual energy x-ray absorptiometry. All scanning and analyses were conducted by the same operator.	Health	A Z-score < -2 indicates a bone density below the expected range for age (Lewiecki et al., 2004), while a Z-score < -1 is lower than expected for a weight-bearing athlete and warrants further investigation (Nattiv et al., 2007). Low bone mineral density may also be a key indicator of low energy availability in the correct clinical context.
Effect of Sodium Bicarbonate on [HCO3–], pH, and Gastrointestinal	Physically active subjects undertook 8 sodium bicarbonate experimental ingestion	Health	Gastrointestinal symptoms are a functional side effect of
Symptoms (Carr et al., 2011)	protocols and 1 placebo protocol. Capillary		sodium bicarbonate ingestion.

 Table 5. Examples demonstrating how different studies would be categorised into our proposed SSSM research themes.

	blood was taken every 30 min and analysed for pH and bicarbonate. Gastrointestinal symptoms were quantified every 30 min via questionnaire.		Although blood pH and bicarbonate concentration (indirect measures) were also measured. The health measure ranks higher in priority on our scale.
Gastrointestinal function during exercise: comparison of water, sports drink, and sports drink with caffeine (Van Nieuwenhoven et al., 2000)	Ten well-trained subjects underwent a rest- cycling-rest protocol three times. Measurements taken were oesophageal motility, gastroesophageal reflux, intragastric pH, orocecal transit time, intestinal permeability, glucose absorption and gastric emptying.	Indirect associations with performance or health	Multiple measures of gastrointestinal function, but without any measurements of functional gastrointestinal symptoms.
Neither Beetroot Juice Supplementation nor Increased Carbohydrate Oxidation Enhance Economy of Prolonged Exercise in Elite Race Walkers (Burke et al., 2021)	Measured carbohydrate and fat oxidation, plasma nitrate and nitrate concentrations, alongside oxygen uptake across a 26 km race walking protocol.	Indirect associations with performance or health	Measured substrate oxidation and exercise economy, with no measure of performance.
A short-term ketogenic diet impairs markers of bone health in response to exercise (Heikura et al., 2019)	Measured markers of bone modelling/remodelling after short-term ketogenic low-carbohydrate, high-fat diet. Markers included serum markers of bone breakdown (cross-linked C-terminal telopeptide of type I collagen), formation (procollagen 1 N-terminal propeptide) and metabolism (osteocalcin).	Indirect associations with performance or health	Only measured acute bone markers with no information on bone density over a chronic period.



Lack of sport/exercise research on women

Systematically and expediently direct the priorities for future research

Figure 1. Concept flowchart illustrating how our "standardized auditing protocol" will be utilized and the associated benefits.



Figure 2. Flowchart to be used in reporting results. Bold text denotes papers, non-bold typeface is used for studies. From section B onwards, all counts and percentages refer to individual studies (not papers). Symbols should be used after the reported percentage in section B to denote studies directly comparing different athletic tiers. Male (M), female (F), menstruating (men), hormonal contraception (HC).