Developing resources for high-quality applied research investigating the effect of female sex hormones on performance

Smith, Eleanor

Smith, E. (2025). Developing resources for high-quality applied research investigating the effect of female sex hormones on performance [PhD Thesis]. Australian Catholic University. https://doi.org/10.26199/acu.919y5

Developing resources for high-quality applied research investigating the effect of female sex hormones on performance

Submitted by

Eleanor Smith

Student ID: S00327548

BSc (Hons) Sport and Exercise Science, MSc Exercise Physiology

A thesis submitted in total fulfilment of the requirements of the degree of

Doctor of Philosophy

Exercise and Nutrition Research Program

Mary McKillop Institute for Health Research

Faculty of Health Sciences

Australian Catholic University

Principal supervisor: Professor Louise Burke

Co-supervisor: Dr Alannah McKay

Co-supervisor: Professor Kirsty Elliott-Sale

Co-supervisor: Dr Trent Stellingwerff

Co-supervisor: Dr Kathryn Ackerman

End-user supervisor: Dr Rachel Harris

3rd January 2025

Declaration of Authorship and Sources

This thesis contains no material that has been extracted in whole or in part from a thesis that I have submitted towards the award of any other degree or diploma in any other tertiary institution.

No other person's work has been used without due acknowledgment in the main text of the thesis.

All research procedures reported in the thesis received the approval of the relevant Ethics/Safety Committees (where required).

Eleanor Smith
03/01/2025

Acknowledgements

Wow, I could not have asked for a better supervisory team to guide me through this PhD experience. When I first applied for this PhD I believed that the supervisory panel was, quite literally, the best I could ever dream of. Somehow, my dreams came true and here we are!

Louise, you supported me through the significant move from the UK, understanding the challenges of being far away from friends and family, and you continued to provide unwavering support throughout. From day one you made me feel part of the group. Even while I was stuck in the UK due to COVID, your guidance and opportunities to contribute to the group from afar made all the difference.

AKA thank you for keeping me on track and answering my stupid questions!

Trent, Kirsty, Kate and Rachel, simply thank you. I feel so lucky to have many diverse perspectives contribute to this work. I can wholeheartedly say it would not have been the same without each of you, and I am excited to hopefully continue to collaborate in the future!

Megan, thank you for being my PhD buddy. I will never forget the countless hours spent at a deserted AIS on the weekend weighing out M&Ms and smashing glass bottles!

To the rest of the **ENRP "sport dream team"**. Thank you for being there throughout this entire process. From stepping in to help with data collection when an extra set of hands was needed, to providing constant support, your contributions have been invaluable.

Thank you to all the **athletes** who have literally given their blood, sweat and tears to make this research possible. Your dedication and commitment were the foundation of this work, and none of it would have been achievable without you.

Finally, to my amazing friends and family, thank you for the ongoing support and many laughs throughout.

Table of Contents

List of Publications

Publications relating to the thesis

Smith, E. S., Kuikman, M. A., Weakley, J., Tee, N., McCormick, R., Ackerman, K. E., Elliott-Sale, K. J., Stellingwerff, T., Harris, R., McKay, A. K., & Burke, L. M. Effects of 24-hour diet- or exercise-induced energy availability manipulations on substrate utilisation and performance. *Medicine and Science in Sports and Exercise*, (aop). https://doi.org/10.1249/MSS.0000000000003608.

Smith, E. S., Weakley, J., McKay, A. K., McCormick, R., Tee, N., Kuikman, M. A., Harris, R., Minahan, C.L., Buxton, S., Skinner, J., Ackerman, K. E., Elliott-Sale, K. J., Stellingwerff, T., & Burke, L. M. (2024). Minimal influence of the menstrual cycle or hormonal contraceptives on performance in female rugby league athletes. *European Journal of Sport Science, 24*(8), 1067-1078. https://doi.org/10.1002/ejsc.12151.

Smith, E. S., McCormick, R., McKay, A. K., Ackerman, K. E., Elliott-Sale, K. J., Stellingwerff, T., Harris, R., & Burke, L. M. Perceived negative menstrual cycle symptoms, but not changes in oestrogen or progesterone, are associated with impaired cycling race performance. *Medicine and Science in Sports and Exercise*, (aop). https://doi.org/10.1249/MSS.0000000000003587.

Smith, E. S., McKay, A. K., Ackerman, K. E., Elliott-Sale, K. J., Stellingwerff, T., Harris, R., & Burke, L. M. Original investigation: Female athletes report positive experiences as research participants. *International Journal of Sport Nutrition and Exercise Metabolism,* (in review).

Additional publications during the candidature

2022 McKay, A. K., Stellingwerff, T., **Smith, E. S.,** Martin, D. T., Mujika, I., Goosey-Tolfrey, V. L., Sheppard, J. & Burke, L. M. (2022). Defining training and performance calibre: a participant classification framework. *International Journal of Sports Physiology and Performance*, *17*(2), 317-331.

2022 **Smith, E. S.,** McKay, A. K., Ackerman, K. E., Harris, R., Elliott-Sale, K. J., Stellingwerff, T., & Burke, L. 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), 114-127.

2022 **Smith, E. S.,** McKay, A. K., Kuikman, M., Ackerman, K. E., Harris, R., Elliott-Sale, K. J., Stellingwerff, T., & Burke, L. M. (2022). Auditing the representation of female versus male athletes in sports science and Sports Medicine research: evidence-based performance supplements. *Nutrients*, 14(5), 953.

2022 **Smith, E. S.,** McKay, A. K., Kuikman, M., Ackerman, K. E., Harris, R., Elliott-Sale, K. J., Stellingwerff, T., & Burke, L. M. (2022). Managing female athlete health: Auditing the representation of female versus male participants among research in supplements to manage diagnosed micronutrient issues. *Nutrients*, 14(16), 3372.

2023 Kuikman, M. A., McKay, A. K., **Smith, E. S**., Ackerman, K. E., Harris, R., Elliott-Sale, K. J., Stellingwerff, T., & Burke, L. M. (2023). Female athlete representation and dietary control methods among studies assessing chronic carbohydrate approaches to support training. *International Journal of Sport Nutrition and Exercise Metabolism,* 33(4), 198-208.

2023 Kuikman, M. A., **Smith, E. S**., McKay, A. K., Ackerman, K. E., Harris, R., Elliott-Sale, K. J., Stellingwerff, T., & Burke, L. M. (2023). Fueling the female athlete: auditing her representation in studies of acute carbohydrate intake for exercise. *Medicine and Science in Sports and Exercise*, 55(3), 569.

2023 Carr, A. J., McKay, A. K., Burke, L. M., **Smith, E. S.**, Urwin, C. S., Convit, L., Jardine, W.T., Kelly, M.K., & Saunders, B. (2023). Use of Buffers in Specific Contexts: Highly trained female athletes, extreme environments and combined buffering agents — a narrative review. *Sports Medicine*, 53(Suppl 1), 25-48.

2024 Kelly, M. K., **Smith, E. S.**, Brown, H. A., Jardine, W. T., Convit, L., Bowe, S. J., Condo, D., Guy, J.H., Burke, L.M., Periard, J.D., Snipe, R.M.J, Snow, R. & Carr, A. J. (2024). Auditing the representation of females versus males in heat adaptation research. *International Journal of Sport Nutrition and Exercise Metabolism*, 34(2), 111-121.

2024 Kuikman, M. A., McKay, A. K., Minahan, C.L., Harris, R., Elliott-Sale, K. J., Stellingwerff, T., **Smith, E. S**., McCormick, R., Tee, N., Skinner, J., Ackerman, K. E., & Burke, L. M. (2024). Effect of menstrual cycle phase and hormonal contraceptives on resting metabolic rate and body composition. *International Journal of Sport Nutrition and Exercise Metabolism*, 34(4), 207-217.

2024 Flood, T.R., Clausen, E., Kuikman, M. A., **Smith, E. S**., McKay, A. K., Burke, L. M., & Elliott-Sale, K. J. (2024). Auditing the representation of hormonal contraceptives in studies assessing exercise performance in women. *Journal of Sport Sciences,* 42(9), 825-839.

2024 **Smith, E. S**., & Burke, L. M. (2024). Have we considered women in current sports nutrition guidelines? *Nutrition Today,* 59(4), 168-176.

2024 Kuikman, M. A., **Smith, E. S**., McKay, A. K., McCormick, R., Ackerman, K. E., Harris, R., Elliott-Sale, K. J., Stellingwerff, T., & Burke, L. M. (2024). Impact of acute dietary and exercise manipulation on RMR measurements and DXA body composition estimates. *Medicine and Science in Sports and Exercise,* (aop). https://10.1249/MSS.0000000000003555

List of Figures

in figures 3.4A and 3.4C represent the standard deviations. *iAUC*; incremental area under the curve, *LEA*, low energy availability; *HEA*, high energy availability; *GEA,* high energy availability for mass gain/growth. **denotes significance *p<0.01.*

Figure 3.5. (A) relative peak power, and (B) relative mean power during a 30 s Wingate performance task. Data is expressed as mean±standard deviation, displayed across conditions. Colours denote sex. *LEA*, low energy availability; *HEA*, high energy availability; *GEA,* high energy availability for mass gain/growth.

Figure 4.1. Study overview. Performance testing occurred at either (A) phases 1 [low oestrogen/progesterone concentration (day 1.8±0.4)], 2 [high oestrogen and low progesterone (day 11.4±1.4)] and 4 [moderate oestrogen and progesterone (day 20.8±1.6)] for athletesNM. The days reported refer to the cycle day on which the test was conducted; (B) three equally spaced timepoints for athletesHC utilising the implant or hormonal injection or, three equally spaced timepoints avoiding the withdrawal bleed for athletesHC using the oral contraceptive pills. It should be noted that the concentration of exogenous progestin following the implant and injection gradually declines with time (Huber, 1998), and hence the exact hormonal profile is dependent on the date of the implant or injection. Time points are displayed according to an idealised 28-day cycle. (C) Performance testing schedule. IMTP; isometric mid-thigh pull, NM; naturally menstruating, HC; hormonal contraceptives, OCP; oral contraceptive pills. *Figure created with BioRender.com*.

Figure 4.2. Serum oestradiol concentration across the three tests in (A) naturally menstruating athletes and (B) athletes using hormonal contraception (n=1 outlier removed during test three). Serum progesterone concentration across the tests in (C) naturally menstruating athletes (n=1 outlier removed during phase 4) and (D) athletes using hormonal contraception. Calculated free testosterone across the tests in (E) naturally menstruating athletes and (F) athletes using hormonal contraception. Total testosterone across the tests in (G) naturally menstruating athletes and (H) athletes using hormonal contraception. Black lines denote mean values. *denotes significance *p<0.05,* **denotes significance *p<0.001.*

Figure 4.3. Performance outcomes at each test, alongside relative mean power during the countermovement jump and impulse at 50 ms during the squat jump, between naturally menstruating athletes (athletesNM) and athletes using hormonal contraception (athletesHC). *denotes significance *p<0.05.*

Figure 5.1. Study protocol. Participants completed a monitoring period (to capture both their menstrual cycle and training) of between four and six weeks prior to, or after if necessary, the commencement of the four-week race series. *The blood sample was collected either the morning of (Thursday) or after (Friday) each race at the same time each week.

Figure 5.2. Participant flow chart from pre-screening to the final sample size included for analysis. *Exclusion criteria: living outside of Australia, exercising <150 min per week, no access to Zwift cycling app, use of hormonal contraceptives within three months of study commencement, currently pregnant or breastfeeding, current amenorrhea.

Figure 5.3. Repeated measures correlation between race completion time and (A) oestradiol concentration, and (B) progesterone concentration. Different colours represent individual participants.

Figure 5.4. Race time, separated by (A) races when athletes were bleeding compared to the mean time across non-bleeding race days (n=24), (B) days when athletes ovulated on race day compared to the mean time across other races (n=9), and (C) between follicular and luteal phases, as separated by ovulation for ovulatory athletes who completed at least one race in each phase (n=31). The different colour lines and symbols represent individual participants: triangle denotes participants present on all three graphs, a circle for two graphs, and a square for one graph.

Figure 5.5. Repeated measures correlation between race completion time and (A) the number of "moderate" severity GI symptoms reported pre-race, (B) the number of "moderate" severity GI symptoms reported post-race, and (C) the total number of MC symptoms recorded on the day of the race. Different colours represent individual participants. *GI;* gastrointestinal, *MC;* menstrual cycle.

Figure 5.6. Repeated measures correlation between race competition time and (A) the change in participant ratings of thermal sensation pre-to-post race, and (B) the change in participant ratings of thermal comfort pre-to-post race. Different colours represent individual participants. *TS*; thermal sensation, *TC;* thermal comfort.

Figure 6.1. Reasons for study participation were rated between 0-100 according to their perceived importance across all four studies, whereby "0" represented that the reason "did not influence their decision to participate at all" and "100" was "very much influenced their decision". The box plot displays the median, first and third quartiles, with symbols displaying outliers outside of the 5-95th percentile. Data is presented for female athletes only, and the total participant number (out of a maximum of n=89) for each reason is provided on the y axis labels (as some aspects were not applicable in all studies). Clear bars represent study aspects only present in one out of four studies [food provision in Study 1 (Smith et al., 2024a) and prize money in Study 3 (Smith et al., 2024b)], dashed bars denote aspects present in two studies, and solid bars represent aspects present in three or four studies. "Living environment" applies to training camp studies where participants stayed on site and includes the accommodation and food provided, while "food provision" applies to food given to participants to eat in free living conditions. "Financial reimbursement" was given to all participants following study completion, whereas "prize money" was only won by certain participants based on their race performance in Study 3.

Figure 6.2. Participant rating of each test/procedure completed, whereby "1" represented "the worst experience" and "10" was "the best experience". The "test experience" refers to the experience of the test itself (e.g., the actual process of undertaking a blood test or using the at-home device to measure sleep), while "test feedback" refers to the individual feedback that participants received post-study (e.g., blood test/ sleep summary results). Data presented as the median, with error bars denoting the interquartile range. Data is presented for female athletes only with the total participant number (out of a maximum of n=89) for each test provided on the y axis labels, as most tests were not present in all studies. Wilcoxon signed-rank tests compared within-participant ratings of the test experience *vs* feedback. **denotes $p<0.01$. *HRV*; heart rate variability, *DXA*; dual x-ray absorptiometry, *MC*; menstrual cycle, *RMR;* resting metabolic rate.

Figure 6.3. Participants' perceived importance of differing reasons for study participation, separated according to males and females in Study 1 only (n=20, 10M, 10F). Reasons for study participation were rated between 0-100 according to their perceived importance, whereby "0" represented that the reason "did not influence their decision to participate at all" and "100" was "very much influenced their decision". Data presented as mean±standard deviation. Independent t-tests compared the perceived importance of each distinct reason between males and females.

Figure 6.4. Participants' perceived importance of differing reasons for study participation, separated according to athletic tier as defined by (McKay et al., 2022b) (n=70 Tier 2-3, n=19 Tier 4-5). Reasons for study participation were rated between 0-100 according to their perceived importance across all four studies, whereby "0" represented that the reason "did not influence their decision to participate at all" and "100" was "very much influenced their decision". Data is presented for female athletes only, the total participant number for each reason is provided on the y axis labels [n=19 Tier 4-5 athletes is consistent across all reasons, while the number of Tier 2-3 athletes varied (out of a maximum of n=70) as some reasons were not applicable in all studies]. Data presented as the median, with error bars denoting the interquartile range. Mann Whitney U tests compared the perceived importance of each distinct reason between athletic tiers. *denotes $p<0.05$.

Figure 7.1. Flowchart detailing the process of implementing each tier of methodological control.

List of Tables

List of Abbreviations

ANOVA	Analysis of variance
BBT	Basal body temperature
CHO	Carbohydrate
CMJ	Countermovement jump
CI	Confidence interval
CV	Coefficient of variation
DSI	Dynamic strength index
DXA	Dual-energy X-ray absorptiometry
EA	Energy availability
EEE	Exercise energy expenditure
EI	Energy intake
EUR	Eccentric utilisation ratio
FATMAX	Exercise intensity at which maximal fat oxidation occurs
FFM	Fat free mass
FSH	Follicle stimulating hormone
GEA	Higher energy availability for mass gain/growth
GI	Gastrointestinal
HC	Hormonal contraception
HEA	High energy availability
HR	Heart rate
iAUC	Incremental area under the curve
IMTP	Isometric mid-thigh pull
IUD	Intrauterine device
IUS	Intrauterine system
LBM	Lean body mass
LEA	Low energy availability
LH	Luteinising hormone
MC	Menstrual cycle
MI	Menstrual irregularities
MMTT	Mixed meal tolerance test
OCP	Oral contraceptive pill
PCOS	Polycystic ovary syndrome
P:E	Progesterone: oestrogen ratio

PFO	Peak fat oxidation
RCT	Randomised control trial
REDs	Relative energy deficiency in sport
RER	Respiratory exchange ratio
RFD	Rate of force development
RMR	Resting metabolic rate
RPE	Rating of perceived exertion
RSI	Reactive strength index
SD	Standard deviation
SES	Sport and exercise science
SJ	Squat jump
TC	Thermal comfort
TS	Thermal sensation
TT	Time trial
TTE	Time to exhaustion
$\dot{V}O_2$	Rate of oxygen utilisation
$\dot{V}O_2max$	Rate of maximal oxygen utilisation

Abstract

Female athletes are substantially underrepresented in the sports science literature, with an acute scarcity of studies examining performance outcomes while appropriately controlling for menstrual status. This lack of high-quality research does a disservice to female athletes, as research conducted in men cannot be directly applied to women without consideration of sexual dimorphisms. As such, a correction of the sex- and gender-biases in the sport and exercise science (SES) literature is of primary importance. This thesis addresses dual aims: examining a range of methodologies and menstrual statuses to develop resources that support researchers in conducting high-quality applied SES research in female athletes, alongside examining the potential influence of ovarian hormones on performance.

Study 1 (chapter 3) examined sex differences in the responses to 24-hour manipulations in energy availability (EA) on substrate oxidation, metabolism, and performance, among 20 cyclists/triathletes [10 females using combined monophasic oral contraceptive pills (OCP) and 10 males]. Five different conditions of EA were implemented through a randomised control trial (RCT) in a Latin square design: low/high/higher EA (LEA/HEA/GEA) at 15/45/75 kcal·kg^{-1} FFM·day^{-1}. Conditions of LEA and HEA were separately achieved by manipulations of either energy intake (EI) or exercise energy expenditure (EEE). The following day fasted peak fat oxidation (PFO) during cycling and two-hour postprandial metabolism were assessed, alongside several physical and cognitive performance tests. Among both sexes, the highest PFO occurred under LEA induced by exercise ($p<0.01$), while postprandial glucose ($p<0.01$) and insulin ($p<0.05$) responses were highest when LEA was induced by diet. Performance during a Wingate test, countermovement jump (CMJ), isometric mid-thigh pull (IMTP), and Stroop Test, did not differ with altered EA ($p>0.05$). Lastly, the response to EA manipulations did not differ between the sexes across any outcome measure.

Study 2 (chapter 4) employed an observational study design in a training camp environment. National Rugby League Indigenous Women's Academy athletes [n=11 naturally menstruating (NM), n=13 using HC] completed performance tests [CMJ, squat jump (SJ), IMTP, 20 m sprint, power pass and Stroop Test] during phases 1,2 and 4 of a single menstrual cycle (MC), or during three weeks of hormonal contraceptive (HC) use. Menstrual status was confirmed through ovulation tests alongside serum oestrogen and progesterone concentrations. MC phase or HC use did not influence jump height, peak force, sprint time, distance thrown or the Stroop effect. However, there were small variations in kinetic and kinematic CMJ/SJ outputs. NM athletes produced greater mean concentric power in MC phase 4 than 1 [+0.41 W·kg^{-1} (+16.8%), $p=0.021$] during the CMJ, alongside greater impulse at 50 ms at MC phase 1 than 4 [+1.7 N·s (+4.7%), $p=0.031$] during the SJ, without differences between tests for HC users.

Study 3 (chapter 5) presents a novel observational study design among 37 cyclists/triathletes not using HC. Participants completed four separate virtual e-cycling races [19.5 km time trial (TT)] across a one-month period. At each race, individual hormonal concentrations were correlated to race completion time, with no relationships observed between either 17-β-oestradiol (r=-0.001, *p=0.992*) nor progesterone (r=-0.023, *p=0.833*) and TT time. Perceived negative symptoms at each race were also recorded. The total number of MC symptoms was positively correlated to a slower TT time [r=0.268 (95% CI 0.056 to 0.457), *p=0.014*], as was the number of GI symptoms of at least "moderate" severity pre-race (r=0.233 [95% CI 0.021 to 0.425], *p=0.031*), but not post-race (r=0.022, *p=0.841*).

Study 4 (chapter 6) collates the experiences of the female athletes (n=70) participating in all three studies (plus an additional 19 elite athletes), to inform future research and increase female participation rates. The majority (81%) had never participated in research before, with a perceived lack of opportunities as the primary barrier (93%). Participants rated an interest in the research outcome as the most important aspect influencing their decision to participate, followed by receiving personalised results and education. Most participants (87%) stated that they would apply the study findings to their sports involvement, and the majority (94%) indicated a willingness to participate in future studies. The research experience was also rated positively at a mean 77 out of 100.

This series of research studies has contributed numerous novel findings to the literature regarding the influence of the female sex hormones on athletic performance, alongside exploring the nuances of female research participation. Collectively, the findings of this thesis can be summarised as follows:

(1) The responses to acute (24 h), severe, LEA did not differ between sexes. Exercise-induced LEA influenced fasted substrate oxidation more than diet-induced LEA, while LEA did not impair strength/power, sprint capacity, or cognitive performance and so could be implemented by both male and female athletes in the immediate 24 h prior to competition. Future research may consider if sex differences in these responses might be observed across longer LEA time periods, or at a higher EA (less severe LEA dose), alongside examining the influence of fluctuations in oestrogen and progesterone on these findings by focussing on women not using HC.

(2) Performance across tests of explosive power and strength do not differ across a MC, or during active HC use. As such, evidence is currently insufficient to justify "menstrual phase- or status-based" testing programs at a group to team-based level. There were small variations in kinetic and kinematic CMJ/SJ outputs in NM athletes, without differences between tests for HC users. However, it could not be determined if the observed alterations exceeded between-day variability. Further research is therefore required to fully understand the effects of

oestradiol and progesterone on performance, alongside examining MC phases 2 and 3 which were not captured in this study.

(3) Fluctuations in oestrogen and progesterone concentrations across a MC do not appear to affect real-world cycling performance, while perceived negative MC and GI symptoms may relate to impaired performance. Therefore, it may be important to focus on the management of negative symptoms to mitigate performance decline. However, a standardised and validated questionnaire to assess the incidence and severity of MC-related symptoms does not currently exist and hence its development should be a priority. Further investigation is also required to understand if specific symptoms are driving an association with performance, and if this relationship persists into other activities beyond cycling.

(4) Female athletes are willing and interested in research participation, but the vast majority (93%) perceive a lack of opportunity as the main barrier to do so. The main motivations for female athletes to participate were an interest in the research outcomes and to receive personalised results. As such, it may be prudent to focus future recruitment efforts on emphasising these study aspects.

Overall, these outcomes demonstrate a lack of association between ovarian hormone fluctuations, either across a MC or with HC use, and the performance measures assessed, either in the laboratory setting as observed during Study 2 or a real-world race as measured during Study 3. As such, an individualised approach to MC monitoring/tracking, with a focus on individual reported symptoms, is likely to represent the current best-practice approach for athletes given the lack of consistent response at the group level. Meanwhile, the overarching aim of this thesis was to examine the practicalities of implementing high quality classification and control of menstrual status in the applied research setting. A range of resources have been developed to aid this pursuit, whilst the methodological decisions and practical learnings from each of the experimental studies are detailed to inform future high-quality research and ultimately facilitate a correction of the sex-bias in SES research.

Introduction and Overview

The Paris 2024 Olympic Games represented a landmark for female athletes: the first Games in history to achieve parity in participation rates, with female athletes contributing exactly half of the total competitors. This equality has taken 124 years to achieve, with just 2% of competitors at the inaugural Olympic Games being women. However, at Paris 2024, there was still fewer total medals handed out to female athletes (152 medal events for females vs. 157 medal events for men). Unfortunately, the elevated participation rates of women in sport has not been mirrored in the SES literature, with numerous reports detailing that women contribute on average between just 11-42% of the total participant count (Brookshire, 2016; Costello et al., 2014; Cowley et al., 2021; Hutchins et al., 2021; Kelly et al., 2024; Kuikman et al., 2023a; Kuikman et al., 2023b; Smith et al., 2022c). The bias against female participants becomes particularly stark within the literature specifically evaluating performance interventions, and those including highly trained elite athletes and studies using best-practice methodologies for assessing participant menstrual status. The lack of female participation is therefore problematic, as conclusions drawn from research in men cannot be directly translated to female athletes without adequate consideration of sexual dimorphisms or event-specific demands.

Interest in the area of female athlete research has recently surged, with the annual outputs pertaining to "female athletes and sports performance" increasing exponentially across the last 15 years. However, a large proportion of current research is unfortunately hindered by poor methodological classification and control of participant menstrual status (Elliott-Sale et al., 2020b; McNulty et al., 2020b). This is important as the female sex hormones, oestrogen and progesterone, are purported to influence numerous physiological systems associated with performance. Several guidelines have been published providing comprehensive recommendations on the appropriate methodologies for the classification and control of participant menstrual status (Elliott-Sale et al., 2021; Janse de Jonge et al., 2019). However, in practice when studying elite athletes, we acknowledge that it can be challenging to achieve such a rigorous level of control and stringent inclusion/exclusion criteria.

This thesis therefore addresses dual aims: developing resources to aid researchers in conducting high-quality studies among female athletes in the applied setting, alongside further elucidating any influence of ovarian hormonal profiles on sports performance. Each study will implement best-practice methods relating to menstrual status classification and control (Elliott-Sale et al., 2021) across a range of distinct study designs, inclusive of women with different menstrual statuses. Our experiences will then be utilised to develop tools to aid the pursuit of future high-quality research. Additionally, each study examines performance outcomes to address the acute lack of research employing high-quality methodological classification and control of menstrual status in this area. **Chapter 1** provides a

literature review outlining the intricacies of sexual dimorphisms, and the current state of the literature regarding our knowledge of the influence of ovarian hormones on athletic performance. **Chapter 2** outlines detailed study methodologies in addition to that provided in the respective chapters, as formatted for publication. **Chapters 3-6** include the research papers aiming to answer the following questions through distinct study designs as described below. **Chapter 7** discusses the novel findings of this thesis, alongside providing resources to aid future high-quality research.

Chapter 3 (*study 1):* *Effects of 24-hour diet- or exercise-induced energy availability manipulations on substrate utilisation and performance.*

- Examines sex differences in substrate oxidation, postprandial metabolism, and performance in response to 24-hour manipulations in EA, induced by manipulations to EI or EEE.
- A traditional laboratory-based RCT, with rigorous study controls and ongoing participant recruitment across an eight-month period. Examined exclusively women using combined monophasic OCPs.

Chapter 4 (*study 2*): *Minimal influence of the menstrual cycle or hormonal contraceptives on performance in female rugby league athletes.*

- Examines measures of explosive power/strength across a single MC and between athletes using HC and those with "natural" cycles.
- An observational study design across a five-week research-embedded training camp, with the MC investigated in terms of discrete phases.

Chapter 5 (*study 3*): *Perceived negative menstrual cycle symptoms, but not changes in oestrogen or progesterone, are associated with impaired cycling race performance.*

- Examines the relationship between oestrogen and progesterone, alongside perceived negative symptoms, on performance during a cycling competition among female athletes not using HC.
- Employed a novel, observational study design with an ecologically valid measurement of real-world performance (virtual e-cycling competition). Used a correlational approach to examine the relationship between oestrogen/progesterone concentration and race performance.

Chapter 6 (*study 4*): *Original investigation: Female athletes report positive experiences as research participants.*

- Collates the experiences of the female athletes who participated in the studies conducted during this thesis (plus an additional 19 elite athletes). Highlights nuances of their participation across a range of studies with different methodological characteristics.

Chapter 1: Review of the literature

The sex-gap in sports science research

Recent decades have seen a surge in the participation of women in high performance sport. Indeed, the Paris 2024 Olympic Games was the first to achieve parity in participation rates, with women accounting for exactly half of total competitors (International Olympic Committee, 2024). This represents an increase from 42% representation in Beijing 2008 and just 30% in Barcelona 1992 (International Olympic Committee, 2021). However, Paris 2024 still handed out fewer medals to female compared to male athletes. Gender equality is a key mission of the International Olympic Committee. As outlined by the Olympic Charter, the International Olympic Committee has a role: "to encourage and support the promotion of women in sport at all levels and in all structures, with a view to implementing the principle of equality of men and women" (International Olympic Committee, 2019). The Paralympic Games, however, trails behind with women accounting for just 45% of competitors during the Paris 2024 Games (International Paralympic Committee, 2024). Nevertheless, in other environments outside of the Olympic movement, women's sport is gaining in popularity across the globe, exemplified by the normalisation of demands for gender equality in pay and media coverage, alongside increases in the number of professional female athletes/teams (Claus, 2020; Douglas, 2018; Townes, 2019).

Unfortunately, the elevated participation rates of women and girls in sport has not been mirrored by an increase in sex-specific research in the SES discipline. There are numerous reports of a conspicuous imbalance of participation rates between men and women across numerous areas of SES research, with various audits reporting that women contribute on average between 11-42% of the total participant count in the literature of interest (Brookshire, 2016; Costello et al., 2014; Cowley et al., 2021; Hutchins et al., 2021; Kelly et al., 2024; Kuikman et al., 2023a; Kuikman et al., 2023b; Smith et al., 2022c). Moreover, of the studies including women, few (4-13%) examine female participants in isolation in comparison to the 18-79% of studies that focus exclusively on men (Brookshire, 2016; Costello et al., 2014; Cowley et al., 2021; Hutchins et al., 2021; Kelly et al., 2024; Kuikman et al., 2023a; Kuikman et al., 2023b; Smith et al., 2022c). The bias against female participants becomes particularly stark within the literature that specifically evaluates interventions or outcomes that affect sports performance, where just 3-16% of the total participants are women (Brookshire, 2016; Kelly et al., 2024; Kuikman et al., 2023a; Kuikman et al., 2023b; Smith et al., 2022c). Moreover, the sex gap is further amplified when examining participation rates among highly trained/elite female athletes. A recent project undertook audits of the SES literature pertaining to performance (Smith et al., 2022c) and medical (Smith et al., 2022a) supplements, carbohydrate (CHO) fuelling strategies (Kuikman et al., 2023a; Kuikman et al., 2023b) and heat adaptation (Kelly et al., 2024) using a standardised tool that included a systematic characterisation of athlete calibre. Across these SES themes, 4-33% of women included in studies were identified as Tiers 3-5 [i.e., at least highly trained/national level athletes, (McKay et al., 2022b)].

Interestingly, although the total number of elite female athletes participating in research is fewer than their elite male counterparts, across many themes, they made up a higher proportion of the participant pool than was seen with male participants (Kelly et al., 2024; Kuikman et al., 2023b; Smith et al., 2022a; Smith et al., 2022c).

Numerous biological, phenotypical, and social characteristics distinguish male and female athletes, alongside differences in sporting event characteristics. Indeed, women's sport commonly differs from those of her male counterparts in terms of both the event demands and typical playing styles, such as lighter equipment or shorter distances for female athletes (Kovalchik & Reid, 2017; Sanders et al., 2019). Moreover, at an early age, girls often face disparities in opportunities, including access to skilled coaches, quality equipment, and adequate training support. The lack of female participation in SES research is therefore problematic, as conclusions drawn from research in men cannot be directly translated to female athletes without adequate consideration of sexual dimorphisms or event-specific demands.

Sexual dimorphisms

Multiple biological and phenotypical differences exist between the sexes, which in turn can influence numerous physiological parameters and ultimately, sports performance (Devries, 2016; Devries et al., 2006; Green et al., 2016; Hunter, 2016; McNulty et al., 2020b; Sheel et al., 2016). Of particular importance are the fluctuations in the ovarian hormones, oestrogen, and progesterone across the lifespan of a woman (Elliott-Sale et al., 2021). A lack of consideration for these cyclical hormonal fluctuations can result in methodologies that are poorly designed and/or executed in studies which include female participants (Elliott-Sale et al., 2021).

The menstrual cycle

The MC is a biological rhythm, characterised by repeating cyclical patterns of production and secretion of the ovarian hormones oestrogen and progesterone. Concentrations of these hormones exhibit substantial inter- and intra-individual variability, both acutely over a given MC, during subsequent MC's and chronically across a lifespan (e.g., prepuberty, pregnancy and menopause) (Elliott-Sale et al., 2021). A eumenorrheic cycle lasts between 21-35 days. In brief (Reed & Carr, 2000), each cycle starts with the development of an ovum inside a follicle within the ovaries (follicular phase). Luteinising hormone (LH) and follicle stimulating hormone (FSH) then trigger the follicle to release the ovum during ovulation. During the luteal phase the corpus luteum develops from the follicle constituents, secreting progesterone, and oestrogen to develop the uterus into an environment to facilitate foetal growth if the ovum is fertilised. Without fertilisation, the corpus luteum breaks down and the uterus lining is shed during menses.

The MC can be broadly considered as two distinct phases, separated by ovulation: the follicular phase which occurs prior to ovulation, and the luteal phase following ovulation. These phases can be further subdivided (Figure 1.1), resulting in four distinct phases (Elliott-Sale et al., 2021) (the days noted in brackets assume an idealised 28-day cycle):

Phase 1. **Early follicular** (days 1-5): signified by the onset of menstrual bleeding, both oestrogen and progesterone are low.

Phase 2. **Late follicular** (14-26 hours prior to ovulation and the LH surge): oestrogen concentration peaks, progesterone concentration remains low.

Phase 3. **Ovulatory** (around days 14-15): indicated by a positive urinary ovulation test, lasts 24-36 hours. LH concentration surges, oestrogen is higher than phase 1, but lower than phases 2 and 4, progesterone concentration starts to rise above the levels observed in phases 1 and 2.

Phase 4. **Mid-luteal** (7 days following ovulation): oestrogen is higher than phase 1 and 3, but lower than phase 2, progesterone reaches its peak concentration.

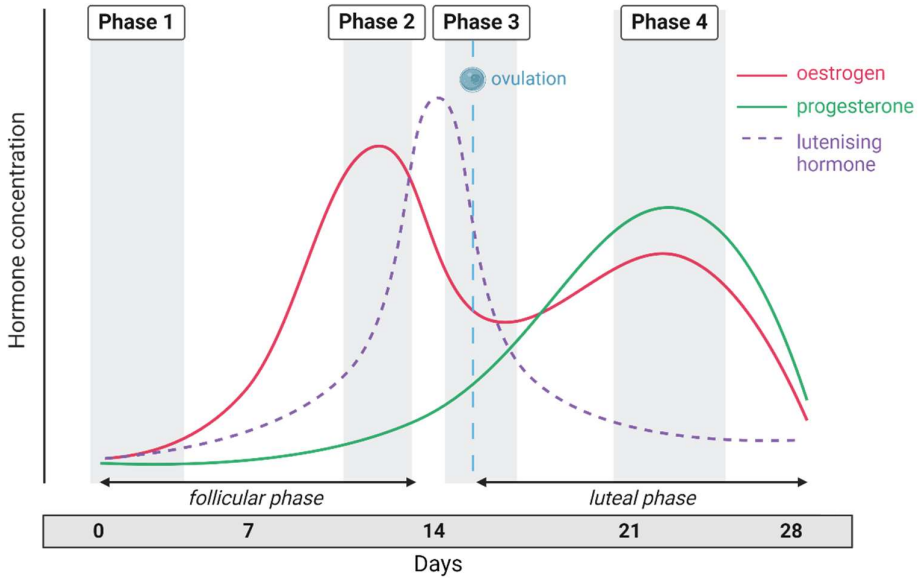


Figure 1.1. A visualisation of the fluctuations in the hormonal profiles of endogenous oestrogen, progesterone, and luteinising hormone across each phase of an idealised 28-day MC, adapted from Elliott-Sale et al. (2021). *Created with BioRender.com.*

In research settings, phases 1, 3 and 4 are relatively easy to identify. Phase 1 is defined as the onset of menstrual bleeding, while Phase 3 involves a positive result on a urinary ovulation kit (or other physical indicators, such as a rise in waking basal body temperature (BBT) or white, slippery vaginal discharge). Phase 4 (seven days following ovulation) is assessed via blood test to confirm a progesterone concentration >16 nmol·L^{-1} (Janse de Jonge et al., 2019), which verifies ovulation (Elliott-Sale et al., 2020a; Elliott-Sale et al., 2021). The identification of Phase 2 however presents a range of challenges:

a short (24-48 hour) duration, absence of a clear, measurable symptom (physical indicator) as is present in phases 1 and 3, and large inter-individual variation in the timing of this phase. As such, the only way to quantify this phase is through prospective daily blood sampling, which is arduous for both researcher and participant. The time-consuming and costly nature of accurately identifying MC phases is undeniable and likely explains their broadly poor consideration in the SES literature. Moreover, studies utilising female participants typically require repeated measures due to fluctuations in ovarian hormonal concentrations within, and between cycles, in contrast to male participants whose sex hormone concentrations remain fairly stable following puberty (Elliott-Sale et al., 2021).

Contraceptives (hormonal and non-hormonal)

HC is an umbrella term describing all medications delivering exogenous steroid hormones (Figure 1.2) that alter an individual's endogenous hormonal profile, thus obscuring a natural MC (Elliott-Sale & Hicks, 2018). These exogenous oestrogens and progestins act in a negative feedback manner on FSH and LH to downregulate the hypothalamic-pituitary-ovarian axis and reduce endogenous oestrogen and progesterone concentrations (Elliott-Sale et al., 2020b). Since ovarian hormonal profiles are meaningfully different between individuals utilising HC, as well as different brands/types of HC, and those who are naturally menstruating, these women should be considered as discrete groups in research and applied settings, with specific and divergent methodological considerations (Elliott-Sale et al., 2021). Importantly, the withdrawal/breakthrough bleeding associated with some forms of HC is physiologically distinct from menses experienced by naturally menstruating women, and therefore cannot be used as a marker of reproductive health (Elliott-Sale et al., 2021). The use of HC can therefore mask some health conditions, such as the menstrual dysfunction that is associated with problematic exposure to LEA, and one of the primary indicators of Relative Energy Deficiency in Sport (REDs) (Mountjoy et al., 2023) which is typically indicated by a loss of menstrual bleeding. Indeed, a breakthrough bleed can still occur with HC use, even if an athlete is experiencing LEA (Elliott-Sale et al., 2021).

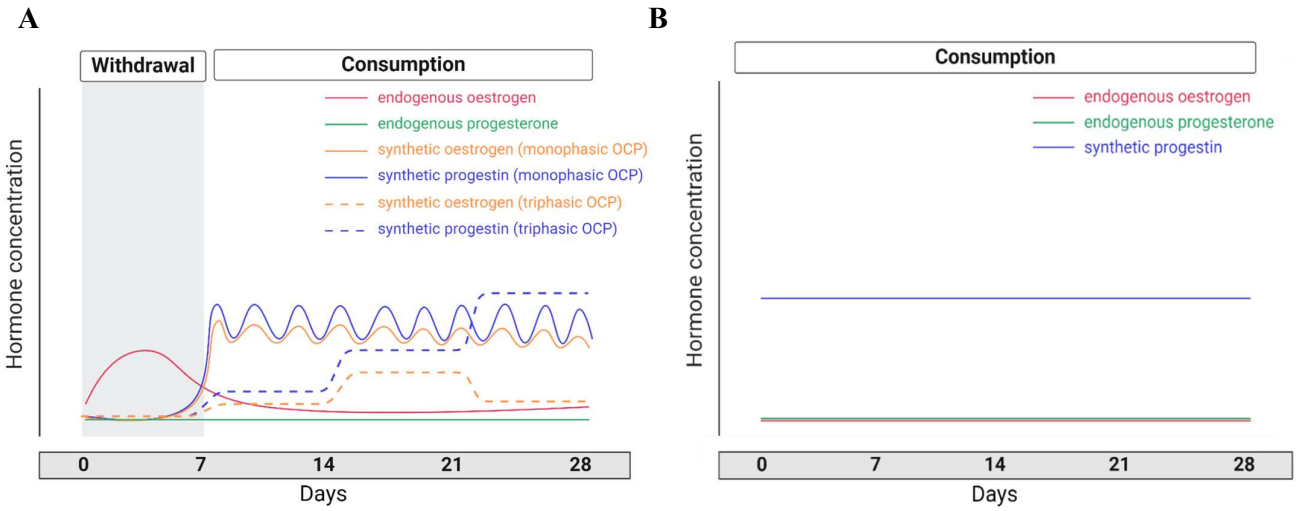


Figure 1.2. Visualisation of the fluctuations in the hormonal profiles of synthetic/exogenous oestrogen and progestins, alongside endogenous oestrogen and progesterone across a 28-day cycle in (A) combined oral contraceptive pill (OCP) users, and (B) progestin-only HC users (e.g., intrauterine system, implant or hormonal injection). It should be noted that the exogenous/synthetic progestin concentration following the hormonal injection and implant gradually declines over time (Huber, 1998), and so the precise hormonal profile at any one timepoint depends on the date of the implant/injection. Adapted from (Chidi-Ogbolu & Baar, 2018; Lewis et al., 2019; Rechichi et al., 2009; Sims & Heather, 2018). *Created with BioRender.com.*

Surveys from around the globe report that around half of the female athlete population uses some form of HC (Brynhildsen et al., 1997; Clarke et al., 2021; Hagmar et al., 2009; Larsen et al., 2020; Martin et al., 2018; Oxfeldt et al., 2020), justifying the need to consider their effects in both research and applied practice. Within the numerous different forms of HC, there are two distinct types: (1) combined hormonal formulations with both an oestrogenic and progestin component, and (2) progestin-only formulations, which may be particularly useful for individuals with contraindicators to oestrogen. An alternative form of contraception is the copper intrauterine device (IUD), which is a non-hormonal method. HCs deliver exogenous doses of synthetic progesterone and/or oestrogen, preventing pregnancy via two key mechanisms of action: thickening of the mucus in the neck of the uterus to hinder sperm penetration, and thinning of the uterus lining to reduce the chance of a fertilised ovum implanting (Elliott-Sale & Hicks, 2018). Moreover some, but not all, types of HC also prevent ovulation (Elliott-Sale & Hicks, 2018). The copper IUD (non-hormonal) acts in a similar mechanism as HC (described above) to prevent pregnancy but does not inhibit ovulation. Moreover, some methods are systemic (contraceptive pills, injections, implants and vaginal ring), delivering synthetic hormones into the bloodstream, while others are local methods [Intrauterine system (IUS), and IUD] that release either

progesterone (IUS) or copper (IUD) locally into the uterus. There are numerous different delivery methods of HC, including pills as well as the more recently developed injectable and implantable methods (Table 1.1).

Table 1.1. An outline of the different types of contraception, including brand examples and approximate rates of usage, among female athletes using contraception across the UK, US and Denmark.

Type	Description	Brand examples	Prevalence (%)
Combined			
Oral contraceptive pills (OCP)	Delivers exogenous oestrogen and progestin in either mono-, bi-, or tri-phasic formulations, whereby the timing of exogenous hormone delivery is altered. Monophasic pills deliver the oestrogenic and progestogenic component in fixed doses, whereas biphasic and triphasic preparations attempt to mimic the conditions of a natural MC by varying the daily hormonal concentrations. Typically taken for 21 consecutive days, followed by a 7-day withdrawal period (characterised by a withdrawal bleed and slight rise in endogenous oestrogen and progesterone concentrations). Some pill types can be taken with no or shorter breaks.	Yasmin, Microgynon, Rigevidon, Cilest, Gedarel	71
Patch	Oestrogen and progestin administered through the skin. Used for three consecutive weeks, followed by a patch-free week which results in a withdrawal bleed for some.	Evra	<1
Vaginal ring	A soft, small, plastic ring inserted inside the vagina. Releases continuous doses of both oestrogen and progestin into the bloodstream. Remains in the vagina for 21 days, it is then removed for 7 days before a new ring is inserted.	NuvaRing	1

Progestin only			
Oral contraceptive pills (OCP)	Delivers daily exogenous progestin. Unlike the combined OCP, there is typically no withdrawal phase. Bleeding becomes irregular/ lighter /more frequent/ longer/ ceases completely.	Cerazetta, Noriday	7
Intrauterine system (IUS)	A small T-shaped plastic device inserted into the uterus. It releases synthetic progesterone locally and directly into the uterine cavity. Although most people continue to ovulate, it can prevent ovulation in some. A long-term contraception option, providing up to five years of birth control.	Mirena, Jaydess, Levosert	9
Injection	Delivers a continuous low dose of progestin into the bloodstream. Effective for 8-13 weeks depending on the brand. Bleeding may become irregular or cease completely.	Depo-Provera, Sayana Press	2
Implant	A small thin device inserted into the arm, delivering a continuous low dose of progestin into the bloodstream. This long-term form of contraception provides up to three-five years of continuous birth control.	Nexplanon, Implanon, Norplant	10
Non-hormonal			
Copper-releasing intrauterine device (IUD)	A small device with a fine copper wire wrapped around a plastic frame that is inserted into the uterus. The device releases copper locally and directly into the uterus. Copper has the same mechanisms of action as exogenous oestrogen and progestin in preventing pregnancy, but without inhibiting ovulation.		<1

Descriptions adapted from Elliott-Sale and Hicks (2018); Elliott-Sale et al. (2021). Other non-hormonal contraceptive methods include condoms, contraceptive gels, and abstinence. However, as these methods are either acutely applied, or applicable to males rather than females, they are not a necessary research consideration. Prevalence statistics combined using a weighted mean, obtained from Martin et al. (2018); 212 UK-based female athletes, Oxfeldt et al. (2020); 106 Danish female athletes and Larsen et al. (2020); 89 female athletes in the United States. Only female athletes of at least Tier 3 [national level (McKay et al., 2022b)] were included.

Menstrual irregularities

Aside from naturally menstruating women and those utilising contraception, the third distinct group of women for consideration are those experiencing menstrual irregularities (MI). There is a higher prevalence of MIs in athletic comparative to general populations, with MI reported to occur in 20-40% of female athletes (Beals & Manore, 2002; Hagmar et al., 2009; Thein-Nissenbaum et al., 2012), with some reporting incidence as high as 86% (Taim et al., 2023) – although the overall rates in elite athletes are not established. There are numerous perturbations to the MC (i.e. alterations to the usual function) (Elliott-Sale et al., 2020a), including symptomatic conditions such as oligomenorrhea, polymenorrhea and amenorrhea alongside those that are asymptomatic including anovulatory cycles and luteal phase deficiency (Table 1.2). There are numerous other MIs such as heavy menstrual bleeding (menorrhagia), dysmenorrhea (painful menstruation), endometriosis and adenomyosis (abnormal growth of endometrial-like tissue), alongside premenstrual syndrome/ pre-menstrual dysphoric disorder (psychological, physical and behavioural symptoms prior to menstruation). As these MIs do not necessarily result in hormonal alterations they will not be discussed in detail. Moreover, although other perturbations to ovarian hormonal profiles occur through pregnancy and menopause, these are beyond the scope of this PhD candidature.

Table 1.2. Some of the most prevalent forms of menstrual irregularities associated with alterations to hormonal profiles, alongside their typical prevalence rates among athletic populations.

Condition	Description	Prevalence (%)
Amenorrhea	Defined as the persistent absence of menses. Characterised by an absence of menstruation or ovulation, alongside downregulated 17-β-oestradiol (the most potent form of oestrogen among pre-menopausal women) and progesterone concentrations (Elliott-Sale et al., 2021; Redman & Loucks, 2005). Can be either primary (absence of menstruation by age 15 when development of secondary sexual characteristics is evident or age 14 when no secondary sexual characteristics are evident) or secondary (absence of at least three consecutive menses by an individual who has previously menstruated).	2-62 *(secondary typically more prevalent)*

Oligomenorrhea	Fewer than eight MCs per year, or cycles longer than 35 days (i.e., infrequent menstruation). MCs otherwise follow the normal eumenorrheic hormonal patterns, but drawn-out over a longer timeframe (Elliott-Sale et al., 2021).	8-61
Polymenorrhea	MCs lasting <21 days (i.e., frequent menstruation). MCs otherwise follow the normal eumenorrheic hormonal patterns but condensed over a shorter timeframe (Elliott-Sale et al., 2021).	~8
Anovulatory cycles	Menstruation is still experienced, but without ovulation, and is therefore an asymptomatic MI. Characterised by an absence of an ovulatory peak in 17-β-oestradiol (Elliott-Sale et al., 2021; Redman & Loucks, 2005).	~25
Luteal phase deficiency	Progesterone secretion in the luteal phase is reduced to <16 $nmol \cdot L^{-1}$, thus preventing successful implantation of a fertilised egg and resulting in infertility (Elliott-Sale et al., 2021; Redman & Loucks, 2005). Asymptomatic MI, and so can only be diagnosed through blood/urine samples to measure ovarian hormonal concentrations.	~27
Polycystic ovary syndrome (PCOS)	Characterised by elevated ovarian production of androgens. A diagnosis requires ≥2 of the following criteria: (a) chronic anovulation or oligo-ovulation, (b) biochemical or clinical evidence of androgen excess, and (c) detection of polycystic ovaries by ultrasound (ESHRE & Group, 2004). Symptoms include menstrual irregularities, signs of androgen excess (e.g., acne, hirsutism), obesity, insulin resistance, elevated serum LH concentration (ESHRE & Group, 2004).	2-44

Prevalence statistics collated from a range of athletes across Tiers 2-5 among numerous countries and sports globally across the last 20 years (Beals & Manore, 2002; Burrows et al., 2003; Coste et al., 2011; Dadgostar et al., 2009; De Souza et al., 2010; Dusek, 2001; Klentrou & Plyley, 2003; Redman & Loucks, 2005; Taim et al., 2023; Thein-Nissenbaum et al., 2012). Studies examining the incidence of anovulation and luteal phase deficiency are particularly scarce, likely due to their asymptomatic nature. Prevalence varies across sports. MC; menstrual cycle, LH; luteinising hormone.

Performance in women: any influence of ovarian hormones?

Despite widespread speculation among both the scientific community (McNulty et al., 2020b) and general population (McCallum, 2022), high-quality evidence of changes in performance in response to fluctuations in ovarian hormonal profiles, across MC phases or between HC users and non-users, is currently lacking. Although their primary role lies in reproduction, fluctuations in oestrogen and progesterone concentrations have the potential to influence multiple physiological systems associated with athletic performance. Broadly speaking, oestrogen and progesterone exhibit opposing physiological effects.

A recent meta-analysis concluded that exercise performance, both endurance and strength, might be trivially reduced during phase 1 (early follicular) of the MC comparative to all other MC phases, with the largest difference in performance occurring between phase 1 (early follicular) and phase 2 (late follicular) (McNulty et al., 2020b). These findings were supported by a more recent review concluding a performance impairment during phase 1 (Carmichael et al., 2021). These results suggest that the low concentrations of oestrogen and progesterone observed during phase 1 may elicit a performance decrement. However, given that findings are highly inconsistent between studies, the magnitude of effect is trivial, and results are confounded by numerous flaws in study quality regarding methodological control and classification of ovarian hormonal profiles, the conclusions from the available literature must be considered weak (McNulty et al., 2020b).

Regarding HC use, a recent meta-analysis reported trivial performance impairments, both endurance and strength-based, among women utilising OCPs (the most common form of HC), compared to naturally menstruating women (Elliott-Sale et al., 2020b). Moreover, performance does not appear to fluctuate across an OCP cycle, including between consumption (active pill-taking) and withdrawal phases (Elliott-Sale et al., 2020b). This suggests that endogenous, rather than exogenous, ovarian hormone concentrations mediate exercise performance. Further, the endogenous hormonal profile in OCP users is comparable to that observed during phase 1 of the MC (low levels of both oestrogen and progesterone). Thus, it could also be hypothesised that the suppression of endogenous ovarian hormone concentrations is responsible for the observed performance decrement (Elliott-Sale et al., 2020c). However, once again any conclusions from this meta-analysis are limited by the poor quality of studies (Elliott-Sale et al., 2020b), and any implications of OCP on exercise performance remain poorly understood, while evidence regarding any influence of other forms of HC is even more sparse.

In contrast to the inconclusive evidence surrounding performance alterations in naturally menstruating women and those utilising HC, there is robust evidence regarding the negative impact of specific MIs (i.e., amenorrhea) on health and performance. Amenorrhea (Table 2) is associated with a reduction in bone mineral density and an increase in stress fracture rates across an athlete's career, alongside serious long-term complications including osteoporosis (Barrack et al., 2014; Brook et al., 2019; Heikura et al.,

2018; Loucks, 2007; Myburgh et al., 1990; Nose-Ogura et al., 2019; Papageorgiou et al., 2018a). It is these health complications that result in a negative influence on performance rather than a direct impact of alterations to hormonal profiles, per se.

Substrate utilisation

Influence of the menstrual cycle

Perhaps the most understood and well-documented effect of ovarian hormones relates to their influence on substrate utilisation during exercise, particularly the substantially greater rates of fat oxidation during exercise among females compared to males (when expressed relative to FFM) (Randell et al., 2017; Venables et al., 2005). Oestrogen and progesterone directly interact with tissues and organs involved in energy metabolism (muscle, liver, adipose tissue). Compared to men, women exhibit a greater reliance on fat oxidation to support the energy demands of low-moderate intensity exercise, as demonstrated by a lower respiratory exchange ratio (RER) and higher PFO when expressed relative to fat free mass (FFM) (Chrzanowski-Smith et al., 2020; Chrzanowski-Smith et al., 2021; Horton et al., 1998; Randell et al., 2017; Tarnopolsky et al., 1990; Venables et al., 2005). This observation is partially attributed to the role of oestrogen in substrate utilisation, elevating lipid synthesis and muscle lipolysis (i.e., elevating the rate of appearance/availability of free fatty acids and glycerol) (Bunt, 1990; Constantini et al., 2005; Isacco & Boisseau, 2017; Nicklas et al., 1989). Concomitantly, oestrogen is associated with decreased hepatic glycogenolysis and gluconeogenesis, increased muscle and hepatic glycogen storage, and reduced peripheral blood glucose uptake. Together, this reduces the rate of appearance and disappearance of blood glucose and ultimately resulting in a decreased CHO utilisation and glycogen-sparing effect (Constantini et al., 2005; Isacco & Boisseau, 2017; Oosthuyse & Bosch, 2010, 2012; Ruby et al., 1997). Progesterone is thought to oppose the actions of oestrogen via a limiting of fat oxidation (Oosthuyse & Bosch, 2010), although it is simultaneously considered to accentuate the CHO-sparing effects of oestrogen (Constantini et al., 2005; D'Eon et al., 2002; Frankovich & Lebrun, 2000; Isacco & Boisseau, 2017).

While the sex differences and the mechanistic actions of oestrogen/progesterone (as determined through animal studies) are more established, information on shifts in substrate utilisation across the phases of the MC within a woman is conflicting. Hypothetically, the increase in both oestrogen and progesterone observed during phase 4 could be associated with an increased reliance on fat oxidation and a decrease in CHO oxidation. Indeed, there is some evidence to support the notion of enhanced fat oxidation and a glycogen-sparing effect during phase 4 (mid luteal), alongside elevated CHO oxidation during phase 1 (early follicular), when oestrogen and progesterone are low (Campbell et al., 2001; Dombovy et al., 1987; Hackney, 1999; Jurkowski et al., 1981; Lavoie et al., 1987; Oosthuyse & Bosch, 2010; Zderic et al., 2001). However, other studies demonstrate no influence of MC phase on substrate utilisation during either short-term/intermittent (Debruynprevost et al., 1984; Lynch & Nimmo, 1998) or endurance

exercise (Bailey et al., 2000; Dombovy et al., 1987; Frandsen et al., 2020; Hackney et al., 1994; Horton et al., 2002; Kanaley et al., 1992), while Lee et al. (2024) observed an increase in CHO oxidation during submaximal cycling during phase 4. It may be that any influence of the MC on substrate utilisation is dependent on exercise intensity and the resulting energy demands, as higher exercise intensities necessitate a higher endogenous glucose production and utilisation. Indeed, Hackney et al. (1994) observed elevated fat oxidation and decreased CHO oxidation during MC phase 4 only during low and moderate exercise intensities [40-60% maximal oxygen uptake ($\dot{V}O_2max$)], while this response disappeared at higher intensities (~70% $\dot{V}O_2max$).

Influence of hormonal contraceptives

The interaction of exogenous hormones with systems regulating energy metabolism is highly complex, and any effects are ill-defined. OCP use has been suggested to increase fat oxidation as illustrated by a decrease in RER, potentially via the action of synthetic progestins in decreasing insulin receptor concentration to decrease glucose flux, resulting in a CHO-sparing effect (Frankovich & Lebrun, 2000; Krauss & Burkman, 1992; Redman & Loucks, 2005). However, most studies show that although OCP use may increase lipolytic activity (e.g., lipid mobilisation) and decrease glucose flux, there is no alteration of substrate utilisation during endurance exercise (Bemben et al., 1992; Bonen et al., 1991; Casazza et al., 2002; Isacco et al., 2012; Suh et al., 2002, 2003). However, the effect of exercise intensity and nutrient status must be considered when interpreting these findings. Indeed, it appears that exercise itself, and the associated increase in lipolytic activity, may mask any greater lipid mobilisation observed postprandially in OCP users (Isacco et al., 2014).

Endurance exercise

Influence of the menstrual cycle

Given the potential effect of oestrogen and progesterone on substrate utilisation, it is possible that ovarian hormones may alter endurance performance/capacity. For example, any glycogen-sparing during phase 4 of the MC may have a beneficial effect on endurance performance by preserving glycogen for use within the muscle contractile units, alongside maintaining plasma glucose availability, during the latter stages of a race when an increase in speed/power is required to win (Hearris et al., 2018). However, the evidence regarding alterations in endurance performance across the MC are conflicting. Indeed, a recent review concluded that aerobic performance appears to be enhanced during phase 1, and impaired during phase 3 of the MC (Carmichael et al., 2021). However, a separate meta-analysis reported endurance performance to be impaired during phase 1 (McNulty et al., 2020b).

Oestrogen is broadly considered to promote improved endurance performance, with progesterone exerting opposing effects (Frankovich & Lebrun, 2000; Oosthuyse & Bosch, 2010). Oestrogen is thought to increase stroke volume and preload, thus enhancing cardiac output during endurance exercise

(Frankovich & Lebrun, 2000). Conversely, progesterone elevates BBT by around 0.3°C–0.5°C during phase 4 (Carpenter & Nunneley, 1988; Constantini et al., 2005; Stephenson et al., 1982), while also augmenting ventilatory responses to exercise; increasing minute ventilation and respiratory drive, and increasing heart rate (Constantini et al., 2005; Dombovy et al., 1987; Schoene et al., 1981; Seebauer et al., 2002). Since an increased breathing rate requires more oxygen, diverting oxygen away from the exercising muscles, an increase in progesterone concentration might result in an elevation of perceived exertion. Indeed, there is some evidence of higher cardiovascular strain during endurance exercise in phase 4 versus phase 2, due to the effects of the peak progesterone concentration versus oestrogen domination, respectively (Dombovy et al., 1987; Janse de Jonge, 2003).

In view of the previously described effects of oestrogen and progesterone on physiological systems underpinning endurance performance, high oestrogen phases might be expected to enhance endurance performance, whereas high progesterone phases might be associated with performance impairment. However, the available evidence fails to support the theory that either $\dot{V}O_2max$, or its determinants (i.e., cardiac output, ventilation, plasma volume, haemoglobin concentration) are altered across MC phases (Bemben et al., 1995; Casazza et al., 2002; Constantini et al., 2005; De Souza et al., 1990; Dombovy et al., 1987; Frandsen et al., 2020; Goldsmith & Glaister, 2020; Janse de Jonge, 2003; Lebrun et al., 2003; Smekal et al., 2007; Stephenson et al., 1982; Vaiksaar et al., 2011). The effects of MC phase on endurance performance itself, or applied physiological outcomes related to aerobic performance, are even less definitive. There is some evidence for improved cycle TT performance during phase 2 (late follicular, coinciding with the pre-ovulatory surge in oestrogen) compared to phase 1 (early follicular) (Oosthuyse et al., 2005). Moreover, there is also evidence for a decline in endurance parameters (running economy, yo-yo intermittent test) during phase 4 (high progesterone) compared to phase 1 (Goldsmith & Glaister, 2020; Julian et al., 2017). However, others report the opposite; improvements in, cycle time to exhaustion (TTE), 1500 m time and running economy during phase 4 (high progesterone) compared to phases 1 and 2 among Tier 0-2 (McKay et al., 2022b) individuals (Dokumaci & Hazir, 2019; Jurkowski et al., 1981; Nicklas et al., 1989; Sutresna, 2016). Moreover, numerous studies report no alteration in endurance performance (TTE, yo-yo intermittent test, velocity at $\dot{V}O_2max$, 2 km rowing TT) across MC phases among a range of athletic calibres (Bailey et al., 2000; Beidleman et al., 1999; Bemben et al., 1995; Burrows & Bird, 2005; De Souza et al., 1990; Dombovy et al., 1987; Guo et al., 2005; Hogwood et al., 2023; Janse de Jonge, 2003; Lebrun et al., 1995; Lebrun, 1991; Tounsi et al., 2018). Although there is some evidence for an increase in body mass and total body water during the luteal phase (Fruzzetti et al., 2007; Stachon, 2016), the related effect on aerobic performance is poorly understood. Moreover, numerous studies demonstrate that body mass (Lebrun et al., 1995; Lebrun, 1991; Rael et al., 2021; Tsampoukos et al., 2010), sum of skinfolds, (Lebrun et al., 1995; Maciejczyk et al., 2014), fat mass (Lebrun et al., 1995; Rael et al., 2021) and total body water (Rael et al., 2021) do not consistently fluctuate over the MC, at least when examining group averages.

In summary, the inconclusive evidence suggests that even if there are alterations to physiological parameters involved with exercise (i.e., substrate metabolism/ ventilatory responses/ cardiac output), this may not translate to alterations in performance.

Influence of hormonal contraceptives

The effects of HC use on endurance performance have been less extensively investigated compared to performance in naturally menstruating women. There is some evidence for a decline in $\dot{V}O_2$max with OCP use (Casazza et al., 2002; Daggett et al., 1983; Lebrun, 1993; Lebrun et al., 2003; Lebrun, 1991; Martin & Buono, 1997; Notelovitz et al., 1987; Redman & Weatherby, 2004; Rogers & Baker, 1996). This purported increased oxygen demand with OCP use may be a result of synthetic progestins increasing ventilation (Frankovich & Lebrun, 2000; Montes et al., 1983). Indeed, there is some evidence for ventilatory increases among OCP users (increased minute ventilation and respiratory drive) (Montes et al., 1983). Although conversely, some studies report a decline in $\dot{V}O_2$max, without associated ventilatory alterations (Joyce et al., 2013; Lebrun, 1991; Notelovitz et al., 1987). A decline in $\dot{V}O_2$max does not appear to translate to impaired endurance performance, however. Joyce et al. (2013) reported that long-term OCP use among Tier 2 (McKay et al., 2022b) women may be detrimental to $\dot{V}O_2$max, but without influencing endurance performance nor associated parameters (minute ventilation, heart rate, blood pressure, blood lactate or cycling economy). Moreover, others demonstrate no effects of OCP use on $\dot{V}O_2$max nor endurance performance (Bryner et al., 1996; Montes et al., 1983), while some studies report enhanced performance during pill consumption compared to pill withdrawal days (Sunderland & Nevill, 2003).

Overall, it appears that OCP users may experience a marginally impaired endurance performance capacity compared to naturally menstruating individuals but, as detailed in a recent meta-analysis (Elliott-Sale et al., 2020b), the magnitude of this effect is likely to be trivial. Moreover, there is no robust evidence to support alterations to endurance performance across an OCP cycle (i.e., consumption vs withdrawal phases) (Elliott-Sale et al., 2020b). It is hypothesised that the chronic downregulation of endogenous hormonal profiles may underlie the observed performance decrement among OCP users (Elliott-Sale et al., 2020b). Indeed, since the endogenous hormonal profile in an OCP user is most similar to that observed in phase 1 of a naturally menstruating individual (i.e., low endogenous oestrogen and progesterone), a performance decrement in OCP users is consistent with observations of reduced performance in phase 1 of the MC compared to all other phases where oestrogen and progesterone levels are comparatively higher (McNulty et al., 2020b). This finding is further supported by the lack of alteration in performance between OCP consumption and withdrawal phases, as endogenous hormonal concentrations remain relatively low.

Short-duration, high-intensity exercise

Influence of the menstrual cycle

Oestrogen exerts anabolic effects on skeletal muscle (Baltgalvis et al., 2010; Lowe et al., 2010), while also altering the secretion and metabolism of growth hormone (Leung et al., 2004). Moreover, both ovarian hormones act as neurosteroids, crossing the blood-brain barrier and potentially influencing maximal neuromuscular performance (Tenan, 2017). Oestrogen elicits neuroexcitatory effects, increasing voluntary activation and reducing inhibition (Smith et al., 2002). On the other hand, progesterone exhibits neuroinhibitory effects, decreasing voluntary activation and increasing inhibition (Smith et al., 2002). Muscle activation is a primary determinant of the rapid force production necessary for explosive movements. As such, oestrogen and progesterone exhibit a positive and negative relationship with force production, respectively (Gordon et al., 2013; Pallavi et al., 2017; Phillips et al., 1996; Smith et al., 2002). It could therefore be hypothesised that strength and power would be greater during high oestrogen phases of the MC when progesterone is also low (i.e., phase 2) and impaired during high progesterone phases. A change in bioavailable testosterone between phases has also been purported as a mechanism for strength alterations across the MC, with increases in salivary and plasma testosterone suggested during phases 3 and 4 (Cook et al., 2018; Lane et al., 2015). However, this mechanism remains under-researched, and it is unknown if any increase in salivary and plasma testosterone translates to an increase in bioavailable testosterone, especially in women (Carmichael et al., 2021).

There is some support for the hypothesis that oestrogen enhances muscle strength and power, assessed by measures of maximal voluntary contractions and force production. Indeed, recent reviews and meta-analyses have demonstrated a minor decrement in strength-based outcomes during phase 1, when oestrogen concentrations are lowest, compared with other phases (Carmichael et al., 2021; McNulty et al., 2020a). Moreover, there is evidence of an increase in strength and maximal voluntary force during phase 2 (late follicular) coinciding with the highest oestrogen concentrations prior to ovulation alongside low progesterone levels (Pallavi et al., 2017; Phillips et al., 1996; Sarwar et al., 1996). Others also observe a decline in power, strength and maximal voluntary contraction during phase 4 (Dam et al., 2022; Graja et al., 2020; Phillips et al., 1996; Rodrigues et al., 2019; Tenan et al., 2016), concomitant with the reduction in oestrogen and elevation in progesterone concentrations at this time. However, other studies present contradictory evidence, observing no influence of MC phase on strength, power, or maximal voluntary contraction, across a range of athletic tiers (Dam et al., 2022; De Jonge et al., 2001; Dragutinovic et al., 2024; Hertel et al., 2006; Kishali et al., 2004; Lebrun et al., 1995; Lebrun, 1991; Otaka et al., 2018; Quadagno et al., 1991; Romero-Moraleda et al., 2019; Wirth & Lohman, 1982). A few studies even oppose the hypothesis that oestrogen enhances muscular strength, reporting an increase in strength during phase 1 when oestrogen is lowest (Davies et al., 1991; Ekenros et al.,

2013; Gordon et al., 2013), while some recent studies have observed negative relationships between oestrogen concentration and explosive performance indices (Dragutinovic et al., 2024; Pessali-Marques et al., 2024). Such inconsistency might be explained by the failure to focus on strength changes between phases 2 and 4, where differences would be amplified by the largest differences between oestrogen and progesterone concentrations across these phases (Carmichael et al., 2021). However, this explanation does not hold true for all studies which have failed to observe a change in strength across the MC.

Evidence of the neuroexcitatory and inhibitory effects of the ovarian hormones is even less conclusive. There is some support for the hypothesised influence of oestrogen and progesterone; an increased voluntary activation and fatiguability during phase 2 when oestrogen levels peak, alongside a decreased voluntary activation and fatigue during phase 4 concurrent with the increase in progesterone concentration (Ansdell et al., 2019; Sarwar et al., 1996). However, other studies report the initial motor unit firing rate to be higher during phase 4, when progesterone is at its peak, comparative to phase 1 (Tenan, 2017; Tenan et al., 2016), and others report poorest muscular endurance during phase 4 (Petrofsky et al., 1976). Meanwhile, it has also been reported that there is no influence of MC phase on muscle fatiguability nor contractile properties (De Jonge et al., 2001; DiBrezzo et al., 1991; Lebrun et al., 1995; Miskec et al., 1997).

Lastly, there is inconsistent evidence for the role of oestrogen and progesterone on short duration (≤3 min) high intensity activities. Hypothetically, an increase in force production and peak power associated with oestrogen might enhance anaerobically-dominate performance events during high oestrogen phases (Smith et al., 2002). Indeed, some studies among Tier 1-3 individuals have reported improvements in performance (cycle sprints, 400 m run) during phases 2 and 3 where oestrogen is elevated (Cook et al., 2018; Shakhlina et al., 2016), or a decline in phase 1 or the end of phase 4 when oestrogen concentration is low (Dam et al., 2022; Lee et al., 2024). On the other hand, the majority of studies report no change in short duration performance indices (repeated sprints, vertical jump, CMJ, Wingate tests) between MC phases among Tier 1-4 (McKay et al., 2022b) individuals across a range of power-based sports (Bemben et al., 1995; Bushman et al., 2006; Dam et al., 2022; Dean et al., 2003; DiBrezzo et al., 1991; Julian et al., 2021; Kishali et al., 2004; Lara et al., 2020a; Lara et al., 2020b; Lebrun et al., 1995; Lebrun, 1991; Miskec et al., 1997; Somboonwong et al., 2015; Štefanovský et al., 2016; Tasmektepligil et al., 2010; Tounsi et al., 2018; Tsampoukos et al., 2010; Wiecek et al., 2016).

Influence of hormonal contraceptives

A recent meta-analysis demonstrated trivial impairments in strength among women utilising OCP compared to naturally menstruating women (Elliott-Sale et al., 2020b). Indeed, while several studies report reduced muscular force and strength in OCP users compared to non-users (Sarwar et al., 1996; Wirth & Lohman, 1982), others report no difference in strength or power measures (Ekenros et al., 2013; Giacomoni et al., 2000; Knowles et al., 2019; Lebrun, 1991; Peters & Burrows, 2006; Rickenlund

et al., 2004). On the contrary, although there is some suggestion that the potential detrimental effect of progesterone on muscle force and endurance is minimised in OCP users due to the chronic suppression of endogenous ovarian hormones (Wirth & Lohman, 1982), this remains unexplored. There appears to be no difference between OCP consumption and withdrawal phases for strength and power, maximal voluntary force, anaerobic performance, or fatigue index (Bushman et al., 2006; Dragutinovic et al., 2024; Ekenros et al., 2013; Giacomoni et al., 2000; Lee et al., 2024; Petrofsky et al., 1976; Phillips et al., 1996; Rechichi & Dawson, 2009; Sarwar et al., 1996). Moreover, as outlined above, it appears that endogenous, rather than exogenous, ovarian hormones are primarily responsible for any alterations in strength performance (Elliott-Sale et al., 2020b)

Field-based/ sport-specific performance measures

Influence of the menstrual cycle

Athletic performance is underpinned by a complex interaction between anatomic, physiological, metabolic, psychological and biomechanical factors, varying substantially between sports (Constantini et al., 2005). Given the lack of consensus regarding any influence of ovarian hormonal concentrations on endurance or strength metrics, it is extremely challenging to determine any effect of MC phase/HC use on real-world sporting performance. For example, the glycogen-sparing effects of oestrogen will be most relevant to endurance sports, while the influence of oestrogen on muscle strength and fatiguability will be more relevant to explosive, power-based sports. However, even if ovarian hormones exert some influence on specific underpinning physiological/ biomechanical parameters, this will not necessarily translate into a meaningful/ measurable impact on performance. Indeed, it is important to note that female athletes have presumably competed successfully at all phases of the MC, winning competitions and setting world records.

Investigating the MC in real-world performance settings (i.e., actual competition) is optimal, given the high ecological validity. However, such studies are challenging to conduct and are therefore lacking. Indeed, most studies of performance across the MC are undertaken in a controlled laboratory environment, in which participants are often blinded, fasted, and lack real-world motivation to perform. Another approach is the use of retrospective, self-reported information, as illustrated by a recent study which reported that the majority (57%) of Tier 2 (McKay et al., 2022b) runners recorded their best marathon time during phase 4 (Greenhall et al., 2021). However, the classification of MC phase was based on self-report data and no attempt was made to verify MC phase through best-practice methods: retrospective blood samples or ovulation testing (Elliott-Sale et al., 2021). Moreover, the MC was separated into a simplistic two-phase model (follicular v luteal) thus ignoring the more complex hormonal profile observed across the four phases.

The few experimental studies of real-world performance measures have provided conflicting outcomes. Quadagno et al. (1991) reported no difference in 100 or 200 m swimming performance across MC phases among Tier 2-3 (McKay et al., 2022b) athletes. However, methodological quality of MC phase classification and control in this study was suboptimal. Although authors compared between three phases (1, 2 and the end of phase 4), across three separate MCs, phases were not confirmed by either of the recommended protocols of retrospective blood samples or ovulation testing (Elliott-Sale et al., 2021). As such, conclusions drawn from this data are limited. In contrast, superior performance during phases 2 and 3 has been reported by two older studies, also with flawed methodological characteristics. Here, Fomin et al. (1989) reported enhanced skiing performance (without confirmation of actual performance times/measures nor any statistical analysis) among Tier 2-3 (McKay et al., 2022b) athletes during phases 2 and 3 (determined via BBT only). Moreover, Bale and Nelson (1985) observed superior sprint swimming performance during phases 2-3, however MC phases were once again not confirmed with objective measures.

Perhaps the most methodologically sound study of real-world performance was conducted by Guo et al. (2005). Authors examined 13 Tier 2 (McKay et al., 2022b) track and field athletes, reporting superior 100 and 200 m track performance in phase 4 compared to 1. This study utilised three separate MCs alongside retrospective blood sampling to confirm phases. However, the absence of confirmed ovulation prohibited the confirmation of eumenorrhea (Elliott-Sale et al., 2021). Moreover, performance was compared across just two phases, ignoring phases 2 and 3. As such, despite increasing popularity among consumers and media (McCallum, 2022), evidence to support alterations in real sports performance across the MC is currently lacking.

Influence of hormonal contraceptives

Research into real-world sporting outcomes among HC users is even more sparse. Crewther et al. (2018) investigated 23 Tier 4 (McKay et al., 2022b) female field hockey athletes across four international matches. They reported no differences between OCP users and non-users in any performance metric (match performance statistics, video-derived positive actions and negative actions, alongside coach and player ratings of performance). However, the authors failed to report details of OCP usage (type, formulation, brand, length of usage) or OCP consumption vs withdrawal days, thus weakening any conclusions from these data. Regarding OCP phase, Rechichi and Dawson (2012) observed no influence of OCP phase (consumption vs withdrawal) on 200 m swim performance among Tier 2-3 (McKay et al., 2022b) swimmers and water polo players consuming the same OCP type (monophasic) and formulation, however OCP brand or length of usage was not reported. Finally, there is almost a complete lack of investigation into HC methods other than OCPs (Flood et al., 2024).

Limitations of the performance literature and future directions

Quality of the literature base

Despite numerous investigations of the influence of ovarian hormonal profiles across a multitude of performance types, consensus regarding if and how ovarian hormones affect performance is currently lacking. This discrepancy may arise from inconsistencies in methodological classification and control of menstrual status/HC use across studies, with a lack of studies employing gold-standard methodologies regarding the classification and control of menstrual status (Elliott-Sale et al., 2021). Indeed, the observation from recent meta-analyses that only 8-17% of studies were considered 'high-quality' in methods regarding menstrual status (Elliott-Sale et al., 2020c; McNulty et al., 2020a), weakens the conclusions from the available evidence base. The accurate classification of MC phase/ HC use is pivotal in supporting causality regarding any influence of ovarian hormones on performance. In light of the poor quality of available evidence regarding menstrual status classification and control, the actual magnitude and/or direction of any effects may be different to those currently reported.

The current evidence base is further weakened by a broad failure to compare performance across multiple MC phases. Studies typically contrast two MC phases, most commonly phase 1 (early follicular, low oestrogen and progesterone) to phase 4 (mid-luteal, high oestrogen and progesterone), as this comparison represents the greatest absolute difference in total ovarian hormone concentration. However, this approach may overlook any effects of oestrogen in isolation, without interference of progesterone, as is observed during phases 2 (late follicular). It is therefore challenging to conclude whether a lack of observed effect of MC phase on performance, based on comparison of two MC phases, represents a true result or the failure to detect real differences in phases that were excluded.

The oestrogen: progesterone (E:P) ratio has also been suggested to influence performance and may to some extent explain the conflicting findings regarding the effects of MC phase on performance. McNulty et al. (2020b) observed the greatest performance difference between phase 1 (poorest performance), where both oestrogen and progesterone concentration are low, and the phase 2 (best performance) when oestrogen is high, and progesterone is low. This method also considers the absolute concentrations of circulating oestrogen/progesterone, rather than grouping into "high" (phase 4) vs "low" (phase 1) hormone phases. It could therefore be suggested that performance is optimised when oestrogen concentration is high, without interference of progesterone. Moreover, a lower E:P ratio in phase 4 may mean that the relatively high progesterone concentration counters/impedes the benefits of oestrogen on performance. Indeed, studies observing improvements in endurance performance during the luteal phase report a higher E:P ratio during this phase (Campbell et al., 2001; Jurkowski et al., 1981; Nicklas et al., 1989) while the studies demonstrating no change in endurance performance across MC phases observed a lower ratio (Bailey et al., 2000; Beidleman et al., 1999; McLay et al., 2007).

Further, it has been suggested that a sufficient elevation in the E:P ratio is required to elicit metabolic alterations (D'Eon et al., 2002).

As individuals are typically habitual HC users or non-users, the majority of evidence pertaining to the influence of HC on performance is drawn from observational, between-group trials, instead of the gold-standard RCT design. Indeed, prescribing or withholding HC from individuals for research purposes poses an ethical challenge, and is infrequently observed in research (Lebrun, 1991). Although a more ethical and practical alternative, observational trials do not permit causal inferences and are also more susceptible to the influence of confounding variables. This hampers the conclusions that are able to be drawn from such study designs, consequently hindering our understanding regarding any effect of HC use on performance. Most studies investigating individuals using HC have examined the most prevalent form, OCPs (Flood et al., 2024). However, the rise in popularity of other methods (such as IUS and implants, Table 1) warrants their investigation.

Other factors influencing performance

It is important to note that performance is a multifaceted complex phenomenon, which may confound any inferences around a direct influence of ovarian hormonal profiles. Moreover, differences in the types of exercise (modality, intensity, and duration), as well as differences in the protocols used to assess outcomes such as endurance capacity and muscle strength/fatigue further prevent robust inferences being drawn from available data. Discrepancies in findings could also be explained by the inter-individual variation observed in response to most interventions (Hopkins, 2015; Swinton et al., 2018), combined with intra- and inter-individual variations in ovarian hormonal profiles (Janse de Jonge, 2003; Oosthuyse & Bosch, 2010).

Other co-founders include differences in participant age and athletic calibre, the circadian influence on physiological variables, alongside pre-trial dietary and exercise controls. For standardisation purposes, many studies have examined performance following an overnight fast; however, findings from such protocols may not hold true under typical competition conditions in which CHO is consumed pre- and during exercise. For example, Campbell et al. (2001) demonstrated an enhanced endurance performance in the follicular, compared to luteal, phase among Tier 2 (McKay et al., 2022b) women following an overnight fast but found that this effect was negated with glucose ingestion. It is therefore possible that a fed or CHO-loaded state may negate any influence of ovarian hormonal concentrations on substrate oxidation or endurance performance. Indeed, a postprandial or CHO-loaded state has been demonstrated to overcome any glycogen-sparing effects of oestrogen during phase 4 (mid-luteal) (Campbell et al., 2001; McLay et al., 2007; Suh et al., 2002, 2003). Importantly, a general lack of pre-trial standardisation of muscle/liver glycogen also limits inferences from these findings, as stored glycogen quantity influences subsequent substrate utilisation (Arkinstall et al., 2004).

Training status may also outweigh any influence of MC phase on performance given the substantial adaptations to the cardiovascular system with exercise. Given the increase in minute ventilation and respiratory drive associated with progesterone, an increase in perceived exertion and oxygen demand combined with impairment of endurance performance may be anticipated during the luteal phase when progesterone concentrations are high (Dombovy et al., 1987; Schoene et al., 1981). Indeed, this hypothesised performance impairment during the luteal phase has been demonstrated among non-athletes, but not Tier 4 athletes (Schoene et al., 1981). Therefore, it is important to examine performance indices across the MCs of trained athletes to account for these physiological adaptations, as well as to take advantage of their superior performance consistency compared to recreational exercisers (Hopkins & Hewson, 2001).

Other factors that may influence performance should also be considered. For example, pre-menstrual symptoms commonly associated with the end of the luteal phase or beginning of the follicular phase may hinder performance, irrespective of any hormonal influences (e.g., cramps, bloating, muscle aches, tiredness, gastrointestinal issues, headaches, poor sleep, and anxiety). These negative symptoms are reportedly experienced by ~60-93% of female athletes (BBC Sport, 2020; Findlay et al., 2020; Martin et al., 2018), with ~50-67% believing that such negative symptoms impair performance (Bruinvels et al., 2017; Findlay et al., 2020; Taim et al., 2024). Concomitantly, a reduction in negative menstrual symptoms with HC use may improve athletic performance, regardless of any alterations in ovarian hormonal concentrations *per se*. However, there are numerous methodological issues and challenges associated with investigating symptoms, making findings challenging to interpret. These include issues with subjectivity and recall in symptom reporting, a lack of validated tools/measures, and the absence of longitudinal datasets. Indeed, to be defined as a "MC symptom", multiple consecutive MCs must be observed to verify the cyclical and repeatable nature of such symptoms exclude other potential causes. Moreover, few studies have directly examined the influence of symptoms on performance, instead providing an indirect link by concentrating on the incidence of symptoms across the MC in conjunction with the athletes' perception of how symptoms influence performance. Lastly, the reporting of symptoms is typically focussed on negative effects. It is possible that recording, and therefore drawing attention to, positive symptoms (e.g., feeling energised) may counteract the potential nocebo effect associated with focusing exclusively on negative symptoms and feelings.

Reasons for the underrepresentation of female athletes in SES research

The underrepresentation of women in the literature is now well established. Moreover, the difficulties in translating and applying male-directed research to female athletes, in light of possible sexual dimorphisms, is understood and appreciated in the scientific community. Numerous recent literature audits have reported an apparent volunteer (or self-selection) bias among women, which leads to their lower participation in studies compared to their male counterparts despite meeting inclusion criteria

(Nuzzo, 2021). Smith et al. (2022c) found that women accounted for just 23% of the participant pool but were included in 34% of all studies, while Cowley et al. (2021) and Costello et al. (2014) reported women to contribute just 34-39% of the total participant count but that 63% of studies included both men and women. The volunteer bias among women, particularly female athletes is highly problematic and undoubtedly contributes to the underrepresentation of high-performance female athletes and warrants investigation.

The underrepresentation of women in SES research is likely driven by a number of factors, including the reluctance of women to participate in studies as well as the decision of researchers to exclude them from their work due to the additional intricacy, time and expense involved in study designs that adequately consider sex-specific needs. Moreover, a recent shift in research focus to include eumenorrheic women for the investigation of MC-dependent outcome measures, by definition excludes over 50% of the athletic population who use HCs. The availability and recruitment of female athletes may also be hampered by a tendency for smaller team sizes, alongside disproportionately low numbers of women within teams/clubs or professional sports who might have the opportunity and availability to participate in research projects (Emmonds et al., 2019). It may also be that study designs do not adequately consider the needs of high-performance women from the athlete's perspective, aside from the aforementioned scientific sex-specific considerations. Ultimately, the underrepresentation of female athletes in SES research undermines a fundamental aim of the sports science/nutrition practitioner in implementing best-practise evidence-based guidelines. As such, a correction of the sex-bias in the SES literature is of primary importance.

While guidelines have been developed to guide researchers in conducting high quality studies in women with adequate methodological consideration of menstrual status (Elliott-Sale et al., 2021; Janse de Jonge et al., 2019), it is acknowledged that these extensive guidelines may appear overwhelming and potentially serve as an unintended barrier for researchers to conduct studies in women. Indeed, it can be challenging to achieve a rigorous level of control, particularly in applied SES settings, given the resource-intensive nature and elevated participant burden. Moreover, while gold-standard study designs are imperative in determining causation, stringent inclusion criteria may limit the applicability of findings to the dynamic training environment. Therefore, the practical translation of these guidelines into high quality research in applied SES practice is also of interest.

Summary of thesis aims

This thesis addresses dual aims: developing resources to aid researchers in conducting high-quality research among female athletes in the applied setting, while simultaneously furthering our understanding of any influence of ovarian hormonal profiles on performance. Each study in this thesis employs a distinct methodological approach, as well as including women of different menstrual statuses and implementing best-practice methodological classification and control of menstrual status. This

approach will encompass a broad range of experiences, allowing us to assess the practicality of implementing best-practice methods related to menstrual status in the applied research environment. Additionally, our insights will help develop resources to support future high-quality research in applied SES settings.

Each study examines performance outcomes to address the underrepresentation of elite female athletes in this area, alongside focussing on key methodological gaps identified among the existing literature: the inclusion of multiple MC phases, analysis of the E:P ratio, alongside the inclusion of trained (≥Tier 2) athletes across all studies. We aim to evaluate both mechanistic and applied performance outcomes between the sexes, across MC phases, and between women who are naturally menstruating and those using HC.

Overview of projects

A range of study designs will be used to explore different methods of conducting high quality research in women and elucidate strengths/limitations of each approach, as follows:

> ***Chapter 3*** uses a RCT in a Latin square design to examine sex differences in the responses to 24 h manipulations in EA on substrate oxidation, metabolism, and explosive performance, among Tier 2-3 (McKay et al., 2022b) cyclists/triathletes.
>
> ***Chapter 4*** employs an observational study design in a training camp environment among Tier 3 (McKay et al., 2022b) National Rugby League Indigenous Women's academy athletes. Measures of performance are assessed across phases 1, 2 and 4 of the MC, alongside comparisons between naturally menstruating athletes, and those using HC.
>
> ***Chapter 5*** is a novel observational study design among Tier 2 cyclists/triathletes, using a live virtual cycling competition to assess real-world competitive performance, examining the relationship between cycling performance and concentrations of oestrogen/progesterone.
>
> ***Chapter 6*** then collates the experiences of athletes currently involved in research to understand the unique perspectives of female athletes, in an attempt to increase female participation rates.

Overall, the studies in this thesis aim to provide knowledge and understanding that will contribute to addressing the sex disparity in the quality and quantity of SES research. Resources will be developed to guide researchers in adopting a high-quality approach for the consideration of participant menstrual status, ultimately improving evidence-based SES guidelines specific to female athletes to maximise their health and performance.

Chapter 2: Methodology

As per Australian Catholic University guidelines, a methodology chapter describing the methods used in each study in full is included. However, each chapter includes its own methods section as written for the respective journal for which the manuscript was published.

2.1 Study 1

Effects of 24-hour diet- or exercise-induced energy availability manipulations on substrate utilisation and performance.

2.1.1 Study design

Twenty endurance trained athletes (ten females and ten males) completed this randomised crossover trial. The study design included a baseline/familiarisation session, followed by five randomised experimental trials, each comprising three consecutive days (15 days of testing per participant; Figure 3.1). Trials were completed in a randomised order in a Latin square design, with an average of eight days separating trials (minimum four days, maximum 33 days). Doses of EA on trial day two were: low (15 kcal·kg^{-1}FFM·day^{-1}; LEA), high (45 kcal·kg^{-1}FFM·day^{-1}; HEA) and higher EA for mass gain/growth (75 kcal·kg^{-1}FFM·day^{-1}; GEA), with conditions of high and low EA separately achieved via manipulations of EEE or EI (Figure 3.1). Participants undertook post-intervention testing on day three at the same time of day (±8 min) across all five trials. Female participants completed each trial during the active (pill-taking) phase of the OCP cycle, when they were not experiencing a withdrawal bleed, to minimize fluctuations in both endogenous and exogenous ovarian hormones across trials. Women were instructed to take their daily OCP after post-intervention testing on day three to minimize effects of a bolus dose of exogenous hormones.

Baseline/Familiarisation. Participants underwent a baseline/familiarisation session 1-2 weeks prior to study commencement, which included a $\dot{V}O_2$max (to calculate subsequent EEE prescription), alongside familiarisation to the FATMAX test and performance measures [Stroop, Wingate, IMTP, CMJ and SJ]. Dual-energy X-ray absorptiometry (DXA) and resting metabolic rate (RMR) measurements occurred to establish EI and EEE prescriptions for subsequent trials. Additionally, participants received an activity tracker [Oura ring (Generation 3, Oura Health, Oulu, Finland)], to wear during each of the five subsequent three-day trial periods to monitor step count and estimated energy expenditure.

Trial day 1 – optimal EA standardisation. Participants consumed a standardised diet, providing 45 kcal·kg^{-1}FFM·day^{-1} (CHO; 4.7±0.5 g·kg^{-1}, protein; 2.1±0.2 g·kg^{-1}, fat; 0.8±0.1 g·kg^{-1}) for 24 hours prior to EA manipulation to ensure each intervention began in a state of optimal EA. Exercise was permitted but was replicated within each participant across all five trials (verified through the Oura ring), and

dietary intake adjusted accordingly. As such, EI and EEE was identical for trial day one across all five trials.

Trial day 2 – EA manipulation. The five EA conditions are outlined in Figure 3.1: LEA [with and without exercise (LEA_{EX} and LEA_{REST})], HEA [with and without exercise (HEA_{EX} and HEA_{REST})], or GEA (without exercise) (Loucks, 2013). For the two conditions involving exercise (LEA_{EX} and HEA_{EX}), participants completed two cycle sessions in the laboratory to achieve a total EEE of 30 $kcal \cdot kg^{-1}FFM \cdot day^{-1}$. Aside from prescribed exercise, participants remained inactive throughout the day, minimising activities of daily living (verified via Oura ring step count). For the three conditions not involving exercise (LEA_{REST}, HEA_{REST} and GEA) participants did not come to the laboratory but adhered to the provided diet and remained inactive (Oura ring verification).

Trial day 3 – post-intervention measures. Upon laboratory arrival in a 10-hour rested and fasted state, body composition was measured via DXA. A cannula was then inserted, and blood sample collected, followed by the FATMAX test (20-30 min). After a mixed meal tolerance test (120 min), physical performance measures were obtained: IMTP, CMJ, SJ and Wingate, alongside a questionnaire regarding perceived muscle soreness (Impellizzeri & Maffiuletti, 2007). Participants then rested for 30 minutes in a quiet, private room with *ad libitum* food, after which they underwent the Stroop Colour and Word Test for cognitive performance. Each individual method is described below in section 2.1.6.

2.1.2 Ethical approval

The study was approved by the Australian Catholic University Human Ethics Research Committee (2022-2561H) and conducted in accordance with the Declaration of Helsinki. All participants provided informed written consent prior to participating.

2.1.3 Participant characteristics

Ten female and ten male Tier 2-3 (McKay et al., 2022b) endurance trained athletes (Table 3.1) participated in this study. Participants were familiar with cycling even if not their primary sport (i.e., cross-training, or regular commuting). A sample size calculation (G*Power 3.1, Dusseldorf, Germany), using data from Chrzanowski-Smith et al. (2021), estimated that 10 males and 10 females were required to detect differences in PFO relative to FFM between the sexes, with 90% statistical power and an alpha of 0.05. To eliminate potential effects of menstrual status/phase, we recruited pre-menopausal females taking a combined OCP (see details in Table S3.1) for >three months prior to study commencement (median usage time was 4 years).

2.1.4 Dietary manipulation

Participants received all food and drink individually pre-packaged and weighed prior to the start of each three-day trial. The standardisation and intervention diets on trial days one and two (totalling a 48 h

period of dietary prescription for each trial) were prepared by an accredited dietitian using Nutritics (Nutritics LTD, Ireland), with consideration to dietary preferences. A food/drink checklist was provided, alongside food preparation information, and participants were instructed to contact the research dietitian immediately in the case of any deviations so that diet plans could be amended to maintain the prescribed EA. Diet prescription is outlined in detail by Kuikman et al. (2024b). In brief, diets prescribed an EI of 15 (LEA_{REST}), 45 ($LEA_{EX,}$ HEA_{REST} [and day one standardisation]) or 75 (HEA_{EX} and GEA) kcal·kg^{-1}FFM·day^{-1}. Macronutrient percentage distribution was equal between all EA conditions (alongside the day one standardisation diet) at 25% of EI from protein, 20% from fat and 55% from CHO (providing an intake of 1.6±0.2 g CHO·kg^{-1}·day^{-1} for LEA_{REST}, 4.7±0.5 g CHO·kg^{-1}·day^{-1} for $LEA_{EX,}$ HEA_{REST} and day one standardisation, and 7.5±1.2 g CHO·kg^{-1}·day^{-1}for HEA_{EX} and GEA). All EA conditions (alongside the day one standardisation diet) provided participants with three meals and three snacks. Participants were instructed to space out meals and snacks by at least one hour, and to consume the last snack 10-hours prior to laboratory arrival on day three. Caffeine consumption was permitted on trial days one and two, but not three, and replicated across each of the five trials. Alcohol was prohibited throughout each three-day trial period. Participants verbally confirmed the consumption of all food/drink upon arrival to the laboratory on trial day three.

2.1.5 Exercise manipulation

For the two EA conditions involving exercise (LEA_{EX} and HEA_{EX}), participants completed two cycle sessions in the laboratory on a stationary load bike (Load Excalibur Sport, Groningen, Netherlands) to achieve an EEE of 30 kcal·kg^{-1}FFM·day^{-1}. The evening session was 60 minutes at 65% $\dot{V}O_2$max (males, 195±46 W; females, 131±19 W), concluding 12 hours prior to next day laboratory arrival. The remaining EEE was completed in the morning at 55% $\dot{V}O_2$max with exercise duration manipulated to achieve 30 kcal·kg^{-1}FFM·day^{-1} EEE (males, 157±40 W for 135±26 min; females, 103±16 W for 163±37 min). The EEE at each cycling intensity was determined from gas exchange data collected during baseline $\dot{V}O_2$max testing. Expired gases were used to calculate substrate oxidation rates and energy expenditure in accordance with the stoichiometric equations outlined by Jeukendrup and Wallis (2005), assuming negligible protein oxidation. An athlete's RMR was then excluded from EE to determine EEE.

2.1.6 Test protocols

DXA and RMR: Both DXA and RMR were assessed at baseline, to calculate EI and EEE prescriptions. RMR was measured first in the 10 h rested and fasted state. Athletes rested in a dark and quiet room for 10 minutes and were then given a one-way mouthpiece for a 15-minute familiarisation period. Data collection then commenced whereby expired air was collected into two separate Douglas bags, with a 10 min data collection period per bag. Bags were analysed using ametek oxygen (O_2) and carbon dioxide (CO_2) analysers (VacuMed, Ventura, CA). Expired air from each bag was sampled for one minute with the gas sampling time and flow rate recorded. A Tissot spirometer was then used to determine the

volume of the remaining expirate via an evacuation pump. The 24-hour RMR (kcal·day^{-1}) was calculated from the concentration of expired O_2 and CO_2 alongside the volume of expired air. Immediately following the RMR measurement, scale body mass was measured and then FFM determined via DXA scan (GE Lunar iDXA, analysed using GE encore) according to the Australian Institute of Sport Best Practice Protocols (Slater et al., 2022). Performing a DXA scan at each laboratory visit allowed results to be normalised to FFM.

FATMAX and $\dot{V}O_2$max: The PFO and exercise intensity eliciting PFO (FATMAX) were assessed by an incremental cycling protocol on a load bike using the measured values approach (Achten et al., 2002). Starting at 30 W for females and 50 W for males, participants completed three-minute stages increasing by 25 W increments until RER >1.0. During familiarisation testing, participants completed an additional maximal exercise bout ($\dot{V}O_2$max test) following completion of the FATMAX protocol. When RER >1.0, instead of ceasing the test, wattage continued to increase in 25 W increments every 60 s until volitional exhaustion, as indicated by the participant. $\dot{V}O_2$max was taken as the highest $\dot{V}O_2$ value observed across a 30 s period. Chest HR (Forerunner, Garmin International) and rating of perceived exertion (RPE, 6–20, Borg Scale) were recorded at the end of each stage. Expired gas was collected and analysed using a custom-built indirect calorimetry system with associated in-house software as previously described (Saunders et al., 2004). The $\dot{V}O_2$ and $\dot{V}CO_2$ values from the last minute of each stage were used to calculate PFO using non-protein RER values (Peronnet & Massicotte, 1991).

Mixed meal tolerance test (MMTT): Participants consumed a breakfast meal of raisin toast, jam, and apple juice (males, 1035±148 kcal; females, 841±172 kcal; 2.00±0.00 g·kg^{-1} CHO; 0.27±0.03 g·kg^{-1} protein; 0.13±0.01 g·kg^{-1} fat) followed by a two-hour resting period. Meal consumption began at 0 min and finished within 15 min.

Blood sampling: At the start of each lab visit, a cannula was inserted into the antecubital vein by a trained phlebotomist while the athlete was in a rested and fasted state. A total of eight 1 ml blood samples were collected per trial: baseline (rested and fasted), alongside the following timepoints during the MMTT: 0 (pre-meal), 15, 30, 45, 60, 90 and 120 min. Blood tubes clotted at room temperature for 30 min before being centrifuged at 2200 G for 10 min at 4°C. The serum was split into aliquots and stored at -80°C until batch analysis. Glucose was measured via an automated colorimetric assay (AU480 chemistry analyser, Beckman Coulter, Brea, California, USA) with intra-assay coefficient of variations (CV) of 1.0%. Insulin was analysed via chemiluminescent immunoassay (Access 2 immunoassay system, Beckman Coulter, Brea, California, USA) with CV of 8.9%. Incremental area under the curve was calculated for glucose and insulin concentration using an automated tool (Narang et al., 2020).

Countermovement/squat jumps and isometric mid-thigh pull: Following a standardised warm-up and wearing the same shoes on all five occasions, participants completed the CMJ, SJ, and IMTP on a dual force plate system sampling at 1000 Hz (0.60 x 0.40 m; Model 10 kN 9286B, Kistler Instrument

AG, Winterthur, Switzerland). Participants first completed three repetitions each of the CMJ and then SJ with ~60 seconds rest between jumps. Participants were instructed to "jump as high and powerfully as possible" with their hands remaining on hips. For the SJ, participants jumped from a 90° squat (or as close as possible) without countermovement. Squat depth was standardised within participants between trials using a plastic pole that participants lowered themselves to, and an additional effort was performed if any countermovement was observed via the force-time trace. The highest jump repetition was analysed; if jump height was equal, then peak power was used to determine the "best" effort. Outcome measures included jump height (calculated through impulse-momentum), mean and peak concentric force, velocity, and power, alongside impulse and rate of force development at 50/100/150/200 ms, as well as contraction time, concentric time, eccentric time, and centre of mass displacement (Weakley et al., 2022). Jump initiation was identified using the criterion method (Owen et al., 2014).

Following two sub maximal warm-up efforts, participants performed two maximal repetitions of the IMTP separated by two minutes rest. Participants pulled at maximal effort for three seconds on an immovable bar fixed to a customised power rack. The bar was set during the familiarisation visit, such that joint angles at the knee and hip were between 125-145° and 140-150°, respectively (Comfort et al., 2019). Participants were instructed to "push the ground away as hard and as fast as possible". Verbal encouragement was maintained throughout. A third effort was performed if: >200 N difference was observed between the peak force of the two efforts; there was variability >50 N in the quiet period; there was a countermovement prior to the lift, excessive pre-tension, or leaning on the bar (Comfort et al., 2019). The highest relative peak force effort was analysed. Pull initiation was identified as the moment when force exceeded five standard deviations (SD) of a participant's body mass (Comfort et al., 2019), established through a one-second stable weighing period. Peak force, time to peak force, rate of force development and impulse at 50/100/150/200/250 ms were calculated.

All ground reaction force-time data for the CMJ, SJ and IMTP were recorded using ForceDecks software (VALD ForceDecks, 2.0.8587) and then exported for analysis via a customised R script. CMJ and SJ jump heights were also used to calculate the eccentric utilisation ratio (EUR) and reactive strength index (RSI), while the dynamic strength index (DSI) was calculated from CMJ peak concentric force and IMTP peak force.

Wingate: Participants performed a five-minute standardised cycling warm-up, which included three six-second sub-maximal sprints. Participants then completed a 30 second all-out cycling effort (Wattbike Pro, Nottingham, England) at maximal speed against a high braking force from a rolling start. Participants were instructed to "pedal as hard as possible from the start without pacing the effort but remaining in the saddle". Verbal encouragement was maintained throughout. Outcome measures were peak power, mean power, and fatigue index.

Stroop Colour and Word Test: Participants were shown coloured words on a laptop and asked to indicate the word's colour (and not it's meaning) by pressing a key as fast as possible whilst minimising errors (Stroop, 1935). Coloured labels were placed on keyboard keys to signify the corresponding colour. Three types of trials were presented: control (coloured rectangles), congruent (words of matched colour and meaning) and incongruent (words with mismatched colour and meaning). A red "X" flashed onto the screen when an incorrect response occurred. Each test had 180 trials, taking approximately three minutes to complete. The Stroop test was administered using Inquisit 6 [6.6.1 64bit, (Windows 10), (2020) retrieved from https://www.millisecond.com] in a quiet, private room. The Stroop effect was calculated as the difference between responses (both the proportion correct/accuracy and reaction time) in the incongruent versus congruent trials.

Muscle soreness. A seven-point Likert scale for lower limb muscle soreness (Impellizzeri & Maffiuletti, 2007) (Table S3.2) was completed at 0, 60 and 120 min during the MMTT, with the mean score used in analysis.

2.1.7 Statistical analysis

Results were compared across EA conditions and between sexes using linear mixed models. Fixed effects were "condition" (LEA_{REST}, LEA_{EX}, HEA_{REST}, HEA_{EX}, GEA), and "sex" (female or male), with "subject identification" as a random effect. Statistical significance of fixed effects occurred using type II Wald tests with Kenward–Roger degrees of freedom. Where significant fixed effects were established, pairwise comparisons were performed with Tukey *post hoc* adjustments. Significance was accepted at $p<0.05$. Data are presented as mean±SD with non-normal data (assessed via histogram inspection) log-transformed prior to analysis. Outliers > three SD beyond the group mean were removed (Howell et al., 1998). Muscle soreness data are missing for one condition (LEA_{REST} and HEA_{REST}) for two females due to a failure in the server administering questionnaires. SJ results are missing for a single condition (LEA_{REST} and HEA_{EX}) for two females due to technical data collection issues. Two females are missing one timepoint in the LEA_{REST} condition (15 and 90 min) for glucose and insulin because of cannula blockage; linear interpolation was used to address this when calculating the iAUC (Narang et al., 2020).

2.2 Study 2

Minimal influence of the menstrual cycle or hormonal contraceptives on performance in female rugby league athletes.

2.2.1 Study design

Twenty-four female Australian National Rugby League's Indigenous Women's Academy athletes attended a five-week residential training camp at the Australian Institute of Sport. The camp duration was selected to facilitate the monitoring of a complete MC in participants with a cycle length less than 42 days. This study implemented an observational design within a training camp environment. The group was initially divided into those reporting the use of HC (athletesHC) and those who were considered by their self-reports as being naturally menstruating (athletesNM) until menstrual status was studied during the project. Following two familiarisation sessions, a battery of performance tests was completed on three separate occasions across each participant's individualised menstrual or HC cycle. For athletesNM, the three phases occurred in a randomised order, determined by the menstrual phase in which they commenced the training camp.

2.2.2 Ethical approval

The study was approved by the Australian Catholic University Human Ethics Research Committee (2021-285H) and conducted in accordance with the Declaration of Helsinki. All participants provided informed written consent prior to participating.

2.2.3 Participants

Athletes were nominated for participation by their National Sporting Organisation. A mixture of athletes with different menstrual statuses (i.e., those using HC and those with a "natural" cycle) was desired, however the recruitment of an authentic squad was considered of greater priority. Accordingly, no exclusion criteria based on MC function were implemented. Athletes were classified as Tier 3 (national level) according to the classification framework outlined by McKay et al. (2022b). Forty-three athletes initially provided written informed consent, however, barriers to participation meant that only 25 athletes commenced the study. One athlete returned home during the study, resulting in 24 full data sets for analysis. The sample size of 24 is reflective of most real-world rugby squads for which a coach or sports scientist may be asked to consider menstrual phase or status-based testing at a group level.

2.2.4 Menstrual status

Menstrual status was tracked in both athletesNM and athletesHC according to best-practice protocols (Elliott-Sale et al., 2021); recording onset of bleeding, performing 16 weeks of MC or HC tracking, using dual hormone urinary ovulation kits, and assessing retrospective serum 17-β-oestradiol (the most

potent form of oestrogen among pre-menopausal women, henceforth referred to as "oestradiol") and progesterone concentration. Prior to camp commencement, athletes completed a preliminary questionnaire regarding their MC history, including any current or previous HC use (type, formulation), the length and frequency of their MC (including determination of any primary and secondary amenorrhea) and prevalence of known menstrual dysfunction diagnoses (e.g., polycystic ovary syndrome [PCOS], endometriosis). One athlete reported a prior diagnosis of PCOS, with no other MC dysfunction reported across the cohort. Subsequently, for a total of 16 weeks (11 weeks prior to the camp plus the five-week training camp), athletes received a daily text message at 8:00 am with a link to an electronic survey. The survey was customised for either athletesNM or athletesHC. All athletes reported the presence and heaviness of menstruation, alongside any symptoms and medication use across the previous 24 h. In addition, athletesNM used a dual hormone ovulation kit (Advanced Digital Ovulation Test, Clearblue, Geneva, Switzerland) from MC day eight until ovulation occurred. If ovulation was not detected, athletes were instructed to stop using ovulation kits on day 17; however, additional days were often requested to increase the chances of capturing ovulation.

Athletes attended the laboratory on three occasions according to their MC function and MC phase. The 11 weeks of MC monitoring prior to camp was used to identify individual potential testing dates to facilitate camp planning, with confirmation of actual testing 24 h prior (when the athlete had completed their MC questionnaire for that day). AthletesNM completed performance testing in MC phase 1 (low oestradiol and low progesterone, phase 2 (elevated oestradiol and low progesterone), and phase 4 (elevated oestradiol and elevated progesterone). Phase 3 was also of interest, however due to athlete availability and to reduce participant burden data collection was limited to three phases. Phase 2 was selected in preference to Phase 3 due to the greater differential between sex hormones (elevated oestradiol and low progesterone) and the scarcity of research examining this phase. Phase 1 testing occurred at day 1.8±0.4, after the athlete reported the commencement of menstruation. Phase 2 occurred on day 11.4±1.4 and was determined by a "flashing smile" icon on the urinary ovulation kits, as reported on the daily MC questionnaires, which indicates an elevated oestrogen metabolite concentration and precedes the LH surge. Serial blood samples were not taken in the days preceding the anticipated Phase 2, and hence the "peak" in oestrogen concentrations could not be instantly confirmed. However, retrospective analysis of venous blood samples verified the absence of late collections (i.e., Phase 3). Phase 4 testing occurred on day 20.8±1.6 and was determined as seven days following ovulation. In situations where ovulation was not detected, an arbitrary "day 21" was instead taken.

AthletesHC completed performance testing at three arbitrary timepoints, separated by 7-10 days to replicate the pattern of blood collection from athletesNM. AthletesHC using OCPs were tested during active pill taking days only and were instructed to take their pill at the same time of day on each testing occasion. Six athletesHC using the contraceptive implant had this inserted between one and three years prior to testing and two athletesHC had this inserted the same month as testing commenced. The

athleteHC using the hormonal injection had her last injection three weeks prior to the first test. Accordingly, all athletesHC completed testing at time points where an assumed, consistent supply of exogenous hormones occurred.

2.2.5 Blood sampling

Prior to performance testing at each visit, a trained phlebotomist collected an 8.5 mL venous blood sample from an antecubital vein into a serum separator tube, while the athlete was in a rested and fasted state. Blood tubes clotted at room temperature for 30 min and were then centrifuged at 2200 G for 10 min at 4°C. The remaining serum was split into aliquots and stored at -80°C until batch analysis. Oestradiol and progesterone were measured via an Access 2 Immunoassay System (Beckman Coulter, Brea, CA, USA) with intra-assay coefficient of variations (CV) 5% and 11% for oestradiol and progesterone, respectively. Total testosterone was analysed using liquid chromatography-tandem mass spectrometry (Waters UPLC-TQX S, Waters Corp., Wilmslow, UK), with a total imprecision CV of 5.8%, and free testosterone was subsequently calculated from total testosterone alongside sex hormone binding globulin and albumin (Vermeulen et al., 1999).

2.2.6 Performance testing protocols

Participants completed a battery of performance tests on three separate occasions across each participant's individualised menstrual or HC cycle. Tests were undertaken at the same time of day (±15 min) across a 90 min period, wearing the same shoes, after completing a standardised warm-up, and adhering to a standardised diet from lunch onwards the day prior to testing (~18h). The warm-up consisted of five minutes cycling on a stationary bike at a perceived "easy" intensity including 3x4 second sprints at 90% of maximal perceived cadence, followed by 10 each of walking lunges, squats, leg swings and calf raises, and concluding with three CMJ each at 70% and 90% of perceived maximal effort.

The CMJ, SJ and IMTP were conducted on a dual force plate system sampling at 1000 Hz (0.60 x 0.40 m; Model 10 kN 9286B, Kistler Instrument AG, Winterthur, Switzerland). Participants were familiarised at two separate sessions to the CMJ, SJ and IMTP protocols, alongside the Stroop Colour and Word Test, during the first two days of the training camp. Specific familiarisation was not undertaken for the power pass or 20 m sprint as these are regularly performed as part of the National Rugby League testing battery. These tests were selected as they represent different domains of performance (James et al., 2023), were familiar to participants, are commonly used throughout the literature with rugby athletes (Owen et al., 2020) and demonstrate acceptable between-day reliability and ecological validity (Weakley et al., 2022).

Countermovement jump and squat jump: Participants completed three repetitions each of the CMJ and SJ with ~60 s rest between jumps (Weakley et al., 2022). Participants were instructed to jump as

high and powerfully as possible with their hands remaining on hips (both CMJ and SJ). For the SJ, participants jumped from a 90° squat (or as close as possible) without any countermovement. An additional effort was performed if any countermovement was observed. Squat depth was standardised within participants between trials using a plastic pole that participants squatted to reach until touching. The highest jump at each test was taken for analysis; if jump height was equal, then peak power was used to determine the "best" effort.

Outcome measures included jump height (calculated through impulse-momentum), mean and peak concentric force, velocity, and power, alongside impulse and rate of force development at 50/100/150/200 ms, as well as contraction time, concentric time, eccentric time, and centre of mass displacement. Jump initiation was identified using the criterion method of taking the instant when vertical force was less or greater than a threshold equal to five times the SD of body mass measured during a one second stable weighing period (Owen et al., 2014). Jump heights in the CMJ and SJ were also used to calculate the EUR and RSI, while the DSI was calculated from CMJ peak concentric force and IMTP peak force.

Isometric mid-thigh pull: Following two-sub maximal warm-up efforts, participants performed two maximal repetitions of the IMTP separated by 2 min rest. Participants pulled as hard as possible for 3 s on an immovable bar fixed to a customised power rack. Participants were instructed to "push the ground away as hard and as fast as possible". Verbal encouragement was maintained throughout. A third effort was performed if: >200 N difference was observed between the peak force of the two efforts, there was variability >50 N in the quiet period, there was a countermovement prior to the lift, excessive pre-tension, or leaning on the bar. The effort with the highest relative peak force was taken for analysis. Initiation of the pull was identified as the moment when force exceeded five SDs of a participant's body mass, established through a one second stable weighing period. Peak force, time to peak force, rate of force development and impulse at 50/100/150/200/250 ms were calculated.

All ground reaction force-time data for the CMJ, SJ and IMTP were recorded using ForceDecks software (VALD ForceDecks, 2.0.8587), and then exported for analysis via a customised R script. The kinetic and kinematic outcome variables were selected as they represented different domains of force expression and also provided information that could provide context in relation to changes in temporal performance and movement strategy. Furthermore, ratio data (e.g., DSI) were provided to give context on whether force expression changed relative to difference strength domains (e.g., isometric vs dynamic strength).

Power pass: Athletes stood with their feet shoulder width apart and pushed a 3 kg med ball from the chest as far as possible into a long-jump pit. Countermovement in the legs was permitted, but feet were not permitted to leave the ground. The throw distance was measured from the back of the imprint left by the ball in the sand to the nearest cm. The furthest throw at each test was used in analysis.

20 m sprint: The 20 m sprint was conducted on an indoor athletics track with four light gates (Fusion SmartSpeed V2) positioned at 0/5/10/20 m, measuring at a height of 57 cm (0 m gate) and 87cm (5/10/20 m gates). From a split-stance position, 10 cm behind the first light gate as marked-up on the track (Weakley et al., 2023), participants sprinted at maximal effort for 20 m. The start was initiated when participants broke the plane of the first light gate. An additional light gate, alongside tape to signify a "finishing line", was placed at ~23 m. Participants were instructed to run through this line to prevent deceleration prior to 20 m. Each participant completed a warmup sprint, followed by two maximal efforts, with the fastest taken for analysis.

Stroop Colour and Word Test: Coloured words were displayed on a laptop and participants were asked to indicate the colour of the word (not it's meaning) by pressing a corresponding key as fast as possible while minimising errors (Stroop, 1935). Coloured labels were placed on keyboard keys to signify the corresponding colour. Three types of trials were presented: control (coloured rectangles), congruent (words of matched colour and meaning) and incongruent (words with mismatched colour and meaning). A red "X" flashed onto the screen in the event of an incorrect response. There were 180 trials for each test, taking approximately three minutes to complete. The Stroop test was administered using Inquisit 6 [6.6.1 64bit, (Windows 10), (2020) retrieved from https://www.millisecond.com] in a quiet, private room. The Stroop effect was calculated as the difference between responses (both the proportion correct/accuracy and reaction time) in the incongruent versus congruent trials.

2.2.7 Statistical analysis

Statistical analyses were performed using R Studio (v3.5.2) with statistical significance accepted at an α level of $p \leq 0.05$. Two separate approaches were taken for statistical analyses. Initially, outcome measures were compared both within individuals (i.e., across menstrual or HC cycle phases) and between individuals (i.e., between athletesNM and athletesHC) – termed "phase-based analysis". Linear mixed models were used to analyse each variable, using "menstrual status" and "cycle phase/test day" as fixed effects, alongside "subject identification" and "test order" as random effects. Statistical significance of fixed effects was identified using type II Wald tests with Kenward–Roger degrees of freedom. Where significant fixed effects were established, pairwise comparisons were identified using Tukey *post hoc* adjustments. Non-normally distributed data were identified through histogram inspection [Stroop outcomes, rate of force development (RFD) and impulse during the IMTP, RFD, flight time (FT): contraction time (CT), CT and concentric time during the CMJ, impulse during the SJ, alongside EUR] and were log transformed prior to statistical analyses. An independent t-test was conducted to compare total training load between groups (athletesNM vs athletes HC).

Following analysis of serum oestradiol and progesterone concentrations, it was determined that a 'true' phase 2 (where oestradiol concentration exceeds that achieved in phase 4) was only achieved in one out of 11 athletesNM (McKay et al., 2024), and results were therefore compared across phases 1 and 4 only.

Three athletesNM were also excluded due to hormonal profiles not meeting the criteria for phase 4 (progesterone >16 $nmol \cdot l^{-1}$). As such, phase-based analyses were performed in n=8 athletesNM. Therefore, a repeated measures correlation was also used to assess associations between performance measures and oestradiol or progesterone concentration, alongside E:P ratio and oestradiol: serum free testosterone ratio (E:T), – termed "correlation analysis". Correlations were conducted among athletesNM exclusively, given that a) only endogenous hormones were measured and b) there was potential for variable results outside of hormonal influences due to the largely unknown effects of the exogenous hormonal milieu in athletesHC. This analysis approach did not require discrete MC phases, and thus "phase 2" results were included, alongside results from athletes with only two out of three completed tests, resulting in n=11 athletesNM. A single progesterone value from the athlete with PCOS was excluded from correlational analysis because it was >2.5 SD above the mean.

2.3 Study 3

Perceived negative menstrual cycle symptoms, but not changes in oestrogen or progesterone, are associated with impaired cycling race performance

2.3.1 Study design

In a novel observational study design thirty-seven female cyclists/triathletes completed a total of four virtual indoor cycling races, one per week across a one-month period. Venous blood samples were collected within 21 h of each race (pre- or post-race) to determine serum 17-β-oestradiol and progesterone concentration. The concentrations of these sex hormones were then matched with the respective race and correlated to each participants' race completion time. The incidence of self-reported MC and GI symptoms on race day were also correlated to race time as a secondary outcome measure.

2.3.2 Ethical approval

The study was approved by the Australian Catholic University Human Ethics Research Committee (2023-3192H) and conducted in accordance with the Declaration of Helsinki. All participants provided informed written consent prior to participating.

2.3.3 Participant characteristics

Thirty-seven female cyclists/triathletes (mean age: 35±6 y, mean body mass: 67.0±10.3 kg, mean training volume: 8.0±3.5 $h \cdot wk^{-1}$, mean age of menarche: 13±3 y) were recruited. Athletes were classified as Tier 2 (trained/developmental level) according to the classification framework outlined by McKay et al. (2022b). Inclusion criteria were: residing in Australia, pre-menopausal (confirmed via ovulation detection), absence of HC use for >three months prior to study commencement, not pregnant or breastfeeding (Figure 5.2). The only exclusion criteria based on MC function was current amenorrhea (absence of a MC for >three months) (Elliott-Sale et al., 2021), given that it results in the suppression of endogenous hormones and hence would prohibit the investigation of the primary outcome. Other menstrual irregularities that do not as severely suppress endogenous hormonal profiles were therefore included to increase study generalisability. Of the 108 athletes who expressed interest in participating, 49 were eligible and consented to study participation. Eight participants then withdrew during the pre-race period of MC tracking, while two participants withdrew during the one-month race period. Finally, two participants were excluded from analysis due to noncompliance with pre-race standardisation and incorrect bike calibration. This totalled 37 participants for analysis.

2.3.4 Menstrual status

Prior to participation, athletes completed an initial questionnaire regarding their menstrual status, including MC length and frequency, prevalence of known MC dysfunction [e.g., PCOS, amenorrhea],

and any current or previous HC use. Three athletes reported a diagnosis of PCOS and two reported a diagnosis of endometriosis. Participants' MCs were then tracked, according to best-practice protocols (Elliott-Sale et al., 2021), across the four weeks of racing – with additional weeks before or after to capture two complete MCs per athlete. Participants completed daily online questionnaires [REDCap (Harris et al., 2019; Harris et al., 2009)] pertaining to presence and heaviness of menstruation, symptom incidence, and medication use in the preceding 24 h (Figure S5.1). Athletes also used dual hormone urinary ovulation kits (Advanced Digital Ovulation Test, Clearblue, Geneva, Switzerland) from MC day 10 until ovulation occurrence (continuing until the next bleed if ovulation was not detected), recording the result on the online questionnaire. Venous blood samples were collected within 21 h of racing (pre- or post-race) to determine progesterone and serum 17β-oestradiol (the most potent form of oestrogen among pre-menopausal women) concentration.

2.3.5 Virtual cycling races

Races were an individual TT format: 19.5 km in length with 32 m elevation, completed on the Zwift online cycling platform (2023 Zwift, Inc. v2.183.0). The race was a private event open only to study participants, whereby participants could see all other competitors in the ride to replicate a real race environment. The Zwift software was programmed to a standardised bike setting, while drafting and powerups were disabled. The indoor trainer (n=33) or stationary bike (n=4) was consistent within each participant across all their races. Participants raced every Thursday evening across four consecutive weeks, commencing at 19:45 AEDT. Participants chose their own warm-up and replicated this each week. The race was completed indoors, with permission to use fans or air conditioning and the use of these recorded.

To enhance the ecological validity and motivation, prize money was available to the top performers. Participants were grouped into categories (A-D) based on ability ($W \cdot kg^{-1}$) as determined by Zwift (2024): Cat A: 3.70-5.0 $W \cdot kg^{-1}$, Cat B: 3.20-3.69 $W \cdot kg^{-1}$, Cat C: 2.50-3.19 $W \cdot kg^{-1}$, Cat D: 1.0-2.49 $W \cdot kg^{-1}$. At each race, participants provided a photo of themselves standing on a scale pre-race to verify body mass. Prize money was awarded separately across each category, such that riders were only directly competing against individuals of a similar ability (although they could see all other riders during the live race). Participants voted on the prize money allocation system, and the number of prizes awarded was adjusted based on the total number of athletes in each category, such that the top 30% of riders in each category were awarded a prize.

2.3.6 Pre-race standardisation

Dietary intake (all food, beverages, and caffeine consumption) was standardised for 36 hours pre-race, with participants allowed to choose their own nutrition strategies but repeat them for each race. Dietary records were maintained to verify compliance with these instructions, via the use of meal photos posted

on the MealLogger app (MealLogger). Alcohol was prohibited throughout both days. Training was permitted the day before the race but was kept consistent every week and recorded on Strava/Garmin for verification. No training was permitted on the day of the race, with the exception of one athlete who completed the same 45 min run on the morning of each race.

2.3.7 Pre- and post-race questionnaires

Before (within 15 mins of race commencement) and immediately after each race, participants completed an online questionnaire [REDCap (Harris et al., 2019; Harris et al., 2009)] regarding GI symptoms and thermal perception [thermal sensation (TS) and thermal comfort (TC)]. Any GI symptoms: nausea, vomiting, belching, bloating, stomach pain, gastric acidosis, constipation, diarrhoea, urge to defecate and gas, were reported on a 10-point Likert scale [1=no symptoms, 10=extreme symptoms resulting in race withdrawal (Table S5.1), with a score ≥5 deemed to be at least "moderate" in severity (Pugh et al., 2019)]. Participant TS and TC were measured on nine-point (very cold to very hot) and six-point (very uncomfortable to very comfortable) Likert scales, respectively [Tables S5.2 and S5.3, (Zhang et al., 2004)]. Visual analogue scales (0-100) measured readiness to race (pre-race only): "how ready to race do you feel?", with 0 representing "not at all ready" and 100 as "the most ready I have ever felt", and race perception (post-race only): "how do you feel like you raced?", with 0 representing "the worst I have ever raced" and 100 as "the best I have ever raced".

2.3.8 Blood sampling

Each week, using pre-organised pathology request slips, participants attended the same commercial pathology branch (Australian Clinical Labs) to have a rested blood sample drawn (total of four samples). An 8.5 mL venous blood sample was collected by a trained phlebotomist into a serum separator tube. Oestradiol and progesterone were measured via a Siemens Atellica IM Analyzer using a direct chemiluminescent immunoassay. Four participants did not reside in the locale of an Australian Clinical Labs centre and therefore attended an alternative pathology centre (Healius Pathology). Participants were advised to complete their blood test the morning prior to the race; however, athletes were not excluded from participation if this was not achievable. Hence, the blood samples were collected either the morning of (74% of participants) or after (26%) each race at the same time each week (±1.3 hours), at a mean time of within 11.5 h of the race start and all samples were collected within 21 h of the race.

2.3.9 Statistical analysis

Statistical analyses were performed using R Studio (v3.5.2) with statistical significance accepted at an α level of $p \leq 0.05$. Data are presented as mean±SD. Hormone concentrations > three SD from the group mean were removed as outliers [three elevated oestradiol measures and one elevated progesterone measure, (Howell et al., 1998)]. Repeated measures correlations assessed associations between race time and oestradiol/progesterone concentration and the progesterone: oestradiol ratio (P:E. as nmol·L^{-}

[1]) as our primary outcome measures, alongside our secondary outcome measures: total perceived MC symptoms, GI symptoms of at least moderate severity, and changes in TC/TS pre- to post-race. Because these secondary outcomes are ordinal measurements, they present some analysis limitations, however a non-parametric alterative to repeated measures correlation does not exist. A one-way ANOVA or paired t-test assessed differences in race completion time and participant weekly training volume across the four races, alongside participant mean weekly training volume during race weeks compared to volume across non-race weeks.

Sub-analyses using paired t-tests were conducted for athletes who experienced: menses during a race (n=24) and/or ovulation within 24 hours of the race (n=9) with race performance during these events compared to the mean performance across other races. Race completion time during follicular vs. luteal phases, as separated by ovulation, was also compared for athletes completing at least one race in each phase (n=31). Finally, sensitivity analyses were performed (Deeks et al., 2019), whereby results were analysed separately excluding athletes with menstrual irregularities [(MI), n=8, 27 races] and races with minor protocol deviations (six races).

2.4 Study 4

Original investigation: Female athletes report positive experiences as research participants

2.4.1 Study design

The experiences of participants who undertook four separate experimental studies were collated (Kuikman et al., 2024a; Smith et al., 2024a; Smith et al., 2024b; Smith et al., 2024c), each employing a unique study design. Upon study completion, participants completed a voluntary questionnaire regarding their experiences of participating in the study.

2.4.2 Ethical approval

Ethical approval was received separately for each study presented.

Smith et al. (2024a): 2022-2561H (Chapter 3)

Smith et al. (2024c): 2021-285H (Chapter 4)

Smith et al. (2024b): 2023-3192H (Chapter 5)

Kuikman et al. (2024a): 2022-2701HC

All studies were conducted in accordance with the Declaration of Helsinki. Specifically, each study included explicit participant consent to complete the questionnaire presented in Study 4, for which participants provided their written consent prior to enrolling.

2.4.3 Participant characteristics

Data from a convenience sample of 89 female athletes (Table 6.1) aged 18-45 from four separate experimental studies conducted in Australia are presented. Athlete performance/fitness status were tiered (Tier 0 = Sedentary; Tier 1 = Recreationally Active; Tier 2 = Trained/Developmental; Tier 3 = Highly Trained/National Level; Tier 4 = Elite/International Level; Tier 5 = World Class) according to McKay et al. (2022b), as follows: n=10 Tier 2-3 cyclists/triathletes [Study 1, (Smith et al., 2024a)], n=22 Tier 2-3 National Rugby League Indigenous Women's Academy players [Study 2, (Smith et al., 2024c)], n=38 Tier 2 cyclists/triathletes [Study 3, (Smith et al., 2024b)], and n=19 Tier 3-5 race walkers [Study 4, (Kuikman et al., 2024a)]. The first study also included 10 male athletes whose data are included separately for between-sex comparisons only. Full participant information and study details, including ethical approval, can be found in the respective publications.

All studies achieved a minimum of "silver" tier regarding the classification and control of menstrual status according to the tiering system outlined by Smith et al. (2022b), which classifies the extent to which best practice guidelines (Elliott-Sale et al., 2021) have been adhered to. A "silver" tier reflects

studies that achieve the majority of recommendations for best-practice, while a "gold" tier study would implement every methodological recommendation. Divergence from the "gold" tier methodological standards occurred around HC use in Studies 1, 2 and 4, where participants using HC were not restricted to a single contraceptive type (i.e., included a mixture of methods such as OCPs or IUS). Furthermore, Studies 2, 3 and 4 examined outcome measures across a single MC only, rather than repeating the outcomes across ≥two separate cycles to verify a consistent response as is recommended.

2.4.4 Questionnaire

Participants independently completed a voluntary electronic questionnaire [(REDCap (Harris et al., 2019; Harris et al., 2009)], distributed electronically upon study completion (Figure S6.1), without input from the principal investigator or data analysis team. Questions were separated into 3 category themes:

A. **Prior participation** in *other* research studies:

   Participants were asked if they had participated in research before, and if not, to state their primary reason for non-participation.

B. Reasons for **participating** in the *current* research study:

   Participants were asked to rate the degree to which different aspects influenced their decision to participate in the study on a visual analogue scale from 0-100, whereby "0" was "did not influence their decision at all" and "100" was "very much influenced their decision".

C. **Experiences** during the *current* research study:

   1) Participants were asked to rate each test completed during the study, whereby the *experience* of the test itself (e.g., the actual process of undertaking a blood test or tracking their MC) was differentiated from the *feedback* received post-study (e.g., blood test results). Participants received individualised results following study completion (prior to full data analysis and publication). Importantly, these results were provided to participants prior to completing the questionnaire. Each test was rated out of 10 on a Likert scale, whereby "1" represented "the worst experience" and "10" was "the best experience".

   2) Participants were asked to rate their overall experience on a visual analogue scale from 0-100, whereby "0" was "the worst study ever" and "100" was "the best study ever", alongside if they would apply the study findings to their sports involvement and if they would be willing to participate in future studies.

2.4.5 Statistical analysis

Statistical analyses were performed using R Studio (v3.5.2) with statistical significance accepted at an α level of $p \leq 0.05$. Only the survey responses from female participants (n=89) are reported across themes A-C. Males are included exclusively for the purpose of sub-analysis between sexes in Study 1 only. The

statistical approach for each section is outlined, with normality assessed through histogram inspection. With the exception of the sub-analyses between participants of different athletic calibre, statistical comparisons were not conducted between studies due to the highly skewed and heterogeneous nature of the data, as well as the substantially different participant demands across the four studies.

Theme A (prior research participation) does not include statistical analysis and presents the proportion of participants who responded "yes" or "no" to prior research participation, alongside their primary reason for this response.

Theme B (reasons for current study participation) presents the median±interquartile range (IQR, due to skewed data) score for each participation reason. Due to non-normality of this variable, Spearman's rank correlation coefficient examined the relationship between perceived importance of prize money, and the actual prize money received in Study 3.

Theme C (1: study procedures) presents the median±IQR rating of each test conducted during the study, due to non-normal ordinal data. The experience of completing the test was compared to the corresponding feedback received post-study (described above) using a Wilcoxon Signed-Rank test (i.e., analysing within-participant differences for test feedback *vs* experience). Since each study procedure was conducted independently, with varying numbers of participants completing each test based on study requirements, tests were not combined into a single model. Only tests performed in more than one study are analysed, with participant responses for the same test pooled across studies. This approach was considered appropriate as all studies were performed in the same laboratory according to identical institutional protocols, ensuring comparability of the same test across studies (e.g., DXA scans in Studies 1, 2, and 4 were conducted identically). The full data set, including all procedures separated by study, is available in the supplementary material (Figure S6.2).

Theme C (2: Overall study experiences) does not include statistics and presents the mean±SD for participants' overall rating of their research experience, due to normal distribution of this variable.

Sub-analyses: comparisons between sexes. Compares the responses of n=10 female and n=10 male participants in Study 1 only. Responses to Themes B and C(2) were normally distributed and therefore reported as mean±SD, with an independent t-test and Cohen's d effect size used to compare the perceived importance of participation reasons between the sexes. Theme C(1) is presented as median±IQR, due to non-normal ordinal data, with a Mann Whitney U test and Rank-Biserial Correlation (r_{rb}) to evaluate experiences of the test procedures between sexes.

Sub-analyses: comparisons between athletic tiers. The responses of highly trained athletes (n=19) in Study 4 were compared to those of lower-calibre athletes (Tier 2-3, n=70) in studies 1-3. Athletes in Study 4 were primarily Tier 4-5 (n=17), with n=2 classified as Tier 3. However, it was deemed appropriate to include these two athletes with the higher calibre group given they were part of the same high-level training environment. Due to aforementioned non-normality of Themes B and C(1), data are reported as median±IQR with comparisons between athletic tiers made using a Mann Whitney U test and Rank-Biserial Correlation (r_{rb}). Theme C(2), however, was normally distributed and is presented as mean±SD, with a Welch's test to account for unequal group sizes used to compare the overall research experience between tiers.

Chapter 3: Effects of 24-hour diet- or exercise-induced energy availability manipulations on substrate utilisation and performance

Publication statement:

This chapter comprises the following paper accepted for publication in *Medicine and Science in Sports and Exercise.* Published online ahead of print.

Smith, E. S., Kuikman, M. A., Weakley, J., Tee, N., McCormick, R., Ackerman, K. E., Elliott-Sale, K. J., Stellingwerff, T., Harris, R., McKay, A. K., & Burke, L. M. Effects of 24-hour diet- or exercise-induced energy availability manipulations on substrate utilisation and performance. *Medicine and Science in Sports and Exercise*, (aop). https://doi.org/10.1249/MSS.0000000000003608.

Data availability: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Ethical approval: Ethical approval was granted by the Australian Catholic University Human Ethics Research Committee (2022-2561H) in accordance with the Declaration of Helsinki.

Conflicts of interest: Authors declare no conflict of interest. The results of the present study do not constitute endorsement by ACSM. Authors declare that the results are presented clearly, honestly, and without fabrication, falsification, or inappropriate data manipulation.

Funding: This project was co-funded by the Australian Catholic University and Wu Tsai Human Performance Alliance.

3.1 Abstract

Purpose

To examine sex-based differences in substrate oxidation, postprandial metabolism, and performance in response to 24-hour manipulations in EA, induced by manipulations to EI or EEE.

Methods

In a Latin Square design, 20 endurance athletes (10 females using monophasic OCPs and 10 males) undertook five trials, each comprising three consecutive days. Day one was a standardized period of high EA; EA was then manipulated on day two; post-intervention testing occurred on day three. Day two EA was low/high/higher EA (LEA/HEA/GEA) at 15/45/75 kcal·kg^{-1}FFM·day^{-1}, with conditions of LEA and HEA separately achieved by manipulations of either EI or EEE (LEA $_{REST/EX}$ vs. HEA$_{REST/EX}$). On day three, fasted PFO during cycling and two-hour postprandial (high carbohydrate and energy meal) metabolism were assessed, alongside several performance tests: Wingate, CMJ, SJ, IMTP, and the Stroop Colour and Word Test.

Results

Highest PFO occurred under LEA induced by exercise ($p<0.01$), with no difference between sexes. Postprandial glucose ($p<0.01$) and insulin ($p<0.05$) responses were highest across both sexes when LEA was induced by diet. Relative peak and mean power throughout the Wingate, alongside CMJ height did not differ between EA conditions ($p>0.05$), while SJ height was lower during GEA than both LEA$_{REST}$ ($p=0.045$) and HEA$_{EX}$ ($p=0.016$). IMTP peak force and the Stroop effect did not change with altered EA ($p>0.05$).

Conclusion

Acute (24-hour) exercise-induced LEA influenced fasted substrate oxidation more than diet-induced LEA, while 24 hours of LEA did not impair strength/power, sprint capacity, or cognitive performance. Finally, the responses to EA manipulations did not differ between sexes.

Keywords: Relative Energy Deficiency in Sport, energy intake, energy expenditure, energy balance, athletes, cycling

3.2 Introduction

EA is defined as the difference between dietary EI and EEE, expressed relative to FFM and represents the energy remaining for physiological functions (Loucks, 2004; Loucks et al., 2013). Originally, short-term (~5-day) studies considered optimal EA to be 45 kcal·kg^{-1}FFM·day^{-1}, while EA <30 kcal·kg^{-1}FFM·day^{-1} was considered low (LEA) and associated with health and performance impairments (Loucks, 2004; Loucks, 2013; Loucks et al., 2013). However, these concepts evolved from laboratory-based studies on a small sample of sedentary women (Ihle & Loucks, 2004; Loucks et al., 1998; Loucks & Thuma, 2003) and were not intended to be rigid or universally applied (Mountjoy et al., 2023). Indeed, the EA "cut off" associated with health and performance impairments is moderated by factors such as sex and training history (Koehler et al., 2016; Loucks, 2006; Mountjoy et al., 2023; Papageorgiou et al., 2017). Athletes commonly undertake both intentional and unintentional reductions in EA to facilitate performance goals. For example, intensified training blocks that increase EEE without a compensatory increase in EI, short periods of energy restriction to achieve optimal competition physique characteristics (Stellingwerff, 2018), alongside athletes in weight division sports who need to reduce body mass to meet competition weigh-in targets (Burke et al., 2021). Therefore, a more contemporary view is that while some short exposure to LEA may cause transient and minor metabolic adjustments and/or be associated with performance benefits ("adaptable" LEA), other LEA exposures are "problematic" because they are associated with negative health and performance outcomes that may result in REDs (Mountjoy et al., 2023). However, characteristics of adaptable versus problematic LEA and moderating factors are not yet fully identified.

Short-term (3-6 day) LEA exposures <30 kcal·kg^{-1}FFM·day^{-1} may alter bone metabolism, reproductive function, metabolic hormones (insulin, leptin), fat oxidation and RMR in some populations (Ihle & Loucks, 2004; Ishibashi et al., 2020; Koehler et al., 2016; Langan-Evans et al., 2020; Loucks & Thuma, 2003; Trexler et al., 2014), but performance effects are either uninvestigated or unclear. This is important for athletes needing to implement acute strategies, as aforementioned. Previous studies have reported maintenance of endurance capacity following three days at an EA of 19 kcal·kg^{-1}FFM·day^{-1} among Tier 2 (McKay et al., 2022b) male runners, despite reductions in muscle glycogen (Kojima et al., 2020). Meanwhile Burke et al. (Burke et al., 2023) demonstrated no impairments to 10,000 m race walk performance among Tier 4 athletes of both sexes following a slightly longer exposure (nine days at 15 kcal·kg^{-1}FFM·day^{-1}), when optimal pre-race fuelling was implemented. However, to our knowledge, no studies have examined the influence of acute (<7 days) LEA exposure on strength/power performance outcomes <60 seconds in duration.

Cognition and decision-making are key aspects of sport. However, few studies have examined the influence of acute LEA on cognitive performance, though preliminary evidence suggests that women experiencing LEA may show some resilience in cognitive function. Martin et al. (Martin et al., 2021)

reported no alterations to cognitive function among women following three days of exercise- (but not diet-) induced LEA (15 kcal·kg^{-1}FFM·day^{-1}). Moreover, Lieberman et al. (Lieberman et al., 2017) observed that the cognitive performance decline observed with two days of severe LEA (-3681±716 kcal·kg^{-1}FFM·day^{-1}) was less pronounced among women than men, although women accounted for only 26% of the sample. However, these studies were all conducted in non-athletic [≤Tier 1 (McKay et al., 2022b)] populations.

Sex may mediate the response to EA manipulations, with women potentially more sensitive to acute LEA, and experiencing negative consequences at a higher EA compared to men (Papageorgiou et al., 2017; Trexler et al., 2014). LEA intervention studies in male athletes (Koehler et al., 2016; McKay et al., 2022a; Papageorgiou et al., 2017) have reported fewer perturbations to body systems (bone metabolism, immune, inflammatory, and iron-regulatory responses and metabolic hormones) than shown in young untrained women (Ihle & Loucks, 2004; Loucks & Thuma, 2003; Papageorgiou et al., 2017). However, since few studies have examined performance indices, particularly among trained [≥Tier 2 (McKay et al., 2022b)] athletes, or with specific designs that can contrast responses between the sexes, robust conclusions regarding sex differences are not possible.

The complete time course over which various maladaptations to LEA manifest requires clarification. It is also unknown whether the method of reducing EA (i.e., dietary restriction or increased EEE) has divergent effects on physiological outcomes. From a health standpoint, it appears that an EA of 15 kcal·kg^{-1}FFM·day^{-1} induced through dietary restriction, but not increased EEE, decreases bone formation; but that both methods alter hormonal profiles (Koehler et al., 2016; Loucks et al., 1998; Papageorgiou et al., 2018b). There is also preliminary evidence that cognition may be more impaired by acute exercise-induced LEA (Martin et al., 2021), however physical performance has never been examined. This could be relevant when tailoring training/nutritional protocols to alter body composition whilst minimising negative performance effects. Lastly, low carbohydrate (CHO) availability, independent of LEA, is associated with perturbations to iron (McKay et al., 2022a) and bone metabolism (Fensham et al., 2022; Heikura et al., 2020), however acute performance effects remain uninvestigated.

We therefore aimed to assess effects of acute (24-hour) manipulations in EA (induced via diet or exercise) on substrate utilisation, postprandial metabolism, and physical/cognitive performance among trained males and females, evaluating results within individuals and between sexes. We hypothesized that increased fat oxidation would occur under LEA, with an augmented response among women, due to the concurrent reductions in CHO availability but the diet/exercise manipulation would be too brief alter the exercise intensity at which maximal fat oxidation occurs (FATMAX) (Achten et al., 2002). We also expected that acute EA manipulations would not alter strength/power, but that men would experience a greater decline in cognitive performance with LEA compared to women.

3.3 Methods

Participants

Ten female and ten male Tier 2-3 (McKay et al., 2022b) endurance trained athletes (Table 3.1) participated in this study approved by the Australian Catholic University Human Ethics Research Committee (2022-2561H) in accordance with the Declaration of Helsinki.

Table 3.1. Participant characteristics.

	Females (n=10)	Males (n=10)
Age (yrs)	33±7	38±9
Athletic tier	Tier 2 (n=8) Tier 3 (n=2)	Tier 2 (n=10)
Primary sport	Cycling (n=3) Mountain bike (n=1) Triathlon (n=4) Running (n=2)	Cycling (n=5) Mountain bike (n=3) Triathlon (n=1) Running (n=1)
Body mass (kg)	65.6±10.9	81.0±12.4
Fat free mass (kg)	47.7±6.1	62.4±9.7
Body fat percentage (%)	26.7±5.6	22.6±8.3
Body mass index	23.6±3.3	25.2±3.4
Absolute $\dot{V}O_2$max ($L \cdot min^{-1}$)	2.8±0.4	4.1±0.8
Relative $\dot{V}O_2$max ($ml \cdot kg \cdot min^{-1}$)	44.4±8.1	50.9±10.8
Absolute Wmax (W)	263±24	350±71
Relative Wmax ($W \cdot kg^{-1}$)	4.1±0.6	4.4±1.0

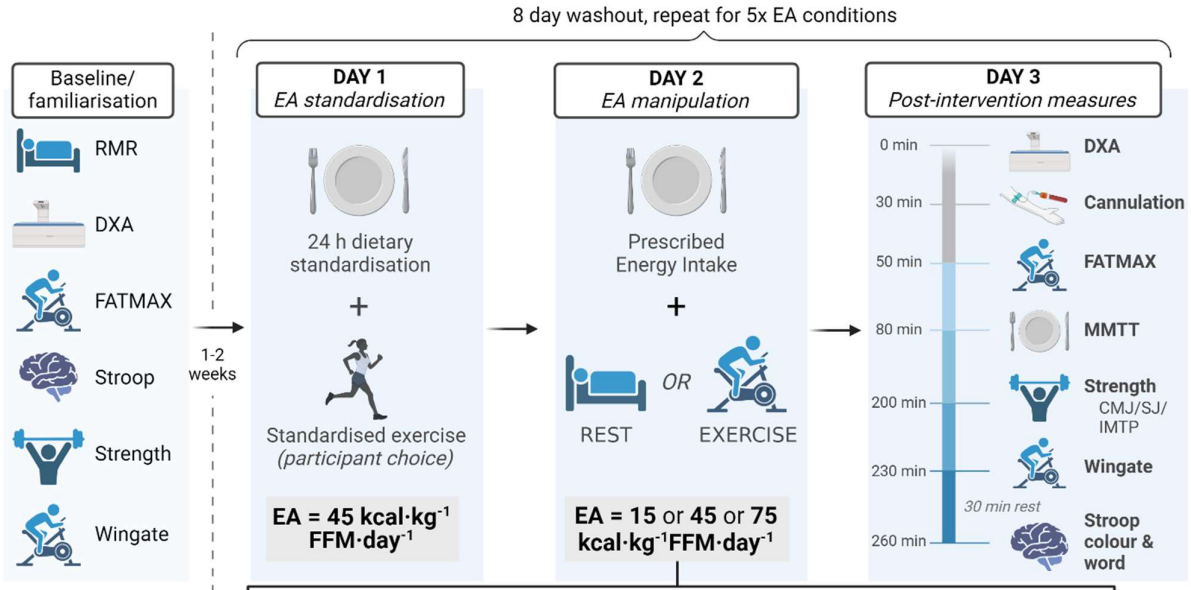
Values displayed as mean±standard deviation. Athletic tier as defined by McKay et al. (2022b) Wmax; maximal power output in Watts

Participants were familiar with cycling even if not their primary sport (i.e., cross-training, or regular commuting). A sample size calculation (G*Power 3.1, Dusseldorf, Germany), using data from Chrzanowski-Smith et al. (2021), estimated that 10 males and 10 females were required to detect differences in PFO relative to FFM between the sexes, with 90% statistical power and an alpha of 0.05. To eliminate potential effects of menstrual status/phase, we recruited pre-menopausal females taking a combined OCP (see details in Table S3.1) for >three months prior to study commencement (median usage time was 4 years). All participants provided informed consent prior to participating.

Experimental overview

The study design included a baseline/familiarisation session, followed by five randomised experimental trials, each comprising three consecutive days (15 days of testing per participant; Figure 3.1). Trials were completed in a randomised order in a Latin square design, with an average of eight days separating trials (minimum four days, maximum 33 days). Doses of EA on trial day two were: low (15 kcal·kg^{-1}FFM·day^{-1}; LEA), high (45 kcal·kg^{-1}FFM·day^{-1}; HEA) and higher EA for mass gain/growth (75 kcal·kg^{-1}FFM·day^{-1}; GEA), with conditions of high and low EA separately achieved via manipulations of EEE or EI (Figure 3.1). Participants undertook post-intervention testing on day three at the same time of day (±8 min) across all five trials. Female participants completed each trial during the active (pill-taking) phase of the OCP cycle, when they were not experiencing a withdrawal bleed, to minimise fluctuations in both endogenous and exogenous ovarian hormones across trials. Women were instructed to take their daily OCP after post-intervention testing on day three to minimise effects of a bolus dose of exogenous hormones.

Baseline/Familiarisation. Participants underwent a baseline/familiarisation session 1-2 weeks prior to study commencement, which included a $\dot{V}O_2$max (to calculate subsequent EEE prescription), alongside familiarisation to the FATMAX test and performance measures [Stroop, Wingate, CMJ, SJ and IMTP)]. Dual-energy X-ray absorptiometry (DXA) and RMR measurements occurred to establish EI and EEE prescriptions for subsequent trials. Additionally, participants received an activity tracker [Oura ring (Generation 3, Oura Health, Oulu, Finland)], to wear during each of the five subsequent three-day trial periods to monitor step count and estimated energy expenditure.



Condition	Energy Intake (kcal·kg⁻¹FFM·day⁻¹)	Energy expenditure (kcal·kg⁻¹FFM·day⁻¹)	Energy availability (kcal·kg⁻¹FFM·day⁻¹)	CHO intake (g·kg⁻¹)
LEA_{REST}	15	0	15	1.6±0.2
LEA_{EX}	45	30	15	4.7±0.5
HEA_{REST}	45	0	45	4.7±0.5
HEA_{EX}	75	30	45	7.8±0.8
GEA	75	0	75	7.8±0.8

Figure 3.1. Project overview. Timings on day three were consistent within participants (±8 min) and varied±1.5 h between participants according to participants' habitual wake time [assessed via the Morningness-Eveningness Questionnaire (Horne & Östberg, 1976)]. *RMR,* resting metabolic rate; *DXA*, dual-energy X-ray absorptiometry; *FATMAX*; exercise intensity at which the maximal rate of fat oxidation occurs; *EA,* energy availability; *EI*, energy intake; *FFM,* fat free mass; *LEA*, low energy availability; *HEA*, high energy availability; *GEA,* high energy availability for mass gain/growth; *MMTT,* mixed meal tolerance test; *IMTP*, isometric mid-thigh pull; *CMJ*, countermovement jump; *SJ*, squat jump; *CHO*, carbohydrate. Figure created with BioRender.com.

Trial day 1 – optimal EA standardisation. Participants consumed a standardised diet, providing 45 kcal·kg^{-1}FFM·day^{-1} (CHO; 4.7±0.5 g·kg^{-1}, protein; 2.1±0.2 g·kg^{-1}, fat; 0.8±0.1 g·kg^{-1}) for 24 hours prior to EA manipulation to ensure each intervention began in a state of optimal EA. Exercise was permitted but was replicated within each participant across all five trials (verified through the Oura ring), and dietary intake adjusted accordingly. As such, EI and EEE was identical for trial day one across all five trials.

Trial day 2 – EA manipulation. The five EA conditions are outlined in Figure 3.1: LEA [with and without exercise (LEA_{EX} and LEA_{REST})], HEA [with and without exercise (HEA_{EX} and HEA_{REST})], or GEA (without exercise) (Loucks, 2013). For the two conditions involving exercise (LEA_{EX} and HEA_{EX}),

participants completed two cycle sessions in the laboratory to achieve a total EEE of 30 kcal·kg^{-1}FFM·day^{-1}. Aside from prescribed exercise, participants remained inactive throughout the day, minimising activities of daily living (verified via Oura ring). For the three conditions not involving exercise (LEA_{REST}, HEA_{REST} and GEA) participants did not come to the laboratory but adhered to the provided diet and remained inactive (Oura ring verification).

Trial day 3 – post-intervention measures. Upon laboratory arrival in a 10-hour rested and fasted state, body composition was measured via DXA. A cannula was then inserted, and blood sample collected, followed by the FATMAX test (20-30 min). After a mixed meal tolerance test (120 min), physical performance measures were obtained: IMTP, CMJ, SJ and Wingate, alongside questionnaire regarding perceived muscle soreness (Impellizzeri & Maffiuletti, 2007). Participants then rested for 30 minutes in a quiet, private room with *ad libitum* food, after which they underwent the Stroop Colour and Word Test for cognitive performance. Each individual method is described below.

Dietary manipulation

Participants received all food and drink individually pre-packaged and weighed prior to the start of each three-day trial. Diet prescription is outlined in detail by Kuikman et al., (Kuikman et al., 2024b). In brief, diets prescribed an EI of 15 (LEA_{REST}), 45 ($LEA_{EX,}$ HEA_{REST} [and day one standardisation]) or 75 (HEA_{EX} and GEA) kcal·kg^{-1}FFM·day^{-1}. Macronutrient percentage distribution was equal between all EA conditions (alongside the day one standardisation diet) at 25% of EI from protein, 20% from fat and 55% from CHO (providing an intake of 1.6±0.2 g CHO·kg^{-1}·day^{-1} for LEA_{REST}, 4.7±0.5 g CHO·kg^{-1}·day^{-1} for $LEA_{EX,}$ HEA_{REST} and day one standardisation, and 7.5±1.2 g CHO·kg^{-1}·day^{-1}for HEA_{EX} and GEA). All EA conditions (alongside the day one standardisation diet) provided participants with three meals and three snacks. Participants were instructed to space out meals and snacks by at least one hour, and to consume the last snack 10-hours prior to laboratory arrival on day three. Caffeine consumption was permitted on trial days one and two, but not three, and replicated across each of the five trials. Alcohol was prohibited throughout each three-day trial period. Participants verbally confirmed the consumption of all food/drink upon arrival to the laboratory on trial day three.

Exercise manipulation

For the two EA conditions involving exercise (LEA_{EX} and HEA_{EX}), participants completed two cycle sessions in the laboratory on a stationary load bike (Load Excalibur Sport, Groningen, Netherlands) to achieve an EEE of 30 kcal·kg^{-1}FFM·day^{-1}. The evening session was 60 minutes at 65% $\dot{V}O_2$max (males, 195±46 W; females, 131±19 W), concluding 12 hours prior to next day laboratory arrival. The remaining EEE was completed in the morning at 55% $\dot{V}O_2$max with exercise duration manipulated to achieve 30 kcal·kg^{-1}FFM·day^{-1} EEE (males, 157±40 W for 135±26 min; females, 103±16 W for 163±37 min). The EEE at each cycling intensity was determined from gas exchange data collected during baseline $\dot{V}O_2$max testing. Expired gases were used to calculate substrate oxidation rates and energy

expenditure (EE) in accordance with the stoichiometric equations outlined by Jeukendrup and Wallis (2005), assuming negligible protein oxidation. An athlete's RMR was then excluded from EE to determine EEE.

Test Protocols

DXA and RMR: Both DXA and RMR were assessed at baseline, to calculate EI and EEE prescriptions. A DXA was performed at each laboratory visit to normalise results to FFM.

FATMAX and $\dot{V}O_2$max: The PFO and exercise intensity eliciting PFO (FATMAX) were assessed by an incremental cycling protocol on a load bike using the measured values approach (Achten et al., 2002). Starting at 30 W for females and 50 W for males, participants completed three-minute stages increasing by 25 W increments until RER >1.0. During familiarisation testing, participants completed an additional maximal exercise bout ($\dot{V}O_2$max test) following completion of the FATMAX protocol. When RER >1.0, instead of ceasing the test, wattage continued to increase in 25 W increments every 60 s until volitional exhaustion, as indicated by the participant. $\dot{V}O_2$max was taken as the highest $\dot{V}O_2$ value observed across a 30 s period. Chest HR (Forerunner, Garmin International) and rating of perceived exertion (RPE, 6–20, Borg Scale) were recorded at the end of each stage. Expired gas was collected and analysed using a custom-built indirect calorimetry system with associated in-house software as previously described (Saunders et al., 2004). The $\dot{V}O_2$ and $\dot{V}CO_2$ values from the last minute of each stage were used to calculate PFO using non-protein RER values (Peronnet & Massicotte, 1991).

Mixed meal tolerance test (MMTT): Participants consumed a breakfast meal of raisin toast, jam, and apple juice (males, 1035±148 kcal; females, 841±172 kcal; 2.00±0.00 g·kg^{-1} CHO; 0.27±0.03 g·kg^{-1} protein; 0.13±0.01 g·kg^{-1} fat) followed by a two-hour resting period. Meal consumption began at 0 min and finished within 15 min.

***Blood sampling*:** At the start of each lab visit, a cannula was inserted into the antecubital vein by a trained phlebotomist while the athlete was in a rested and fasted state. A total of eight 1 ml blood samples were collected per trial: baseline (rested and fasted), alongside the following timepoints during the MMTT: 0 (pre-meal), 15, 30, 45, 60, 90 and 120 min. Blood tubes clotted at room temperature for 30 min before being centrifuged at 2200 G for 10 min at 4°C. The serum was split into aliquots and stored at -80°C until batch analysis. Glucose was measured via an automated colorimetric assay (AU480 chemistry analyser, Beckman Coulter, Brea, California, USA) with intra-assay coefficient of variations (CV) of 1.0%. Insulin was analysed via chemiluminescent immunoassay (Access 2 immunoassay system, Beckman Coulter, Brea, California, USA) with CV of 8.9%. Incremental area under the curve was calculated for glucose and insulin concentration using an automated tool (Narang et al., 2020).

Countermovement/squat jumps and isometric mid-thigh pull*:* Following a standardised warm-up and wearing the same shoes on all five occasions, participants completed the CMJ, SJ, and IMTP on a dual

force plate system sampling at 1000 Hz (0.60 x 0.40 m; Model 10 kN 9286B, Kistler Instrument AG, Winterthur, Switzerland). Participants first completed three repetitions each of the CMJ and then SJ with ~60 seconds rest between jumps. Participants were instructed to "jump as high and powerfully as possible" with their hands remaining on hips. For the SJ, participants jumped from a 90° squat (or as close as possible) without countermovement. Squat depth was standardised within participants between trials using a plastic pole that participants lowered themselves to, and an additional effort was performed if any countermovement was observed via the force-time trace. The highest jump repetition was analysed; if jump height was equal, then peak power was used to determine the "best" effort. Outcome measures included jump height (calculated through impulse-momentum), mean and peak concentric force, velocity, and power, alongside impulse and rate of force development at 50/100/150/200 ms, as well as contraction time, concentric time, eccentric time, and centre of mass displacement (Weakley et al., 2022). Jump initiation was identified using the criterion method (Owen et al., 2014).

Following two sub maximal warm-up efforts, participants performed two maximal repetitions of the IMTP separated by two minutes rest. Participants pulled at maximal effort for three seconds on an immovable bar fixed to a customised power rack. The bar was set during the familiarisation visit, such that joint angles at the knee and hip were between 125-145° and 140-150°, respectively (Comfort et al., 2019). Participants were instructed to "push the ground away as hard and as fast as possible". Verbal encouragement was maintained throughout. A third effort was performed if: >200 N difference was observed between the peak force of the two efforts; there was variability >50 N in the quiet period; there was a countermovement prior to the lift, excessive pre-tension, or leaning on the bar (Comfort et al., 2019). The highest relative peak force effort was analysed. Pull initiation was identified as the moment when force exceeded five SD of a participant's body mass (Comfort et al., 2019), established through a one-second stable weighing period. Peak force, time to peak force, rate of force development and impulse at 50/100/150/200/250 ms were calculated.

All ground reaction force-time data for the CMJ, SJ and IMTP were recorded using ForceDecks software (VALD ForceDecks, 2.0.8587) and then exported for analysis via a customised R script. CMJ and SJ jump heights were also used to calculate the EUR and RSI, while the DSI was calculated from CMJ peak concentric force and IMTP peak force.

Wingate: Participants performed a five-minute standardised cycling warm-up, which included three six-second sub-maximal sprints. Participants then completed a 30 second all-out cycling effort (Wattbike Pro, Nottingham, England) at maximal speed against a high braking force from a rolling start. Participants were instructed to "pedal as hard as possible from the start without pacing the effort but remaining in the saddle". Verbal encouragement was maintained throughout. Outcome measures were peak power, mean power, and fatigue index.

***Stroop Colour and Word Test*:** Participants were shown coloured words on a laptop and asked to indicate the word's colour (and not it's meaning) by pressing a key as fast as possible whilst minimising errors (Stroop, 1935). Coloured labels were placed on keyboard keys to signify the corresponding colour. Three types of trials were presented: control (coloured rectangles), congruent (words of matched colour and meaning) and incongruent (words with mismatched colour and meaning). A red "X" flashed onto the screen when an incorrect response occurred. Each test had 180 trials, taking approximately three minutes to complete. The Stroop test was administered using Inquisit 6 (6.6.1 64bit, [Windows 10], (2020) Retrieved from https://www.millisecond.com). The Stroop effect was calculated as the difference between responses (proportion correct and reaction time) in the incongruent versus congruent trials.

Muscle soreness. A seven-point Likert scale for lower limb muscle soreness (Impellizzeri & Maffiuletti, 2007) (Table S3.2) was completed at 0, 60 and 120 min during the MMTT, with the mean score used in analysis.

Statistical analyses

Results were compared across EA conditions and between sexes using linear mixed models. Fixed effects were "condition" (LEA_{REST}, LEA_{EX}, HEA_{REST}, HEA_{EX}, GEA), and "sex" (female or male), with "subject identification" as a random effect. Statistical significance of fixed effects occurred using type II Wald tests with Kenward–Roger degrees of freedom. Where significant fixed effects were established, pairwise comparisons were performed with Tukey *post hoc* adjustments. Significance was accepted at $p<0.05$. Data are presented as mean±SD with non-normal data (assessed via histogram inspection) log-transformed prior to analysis. Outliers >three SD beyond the group mean were removed (Howell et al., 1998). Muscle soreness data are missing for one condition (LEA_{REST} and HEA_{REST}) for two females due to a failure in the server administering questionnaires. SJ results are missing for a single condition (LEA_{REST} and HEA_{EX}) for two females due to technical data collection issues. Two females are missing one timepoint in the LEA_{REST} condition (15 and 90 min) for glucose and insulin because of cannula blockage; linear interpolation was used to address this when calculating iAUC (Narang et al., 2020).

3.4 Results

Energy availability: As intended, the LEA_{REST} and LEA_{EX} conditions, alongside HEA_{REST} and HEA_{EX} conditions were matched for EA (all $p=1.000$), with differences observed between all other conditions (all $p<0.001$, Table 3.2). Similarly, LEA_{EX} and HEA_{REST}, alongside HEA_{EX} and GEA, were matched for dietary EI, CHO, protein, and fat (all $p>0.050$). EEE was also matched for LEA_{EX} and HEA_{EX} within sexes ($p=1.000$). Males had a higher EEE than females in both exercise conditions ($p<0.001$) and higher EI than females in HEA_{EX} and GEA ($p<0.001$). There was no difference in step counts between trials ($p=0.128$).

Table 3.2. Energy availability, energy intake, exercise energy expenditure, and macronutrient composition of all five experimental conditions, for both males and females.

	Males (n=10)					Females (n=10)				
	LEA_{REST}	LEA_{EX}	HEA_{REST}	HEA_{EX}	GEA	LEA_{REST}	LEA_{EX}	HEA_{REST}	HEA_{EX}	GEA
Energy availability ($kcal \cdot kg^{-1}FFM \cdot day^{-1}$)	15.0±0.2^	15.0±0.4^	44.8±0.8°	44.7±0.9°	74.0±1.1*	15.1±0.3^	15.1±0.4^	45.0±0.9°	44.9±0.6°	75.1±0.7*
Energy intake (kcal)	934±139*	2799±436†	2799±436†	4678±733$^{\&\#}$	4679±733$^{\&\#}$	712±93*	2145±269†	2145±269†	3578±466$^{\&\#}$	3576±465$^{\&\#}$
Exercise energy expenditure (kcal)	-	1867±282$^{\#}$	-	1867±282$^{\#}$	-	-	1430±184$^{\#}$	-	1430±184$^{\#}$	-
Carbohydrate intake ($g \cdot kg^{-1}$)	1.6±0.2*	4.8±0.5†	4.8±0.5†	7.4±0.9$^{\&}$	7.9±0.9$^{\&}$	1.6±0.3*	4.6±0.4†	4.6±0.4†	7.1±1.9$^{\&}$	7.5±0.6$^{\&}$
Protein intake ($g \cdot kg^{-1}$)	0.8±0.1*	2.2±0.2†	2.3±0.3†	3.6±0.4$^{\&}$	3.6±0.4$^{\&}$	0.7±0.1*	2.1±0.2†	2.1±0.2†	3.4±0.3$^{\&}$	3.4±0.3$^{\&}$
Fat intake ($g \cdot kg^{-1}$)	0.2±0.0*	0.8±0.1†	0.8±0.1†	1.3±0.1$^{\&}$	1.2±0.1$^{\&}$	0.2±0.0*	0.7±0.1†	0.7±0.1†	1.2±0.1$^{\&}$	1.3±0.1$^{\&}$

Values displayed as mean±standard deviation. LEA, low energy availability; *HEA*, high energy availability; *GEA,* high energy availability for mass gain/growth. **significantly different from all other conditions. ^significantly different from HEA_{REST}, HEA_{EX}, and GEA, °significantly different from LEA_{REST}, LEA_{EX}, and GEA, †significantly different from LEA_{REST}, HEA_{EX}, and GEA, $^{\&}$significantly different from LEA_{REST}, LEA_{EX} and HEA_{REST}. $^{\#}$significantly different from same condition in the opposite sex.*

Day two exercise during LEA_{EX} and HEA_{EX} conditions: Heart rate and RPE were higher during the second exercise bout at 65% $\dot{V}O_2$max (144±13 b·min^{-1} and 14±1) than the first bout at 55% $\dot{V}O_2$max (131±15 b·min^{-1} and 11±2, both *p<0.001*). There were no differences between sexes or condition, nor any interactions (all *p>0.050*). Self-reported/ perceived muscle soreness on trial day three was higher in the exercise conditions (LEA_{EX} and HEA_{EX}) than all other conditions (*p<0.010*, Table 3.3), with no difference between sexes (*p=0.668*).

Table 3.3. Self-reported/ perceived muscle soreness on trial day three averaged across all timepoints for each condition, where "0" represents a complete absence of soreness, and "6" denotes a severe muscle soreness, stiffness or weakness that limits the ability to move.

Condition	Soreness (0-6)	*p value*	
		vs LEA_{EX}	vs HEA_{EX}
LEA_{REST}	1.1 + 1.4	***<0.0001****	***<0.0001****
LEA_{EX}	1.9 + 1.5	-	*0.993*
HEA_{REST}	1.2 + 1.4	***0.0003****	***0.0001****
HEA_{EX}	2.0 + 1.6	*0.993*	-
GEA	0.8 + 1.1	***<0.0001****	***<0.0001****

Values displayed as mean±SD. LEA, low energy availability; *HEA*, high energy availability; *GEA,* high energy availability for mass gain/growth. ***denotes significance p<0.01,* ****denotes significance p<0.0001.*

Fat oxidation: There was a main effect of condition (*p<0.001*) but not sex or interaction *(*all *p>0.050)* for both absolute PFO (Figure 3.2A) and PFO relative to FFM (Figure 3.2B, FFM reported in Table S3.3). Absolute PFO was greatest in LEA_{EX} (0.60±0.17 g·min^{-1}) and lowest in GEA (0.37±0.13 g·min^{-1}). Differences between conditions were the same for absolute and relative PFO. Relative PFO peaked under LEA_{EX} (10.9±2.5 mg·(kg·FFM^{-1})·min^{-1}), 48% greater than GEA (6.7±2.0 mg·(kg·FFM^{-1})·min^{-1}, *p<0.001*), 34% greater than HEA_{REST} (7.7±2.1 mg·(kg·FFM^{-1})·min^{-1}, *p<0.001*) and 17% greater than LEA_{REST} (9.2±2.2 mg·(kg·FFM^{-1})·min^{-1}, *p=0.006*). Relative PFO was lower in GEA than HEA_{EX} (37%, 9.7±2.3 mg·(kg·FFM^{-1})·min^{-1}, *p<0.001*) and LEA_{REST} (31%, *p<0.001*), as well as in HEA_{REST} compared to HEA_{EX} (23%, *p<0.001*) and LEA_{REST} (18%, *p=0.022*). There was no difference in relative (or absolute) PFO between exercising conditions (LEA_{EX} and HEA_{EX}, *p=0.092*), nor between LEA_{REST} and HEA_{EX} (*p=0.846*) or GEA and HEA_{REST} (*p=0.218*). Mean total CHO oxidation across all conditions

during the exercise test was 29.5±14.0 g·min^{-1}, and when expressed relative to FFM, did not differ between conditions (*p=0.459*) or sexes (*p=0.065,* Table S3.4).

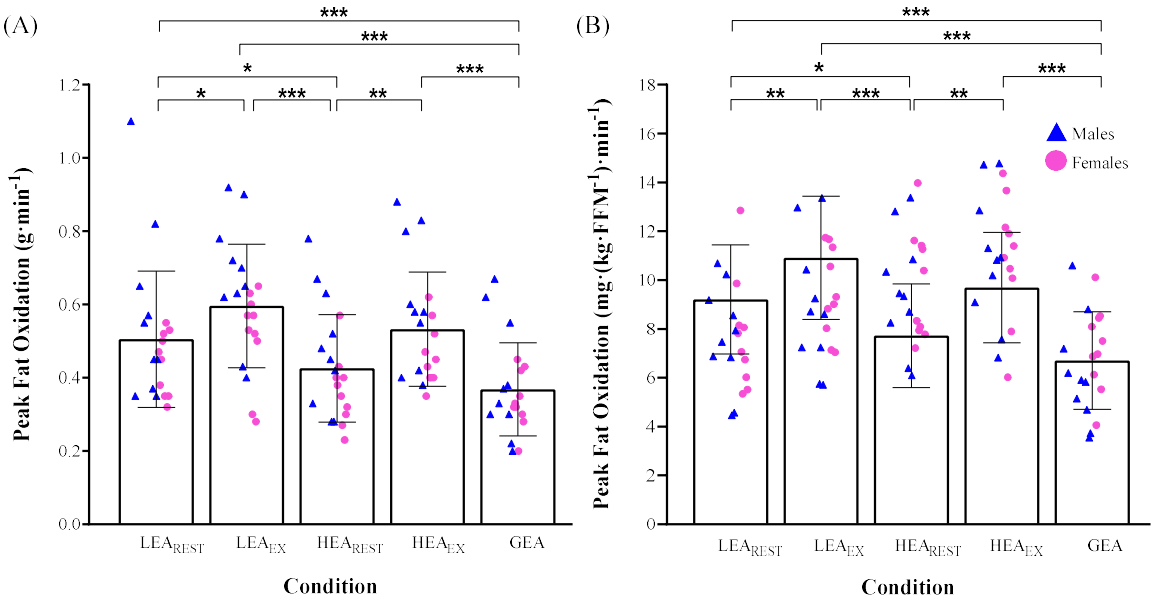


Figure 3.2. Peak fat oxidation as (A) absolute values and, (B) values expressed relative to FFM. Colours denote sex. *FFM,* fat free mass; *LEA*, low energy availability; *HEA*, high energy availability; *GEA,* high energy availability for mass gain/growth. *denotes significance *p<0.05,* **denotes significance *p<0.01,* ***denotes significance *p<0.0001.*

FATMAX, both absolute and relative to body mass, did not differ between EA conditions (*p>0.050,* Figures 3.3A and 3.3B). There was a main effect of sex for absolute FATMAX *(p=0.022,* Figure 3.3A*)*, with males reaching FATMAX at a higher power than females (150±75 W *vs* 105±75 W). However, when expressed relative to body mass, there was no effect of sex *(p=0.119,* Figure 3.3B*)*. The HR at FATMAX, as a percentage of maximal HR, was higher in LEA_{EX} (72±8%, Figure 3.3C) versus HEA_{REST} (64±10%, *p=0.016)* and GEA (64±9%, *p=0.033*). The RPE at FATMAX was lower during GEA (9±2, Figure 3.3D) than LEA_{EX} (11±2, *p=0.006*), LEA_{REST} (11±2, *p=0.005*) and HEA_{EX} (11±2, *p=0.010*). The $\dot{V}O_2$ at FATMAX, as a percentage of $\dot{V}O_2$max, was lower during HEA_{REST} (49±16%, Figure 3.3E) than LEA_{EX} (58±12%, *p=0.008*), LEA_{REST} (56±11%, *p=0.047*) and HEA_{EX} (56±10%, *p=0.039*), and was also lower during GEA (48±11%) than LEA_{EX} *(p=0.026*). Time to reach FATMAX was lower in HEA_{REST} (11.0±7.2 min, Figure 3.3F) versus LEA_{REST} (13.8±5.5, *p=0.027*) and LEA_{EX} (14.0±5.8, *p=0.024*). There was no effect of sex or sex*condition interaction on HR, RPE or time to FATMAX *(p>0.050).*

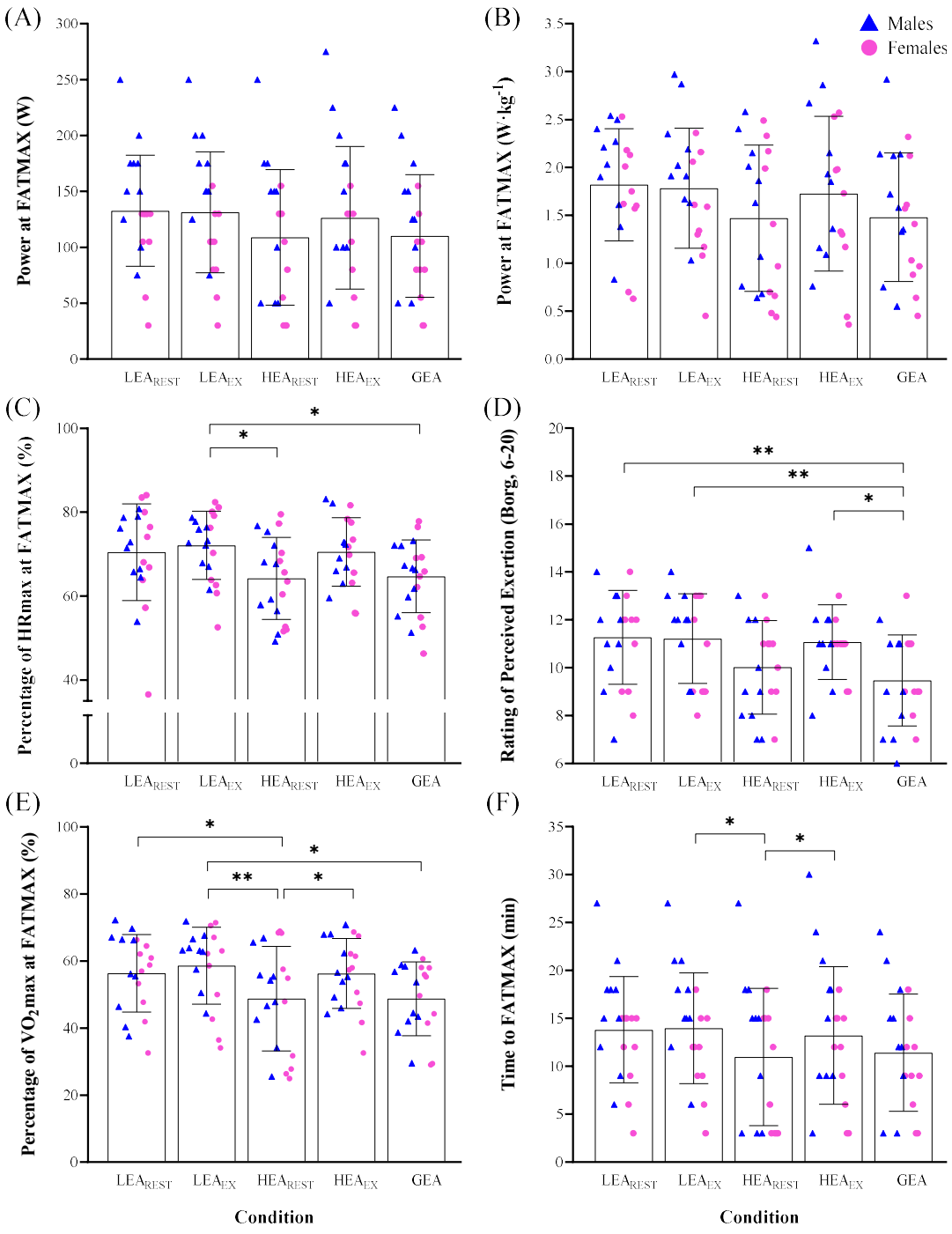


Figure 3.3. (A) Absolute power output at FATMAX, (B) power output at FATMAX relative to body mass, (C) heart rate at FATMAX as a percentage of HRmax, (D) RPE at FATMAX, (E) $\dot{V}O_2$ at FATMAX as a percentage or $\dot{V}O_2$max, and (F) time to reach FATMAX from the start of exercise. *FATMAX,* the exercise intensity at which maximal fat oxidation occurs; *HRmax,* maximal heart rate*; LEA*, low energy availability; *HEA*, high energy availability; *GEA,* high energy availability for mass gain/growth. Colours denote sex. *denotes significance $p<0.05$, **denotes significance $p<0.01$.

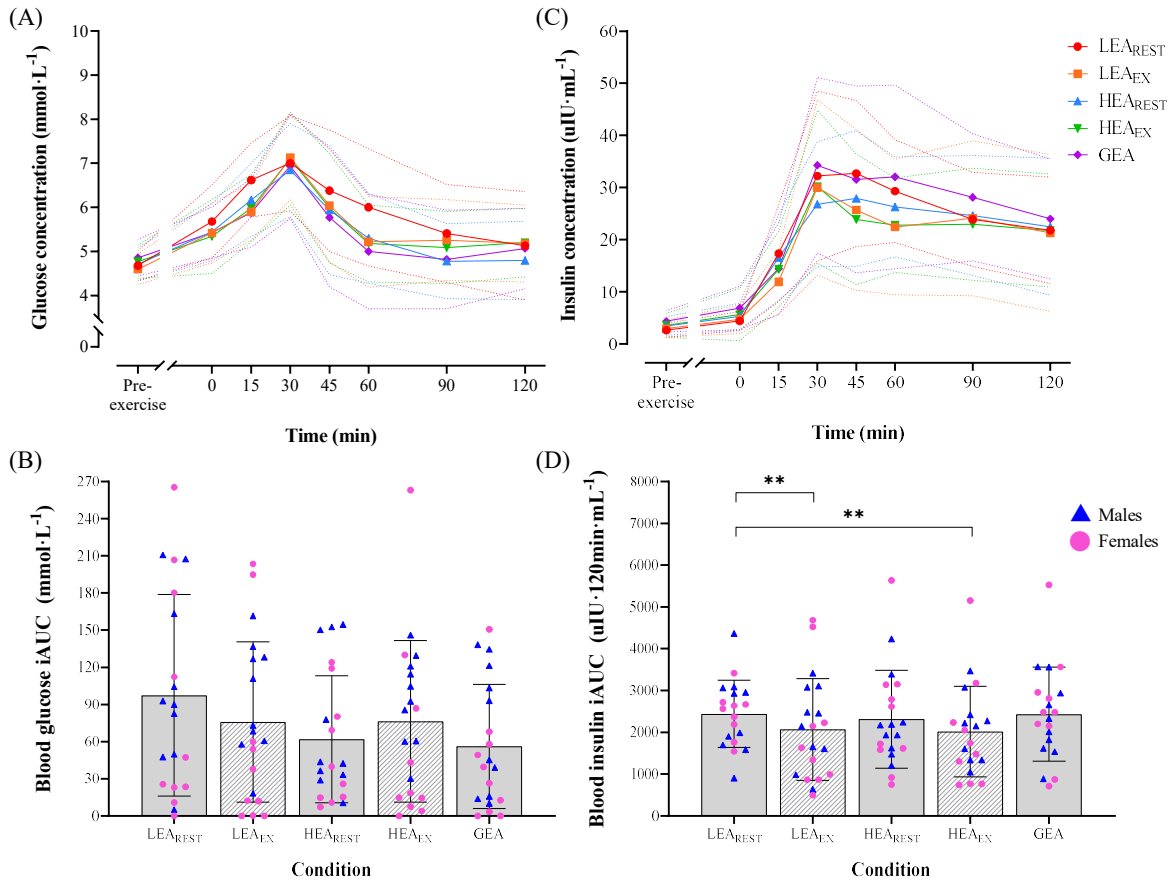


Figure 3.4. (A) Postprandial serum glucose concentration, (B) serum glucose incremental area under the curve, (C) postprandial serum insulin concentration, (D) serum insulin incremental area under the curve during the mixed meal tolerance test. Data is expressed as mean±standard deviation. Dashed lines in figures 3.4A and 3.4C represent the standard deviations. *iAUC*; incremental area under the curve, *LEA*, low energy availability; *HEA*, high energy availability; *GEA,* high energy availability for mass gain/growth. **denotes significance $p<0.01$.

MMTT: There was no clear difference in rested, fasted glucose concentration between conditions ($p=0.050$), but baseline fasted insulin concentration was higher in GEA (4.36±2.04 uIU·mL^{-1}) than LEA_{REST} (2.66±1.32 uIU·mL^{-1}, $p<0.001$), LEA_{EX} (2.96±1.66 uIU·mL^{-1}, $p<0.001$) and HEA_{EX} (3.24±1.72 uIU·mL^{-1}, $p=0.007$). Mean postprandial glucose concentration was higher in LEA_{REST} (6.01±1.27 mmol·L^{-1}, Figure 3.4A) than GEA (5.56±1.28 mmol·L^{-1}, $p<0.001$) and HEA_{REST} (5.60±1.20 mmol·L^{-1}, $p=0.007$). There was no alteration in glucose iAUC or maximum/minimum glucose concentration across conditions (all $p>0.050$, Figure 3.4B). Mean postprandial insulin concentration was lower in LEA_{EX} (20.0±15.1 uIU·mL^{-1}, Figure 3.4C) than GEA (24.4±16.3 uIU·mL^{-1}, $p=0.013$) and LEA_{REST} (23.0±14.2 uIU·mL^{-1}, $p=0.035$). Minimum insulin concentration was also lower in LEA_{EX} (4.16±2.43 uIU·mL^{-1}) than HEA_{REST} (5.30±2.46 uIU·mL^{-1}, $p=0.045$) and GEA (5.59±2.56 uIU·mL^{-1},

p=0.011). Insulin iAUC was higher in LEA_{REST} (2440±803 uIU·120min·mL^{-1}, Figure 3.4D) than LEA_{EX} (2068±1214 uIU·120min·mL^{-1}, *p=0.003*) and HEA_{EX} (2018±1086 uIU·120min·mL^{-1}, *p=0.007*). There were no alterations in peak insulin concentration between conditions (*p>0.050*) and no differences between sexes for glucose or insulin responses.

CMJ: There was a main effect of sex for jump height, take-off velocity, peak and mean velocity and relative power, CM displacement, RFD at 200ms, impulse at 50-200ms, and total impulse (all *p<0.050*, Table S3.5). Males jumped higher and produced greater velocity, power, CM displacement, RFD and impulse. There was also a main effect of sex for RSI (*p=0.045*), calculated from CMJ and SJ height, with males displaying higher values than females. The CM displacement was lower during GEA than HEA_{REST} (*p=0.007*) and HEA_{EX} (*p=0.025*), and also lower during LEA_{EX} (*p=0.043*) than HEA_{REST}. Mean velocity was lower during GEA than LEA_{REST} (*p=0.041*). There was no sex*condition interaction for any outcome measure.

SJ: There was a main effect of sex for jump height, velocity at take-off, velocity, power, RFD at 150 and 200ms, and impulse at 50-200ms and total impulse (all *p<0.050,* Table S3.6), with males jumping higher and producing greater velocity, power, RFD, and impulse. During GEA, jump height and peak velocity were both lower compared to LEA_{REST} (*p=0.045* and *p=0.043*) and HEA_{EX} (*p=0.016* and *p=0.023*), whilst take-off velocity was also lower compared to HEA_{EX} (*p=0.040*). There was no sex*condition interaction for any variable.

IMTP: There was a main effect of sex for impulse between 50-250ms (all *p<0.010*) and RFD between 150-250ms (all *p<0.050*), with males producing a greater RFD and impulse (Table S3.7). There was no main effect of condition, nor sex*condition interaction for any outcome.

Wingate: There was a main effect of sex for relative peak power (*p=0.001*) and relative mean power *p=0.007*), with males producing a peak power 26% greater than females (12.0±2.1 W·kg^{-1} *vs* 8.9±1.6 W·kg^{-1}, Figure 3.5A) and a mean power 21% greater (7.8±1.5 W·kg^{-1} *vs* 6.2±0.9 W·kg^{-1}, Figure 3.5B). There was no main effect of condition, nor condition*sex interaction for either relative peak or mean power (all *p>0.050*). Fatigue index was also not different between sexes or across conditions (*p>0.050*).

Stroop Colour and Word: There was no effect of sex, condition, nor their interaction on the Stroop effect (either proportion of correct responses or reaction time, all *p>0.050*, Table S3.8)

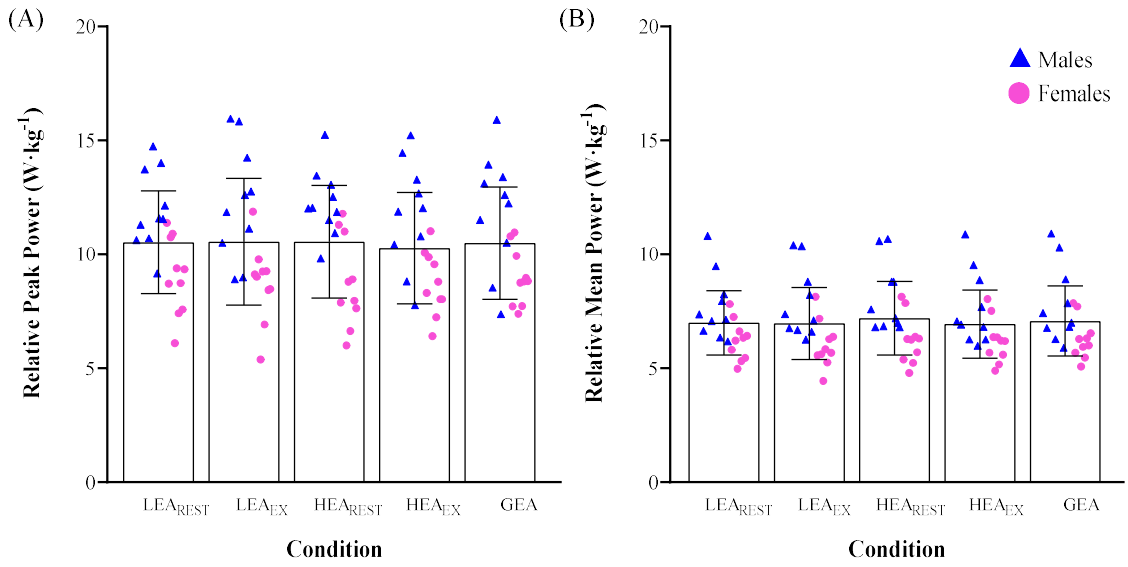


Figure 3.5. (A) relative peak power, and (B) relative mean power during a 30 s Wingate performance task. Data is expressed as mean±standard deviation, displayed across conditions. Colours denote sex. *LEA*, low energy availability; *HEA*, high energy availability; *GEA,* high energy availability for mass gain/growth.

3.5 Discussion

This study examined the effects of an acute 24-hour EA manipulation, induced by either diet or exercise, on substrate oxidation, postprandial metabolism and physical/cognitive performance among endurance trained males and females. Our primary findings demonstrate that the highest rate of PFO occurred under conditions of LEA induced by exercise (LEA_{EX}), with no differences in FATMAX between EA conditions. However, HR and RPE at FATMAX were both lowest under conditions of high EA for mass gain (GEA). Postprandial mean glucose concentration was higher in LEA_{REST} than other resting conditions (GEA and HEA_{REST}), and the insulin response was augmented in LEA_{REST} compared to the exercise conditions (LEA_{EX}/HEA_{EX}). There was no effect of EA manipulation on performance during the CMJ/IMTP, Wingate or Stroop Test, although SJ height was impaired with GEA. Lastly, the response to EA manipulations did not differ between sexes. Our findings therefore suggest that 24 hours of LEA is not a sufficient exposure to impair strength/power, sprint capacity, or cognitive performance, at least when measured post-prandially, but that 24 hours of exercise-induced LEA appears to influence substrate oxidation more than LEA induced by diet alone.

The elevated PFO and increased reliance on fat oxidation following LEA supports the predicted outcome of reduced EI and resultant decline in CHO availability, noting a three-fold decrease in CHO intake with LEA_{REST} (compared to LEA_{EX} and HEA_{REST}) and increased CHO utilisation with the increased EEE during LEA_{EX}. Conversely, the increased CHO availability in the HEA_{REST} and GEA

conditions decreased PFO and increased CHO oxidation. Interestingly, both conditions involving exercise (LEA_{EX} and HEA_{EX}) elicited a higher PFO compared to their respective EA-matched conditions without exercise (LEA_{REST} and HEA_{REST}), while PFO did not differ between LEA_{EX} and HEA_{EX}. This increase in PFO following prior day exercise is likely underpinned by an acute decline in muscle and liver glycogen rather than a chronic adaptation from muscle retooling to increase fat oxidation via changes in fat mobilisation and transport as observed following a ketogenic diet, since this requires >three-six days of exposure (Burke, 2021). Indeed, depleted muscle glycogen has been demonstrated following three days of 19-20 kcal·kg^{-1}FFM·day^{-1} among Tier 2 male runners (Ishibashi et al., 2020; Kojima et al., 2020).

Our data suggest that 24 hours of LEA derived from energy/CHO restriction depletes muscle glycogen by preventing restoration from prior exercise, but this effect is amplified when further exercise contributes to the manipulation of EA. Thus, despite controlling for the overall EA reduction, acute achievement of the energy mismatch via exercise has a greater effect on CHO availability than energy restriction. In contrast, previous work by Loucks et al. (Loucks et al., 1998) in young sedentary females demonstrated a decline in CHO oxidation during exercise following five days of LEA compared to optimal EA (10-15 vs 45-50 kcal·kg^{-1}FFM·day^{-1}, both 55% CHO), with both conditions involving exercise (30 kcal·kg^{-1}LBM·day^{-1}). Moreover, Loucks et al. (Loucks et al., 1998) observed greater perturbations to luteinising hormone pulsatility with diet- versus exercise-induced LEA. Here, exercise achieved an augmented decline in relative CHO availability which was attributed to a within-exercise glycogen-sparing achieved with a longer LEA exposure of four days. The influence of CHO availability is important; other work has demonstrated that low CHO availability, independent of LEA, is associated with perturbations to iron (McKay et al., 2022a) and bone metabolism (Fensham et al., 2022; Heikura et al., 2020). Therefore, more research is needed to differentiate the effects of exercise increase and dietary restriction on CHO availability, independently of EA outcomes. At present, the disparity between our results regarding CHO availability and that of Loucks et al. (Loucks et al., 1998; Loucks & Heath, 1994) might be explained by a higher training status of participants and shorter duration of EA manipulation in the current study.

The metabolic response to breakfast was associated with a 0.41-0.45 mmol·L^{-1} (7-8%) increase in mean postprandial glucose concentration in LEA_{REST} comparative to the other resting conditions (GEA and HEA_{REST}). Insulin iAUC and/or mean postprandial insulin concentrations were also 14-19% higher in LEA_{REST} compared to LEA_{EX} and HEA_{EX}. This increase in both postprandial glucose and insulin concentration under LEA_{REST} may suggest an over-compensatory metabolic response to the first high CHO meal following 24 hours of substantial underfeeding. Nevertheless, the mean postprandial glucose concentration in LEA_{REST} was 6.01±1.27 mmol·L^{-1} and glucose concentration returned to baseline by 90 minutes, indicating excellent glucose control (DiabetesAustralia, 2023). The elevated insulin concentration in LEA_{REST} versus exercise conditions (LEA_{EX} and HEA_{EX}) may reflect the influence of

exercise in increasing insulin sensitivity (Bird & Hawley, 2016). As insulin-stimulated peripheral tissue glucose uptake is considered the primary driver of post-prandial glucose tolerance (Meyer et al., 2002), an increased insulin sensitivity following prior day exercise in LEA_{EX} and HEA_{EX} may result in a lower insulin concentration required to regulate postprandial glucose control compared to LEA_{REST}.

Total CHO oxidation during the FATMAX test immediately before the MMTT was identical across conditions (Table S3.4) and is therefore unlikely to explain the altered postprandial response. Due to increased hepatic glucose output during exercise, both glucose and insulin concentrations increased following the FATMAX test across all conditions. However, the magnitude of this increase was 4-10% larger in LEA_{REST} versus other conditions. It is possible that an elevated glucose/insulin concentration immediately prior to the MMTT during LEA_{REST} somewhat accounted for the higher mean postprandial glucose/insulin concentrations. There were no differences in blood glucose concentration in the rested and fasted state between conditions, whereas baseline insulin concentration was 0.9-1.7 $uIU \cdot mL^{-1}$ lower in LEA_{REST}, LEA_{EX} and HEA_{EX} versus GEA. Lower fasted blood glucose/insulin has been reported following five days of LEA <20 $kcal \cdot kg^{-1}FFM \cdot day^{-1}$ (Koehler et al., 2016; Loucks & Thuma, 2003). Our results suggest that just 24 hours of LEA, in addition to independent effects of exercise, may lead to lower next day fasted insulin.

There were no alterations to peak or mean power output during the Wingate across conditions, and no effect of EA on CMJ height or IMTP peak force. During the SJ, participants jumped an average of 1.2-1.3 cm (5%) lower under conditions of GEA versus LEA_{REST} and HEA_{EX}. When considering kinetic/kinematic outcomes, there was little alteration across EA conditions. The exceptions were a 2.9-3.3 cm (~10%) lower CM displacement in the CMJ under GEA versus HEA_{REST} and HEA_{EX}, alongside a 2.7 cm (~9%) decline in CM displacement in LEA_{EX} versus HEA_{REST}. Mean velocity was also 0.04 $m \cdot s^{-1}$ (3%) lower during the CMJ in GEA versus LEA_{REST}. The SJ peak velocity was 0.06 $m \cdot s^{-1}$ (3%) lower in GEA versus both LEA_{REST} and HEA_{EX}, while take-off velocity was also 0.06 $m \cdot s^{-1}$ (3%) lower in GEA compared to HEA_{EX}. The decreased jump height, CM displacement, and velocity with GEA may be the result of an elevated body mass in GEA (Table S3.3). A lack of change in overall performance (jump height, power output, peak force production) following LEA suggests a 24-hour exposure is too brief to impair strength/power performance. However, as these measures were not a primary outcome variable, we may lack statistical power. Additionally, while lower limb muscle soreness on day three was higher in exercise versus non-exercise conditions, the absolute soreness ratings were low for both (“two” out of six; moderate soreness/slight persistent ache vs “one”; light soreness/ vague ache, respectively). As physical performance was not altered between conditions of rest and exercise, it appears that prior day exercise-induced muscle soreness was not substantial to influence overall performance.

To our knowledge, this is the first study to examine acute (24-hour) manipulation in EA on strength/power or sprint capacity. Acute (five day) periods of rapid weight loss can impair dynamic force expression (e.g., punching force) (Smith et al., 2001). However, such studies are typically free living, examining EI/EEE through dietary recall/training logs without directly manipulating EA, challenging comparisons to our study, which implemented a high degree of control. A longer (14 day) period of exercise-induced EAs ranging from 9-22 $kcal \cdot kg^{-1}FFM \cdot day^{-1}$ among Tier 2-3 males all reduced CMJ height by ~3 cm (Jurov et al., 2022a; Jurov et al., 2022b). However, EA was altered in the field by monitoring participant self-selected EI and then altering EEE accordingly, rather than the more precise method of prescribing a standardised EI/EEE as in the present study.

Manipulations to EA did not influence the Stroop effect (proportion of correct responses or reaction time). Interestingly, prior work has observed declines in other aspects of cognitive function when exercise (but not diet) is used to induce LEA (Lieberman et al., 2017; Lieberman et al., 2008; Martin et al., 2021). Moreover, we observed no differences to cognitive performance between the sexes, in contrast to Lieberman et al. (Lieberman et al., 2017), who observed a decline in cognitive function among women, but not men, experiencing LEA. However, women accounted for only 26% (n=6) of the sample in this study, and EI was severely restricted [266 $kcal \cdot day^{-1}$ combined with a 4-hour exercise bout (Lieberman et al., 2017)], which is unlikely to represent practices in elite athletes.

Sex is thought to moderate the response to EA manipulations, with women potentially more sensitive to acute LEA, and experiencing negative consequences at a higher EA compared to men (Papageorgiou et al., 2017). Our findings do not support this theory; we observed no differences in the response to EA manipulations between the sexes in performance, substrate oxidation, or postprandial metabolism. Given innate differences in substrate oxidation, with women demonstrating a lesser reliance on whole-body CHO oxidation to support fuel requirements for endurance activities (Devries, 2016), and reaching FATMAX at greater exercise intensities than males (Chrzanowski-Smith et al., 2021), it may be that reduced muscle glycogen influences substrate oxidation in males more than in females. However, the response to substrate oxidation with altered EA did not differ between the sexes, whilst there was also no difference in the exercise intensity eliciting FATMAX. This may be explained in part by a lower $\dot{V}O_2max$ in our female participants compared to males (although not statistically different *p=0.145*), perhaps suggesting a reduced capability of our female athletes to oxidise fats as a fuel source relative to more endurance-trained athletes (Melanson et al., 2009), however this is speculative.

Sex-based differences in substrate oxidation/performance are hypothesised to be mediated by endogenous oestrogen concentrations (Devries, 2016). Therefore, a lack of differential response may be a result of studying females using OCP, eliminating the cyclical fluctuations in endogenous oestrogen observed in naturally menstruating female athletes. Indeed, there are some reports of altered CHO/fat oxidation during 45-90 min submaximal cycling, alongside power output across repeated sprints,

between the follicular and luteal phases (Cook et al., 2018; Devries, 2016; Devries et al., 2006; Lee et al., 2024), with no such effects observed between active and withdrawal pill-taking phases among OCP users (Lee et al., 2024). However, the directionalities of such alterations are conflicting. Moreover, because adequate classification and control of menstrual status was implemented in only one of these studies (Lee et al., 2024), conclusions regarding an effect of menstrual status on glycogen utilisation and muscular power are difficult. A lack of sex-based differences may also be due to a shorter, more severe EA restriction in the present study, with conditions of LEA providing 15 kcal·kg^{-1}FFM·day^{-1}. It is possible that different responses would be observed across longer time periods or at a higher LEA threshold (~20-25 kcal·kg^{-1}FFM·day^{-1}). However, this remains speculative.

Our findings should be considered in light of potential limitations. First, because we recruited female athletes utilising OCPs to facilitate the standardisation of ovarian hormones, our results directly apply only to a subset of women. Second, during the conditions involving exercise (LEA_{EX} and HEA_{EX}), participants were not provided explicit instructions about the timing of food intake around exercise. Given the effect of nutrient timing on post-exercise muscle glycogen resynthesis (Alghannam et al., 2018; Burke et al., 2017), there may have been small differences in post-exercise muscle glycogen repletion that may have consequently altered next-day substrate oxidation. Finally, only whole-body substrate oxidation was measured and future research examining the rate of appearance and disappearance of glucose and lipids would provide greater mechanistic detail.

3.6 Conclusions and future research

A brief (24-hour) period of LEA appeared not to impair strength/power, sprint capacity, or cognitive performance, at least when measured post-prandially. However, in trained individuals, a 24-hour LEA exposure induced by exercise appears to influence substrate oxidation to a greater extent that LEA induced by diet alone. Whether this translates to more prolonged EA manipulation is of interest, as is the potential for independent alterations to CHO availability to affect various body systems. Lastly, future research may consider examining LEA between ~20-25 kcal·kg^{-1}FFM·day^{-1} to elucidate potential sex differences that may occur.

Interlinking chapter

Study 1 (*chapter 3*) examined sex-differences in measures of explosive power/strength performance in a highly controlled laboratory-based study. However, due to exclusively examining women using HC, it was not possible to investigate any influence of oestrogen or progesterone on these outcome measures. Study 2 (*chapter 4)* therefore examined fluctuations in explosive power/strength performance across the MC, and between athletes using HC and those with "natural" MCs. Unlike the rolling-recruitment model used in Study 1 where testing was able to be scheduled in advance, the observational research-embedded training camp design of study 2 facilitated access to participants and the capacity for testing at short notice. We were therefore able to test athletes in MC phases 1, 2 and 4, via the use of daily questionnaires and urinary ovulation measures to track the progress of each MC and quick-response scheduling of laboratory testing.

Chapter 4: Minimal influence of the menstrual cycle or hormonal contraceptives on performance in female rugby league athletes

Publication statement:

This chapter comprises the following paper published in the *European Journal of Sport Science.*

Smith, E. S., Weakley, J., McKay, A. K., McCormick, R., Tee, N., Kuikman, M. A., Harris, R., Minahan, C.L., Buxton, S., Skinner, J., Ackerman, K. E., Elliott-Sale, K. J., Stellingwerff, T., & Burke, L. M. (2024). Minimal influence of the menstrual cycle or hormonal contraceptives on performance in female rugby league athletes. *European Journal of Sport Science, 24*(8). 1067-1078. https://doi.org/10.1002/ejsc.12151.

Data availability: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Ethical approval: Ethical approval was granted by the Australian Catholic University Human Ethics Research Committee (2021-285H) in accordance with the Declaration of Helsinki.

Conflicts of interest: Authors declare no conflict of interest.

Funding: This project was co-funded by the Australian Catholic University, Australian Institute of Sport, and Wu Tsai Human Performance Alliance.

4.1 Abstract

We examined performance across one MC and three weeks of HC use to identify whether known fluctuations in oestrogen and progesterone/progestin are associated with functional performance changes. National Rugby League Indigenous Women's Academy athletes [n=11 naturally menstruating (NM), n=13 using HC] completed performance tests [CMJ, SJ, IMTP, 20 m sprint, power pass and Stroop test] during three phases of a MC or three weeks of HC usage, confirmed through ovulation tests alongside serum oestrogen and progesterone concentrations. MC phase or HC use did not influence jump height, peak force, sprint time, distance thrown or Stroop effect. However, there were small variations in kinetic and kinematic CMJ/SJ outputs. NM athletes produced greater mean concentric power in MC phase 4 than 1 [+0.41 $W \cdot kg^{-1}$ (+16.8%), *p=0.021*] during the CMJ, alongside greater impulse at 50ms at phase 1 than 4 [+1.7 N·s (+4.7%), *p=0.031*] during the SJ, without differences between tests for HC users. Among NM athletes, oestradiol negatively correlated with mean velocity and power (r=–0.44 to –0.50, *p<0.047*), progesterone positively correlated with contraction time (r=0.45, p=0.045), and both negatively correlated with rate of force development and impulse (r=–0.45 to –0.64, *p<0.043*) during the SJ. During the CMJ, oestradiol positively correlated to 200ms impulse (r=0.45, *p=0*.049) and progesterone to mean power (r=0.51, *p=0.021*). Evidence of changes in testing performance across a MC, or during active HC use, is insufficient to justify "phase-based testing", however kinetic or kinematic outputs may be altered in naturally menstruating athletes.

Keywords: sex hormones, oestrogen, progesterone, female, women, strength

Key Points

1. Evidence of changes in testing performance across a MC, or during active HC use, is insufficient to justify "phase-based testing" at a group or team-based level among female rugby league athletes.
2. Kinetic or kinematic outputs in jumping movements may be altered in naturally menstruating athletes, however it could not be determined if the observed alterations exceeded between-day variability.

4.2 Introduction

Cyclical fluctuations in oestrogen and progesterone across the MC potentially influence multiple biological systems associated with athletic performance. Indeed, both sex hormones may influence force development through alterations to muscle contractile properties. Oestrogen has been shown to elicit neuroexcitatory effects resulting in increased voluntary activation and reduced inhibition, while progesterone has been shown to exhibit neuroinhibitory effects (Smith et al., 2002). Accordingly, if oestrogen and progesterone augment and attenuate force production (Pallavi et al., 2017; Smith et al., 2002) then physical performance may be enhanced when oestrogen is elevated and impaired when oestrogen is suppressed, with the reverse for progesterone. There is some (albeit predominately low-quality) evidence for improved force and power outcomes during phases 2 and 4 of the MC (when oestrogen concentration is high and moderate, respectively), alongside a trivial performance reduction during phase 1 (when oestrogen is low) (McNulty et al., 2020b). For women using typical HC, exogenous oestrogen and progestin are supplemented on 21 continuous days and endogenous oestrogen and progesterone are therefore suppressed, comparable to the low endogenous hormonal profiles observed during phase 1 of the MC. Thus, in HC users, there may be marginal performance impairments compared to naturally menstruating women because of such endogenous oestrogen suppression regardless of the daily exogenous oestrogen supplementation (Elliott-Sale et al., 2020b).

Cognition is a key aspect of performance in numerous sports, particularly team events that require continuous rapid and accurate decision making. There is a hypothetical role for oestrogen and progesterone in cognitive performance, based on their entry through the blood-brain barrier and the presence of receptors in multiple brain regions (Brinton et al., 2008; Hara et al., 2015). Indeed, enhanced cognitive performance during MC phase 1 (low oestrogen and progesterone concentrations) has been reported in comparison to other phases involving elevated hormones (Barel et al., 2019; Šimić & Santini, 2012), which may have relevance to team sports. However, this finding is not consistent with other studies reporting no alterations across the MC (Hampson, 1990; Kozaki & Yasukouchi, 2009).

Our understanding of any influence of oestrogen or progesterone on physical and/or cognitive performance, through MC phases or with HC use, are inconclusive. This uncertainty partially stems from the broad failure of studies to achieve sufficient methodological classification and control of hormonal profiles (Elliott-Sale et al., 2020b; McNulty et al., 2020b). Accurate and purposeful classification of MC phase and HC use is necessary to support causality regarding any influence of oestrogen and progesterone on performance. Accordingly, the aim of this study was to examine performance across the MC and between athletes using HC and those with “natural” cycles, employing gold standard protocols regarding the classification and control of participant menstrual status.

4.3 Materials & methods

A comprehensive methodological overview including participant recruitment, study design, and MC tracking is detailed elsewhere (McKay et al., 2024). Only information specific to this study is detailed below.

Twenty-four female Tier 3 (national level) (McKay et al., 2022b) Australian National Rugby League's Indigenous Women's Academy athletes attended a five-week residential training camp at the Australian Institute of Sport. This sample size is reflective of most real-world rugby squads for which a coach or sports scientist may be asked to consider menstrual phase or status-based testing at a group level. The group was initially divided into those reporting the use of HC (athletesHC) and those who were considered by their self-reports as being naturally menstruating (athletesNM) until menstrual status was studied during the project. The actual menstrual status of athletes and their baseline characteristics are summarised in Table 4.1. This study implemented an observational design within a training camp environment. Following two familiarisation sessions, a battery of performance tests was completed on three separate occasions across each participant's individualised menstrual or HC cycle (Figure 4.1). Participants undertook these tests at the same time of day (±15 min) across a 90 min period, wearing the same shoes, after completing a standardised warm-up, and adhering to a standardised diet from lunch onwards the day prior to testing (~18h). The warm-up consisted of five minutes cycling on a stationary bike at a perceived "easy" intensity including 3x4 second sprints at 90% of maximal perceived cadence, followed by 10 each of walking lunges, squats, leg swings and calf raises, and concluding with three CMJ each at 70% and 90% of perceived maximal effort. For athletesNM, the three phases occurred in a randomised order, determined by the menstrual phase in which they commenced the training camp.

Menstrual status

Menstrual status was tracked in both athletesNM and athletesHC according to best-practice protocols (Elliott-Sale et al., 2021); recording onset of bleeding, performing 16 weeks of MC or HC tracking, using dual hormone urinary ovulation kits, and assessing retrospective serum 17-β-oestradiol (the most potent form of oestrogen among pre-menopausal women, henceforth referred to as "oestradiol") and progesterone concentration. Performance testing was completed at MC phases 1 (day 1.8±0.4), 2 (day 11.4±1.4), and 4 (day 20.8±1.6) for athletesNM, and three equally spaced time points for athletesHC (Figure 4.1). AthletesHC using OCPs were tested during pill taking days only and were instructed to take their pill at the same time of day on each testing occasion. As such, AthletesHC were all tested during active HC usage (Test 1, Test 2, and Test 3). Six athletesHC using the contraceptive implant had this inserted between one and three years prior to testing and two athletesHC had this inserted the same month as testing commenced. The athleteHC using the hormonal injection had her last injection three weeks prior to the first test.

Table 4.1. Participant baseline characteristics.

	"Naturally menstruating" (non-hormonal contraceptive using) athletes (n=11)	Athletes using hormonal contraception (n=13)
Age (yrs)	21±3	22±4
Actual Menstrual characteristics	Eumenorrheic (n=1) Naturally menstruating (n=4) Polycystic ovary syndrome (n=1) Oligomenorrheic (n=3) Anovulatory (n=1) Luteal phase deficiency (n=1)	Contraceptive implant (n=8) *[Implanon]* Hormonal injection (n=1) *[Depo Provera]* Combined oral contraceptive pill (n=4) *[Evelyn 150/30 ED:30 µg ethinyloestradiol, 150 µg levonorgestrel, Femme-Tab 20/100 ED: 20 µg ethinyloestradiol, 100 µg levonorgestrel, Lenest 30 ED: 30 µg ethinyloestradiol, 150 µg levonorgestrel, Yasmin: 30 µg ethinyloestradiol, 3mg drospirenone]*
Age at menarche (yrs)	13±2	13±2
Body mass (kg)	71.7±8.4	80.1±13.6
Body mass index ($kg \cdot m^2$)	27.1±3.4	28.8±4.7

Data presented as mean±standard deviation. Comprehensive menstrual characteristics are detailed in McKay et al. (2024) Menstrual status was defined according to Elliott-Sale et al. (2021) – eumenorrhea: "menstrual cycle length ≥ 21 days and ≤ 35 days resulting in 9 or more consecutive periods per year, plus evidence of LH surge, plus correct hormonal profile, plus no HC use 3 months prior to recruitment", naturally menstruating "experience menstruation, with menstrua cycle lengths ≥ 21 days and ≤ 35 days, but without confirmed ovulation [ovulation was not confirmed by urinary LH surge or verified by hormone concentrations via blood sample analysis]", oligomenorrhea: "cycle length >35 days", anovulatory: "those who experience menstruation but do not ovulate (ovulation cannot be detected by urinary LH surge or confirmed by hormone concentrations via blood sample analysis)", luteal phase deficiency: "cycles with less than 16 $nmol \cdot L^{-1}$ of progesterone, when a single luteal phase progesterone measurement is taken"

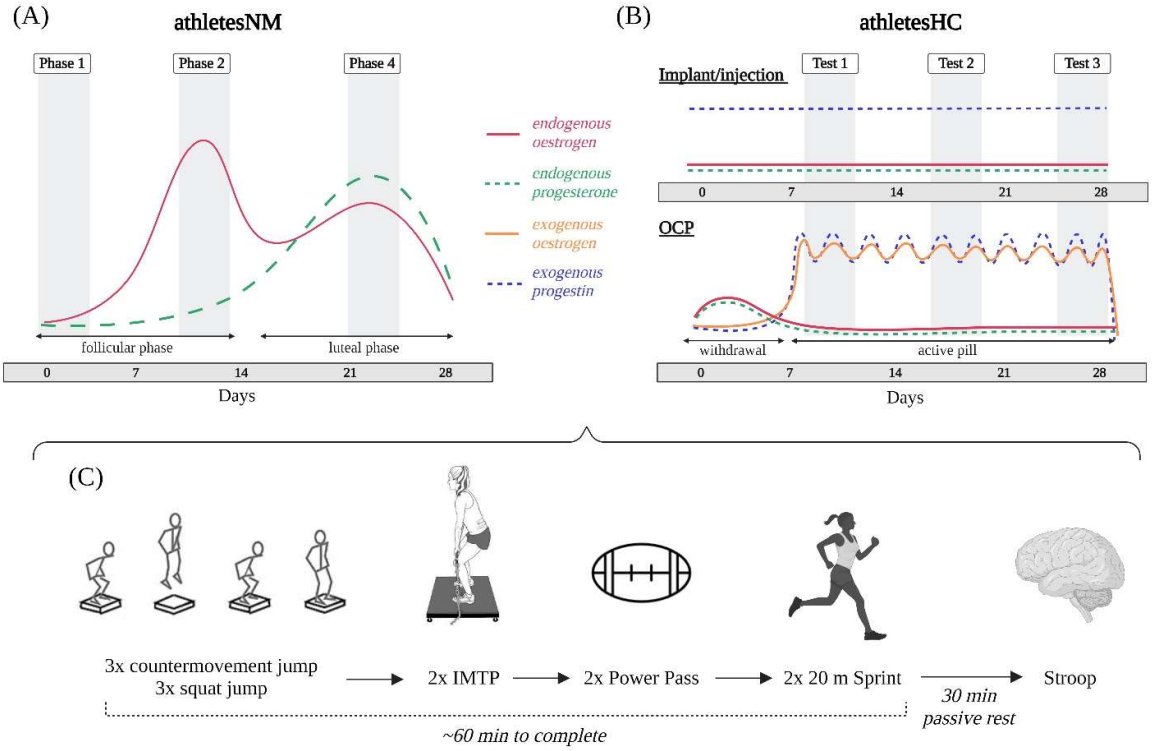


Figure 4.1. Study overview. Performance testing occurred at either (A) phases 1 [low oestrogen/progesterone concentration (day 1.8±0.4)], 2 [high oestrogen and low progesterone (day 11.4±1.4)] and 4 [moderate oestrogen and progesterone (day 20.8±1.6)] for athletesNM. The days reported refer to the cycle day on which the test was conducted; (B) three equally spaced timepoints for athletesHC utilising the implant or hormonal injection or, three equally spaced timepoints avoiding the withdrawal bleed for athletesHC using the oral contraceptive pill. It should be noted that the concentration of exogenous progestin following the implant and injection gradually declines with time (Huber, 1998), and hence the exact hormonal profile is dependent on the date of the implant or injection. Time points are displayed according to an idealised 28-day cycle. (C) Performance testing schedule. *IMTP*; isometric mid-thigh pull, *NM*; naturally menstruating, *HC*; hormonal contraceptives, *OCP*; oral contraceptive pills. *Created with BioRender.com*.

Blood sampling

Prior to performance testing at each visit, a trained phlebotomist collected an 8.5 mL venous blood sample from an antecubital vein into a serum separator tube, while the athlete was in a rested and fasted state. Blood tubes clotted at room temperature for 30 min and were then centrifuged at 2200 G for 10 min at 4°C. The remaining serum was split into aliquots and stored at -80°C until batch analysis. Oestradiol and progesterone were measured via an Access 2 Immunoassay System (Beckman Coulter, Brea, CA, USA) with intra-assay coefficient of variations (CV) 5% and 11% for oestradiol and

progesterone, respectively. Total testosterone was analysed using liquid chromatography-tandem mass spectrometry (Waters UPLC-TQX S, Waters Corp., Wilmslow, UK), with a total imprecision CV of 5.8%, and free testosterone was subsequently calculated from total testosterone alongside sex hormone binding globulin and albumin (Vermeulen et al., 1999).

Performance testing protocols

The CMJ, SJ and IMTP were conducted on a dual force plate system sampling at 1000 Hz (0.60 x 0.40 m; Model 10 kN 9286B, Kistler Instrument AG, Winterthur, Switzerland). Participants were familiarised at two separate sessions to the CMJ, SJ and IMTP protocols, alongside the Stroop Colour and Word Test, during the first two days of the training camp. Specific familiarisation was not undertaken for the power pass or 20 m sprint as these are regularly performed as part of the National Rugby League testing battery. These tests were selected as they represent different domains of performance (James et al., 2023), were familiar to participants, are commonly used throughout the literature with rugby athletes (Owen et al., 2020), and demonstrate acceptable between-day reliability and ecological validity (Weakley et al., 2022).

Countermovement jump and squat jump: Participants completed three repetitions each of the CMJ and SJ with ~60 s rest between jumps (Weakley et al., 2022). Participants were instructed to jump as high and powerfully as possible with their hands remaining on hips (both CMJ and SJ). For the SJ, participants jumped from a 90° squat (or as close as possible) without any countermovement. An additional effort was performed if any countermovement was observed. Squat depth was standardised within participants between trials using a plastic pole that participants squatted to reach until touching. The highest jump at each test was taken for analysis; if jump height was equal, then peak power was used to determine the "best" effort.

Outcome measures included jump height (calculated through impulse-momentum), mean and peak concentric force, velocity, and power, alongside impulse and rate of force development at 50/100/150/200 ms, as well as contraction time, concentric time, eccentric time, and centre of mass displacement. Jump initiation was identified using the criterion method of taking the instant when vertical force was less or greater than a threshold equal to five times the SD of body mass measured during a one second stable weighing period (Owen et al., 2014). Jump heights in the CMJ and SJ were also used to calculate the EUR and RSI, while the DSI was calculated from CMJ peak concentric force and IMTP peak force.

Isometric mid-thigh pull: Following two-sub maximal warm-up efforts, participants performed two maximal repetitions of the IMTP separated by 2 min rest. Participants pulled as hard as possible for 3 s on an immovable bar fixed to a customised power rack. Participants were instructed to "push the ground away as hard and as fast as possible". Verbal encouragement was maintained throughout. A third effort

was performed if: >200 N difference was observed between the peak force of the two efforts, there was variability >50 N in the quiet period, there was a countermovement prior to the lift, excessive pre-tension, or leaning on the bar. The effort with the highest relative peak force was taken for analysis. Initiation of the pull was identified as the moment when force exceeded five SDs of a participant's body mass, established through a one second stable weighing period. Peak force, time to peak force, rate of force development and impulse at 50/100/150/200/250 ms were calculated.

All ground reaction force-time data for the CMJ, SJ and IMTP were recorded using ForceDecks software (VALD ForceDecks, 2.0.8587), and then exported for analysis via a customised R script. The kinetic and kinematic outcome variables were selected as they represented different domains of force expression and also provided information that could provide context in relation to changes in temporal performance and movement strategy. Furthermore, ratio data (e.g., DSI) were provided to give context on whether force expression changed relative to difference strength domains (e.g., isometric vs dynamic strength).

Power pass: Athletes stood with their feet shoulder width apart and pushed a 3 kg med ball from the chest as far as possible into a long-jump pit. Countermovement in the legs was permitted, but feet weren't permitted to leave the ground. The throw distance was measured from the back of the imprint left by the ball in the sand to the nearest cm. The furthest throw at each test was used in analysis.

20 m sprint: The 20 m sprint was conducted on an indoor athletics track with four light gates (Fusion SmartSpeed V2) positioned at 0/5/10/20 m, measuring at a height of 57 cm (0 m gate) and 87cm (5/10/20 m gates). From a split-stance position, 10 cm behind the first light gate as marked-up on the track (Weakley et al., 2023), participants sprinted at maximal effort for 20 m. The start was initiated when participants broke the plane of the first light gate. An additional light gate, alongside tape to signify a "finishing line", was placed at ~23 m. Participants were instructed to run through this line to prevent deceleration prior to 20 m. Each participant completed a warmup sprint, followed by two maximal efforts, with the fastest taken for analysis.

Stroop Colour and Word Test: Coloured words were displayed on a laptop and participants were asked to indicate the colour of the word (not it's meaning) by pressing a corresponding key as fast as possible while minimising errors (Stroop, 1935). Coloured labels were placed on keyboard keys to signify the corresponding colour. Three types of trials were presented: control (coloured rectangles), congruent (words of matched colour and meaning) and incongruent (words with mismatched colour and meaning). A red "X" flashed onto the screen in the event of an incorrect response. There were 180 trials for each test, taking approximately three minutes to complete. The Stroop test was administered using Inquisit 6 [6.6.1 64bit, (Windows 10), (2020) retrieved from https://www.millisecond.com] in a quiet, private room. The Stroop effect was calculated as the difference between responses (both the proportion correct/accuracy and reaction time) in the incongruent versus congruent trials.

Statistical analyses

Statistical analyses were performed using R Studio (v3.5.2) with statistical significance accepted at an α level of $p \leq 0.05$. Two separate approaches were taken for statistical analyses; participant numbers reported for each outcome measure are displayed in Figure S4.1. Initially, outcome measures were compared both within individuals (i.e., across menstrual or HC cycle phases) and between individuals (i.e., between athletesNM and athletesHC) – termed "phase-based analysis". Linear mixed models were used to analyse each variable, using "menstrual status" and "cycle phase/test day" as fixed effects, alongside "subject identification" and "test order" as random effects. Statistical significance of fixed effects was identified using type II Wald tests with Kenward–Roger degrees of freedom. Where significant fixed effects were established, pairwise comparisons were identified using Tukey *post hoc* adjustments. Non-normally distributed data were identified through histogram inspection [Stroop outcomes, rate of force development (RFD) and impulse during the IMTP, RFD, FT:CT contraction time and concentric time during the CMJ, impulse during the SJ, alongside EUR] and were log transformed prior to statistical analyses. An independent t-test was conducted to compare total training load between groups.

Following analysis of serum oestradiol and progesterone concentrations, it was determined that a 'true' phase 2 was only achieved in one out of 11 athletesNM (McKay et al., 2024) (Figure 4.2A), and results were therefore compared across phases 1 and 4 only. Three athletesNM were also excluded due to hormonal profiles not meeting the criteria for phase 4 (progesterone >16 $nmol \cdot l^{-1}$, Figure 4.2C). As such, phase-based analyses were performed in n=8 athletesNM. Therefore, a repeated measures correlation was also used to assess associations between performance measures and oestradiol or progesterone concentration, alongside oestradiol: progesterone ratio (E:P) and oestradiol: serum free testosterone ratio (E:T), – termed "correlation analysis". Correlations were conducted among athletesNM exclusively, given that a) only endogenous hormones were measured and b) there was potential for variable results outside of hormonal influences due to the largely unknown effects of the exogenous hormonal milieu in athletesHC. This analysis approach did not require discrete MC phases, and thus "phase 2" results were included, alongside results from athletes with only two out of three completed tests, resulting in n=11 athletesNM. A single progesterone value from the athlete with PCOS was excluded from correlational analysis because it was >2.5 SD above the mean.

4.4 Results

Hormonal profiles: Among athletesNM, oestradiol concentration increased 3-fold from to "phase 2" ($p=0.064$) and 4.3-fold from phase 1 to 4 ($p=0.001$, Figure 4.2A). As a result of the oestradiol changes between phases 1 and 2, phase 2 was only truly captured in one of 11 athletesNM (McKay et al., 2024). Thus, only phases 1 and 4 were analysed and reported for phase-based analysis. Progesterone concentration was constant between phases 1 and "2" ($p=0.999$), and then increased 8-fold during phase 4 ($p<0.001$, Figure 4.2C). For athletesHC, endogenous oestradiol and progesterone concentrations remained constant across tests (all $p>0.05$, Figures 4.2B, 4.2D). Both free and total testosterone concentrations were stable across all tests for both groups and did not differ between athletesNM and athletesHC (all $p>0.05$, Figures 4.2E-H). The athlete with the highest oestradiol concentration (Figure 4.2A) was not the same as that with the highest total testosterone concentration (Figure 4.2G).

Performance tests

There was no change in CMJ or SJ height, IMTP peak force, distance thrown in the power pass, fastest sprint time and the Stroop effect between MC phases 1 and 4, or between tests for athletesHC (all $p>0.05$, Figure 4.3), nor any correlation between these outcome measures and oestradiol or progesterone concentration among athletesNM. There were also no differences between groups (athletesNM *vs* athletesHC) for any performance outcome measure (all $p>0.05$). While overall physical performance outcomes were unchanged, there were some small variations in kinetic and kinematic outputs detected during the CMJ and SJ, detailed below. All outcome measures are displayed in the supplementary material (Table S4.1).

Countermovement jump and squat jump – kinetic and kinematic outcome measures: All outcome measures are displayed in the supplementary material (Table S4.1). Phase-based analysis revealed that relative mean concentric power was 16.8% greater in MC phase 4 than 1 ($p=0.021$) among athletesNM during the CMJ (Figure 4.3), while this remained unchanged between tests in athletesHC ($p=1.000$). Additionally, athletesNM produced a 4.7% greater impulse at 50 ms in phase 1 than 4 ($p=0.031$) during the SJ (Figure 4.3), with no change between tests among athletesHC ($p=0.999$). There were no differences between MC or HC phase for any other outcome measure (all $p>0.05$), nor any difference in calculated metrics (EUR, RSI, or DSI).

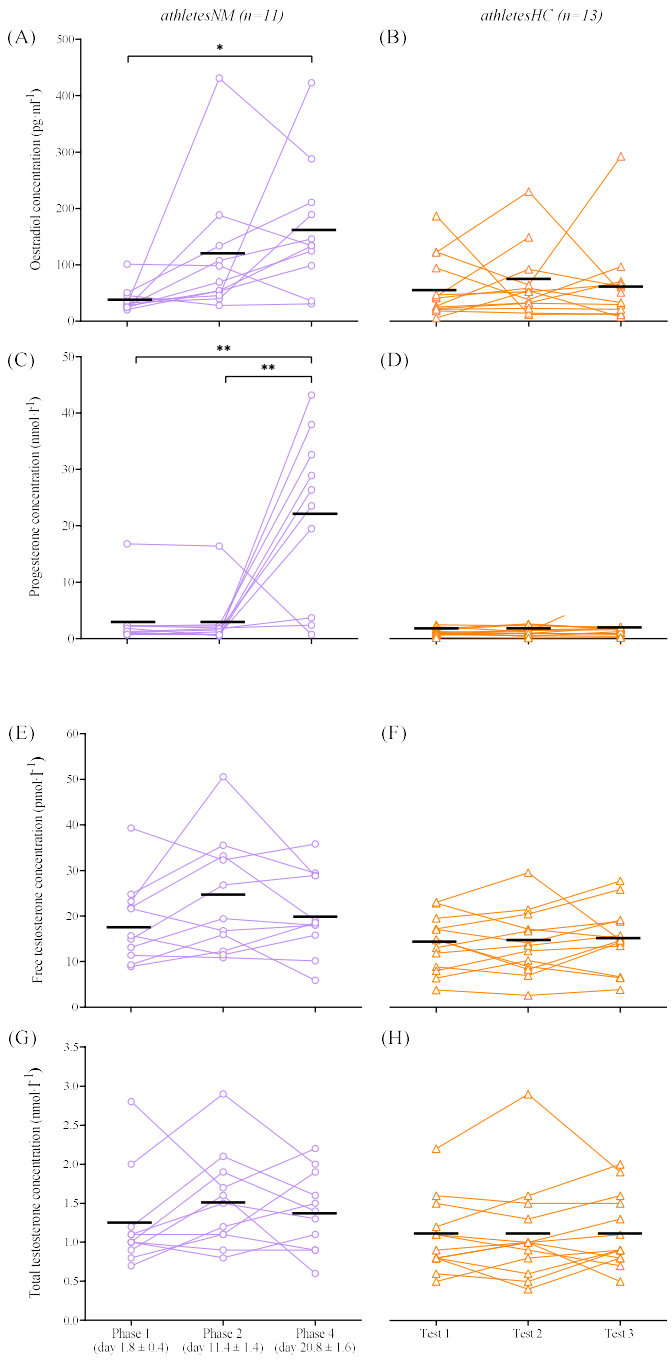


Figure 4.2. Serum oestradiol concentration across the three tests in (A) naturally menstruating athletes and (B) athletes using hormonal contraception (n=1 outlier removed in test three). Serum progesterone concentration across the tests in (C) naturally menstruating athletes (n=1 outlier removed during phase 4) and (D) athletes using hormonal contraception. Calculated free testosterone across the tests in (E) naturally menstruating athletes and (F) athletes using hormonal contraception. Total testosterone across the tests in (G) naturally menstruating athletes and (H) athletes using hormonal contraception. Black lines denote mean values. *denotes significance $p<0.05$, **denotes significance $p<0.001$.

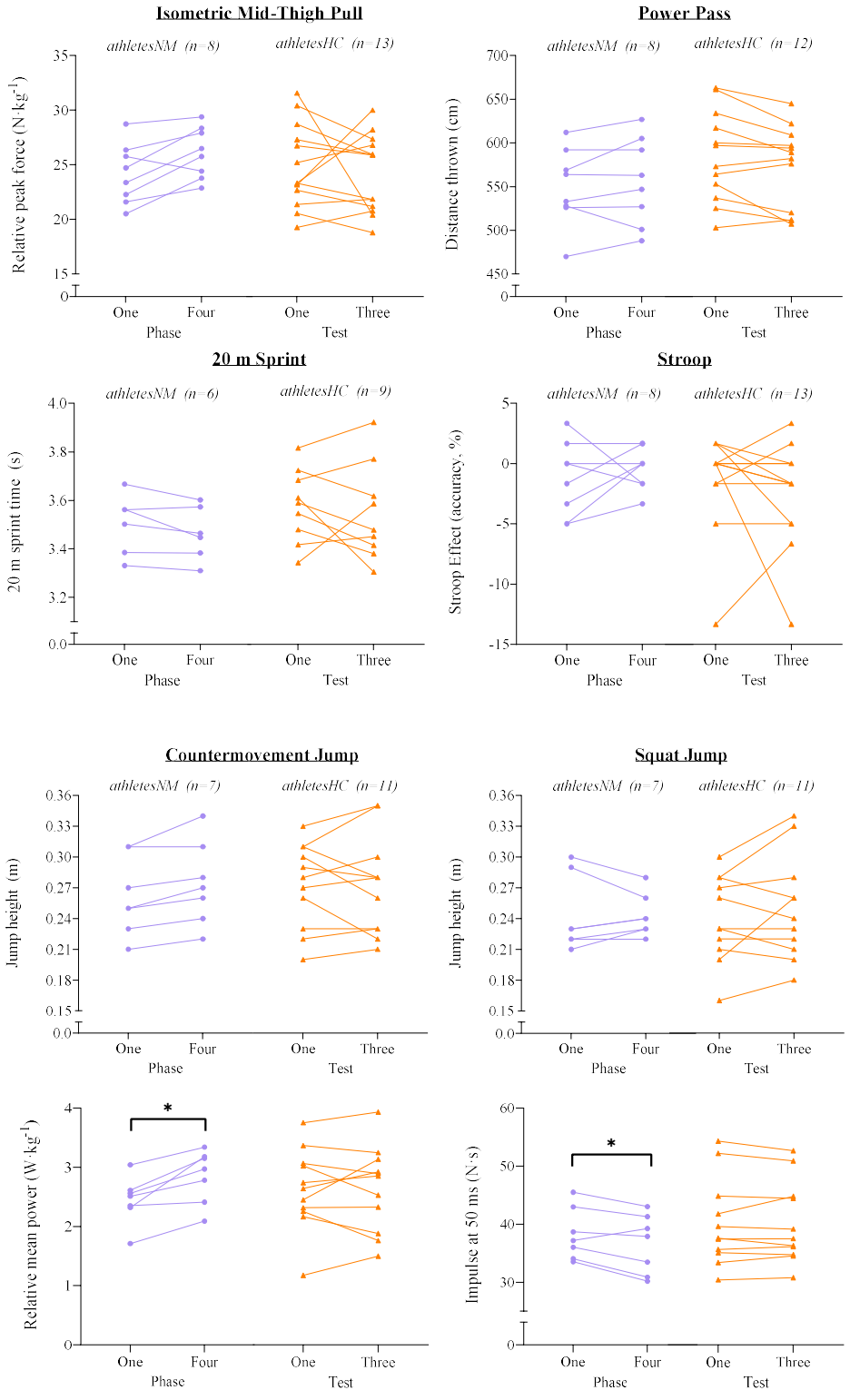


Figure 4.3. Performance outcomes at each test, alongside relative mean power during the countermovement jump and impulse at 50 ms during the squat jump, between naturally menstruating athletes (athletesNM) and athletes using hormonal contraception (athletesHC). *significance $p<0.05$.

During the SJ, there were negative correlations between oestradiol and RFD at 50 ms (Figure S4.2A) and 100 ms (r=–0.45, *p=0.043*), as well as between progesterone and RFD at 50 ms (Figure S4.2B) and 100 ms (r=–0.50, *p=0.026*), but not 150/200 ms (*p>0.05*). There were also negative correlations between oestradiol and impulse at 50 ms (Figure S4.2C), 100 ms (r=–0.50, *p=0.022*) and 150 ms (r=–0.49, *p=0.024*), but not 200 ms (*p=0.065*) alongside progesterone and impulse at 50 ms (Figure S4.2D), 100 ms (r=–0.56, *p=0.011*), 150 ms (r=–0.53, *p=0.016*), but not 200 ms (*p=0.067*). In addition, there was a negative correlation between oestradiol and both mean velocity and relative mean power (Figures 4.S2E and 4.S2G), and a positive correlation between progesterone and contraction time (Figure S4.2F). There were negative correlations between E:T and RFD at 50 ms (r= –0.49, *p=0.023*), and impulse at 50 ms (r= –0.61, *p=0.003*), 100 ms (r= –0.50, *p=0.021*), and 150 ms (r= –0.46, *p=0.036*). During the CMJ, we observed positive correlations between oestradiol and impulse at 200 ms (Figure S4.3A), and between progesterone and relative mean power (Figure S4.3B).

Training load: There were no statistical differences between groups in total training load across the five weeks (Table S4.2).

4.5 Discussion

This study was one of the first to assess a range of performance measures in well-trained athletes across a MC or during HC use. Our findings indicate that overall physical and cognitive performance outcomes were not statistically different between MC phases 1 and 4 in the athleteNM group (n =8), nor across ~3 weeks within athletesHC (n=13). Furthermore, there were no detectable performance differences between the athleteNM and athleteHC groups. There was also no relationship between overall performance outcomes and oestradiol or progesterone concentration among athletesNM. However, despite overall physical performance outcomes being unchanged, some small variations in kinetic and kinematic outputs were detected among athletesNM across the MC in the CMJ and SJ.

While there was no change in jump height, among athetesNM, we observed a 0.41 $W \cdot kg^{-1}$ (16.8%) greater mean concentric power during the CMJ, alongside a 1.7 N·s (4.7%) reduction in impulse at 50 ms during the SJ in phase 4 compared to phase 1 (Figure 4.3), while power and impulse were unchanged between tests for athletesHC. These differences are larger than the in intra-phase CV among athletesNM (9.4% and 2.8% for mean concentric power and impulse at 50 ms, respectively, Table S4.3), suggesting a true difference in these outcomes between phases. However, these differences were less than the inter-test CV observed among athletesHC (28.1% and 17.2%, Table S4.3), and therefore may be attributed to between-day variability. There were also no differences in performance outcomes between athletesNM and athletesHC. Conversely, a recent meta-analysis reported trivial strength impairments among women utilising OCP comparative to naturally menstruating women (Elliott-Sale et al., 2020b). It is possible that differences in the type and mode of hormone delivery of HC used by athletes in the present study [69% using progesterone-only local HC methods (i.e., implant and injection)] compared to the OCPs

examined by Elliott-Sale et al. (2020b) may account for some of this disparity; the effects of different exogenous hormones and absorption routes are largely unknown. If any difference between athletesNM and athletesHC is trivial in magnitude, it may be that the sample size in the present study was too small to detect such differences, or that the differences were too subtle to distinguish, despite testing in an athletic population very familiar with the performance tasks. Indeed, the intra-test CV for overall performance outcomes among athletesHC ranged from 3.1-20.7% (Table S4.3), and it may be that any small performance differences may have been outweighed by day-to-day variability. Taken together, our results currently suggest a lack of justification in the context of altering athlete testing at a group level based solely on menstrual phase or HC use.

In athletesNM, the correlations between oestradiol and progesterone concentration and kinetic and kinematic outcomes during the CMJ and SJ are conflicting. Some observations support the hypothesised roles of oestradiol and progesterone in augmenting and attenuating neuromuscular function, respectively (Pallavi et al., 2017; Smith et al., 2002). Indeed, during the CMJ, an increase in oestradiol concentration was associated with increased impulse at 200 ms (Figure S4.3A), while during the SJ, elevated progesterone was associated with a decline in RFD and impulse and an increase in contraction time (Figures S4.2B, S4.2D and S4.2F). However, in contrast to their hypothesised role, increases in oestradiol were simultaneously correlated with a decline in RFD, impulse, mean velocity, and relative mean concentric power during the SJ (Figures S4.2A, S4.2C, S4.2E, S4.2F), and a progesterone increase was associated with an elevated relative mean power during the CMJ (Figure S4.3B). The influence of oestradiol or progesterone therefore cannot be confirmed. It should also be noted that RFD was highly variable both intra-phase and inter-test (Table S4.3). A change in bioavailable testosterone between MC phases has also been purported to alter strength/power (Cook et al., 2018), however free testosterone did not differ across phases among athletesNM. Hence, it appears that the negative relationships between E:T and both early phase RFD and impulse during the SJ are driven by fluctuations in oestradiol, and not testosterone, and therefore mirror the negative correlations between these outcome measures and oestradiol in isolation.

The lack of change in overall performance outcomes in athletesNM, or between athletesNM and athletesHC, combined with an inconclusive influence of oestradiol and progesterone, suggests that fluctuations in sex hormones may not alter performance outcomes in our population of Tier 3 female athletes. Earlier research surrounding the effect of MC phase or HC use on performance is highly heterogeneous. Numerous studies support our findings, demonstrating no influence of MC phase on measures of strength, power, or velocity (De Jonge et al., 2001; Lebrun et al., 1995; Romero-Moraleda et al., 2019). Indeed similar to the present study, both Pessali-Marques et al. (2024) and Thompson et al. (2021) observed no alteration in CMJ or SJ height between MC phases, but did report correlations between both oestrogen and progesterone various musculoskeletal parameters (Pessali-Marques et al., 2024), alongside an enhanced CMJ flight time during phase 4 compared to 2 (Thompson et al., 2021).

However, there are other reports of improvements in these indices during phases 2 and 3 of the MC (Ansdell et al., 2019; Pallavi et al., 2017), as well as a decline in strength-based outcomes during phase 1 (Dam et al., 2022; Gordon et al., 2013; McNulty et al., 2020b), alongside studies reporting the opposite (Davies et al., 1991; Phillips et al., 1996). Studies examining cognitive performance are similarly inconclusive, with some prior work supporting our lack of relationship between oestradiol and progesterone and cognition (Hampson, 1990; Kozaki & Yasukouchi, 2009), while others report alterations across the MC (Barel et al., 2019; Šimić & Santini, 2012). Therefore, our study of an authentic training squad revealing no detectable differences in overall cognitive or physical performance within or between athletesNM and athletesHC suggests that the logistical difficulties with altering "real-world" team testing according to MC phase are not justified.

The majority of previous investigations are confounded by a lack of hormonal verification of MC phase or confirmed ovulation (McNulty et al., 2020b). This lack of verification hinders the confidence in findings, as the actual phase and hormonal profile at which a measurement has occurred is unknown. Indeed, many studies use the calendar-based counting approach to classify MC phases, which is demonstrated to be inadequate since it assumes ovulation is exactly mid-cycle and involves no luteal phase and ovulation assessment (Elliott-Sale et al., 2021). Moreover, due to intra-individual MC variability, a particular cycle day is not guaranteed to be the same phase in different cycles in the same individual (Elliott-Sale et al., 2021). Prior studies also examined different combinations of "phases" (e.g., two *vs* four, follicular *vs* luteal) consequently hindering the ability to compare findings across studies.

The methodological quality of MC control and phase verification may influence study findings. The meta-analysis by McNulty et al. (2020b) reported that the majority of papers (12 out of 13) demonstrating differences in strength between MC phases were of low quality, while those studies identified as moderate-to-high-quality trended towards no differences between MC phase (nine out of 10). Training status may also impact any influence of MC phase on performance: differences on performance indices examined may be too subtle to detect in an athletic population already highly trained in the performance indices examined; hormonal influence may not exceed typical day-to-day performance variability; or differences are masked by high training volumes. Indeed, prior studies examining participants ≥Tier 2 (McKay et al., 2022b) in combination with some MC phase verification (retrospective serum oestradiol and progesterone and/or confirmed ovulation) have typically trended towards null findings pertaining to alterations in strength/power/speed across MC phases (Julian et al., 2017; Lebrun et al., 1995; Romero-Moraleda et al., 2019; Vaiksaar et al., 2011). In addition, prior studies have typically examined performance tasks that lack applicability to a high-performance sporting environment, such as single-limb exercises (McNulty et al., 2020b), whereas our study utilised common performance measures, including those utilised in the National Rugby League testing battery. It may be that the higher performance variability in the dynamic sport-specific tests examined in the present study

versus controlled or lab-based tasks also outweighed small differences in performance across the MC. The higher athletic calibre of our participants, the sport-specific ecological validity, combined with gold standard classification and control of menstrual status, may therefore help to explain the lack of performance differences between phases.

Other factors that may influence performance should also be considered. For example, pre-menstrual symptoms commonly associated with the end of phase 4 or beginning of phase 1 may alter performance, irrespective of any hormonal influences (e.g., cramps, bloating, tiredness, gastrointestinal issues, and poor sleep). These negative symptoms are reportedly experienced by ~60-93% of female athletes (Findlay et al., 2020; Martin et al., 2018), with ~50-67% believing that such symptoms impair performance (Bruinvels et al., 2017; Findlay et al., 2020). However, we observed a low frequency of symptoms throughout the duration of the training camp, as assessed through daily online questionnaires reporting symptom presence (McKay et al., 2024). Thus, MC symptoms appear unlikely to have influenced performance. However, symptom severity was not recorded and so presents an area for future study.

Study findings should be considered in light of potential limitations. Phase 2 could only be confirmed in one out of 11 athletesNM. While highlighting the complexities of research among women, this also meant that a correlational approach was taken to facilitate the inclusion of "phase 2" data, which is unable to determine causality. Measurements of serum oestradiol and progesterone were collected at a single timepoint on the day of testing, meaning it was not possible to determine if the hormonal concentration was rising or falling. Moreover, diurnal variation in endogenous oestradiol and progesterone concentrations were also not considered. While we acknowledge our study, with its small participant number, may be underpowered to detect marginal differences in our chosen performance tests, this is one of the first studies using sport-specific performance tasks among well-trained (Tier 3) athletes with a gold-standard approach to MC classification and control, thus improving the robustness and ecological-validity of our findings to the athlete-specific literature. Given the well-trained nature of the population in a training camp environment, it was not possible to control training load in the days preceding testing, which may have masked the ability to detect any small performance alterations. Additionally, testing occurred across a single MC among athletesNM, so we could not determine if any observed effects prevailed during another MC. Since such limitations are also present in the real world, when coaches or performance scientists are asked to consider regimens involving menstrual phase or status-based testing at a group level for a squad, particularly in a national team camp environment, we feel that our study outcomes are still able to inform a decision regarding phase-based testing. The participant cohort also presented a heterogenous mixture of hormonal profiles, with the athletesHC group using a variety of HC types and menstrual irregularities detected among six athletesNM, which may also have obstructed the detection of any minor performance alterations. However, the divide between athletesHC (54%) and athletesNM (46%) is similar to reported prevalence rates among athletes

(Martin et al., 2018) and therefore reflective of heterogeneity within a real-world training squad for which a coach might be asked to consider “menstrual phase or status” testing programs. Our findings suggests that such an approach is not justified at the group level. However, in the applied setting it may be beneficial to undertake long-term MC tracking on an individual athlete basis to identify any performance alterations with menstrual status, although such repeated and longitudinal measures were beyond the scope of the present study.

4.6 Conclusions and future research

Our findings demonstrate no detectable influence of MC phase or HC use on overall physical and cognitive performance outcomes among rugby league athletes. Some kinetic or kinematic outputs during jumping movements may be altered, however it could not be determined if the observed alterations exceeded between-day variability. Further research is required to determine causality and fully understand the effects of oestradiol and progesterone on performance, alongside underpinning mechanisms. In the meantime, our study represents a real-world training squad for which a coach might be asked to consider “menstrual phase or status” based testing programs and fails to provide evidence that such an approach is justified at a team-based level.

Interlinking chapter

Study 2 (*chapter 4*) set out to examine performance across MC phases 1, 2 and 4. However, retrospective analysis of serum hormone measurements revealed that a true MC phase 2 was only achieved in one out of the 11 athletes not using HC. As such, phase 2 data was removed from analysis, consequently excluding almost one third of the collected performance data from phase-based analysis. Therefore, Study 3 (*chapter 5*) set out to use a correlational model to explore the relationship between ovarian hormonal fluctuations across the MC and performance, without seeking to prospectively test in specific MC phases, in an attempt to maximise data retention. This approach also did not restrict inclusion criteria to exclusively naturally menstruating athletes, and therefore facilitated the inclusion of participants with a range of ovarian hormonal profiles (e.g., oligomenorrhea), provided fluctuations in oestrogen and progesterone were present, to increase sample size and ecological validity.

Chapter 5: Perceived negative menstrual cycle symptoms, but not changes in oestrogen or progesterone, are associated with impaired cycling race performance

Publication statement:

This chapter comprises the following paper accepted for publication in *Medicine and Science in Sports and Exercise.* Published online ahead of print.

Smith, E. S., McCormick, R., McKay, A. K., Ackerman, K. E., Elliott-Sale, K. J., Stellingwerff, T., Harris, R., & Burke, L. M. Perceived negative menstrual cycle symptoms, but not changes in oestrogen or progesterone, are associated with impaired cycling race performance. *Medicine and Science in Sports and Exercise*, (aop). https://doi.org/10.1249/MSS.0000000000003587.

Data availability: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Ethical approval: Ethical approval was granted by the Australian Catholic University Human Ethics Research Committee (2023-3192H) in accordance with the Declaration of Helsinki.

Conflicts of interest: Authors declare no conflict of interest. The results of the present study do not constitute endorsement by ACSM. Authors declare that the results are presented clearly, honestly, and without fabrication, falsification, or inappropriate data manipulation.

Funding: This project was funded by the Australian Catholic University.

5.1 Abstract

Purpose

To examine the relationship between MC phase-dependent fluctuations of oestrogen and progesterone and virtual cycling race performance, with a secondary aim of correlating perceived MC-related symptoms with performance.

Methods

In a novel observational study design, thirty-seven female cyclists/triathletes not using any HC completed one virtual cycling race (19.5 km TT) per week across a one-month period (totalling four races). Participants completed MC characterisation and tracking, including urinary ovulation kits, across two complete MCs. Venous blood samples were collected within 21 h of racing to determine serum 17-β-oestradiol and progesterone concentrations, as well as an assessment of self-reported, perceived race-day MC and gastrointestinal (GI) symptoms, which were all then correlated to race performance.

Results

There was no relationship between race completion time and individual oestradiol (r=-0.001, *p=0.992*) or progesterone (r=-0.023, *p=0.833*) concentrations. There was no difference between race time between MC phases (follicular/luteal, *p=0.238*), whether MC bleeding or not bleeding (*p=0.619*) and ovulating or not ovulating (*p=0.423*). The total number of perceived MC symptoms recorded on race day was positively correlated to increased race time [r=0.268 (95% CI 0.056 to 0.457), *p=0.014*], as was the number of GI symptoms of at least "moderate" severity before the race (r=0.233 [95% CI 0.021 to 0.425], *p=0.031*), but not post-race (r=0.022, *p=0.841*).

Conclusions

When implementing a novel, virtual cycling race, fluctuations in ovarian hormone concentrations across the MC do not appear to affect real-world cycling performance among trained cyclists, while perceived negative MC and GI symptoms may relate to impaired performance. Therefore, the management of negative MC and GI symptoms appears important for athletic performance enhancement or to mitigate performance decline.

Keywords: women, female, menstrual cycle, ovarian hormones, contraception

5.2 Introduction

Despite much discussion among the scientific community (McNulty et al., 2020b) and general population (McCallum, 2022), consistent and high-quality evidence of changes in performance in response to fluctuations in the female sex-steroid hormones across the MC is lacking. Alterations in oestrogen and progesterone concentrations have the potential to influence multiple physiological systems associated with athletic performance, such as substrate utilisation (Hackney et al., 1994; Oosthuyse & Bosch, 2010) or force production (Pallavi et al., 2017; Smith et al., 2002). Understanding if performance is systematically altered in response to changes in oestrogen and/or progesterone concentrations is important for competitive athletes.

According to a recent meta-analysis, exercise performance might be trivially reduced during phase 1 (early follicular, low oestrogen and progesterone concentration, begins at menstruation) of the MC comparative to all other MC phases, with the largest difference in performance between phase 1 (early follicular) and phase 2 (late follicular, highest oestrogen and low progesterone concentrations) (McNulty et al., 2020b). These results suggest that the low concentrations of oestrogen and progesterone observed during phase 1 may elicit a performance decrement, while elevated oestrogen may be performance-enhancing. However, given that findings are highly inconsistent between studies, the magnitude of effect is trivial, and results are confounded by study quality variability (particularly regarding flaws in methodological control and classification of ovarian hormones), the conclusions from the available literature are considered weak (McNulty et al., 2020b). However, if trivial but consistent differences do exist, these may be important for elite athletes, for whom marginal gains are of consequence.

Other MC-related factors that may influence performance should also be considered. Indeed, symptoms (e.g., bloating, muscle aches, fatigue, gastrointestinal (GI) issues, headaches, poor sleep, and anxiety) commonly associated with the end of the luteal phase or beginning of the follicular phase (during menses) may hinder performance. These negative symptoms are reportedly experienced by ~60-93% of naturally menstruating female athletes (Armour et al., 2020; BBC Sport, 2020; Findlay et al., 2020; Martin et al., 2018; McNamara et al., 2022; Oxfeldt et al., 2020), with ~50-67% believing that such symptoms impair performance (Armour et al., 2020; Bruinvels et al., 2017; Findlay et al., 2020; McNamara et al., 2022). Disparities may further exist between trained [Tier 2 (McKay et al., 2022b)] and recreational (Tier 1) athletes, with the possibility of detecting minor performance changes as a result of MC-related hormone fluctuations potentially greater in the former, given their superior performance consistency compared to recreational athletes (Hopkins & Hewson, 2001). Therefore, it is important to examine performance indices across the MCs of trained athletes to account for these adaptations and for whom performance is more consistent; hence any small influence of the MC may more likely be detected.

Investigations of the MC in real-world performance settings (i.e., actual competition) are ideal given the high ecological validity but are simultaneously difficult to conduct and are therefore lacking. Indeed, most studies of performance across the MC are undertaken in a controlled laboratory environment (Jurkowski et al., 1981; Nicklas et al., 1989), in which participants are often blinded, fasted, and lack real-world motivation to perform. Furthermore, the few studies that have attempted to examine real-world performance have typically been hampered by self-reported MC phases and failure to verify ovulation and/or serum oestrogen/progesterone concentrations (Fomin et al., 1989; Quadagno et al., 1991). Unsurprisingly, the outcomes from such studies are conflicting and highly variable (Bale & Nelson, 1985; Fomin et al., 1989; Guo et al., 2005; Julian et al., 2021; Quadagno et al., 1991).

Accordingly, the aim of this study was to examine the effect of oestrogen and progesterone on sports performance among female athletes not using HC, employing robust methodological control of menstrual status, the recruitment of trained athletes, and an ecologically valid measurement of performance using an online virtual (Zwift) competition. A secondary aim of the project was to examine the effect of perceived MC- and GI-related symptoms on performance.

5.3 Methods

Experimental overview

In a novel observational study design (Figure 5.1), participants completed one virtual indoor cycling race per week across a one-month period (totalling four races) using the Zwift online cycling platform (2023 Zwift, Inc. v2.183.0). Venous blood samples were collected within 21 h of racing (pre- or post-race) to determine serum 17-β-oestradiol and progesterone concentration. The concentrations of these sex hormones were then matched with the respective race and correlated to each participants' race completion time. The incidence of MC and GI self-reported symptoms on race day were also correlated to race time as a secondary outcome measure. The study was approved by the Australian Catholic University Human Ethics Research Committee (2023-3192H) and conducted in accordance with the Declaration of Helsinki. All participants provided informed consent prior to participating.

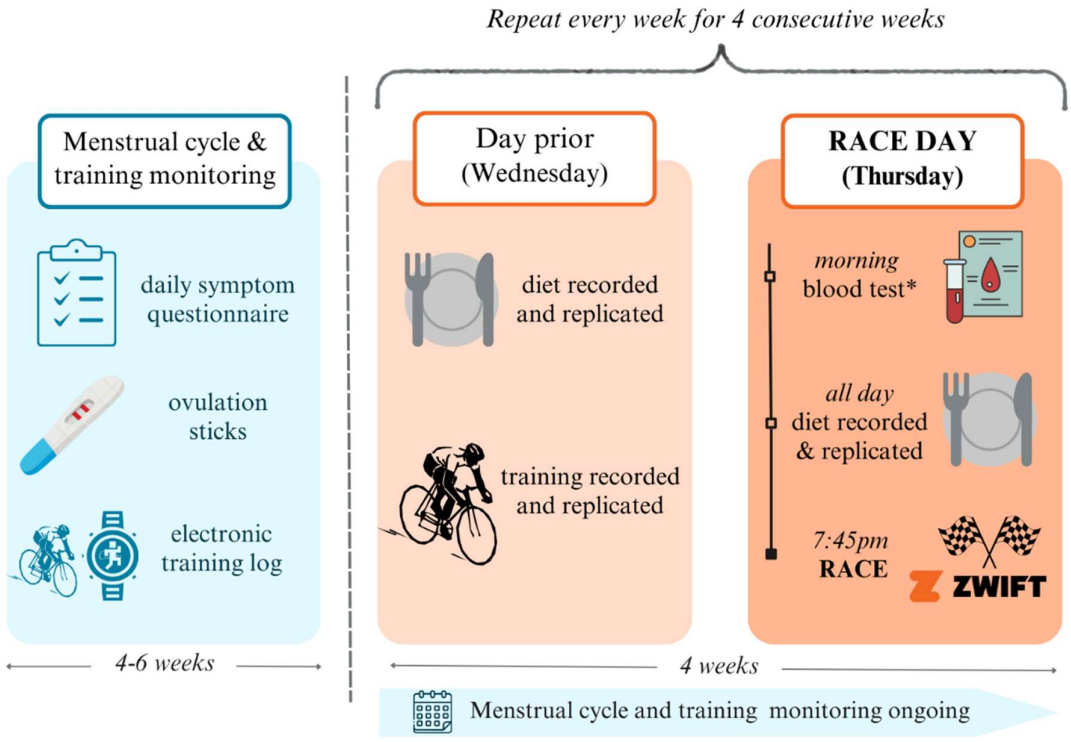


Figure 5.1. Study protocol. Participants completed a monitoring period (to capture both their menstrual cycle and training) of between four and six weeks prior to, or after if necessary, the commencement of the four-week race series. *The blood sample was collected either the morning of (Thursday) or after (Friday) each race at the same time each week.

Participants

Thirty-seven Tier 2 (McKay et al., 2022b) female cyclists/triathletes (mean age: 35±6 y mean body mass: 67.0±10.3 kg, mean training volume: 8.0±3.5 h/wk, mean age of menarche: 13±3 y) were recruited. Inclusion criteria were: residing in Australia, pre-menopausal (confirmed via ovulation detection), absence of HC for >three months prior to study commencement, not pregnant or breastfeeding (Figure 5.2). The only exclusion criteria based on MC function was current amenorrhea (absence of a MC for >three months) (Elliott-Sale et al., 2021), given that it results in the suppression of endogenous hormones and hence would prohibit the investigation of the primary outcome. Other menstrual irregularities that do not as severely suppress endogenous hormonal profiles were therefore included to increase study generalisability.

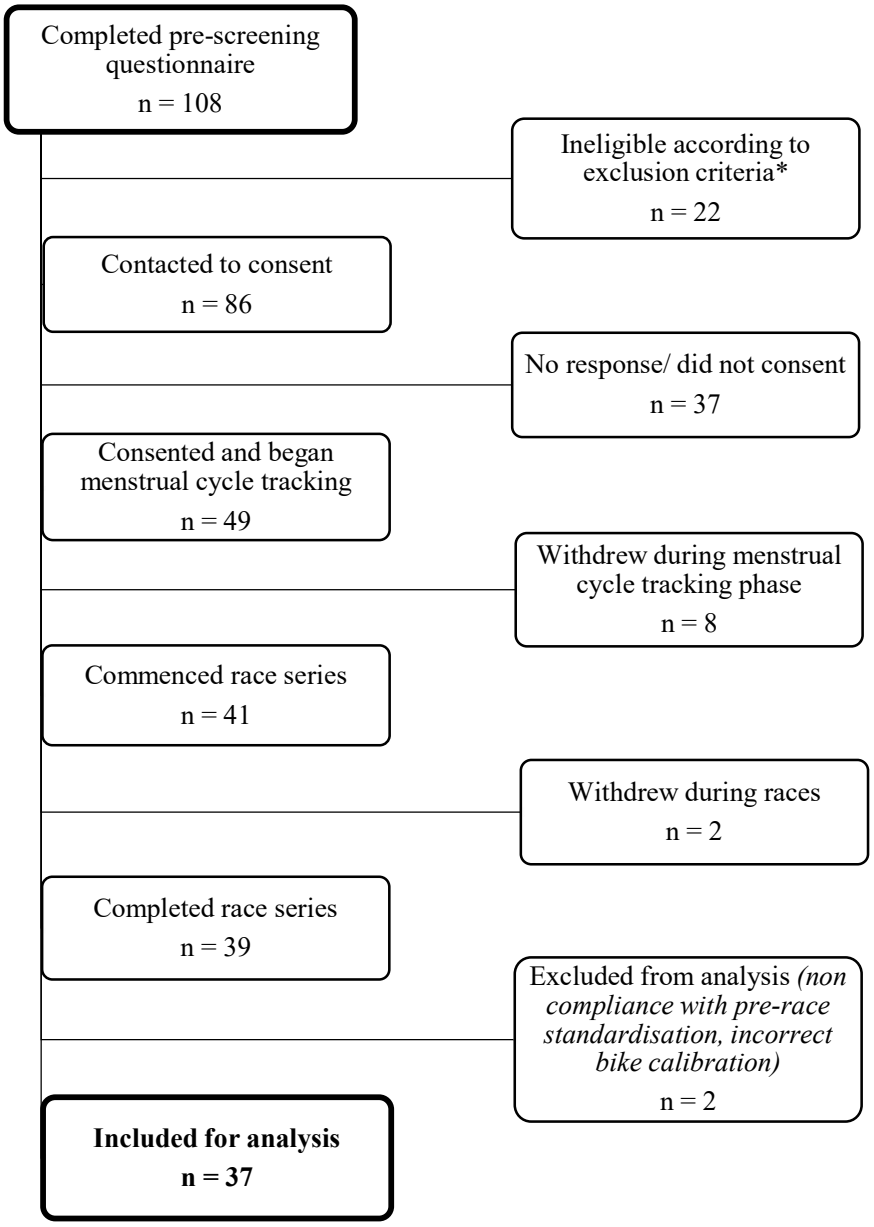


Figure 5.2. Participant flow chart from pre-screening to the final sample size included for analysis. *Exclusion criteria: living outside of Australia, exercising <150 min per week, no access to Zwift cycling app, use of hormonal contraceptives within three months of study commencement, currently pregnant or breastfeeding, current amenorrhea.

Menstrual cycle monitoring

Prior to participation, athletes completed an initial questionnaire regarding their menstrual status, including MC length and frequency, prevalence of known MC dysfunction [e.g., PCOS, amenorrhea],

and any current or previous HC use. Participants' MCs were then tracked, according to best-practice protocols (Elliott-Sale et al., 2021), across the four weeks of racing – with additional weeks before or after to capture two complete MCs per athlete. Participants completed daily online questionnaires [REDCap (Harris et al., 2019; Harris et al., 2009)] pertaining to presence and heaviness of menstruation, symptom incidence, and medication use in the preceding 24 h (Figure S5.1). Athletes also used dual hormone urinary ovulation kits (Advanced Digital Ovulation Test, Clearblue, Geneva, Switzerland) from MC day 10 until ovulation occurrence (continuing until the next bleed if ovulation was not detected), recording the result on the online questionnaire. Venous blood samples were collected within 21 h of racing (pre- or post-race) to determine progesterone and serum 17β-oestradiol (the most potent form of oestrogen among pre-menopausal women, henceforth referred to as "oestradiol") concentration.

Training monitoring

Alongside MC tracking, participants logged all training via Strava or Garmin, and reported the type, duration, and intensity of sessions via the daily questionnaire, together with the presence of any injuries.

Zwift races

Races were an individual TT format: 19.5 km in length with 32 m elevation. The race was a private event open only to study participants, whereby participants could see all other competitors in the ride to replicate a real race environment. The Zwift software was programmed to a standardised bike setting, while drafting and powerups were disabled. The indoor trainer (n=33) or stationary bike (n=4) was consistent within each participant across all their races. Participants raced every Thursday evening across four consecutive weeks, commencing at 19:45 AEDT. Participants chose their own warm-up and replicated this each week. The race was completed indoors, with permission to use fans or air conditioning.

To enhance the ecological validity and motivation, prise money was available to the top performers. Participants were grouped into categories (A-D) based on ability ($W \cdot kg^{-1}$) (Zwift, 2024). At each race, participants provided a photo of themselves standing on a scale pre-race to verify body mass. Prise money was awarded separately across each category, such that riders were only directly competing against individuals of a similar ability. Participants voted on the prize money allocation system, and the number of prizes awarded was adjusted based on the total number of athletes, such that the top 30% of riders in each category were awarded a prize.

Pre-race standardisation

Dietary intake (all food, beverages, and caffeine consumption) was standardised for 36 hours pre-race, with participants allowed to choose their own nutrition strategies but repeat them for each race. Dietary records were maintained to verify compliance with these instructions, via the use of meal photos posted on the MealLogger app (MealLogger). Alcohol was prohibited throughout both days. Training was

permitted the day before the race but was kept consistent every week and recorded on Strava/Garmin for verification. No training was permitted on the day of the race, with the exception of one athlete who completed the same 45 min run on the morning of each race.

Pre- and post- race questionnaires

Before (within 15 mins of race commencement) and immediately after each race, participants completed an online questionnaire [REDCap (Harris et al., 2019; Harris et al., 2009)] regarding GI symptoms (Pugh et al., 2019), (a score ≥5 was deemed at least "moderate" severity, Table S5.1), thermal perception [thermal sensation (TS) and thermal comfort (TC), Tables S5.2 and S5.3 (Zhang et al., 2004)]. Visual analogue scales (0-100) measured readiness to race (pre-race only): "how ready to race do you feel?", with 0 representing "not at all ready" and 100 as "the most ready I have ever felt", and race perception (post-race only): "how do you feel like you raced?", with 0 representing "the worst I have ever raced" and 100 as "the best I have ever raced".

Blood sampling

Each week, using pre-organised pathology request slips, participants attended the same commercial pathology branch (Australian Clinical Labs) to have a rested blood sample drawn (total of four samples). An 8.5 mL venous blood sample was collected by a trained phlebotomist into a serum separator tube. Oestradiol and progesterone were measured via a Siemens Atellica IM Analyzer using a direct chemiluminescent immunoassay. Four participants did not reside in the locale of an Australian Clinical Labs centre and therefore attended an alternative pathology centre (Healius Pathology). Participants were advised to complete their blood test the morning prior to the race; however, athletes were not excluded from participation if this was not achievable. Hence, the blood samples were collected either the morning of (74% of participants) or after (26%) each race at the same time each week (±1.3 hours), at a mean time of within 11.5 h of the race start and all samples were collected within 21 h of the race.

Statistical analysis

Statistical analyses were performed using R Studio (v3.5.2) with statistical significance accepted at an α level of $p \leq 0.05$. Data are presented as mean±SD. Hormone concentrations >three SD from the group mean were removed as outliers [three elevated oestradiol measures and one elevated progesterone measure, (Howell et al., 1998)]. Repeated measures correlations assessed associations between race time and oestradiol/progesterone concentration and the progesterone: oestradiol ratio (P:E. as $nmol \cdot L^{-1}$) as our primary outcome measures, alongside our secondary outcome measures: total perceived MC symptoms, GI symptoms of at least moderate severity, and changes in TC/TS pre- to post-race. Because these secondary outcomes are ordinal measurements, they present some analysis limitations, however a non-parametric alterative to repeated measures correlation does not exist. A one-way ANOVA or paired t-test assessed differences in race completion time and participant weekly training volume across

the four races, alongside participant mean weekly training volume during race weeks compared to volume across non-race weeks.

Sub-analyses using paired t-tests were conducted for athletes who experienced: menses during a race (n=24) and/or ovulation within 24 hours of the race (n=9) with race performance during these events compared to the mean performance across other races. Race completion time during follicular vs. luteal phases, as separated by ovulation, was also compared for athletes completing at least one race in each phase (n=31). Finally, sensitivity analyses were performed (Deeks et al., 2019), whereby results were analysed separately excluding athletes with MI (n=8, 27 races) and races with minor protocol deviations (six races).

5.4 Results

Thirty-seven participants competed in the race series, with n=19 cyclists completing all four races, n=15 completing three and n=3 completing two. Six participants had minor protocol deviations on one race occasion, including training prior to the blood test (n=4), lack of dietary replication (n=1) and lack of prior day training replication (n=1). However, sensitivity analyses removing these six races did not affect the results. In total, 127 individual races were completed, with five individual races excluded due to technical issues during the race while a single race was excluded due to a missed blood test. This totalled 121 individual races for final analysis. Weekly training time in the weeks before/after the races (497±211 min·week^{-1}) did not differ from the training time completed during the four-week race period (484±273 min·week^{-1}, *p=0.714*); weekly training time also did not differ between the four race weeks (*p=0.426*).

Menstrual characteristics: A total of 2,493 questionnaires were completed across the study duration (four racing weeks, plus additional weeks pre- or post-racing weeks, to capture data for two complete MCs) with a compliance rate of 98%. Each participant recorded two complete MCs, with the exception of one athlete who, due to later study enrolment and long cycle length (41 days), only had complete data for a single MC. Our cohort had a MC length of 28±4 days, with 5±1 bleeding days and ovulation occurring on day 15±3. Ovulation was detected in only one of two monitored MCs for five athletes, and in both cycles for 30 athletes, while suspected anovulation was detected in two athletes (aged 28 and 29, with ovulation detected in all older athletes, confirming pre-menopausal inclusion criteria). Therefore, ovulation was detected in 90% of MCs observed across the 37 athletes. The two athletes with suspected anovulation were in the top 25% with regards to weekly training volume (>10 hours per week), however there was no difference in weekly training volume for athletes who displayed two ovulatory MCs (n=30, 468±218 min·week^{-1}) compared to those with disturbances to ovulation as outlined above (n=7, 577±184 min·week^{-1}, *p=0.232*).

Participants' menstrual status was retrospectively classified through calendar counting, urinary ovulation, and serum hormone measurements (Table S5.4): n=18, eumenorrheic; n=11, naturally menstruating; n=8, with MI. Prior to MC monitoring, five athletes reported diagnoses of menstrual dysfunction: PCOS (n=3) and endometriosis (n=2). Following MC monitoring, we identified a further four MIs: suspected anovulation (n=2), oligomenorrhea (n=2, one of whom also had PCOS) and polymenorrhea (n=1, who also was anovulatory). Moreover, based on initial data, two athletes had prior diagnosed primary amenorrhea, while a further eight reported onset of menses at ≥15 years of age. However, all participants were regularly menstruating for at least three years prior to and throughout the study. Sensitivity analyses, removing the eight athletes with MIs did not alter results, and hence they were included for analyses. Therefore, unless otherwise stated, results are presented for 37 athletes across 121 races.

Race performance: There was no correlation between race completion time and oestradiol or progesterone concentration (Figures 5.3A and 5.3B), nor the P:E ratio (r=-0.024, *p=0.834*). Mean race completion time was 31:13±03:04 (mm:ss) and not differ between the four races (*p=0.458*). Performance variability between races was 58±51 s (3%)

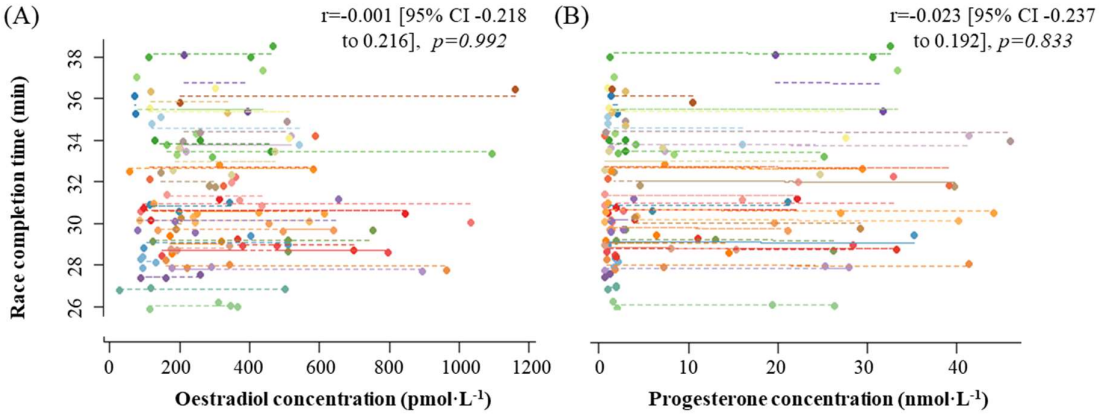


Figure 5.3. Repeated measures correlation between race completion time and (A) oestradiol concentration, and (B) progesterone concentration. Different colours represent individual participants.

There was no difference in race completion time on days when athletes were bleeding, comparative to non-bleeding days (n=24, Figure 5.4A), nor any difference between race performance on days when athletes were ovulating compared to other races (n=9, Figure 5.4B). Race performance was also not different in the follicular comparative to luteal phase, as separated by ovulation (n=31, Figure 5.4C).

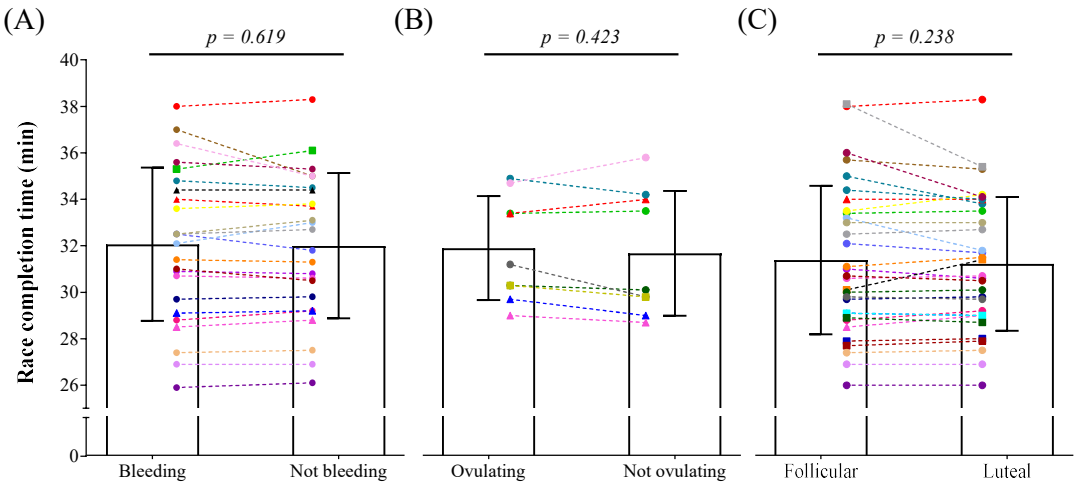


Figure 5.4. Race time, separated by (A) races when athletes were bleeding compared to the mean time across non-bleeding race days (n=24), (B) days when athletes ovulated on race day compared to the mean time across other races (n=9), and (C) between follicular and luteal phases, as separated by ovulation for ovulatory athletes who completed at least one race in each phase (n=31). The different colour lines and symbols represent individual participants: triangle denotes participants present on all three graphs, a circle for two graphs, and a square for one graph.

Symptomology: Most (92%, n=34) athletes reported at least one perceived MC symptom on at least one race day. Of these, n=10 had >3 symptoms, n=24 had 1-3 symptoms, and 82% (n=30) recorded symptoms across multiple races. Bloating was the most common self-reported MC symptom (17% of all symptoms reported), followed by fatigue (14%), abdominal cramps (9%) and appetite changes (9%).

Regarding GI symptoms specifically, 46% (n=17) of athletes reported symptoms considered to be at least "moderate" in severity pre-race on at least one occasion; 27% (n=10) experienced moderate symptoms before several races, and 62% (n=23) experienced moderate symptoms post-race at least once. The most common GI symptom pre- and post-race was bloating, accounting for 47% of all moderate severity symptoms pre-race and 28% post-race; with nausea (16%) and urge to vomit (16%) also common post-race.

The number of GI symptoms of at least "moderate" severity pre- but not post-race positively correlated with race time (Figures 5.5A and 5.5B), as did the total number of perceived MC symptoms recorded on race day (Figure 5.5C). The total number of self-reported MC symptoms on race day negatively correlated with race perception [r=-0.307 (95% CI -0.490 to -0.097), *p=0.005*], but not readiness to race (r=-0.102, *p=0.363*). There was a negative relationship between oestradiol concentration and the total number of perceived MC symptoms reported on race day [r=-0.267 (95% CI -0.459 to -0.052), *p=0.016*] but no relationship between total perceived MC symptoms and progesterone (r=-0.047, *p=0.672*) nor

P:E (r=-0.068, *p=0.549*). There was also no correlation between oestradiol and progesterone, nor their ratio, and moderate GI symptoms pre- or post-race (all *p>0.050*).

Where follicular and luteal phase could be verified through ovulation, 14 instances of moderate GI symptoms pre-race occurred in the follicular phase, and 15 instances in the luteal. There were 35 instances of perceived MC-related symptoms on race day in the follicular phase, and 32 instances during luteal. For athletes reporting bloating as a GI symptom pre-race, there were 19 instances during the follicular phase, and 16 occasions during the luteal phase. There was no difference in body mass on occasions athletes reported "moderate" bloating compared to no bloating or that of less than "moderate" severity (*p=0.476*). Body mass also did not differ between follicular and luteal phases (*p=0.488*).

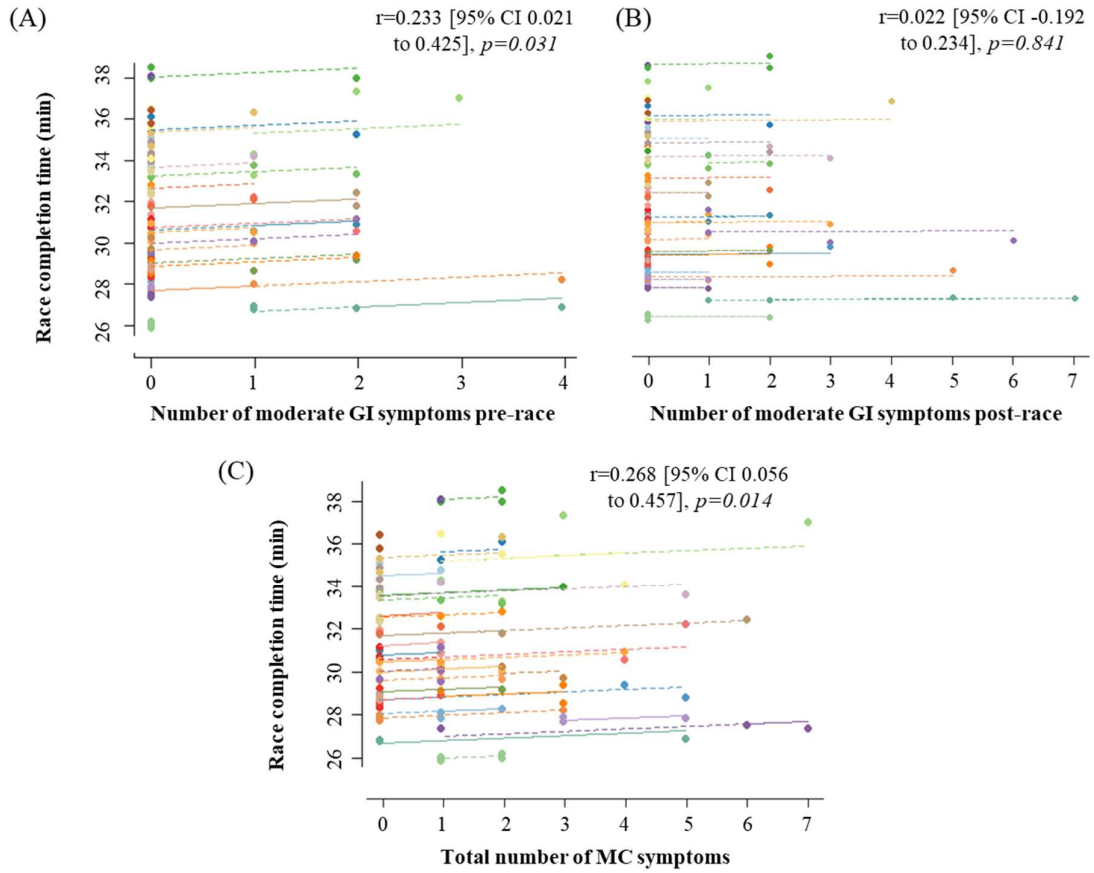


Figure 5.5. Repeated measures correlation between race completion time and (A) the number of "moderate" severity GI symptoms reported pre-race, (B) the number of "moderate" severity GI symptoms reported post-race, and (C) the total number of MC symptoms recorded on the day of the race. Different colours represent individual participants. *GI;* gastrointestinal, *MC;* menstrual cycle.

Thermal Perception: There was a negative relationship between both the change in TS (Figure 5.6A) and TC (Figure 5.6B) pre- to post-race and race completion time.

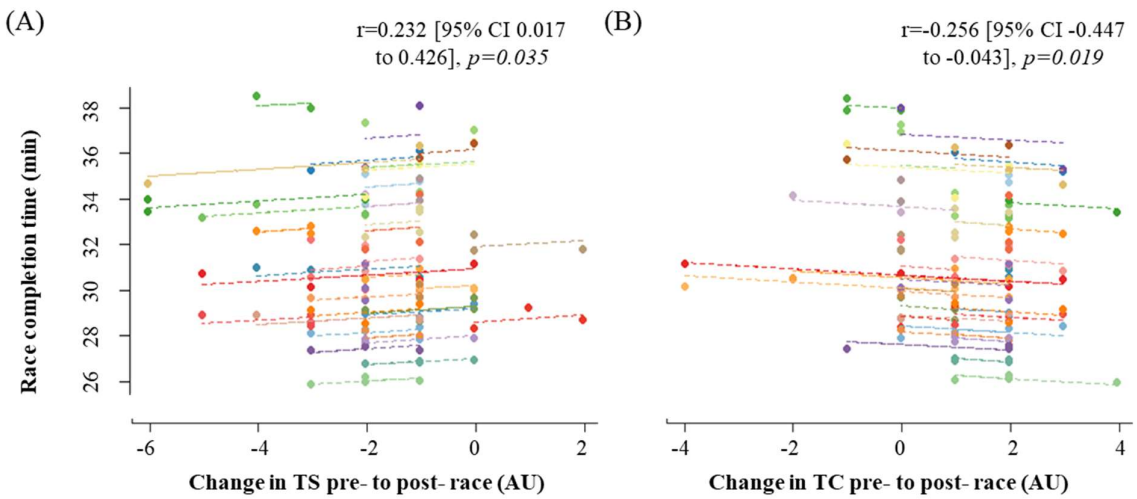


Figure 5.6. Repeated measures correlation between race competition time and (A) the change in participant ratings of thermal sensation pre-to-post race, and (B) the change in participant ratings of thermal comfort pre-to-post race. Different colours represent individual participants. *TS*; thermal sensation, *TC;* thermal comfort.

5.5 Discussion

This study used a novel protocol to investigate whether real-world cycling competition performance was associated with fluctuations in MC phases and associated sex-steroid hormones oestradiol and progesterone. Although we observed that oestradiol and progesterone concentrations were not related to race completion time, small relationships were observed between race performance and the total number of negative symptoms the riders associated with their MC phase, as well as the number of pre-race GI symptoms of moderate severity. Our findings suggest that fluctuations in ovarian hormone concentrations across the MC are not associated with real-world cycling performance, but perception of negative self-reported MC or GI symptoms may have a greater effect. We also note our experiences with a research protocol that offers the potential to increase validity and flexibility with participant recruitment, to enable others to utilise and refine this type of research study design.

Our results agree with prior studies that have incorporated verified serum oestradiol and progesterone concentrations and have failed to observe alterations in cycling TT (16-30 km) performance across MC phases in Tier 2 cyclists (McLay et al., 2007; Oosthuyse et al., 2005). Here we note the considerably larger cohort in the present study (n=37) compared to those investigations (n=5-9) (McLay et al., 2007; Oosthuyse et al., 2005). In addition, these studies either prohibited or failed to consider caffeine intake, and relied on TTs conducted in a laboratory environment, thus providing a less ecologically valid race scenario. Other studies of real-world performance have demonstrated a lack of group level/systematic alterations in football match metrics among Tier 3-4 athletes when serum oestradiol and progesterone concentrations were confirmed (Julian et al., 2021). Moreover, our findings may also explain recent

work suggesting that fluctuations in sex hormones across the MC do not contribute largely to performance changes when compared to the potential effects of individual or day-to-day variation (Julian et al., 2021; Smith et al., 2024c), however these studies have examined team sports which entail different physiological demands to cycling.

By contrast, some studies verifying oestradiol and progesterone concentrations have reported a decline in laboratory measures of endurance capacity during the luteal phase (when progesterone is elevated and oestradiol is moderate) (Julian et al., 2017; Lebrun et al., 1995). A rationale to explain the divergent results includes the recognition of a real-life competition as a dynamic environment in which potential subtle or trivial changes in performance due to ovarian hormone fluctuations may be outweighed by overriding factors such as day-to-day performance variability. This is of real-world significance given the dynamic nature of competitions in which athletes participate. The performance variability between the four races was 3% (58 s): 1% (20 s) for riders in Category A (n=7), 3% (51 and 48 s) for category B (n=6) and C (n=9), and 4% (84 s) for category D (n=15). This aligns with our knowledge that elite athletes demonstrate lesser variability in performance comparative to those less highly trained (Hopkins & Hewson, 2001) . It is worth noting that even when a sub-analysis examining the relationship between ovarian hormones and race time was conducted among only the more highly trained athletes (categories A-C, n=22), no relationship was detected. This suggests that any variability due to hormonal influences is perhaps too small to be detected and potentially outweighed by intrinsic performance variability >1%. Another explanation for the divergent results may also be pre-race fuelling in the present study, which may override any influence of oestradiol or progesterone on performance (Campbell et al., 2001). Indeed, many studies reporting alterations to endurance performance/capacity have been conducted in the fasted state (Campbell et al., 2001; Lebrun et al., 1995; Lee et al., 2024), which is not reflective of real-world pre-competition practices.

While we did not detect an alteration in performance with physiological fluctuations in oestradiol or progesterone concentration, negative symptomology (both MC and GI) was related to a slower race performance (with bloating the most prevalent symptom). It is possible that athlete perception or subjective feelings are potentially more influential than physiological variations *per se*. It is common for women to perceive an impairment in training and competition performance during MC phases 1 and 4, in association with negative symptoms (Armour et al., 2020; Carmichael et al., 2021; Dam et al., 2022; Findlay et al., 2020; Oester et al., 2024). Indeed, some studies that report no performance alterations with hormonal fluctuations across the MC have observed performance changes related to psychological well-being (Dam et al., 2022) and negative MC symptoms (Giacomoni et al., 2000). However, few studies have directly examined the influence of symptoms on performance, instead providing an indirect link by concentrating on the incidence of symptoms across the MC in conjunction with the athletes' perception of how symptoms influence performance. Future research should undertake a more direct investigation of this association, including the pre-tracking of perceived MC-related

symptoms prior to examination under experimental conditions, as well as pre-trial assessment of athletes' personal beliefs around the impact that MC and/or GI will have on performance. We also note that the subjective collection of symptoms may be biased by the presentation of exclusively negative perceived MC symptoms. It is possible that recording, and therefore drawing attention to, positive symptoms (e.g., feeling energised) may counteract the reduced well-being that may be enhanced by focusing only on negative symptoms and feelings. Future research should investigate the potential performance benefits of positive MC symptomatology. Lastly, while we measured the severity of GI symptoms, other perceived MC-related symptoms were reported only in terms of incidence, and hence the influence of MC symptom severity on performance was not able to be examined. This therefore warrants future investigation with greater granularity.

It is important to consider that the correlation between symptoms (both MC and GI) and performance was weak in magnitude. Indeed, although the incidence of athletes self-reporting numerous MC symptoms on a single day was relatively low, occurring at 24 of 121 race occasions (20%), others have reported that athletes who identify three or more symptoms are twice as likely to state they are affected by their MC (McNamara et al., 2022). The presence of multiple GI symptoms at a single race was also relatively low. Of the 17 athletes reporting moderate symptoms before at least one race, just over half (59%, n=10) experienced symptoms at numerous races, while 20 athletes did not experience any GI symptoms of moderate severity prior to racing. It is possible that if more symptoms had been reported in our cohort, perhaps a stronger relationship to performance may have been observed. However, most (81%, n=30) athletes did report perceived MC symptoms on at least two separate races, while just four had symptoms at one race and three did not report any perceived MC symptoms across race days. Of course, symptoms typically associated with the MC (e.g., breast pain, bloating, abdominal cramps) have a range of other causes, and in the absence of exploring a differential diagnosis we cannot unequivocally attribute the reported symptoms to the MC. However, athletes who reported illness on race day were excluded from that race. Separately, medications were not restricted during this study and perhaps individuals experiencing negative perceived MC symptoms utilised analgesics or other relevant medications. However, of those experiencing symptoms, only 17% reported the use of paracetamol or ibuprofen on these days. Interestingly, of the occasions where paracetamol or ibuprofen was used, only 50% of these coincided with menstruation.

Of the 37 menstruating females in this study, only 48% (n=18) were classified as eumenorrheic, 30% were naturally menstruating, and 22% had MI, of which two athletes (5% of the 37) did not ovulate, despite menstruating. Further, despite a mean MC length of 28±4 days, only four out of 37 athletes (11%) had a "typical" 28-day MC, with considerable variability in cycle length within individual athletes (Table S5.4). This further supports previous work demonstrating that methods to assess MC characteristics that do not include ovulation confirmation are inadequate to sufficiently characterize menstrual status (Elliott-Sale et al., 2021; McKay et al., 2024). Given the prevalence of anovulation in

our Tier 2 athlete cohort, who average just eight hours of training per week, there is a need for future work to establish anovulation prevalence among elite athlete cohorts.

Finally, we implemented a novel protocol: remotely recruiting and managing participants and using a virtual cycling race for performance measurement. A live race protocol provided a competitive environment alongside an opportunity to engage a larger sample size than that typically employed in SES research. On the contrary, it was not possible to replicate the dynamic tactical aspects, such as pacing strategies, of a live race between weeks. The controls present in a laboratory environment were also not possible, and we were unable to collect additional data such as heart rate or substrate oxidation which would have provided greater mechanistic detail. Further, the "live" nature of the races meant that rescheduling trials in the event of illness/injury was not possible. Nevertheless, standardisation was implemented where possible, including racing on an identical Zwift course each week, using the same pathology lab for 33 out of 37 participants (as described above) and confirming participant compliance with study protocols. We encourage this novel "virtual" study design for future use/investigation given the opportunities presented by conducting a study in locations remote to the researchers.

Our findings must however consider potential limitations. Races were conducted in the evening, whereas morning exercise is typically used in most studies occurring in the fasted state. However, several competitions (including the Olympic Games) occur in the evening due to broadcasting requirements. Due to the real-world nature of this study, whereby participants competed in live races at specified times, we were unable to test specific MC phases and instead correlated hormonal concentrations to performance, which prevents determining causality. The mean time between race and blood sample was 11.5 hours, and therefore measured hormonal profiles may not be fully reflective of the hormonal milieu exactly at race time, although this is unlikely. Moreover, while participants were instructed to complete their blood sample pre-race, post-race was permitted, if necessary (with between seven and ten athletes completing the blood test post-race each week), and hence we cannot exclude the possibility of an altered sex-hormonal profiles associated with post-race stress in these instances. Finally, although we tracked participants' MCs according to best-practice protocols, and hence were able to identify eumenorrhea, some methodological considerations (Elliott-Sale et al., 2021) required to achieve a "gold" (Smith et al., 2022b) standard were not achieved. A minimum of two MCs were tracked per athlete; however, guidelines stipulate tracking for at least two months prior to testing. Moreover, races were conducted across a four-week period, and hence were not repeated across a second MC. These decisions were taken to reduce participant burden and increase adherence. Lastly, as explained above, we included individuals with MI (except amenorrhea) to increase generalisability and maximize data retention. Sensitivity analyses indicate that individuals with MIs did not affect the results.

5.6 Conclusions and future research

Cycling race performance in a virtual competition setting appears not to be systemically altered with fluctuations in oestradiol or progesterone across the MC in trained cyclists, but performance may be influenced by negative MC/GI symptoms. Therefore, an individualised approach, including monitoring and managing any negative symptoms, may be better for uncovering any links to individual athlete performance or mitigating performance decline. Future research should seek further understanding of the relationship between symptoms and performance, both examining if specific symptoms are driving an association, and if this relationship persists into other activities beyond cycling.

Interlinking chapter

Study 4 (*chapter 6*) collates the experiences of participants across the three experimental studies presented in this thesis. Study 4 also includes data from participants in an additional study that was not directly part of this PhD but was conducted simultaneously by our laboratory (Kuikman et al., 2024a). This study included a cohort of 19 elite (Tier 4-5) female athletes and was included to provide additional data regarding the experiences of the most highly trained athletes participating in research studies, therefore providing a broader range and depth of information.

Chapter 6: Original investigation: Female athletes report positive experiences as research participants

Publication statement:

This chapter comprises the following paper currently under review for publication in *International Journal of Sport Nutrition and Exercise Metabolism.*

Smith, E. S., McKay, A. K., Ackerman, K. E., Elliott-Sale, K. J., Stellingwerff, T., Harris, R., & Burke, L. M. Original investigation: Female athletes report positive experiences as research participants. *International Journal of Sport Nutrition and Exercise Metabolism,* (in review).

Data availability: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Ethical approval: This study collates findings from other previously published investigations, ethical approval can be found in the respective publications.

Conflicts of interest: Authors declare no conflict of interest.

Funding: No direct funding was received for this work. We acknowledge support of the Wu Tsai Human Performance Alliance and the Joe and Clara Tsai Foundation in conducting the studies included in this analysis.

6.1 Abstract

Given the underrepresentation of women in sports and exercise science (SES) research, we sought to understand the experiences of female athletes currently involved in applied SES research to inform future studies and potentially increase participation rates. Accordingly, we investigated the experiences of 89 female athletes (n=48 cyclists/triathletes, n=19 race walkers, n=22 National Rugby League Indigenous Women's Academy players) who participated in four separate studies of sports performance with different methodological characteristics. Participants completed a questionnaire upon study completion that queried prior research participation, reasons for participating and experiences during the current study. Across all four studies, 81% of female athletes were first-time research participants, with the primary barrier cited as a perceived lack of opportunities or being unaware of opportunities (93%). Participants rated an interest in the research outcome as the most important aspect influencing their decision to participate [90±14 (out of 100)], followed by the opportunities to receive personalized results (84±20) and education (78±27). Most participants (87%) stated that they would apply the study findings to their sports involvement, while the remaining 13% reported that they required support to understand the application of results. The majority (94%) of participants indicated a willingness to participate in future studies, while the research experience was rated positively at a mean 77 out of 100. Ultimately, our findings uncovered a perceived lack of opportunity as the primary barrier to female athlete research participation. As such, opportunities for women to participate in high quality studies should be prioritized.

Keywords: women, participation bias, sports science studies, participation barriers

6.2 Introduction

The lack of female participants in sports and exercise science (SES) research is well-established (Bruinvels et al., 2017; Costello et al., 2014; Cowley et al., 2021), with a particular scarcity of the involvement of female athletes of at least trained/development level [≥Tier 2, i.e., those who identify and compete in a specific sport, in contrast to recreational exercisers (Tier 1) or sedentary (Tier 0) individuals; (McKay et al., 2022b)] in studies examining performance outcomes (Kelly et al., 2024; Kuikman et al., 2023a; Kuikman et al., 2023b; Smith et al., 2022c). Moreover, since a high proportion of research conducted among women is of poor quality, especially methodological classification and control of menstrual status (Elliott-Sale et al., 2020b; McNulty et al., 2020b), it is difficult to draw robust conclusions from a literature base that is both of relatively low quantity and quality. As has been comprehensively discussed (Bruinvels et al., 2017; Costello et al., 2014; Cowley et al., 2021; Smith et al., 2022b), findings from research conducted in men cannot be directly applied to women without consideration of sexual dimorphisms. To address the sex-bias in the literature and ensure true "evidence-based" guidelines for female athletes, more high-quality research pertaining specifically to elite female athletes is needed.

Volunteer bias occurs when a certain sub-section of potential eligible study participants is less willing or available to volunteer to participate in research studies, thus potentially biasing the research results towards just those that participated. There is some evidence to suggest that women may have greater volunteer bias (i.e., are less likely to participate) than their male counterparts (Costello et al., 2014; Cowley et al., 2021; Nuzzo, 2021; Smith et al., 2022b), and this may be a contributor to the sex-bias in SES research. A better understanding of the experiences of female athletes involved in applied SES research may assist with strategies to inform future studies and increase participation. Accordingly, we investigated the experiences of female athletes [≥Tier 2, (McKay et al., 2022b)] who were recruited on the basis of biological sex, with studies conducted in cis-gender women (assigned female at birth). These female athletes participated in a range of studies of sports performance with different methodological characteristics (Kuikman et al., 2024a; Smith et al., 2024a; Smith et al., 2024b; Smith et al., 2024c), with all studies undertaking a robust classification of menstrual status.

6.3 Methods

Participants

Data from a convenience sample of 89 female athletes (Table 6.1) aged 18-45 from four separate experimental studies conducted in Australia are presented. Athlete performance/fitness status were tiered (Tier 0 = Sedentary; Tier 1 = Recreationally Active; Tier 2 = Trained/Developmental; Tier 3 = Highly Trained/National Level; Tier 4 = Elite/International Level; Tier 5 = World Class) according to

McKay et al. (2022b), as follows: n=10 Tier 2-3 cyclists/triathletes [Study 1, (Smith et al., 2024a)], n=22 Tier 2-3 National Rugby League Indigenous Women's Academy players [Study 2, (Smith et al., 2024c)], n=38 Tier 2 cyclists/triathletes [Study 3, (Smith et al., 2024b)], and n=19 Tier 3-5 race walkers [Study 4, (Kuikman et al., 2024a)]. The first study also included 10 male athletes whose data are included separately for between-sex comparisons only. Full participant information and study details, including ethical approval, can be found in the respective publications. All studies were conducted in accordance with the declaration of Helsinki, with separate ethical approval received for each of the four studies presented. Specifically, each study included explicit participant consent to complete the questionnaire presented in the present study, for which participants provided their written consent prior to enrolling.

All studies achieved a minimum of "silver" tier regarding the classification and control of menstrual status according to the tiering system outlined by Smith et al. (2022b), which classifies the extent to which best-practice guidelines (Elliott-Sale et al., 2021) have been adhered to. A "silver" tier reflects studies that achieve the majority of recommendations for best-practice, while a "gold" tier study would implement every methodological recommendation. Divergence from the "gold" tier methodological standards occurred around HC use in Studies 1, 2 and 4, where participants using HC were not restricted to a single contraceptive type (i.e., included a mixture of methods such as OCPs or IUS). Furthermore, Studies 2, 3 and 4 examined outcome measures across a single MC only, rather than repeating the outcomes across ≥two separate cycles to verify a consistent response as is recommended.

Questionnaire

Participants independently completed a voluntary electronic questionnaire [(REDCap (Harris et al., 2019; Harris et al., 2009)], distributed electronically upon study completion (Figure S6.1), without input from the principal investigator or data analysis team. Questions were separated into 3 category themes:

A. **Prior participation** in *other* research studies:

   Participants were asked if they had participated in research before, and if not, to state their primary reason for non-participation.

B. Reasons for **participating** in the *current* research study:

   Participants were asked to rate the degree to which different aspects influenced their decision to participate in the study on a visual analogue scale from 0-100, whereby "0" was "did not influence their decision at all" and "100" was "very much influenced their decision".

C. **Experiences** during the *current* research study:

   1) Participants were asked to rate each test completed during the study, whereby the *experience* of the test itself (e.g., the actual process of undertaking a blood test or tracking their MC) was differentiated from the *feedback* received post-study (e.g., blood test results).

Participants received individualized results following study completion (prior to full data analysis and publication). Importantly, these results were provided to participants prior to completing the questionnaire. Each test was rated out of 10 on a Likert scale, whereby "1" represented "the worst experience" and "10" was "the best experience".

2) Participants were asked to rate their overall experience on a visual analogue scale from 0-100, whereby "0" was "the worst study ever" and "100" was "the best study ever", alongside if they would apply the study findings to their sports involvement and if they would be willing to participate in future studies.

Statistical analysis

Statistical analyses were performed using R Studio (v3.5.2) with statistical significance accepted at an α level of $p \leq 0.05$. Only the survey responses from female participants (n=89) are reported across themes A-C. Males are included exclusively for the purpose of sub-analysis between sexes in Study 1 only. The statistical approach for each section is outlined, with normality assessed through histogram inspection. With the exception of the sub-analyses between participants of different athletic calibre, statistical comparisons were not conducted between studies due to the highly skewed and heterogeneous nature of the data, as well as the substantially different participant demands across the four studies.

Theme A (prior research participation) does not include statistical analysis and presents the proportion of participants who responded "yes" or "no" to prior research participation, alongside their primary reason for this response.

Theme B (reasons for current study participation) presents the median±interquartile range (IQR, due to skewed data) score for each participation reason. Due to non-normality of this variable, Spearman's rank correlation coefficient examined the relationship between perceived importance of prize money, and the actual prize money received in Study 3.

Theme C (1: study procedures) presents the median±IQR rating of each test conducted during the study, due to non-normal ordinal data. The experience of completing the test was compared to the corresponding feedback received post-study (described above) using a Wilcoxon Signed-Rank test (i.e., analysing within-participant differences for test feedback *vs* experience). Since each study procedure was conducted independently, with varying numbers of participants completing each test based on study requirements, tests were not combined into a single model. Only tests performed in more than one study are analysed, with participant responses for the same test pooled across studies. This approach was considered appropriate as all studies were performed in the same laboratory according to identical institutional protocols, ensuring comparability of the same test across studies (e.g., DXA scans in Studies 1, 2, and 4 were

conducted identically). The full data set, including all procedures separated by study, is available in the supplementary material (Figure S6.2).

Theme C (2: Overall study experiences) does not include statistics and presents the mean±SD for participants' overall rating of their research experience, due to normal distribution of this variable.

Sub-analyses: comparisons between sexes. Compares the responses of n=10 female and n=10 male participants in Study 1 only. Responses to Themes B and C(2) were normally distributed and therefore reported as mean±SD, with an independent t-test and Cohen's d effect size used to compare the perceived importance of participation reasons between the sexes. Theme C(1) is presented as median±IQR, due to non-normal ordinal data, with a Mann Whitney U test and Rank-Biserial Correlation (r_{rb}) to evaluate experiences of the test procedures between sexes.

Sub-analyses: comparisons between athletic tiers. The responses of highly trained athletes (n=19) in Study 4 were compared to those of lower-calibre athletes (Tier 2-3, n=70) in studies 1-3. Athletes in Study 4 were primarily Tier 4-5 (n=17), with n=2 classified as Tier 3. However, it was deemed appropriate to include these two athletes with the higher calibre group given they were part of the same high-level training environment. Due to aforementioned non-normality of Themes B and C(1), data are reported as median±IQR with comparisons between athletic tiers made using a Mann Whitney U test and Rank-Biserial Correlation (r_{rb}). Theme C(2), however, was normally distributed and is presented as mean±SD, with a Welch's test to account for unequal group sizes used to compare the overall research experience between tiers.

Table 6.1. Participant characteristics and study features.

		Participant characteristics				Study features		
	Reference	Participant number	Age	Sport	Athletic tier	Study design	Laboratory tests	Free living tests
Study 1	(Smith et al., 2024a)	10	33±7	Cycling (n=4) Triathlon (n=4) Running (n=2)	Tier 2 (n=8) Tier 3 (n=2)	Free living, five laboratory trials	DXA, force plate testing (CMJ/SJ/ IMTP), surveys, Stroop colour and word, blood tests, RMR, FATMAX, $\dot{V}O_2$max, MMTT, Wingate	HRV, sleep assessment
Study 2	(Smith et al., 2024c)	22	22±3	National Rugby League	Tier 3	Residential training camp, three laboratory visits	Drop jumps, DXA, force plate testing, 20 m sprint, blood tests, RMR	HRV, daily MC surveys, sleep assessment
Study 3	(Smith et al., 2024b)	38	35±6	Cycling (n=14) Triathlon (n=24)	Tier 2	Remote free living, no laboratory visits	Blood tests (at external pathology labs)	Zwift races, daily MC/ training tracking, pre/post-race questionnaires, food logging

Study 4	(Kuikman et al., 2024a)	19	27±7	Race walking	Tier 5 (n=1) Tier 4 (n=16) Tier 3 (n=2)	Residential training camp, multiple laboratory visits	DXA, $\dot{V}O_2$max, Haemoglobin mass, blood tests, daily MC surveys, RMR	10 km track race

DXA, dual-energy X-ray absorptiometry; *CMJ*, countermovement jump; *SJ*, squat jump; *IMTP*, isometric mid-thigh pull; *RMR,* resting metabolic rate; *FATMAX*; exercise intensity at which the maximal rate of fat oxidation occurs; *$\dot{V}O_2max$*, maximal rate of oxygen consumption ; *MMTT,* mixed meal tolerance test; *HRV*, heart rate variability; *MC*, menstrual cycle.

6.4 Results

Theme A: Prior research participation: Across all studies, 81% (n=72) of female athletes had never previously participated in research. The majority (79%, n=57) stated that this was primarily due to a lack of opportunities, with other reasons including being unaware of opportunities (14%), a lack of financial support (3%), time demands (1%), lack of interest (1%) and distance to the study location (1%).

Theme B: Reasons for current study participation: The research outcome was ranked as most important [98 (out of 100) ±17, Figure 6.1], followed by the opportunity to receive personalized results (96±27), and educational opportunities (88±32). Reasons ranked as less important were networking opportunities (47±90) and financial support via direct reimbursements (36±66). The five individuals who cited a lack of financial support as a reason for not participating in prior studies did not rank financial support higher (32±74) than the group median (36±66). Other study-specific aspects rated as important during Studies 2 and 4 included the training support, such as the presence of training partners, facilities, and coaches (96±20) alongside the living environment (72±39). The provision of food in Study 1 was not highly valued (29±53), while participants in Study 3 did not perceive opportunities to win prize money as important (22±63). Moreover, the correlation between actual amount of prize money received during Study 3, and the perceived importance of prize money was small-trivial and non-significant (r_s = 0.269, *p=0.102*).

Theme C (1): Study procedures: Assessments of heart rate variability [10 (out of 10) ±1], body composition (9±2), strength and power (9±4) alongside competitive races (9±3) were ranked among the most favoured tests to undertake, while sleep assessments (7.5±4) and measures of resting metabolic rate (7±4) were least popular (Figure 6.2). However, all tests rated at a seven or above. The feedback received was rated highest for heart rate variability (10±2), while the majority of tests (8 out of 9) were rated at a median of eight or above, with the feedback for surveys scoring the lowest (7±4). When comparing the experience of the test itself to the feedback received, the feedback was rated higher than the experience of the test itself for both blood sampling (*P=0.009*) and maximal exercise tests (*P=0.004*) while participants rated the feedback received from surveys less highly than the experience of completing the surveys themselves (*P=0.001*).

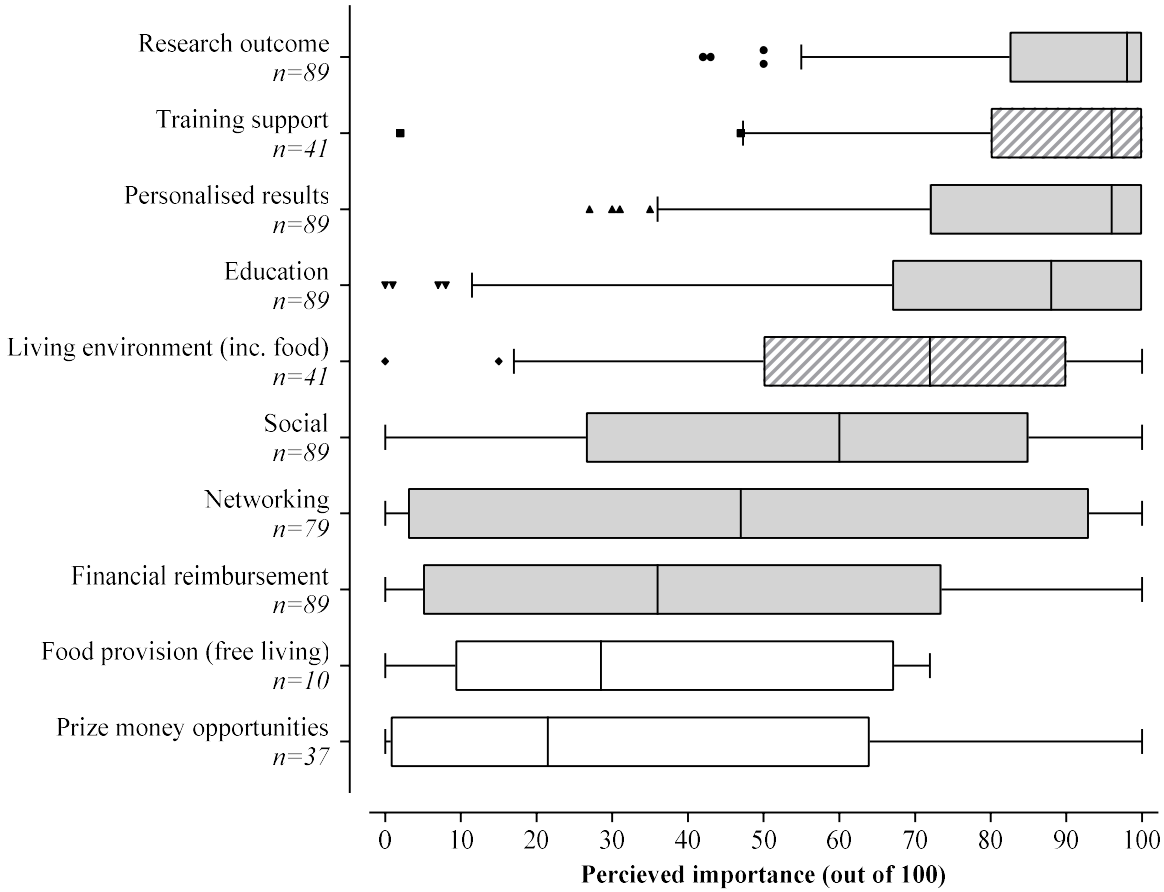


Figure 6.1. Reasons for study participation were rated between 0-100 according to their perceived importance across all four studies, whereby "0" represented that the reason "did not influence their decision to participate at all" and "100" was "very much influenced their decision". The box plot displays the median, first and third quartiles, with symbols displaying outliers outside of the 5-95th percentile. Data is presented for female athletes only, and the total participant number (out of a maximum of n=89) for each reason is provided on the y axis labels (as some aspects were not applicable in all studies). Clear bars represent study aspects only present in one out of four studies [food provision in Study 1 (Smith et al., 2024a) and prize money in Study 3 (Smith et al., 2024b)], dashed bars denote aspects present in two studies, and solid bars represent aspects present in three or four studies. "Living environment" applies to training camp studies where participants stayed on site and includes the accommodation and food provided, while "food provision" applies to food given to participants to eat in free living conditions. "Financial reimbursement" was given to all participants following study completion, whereas "prize money" was only won by certain participants based on their race performance in Study 3.

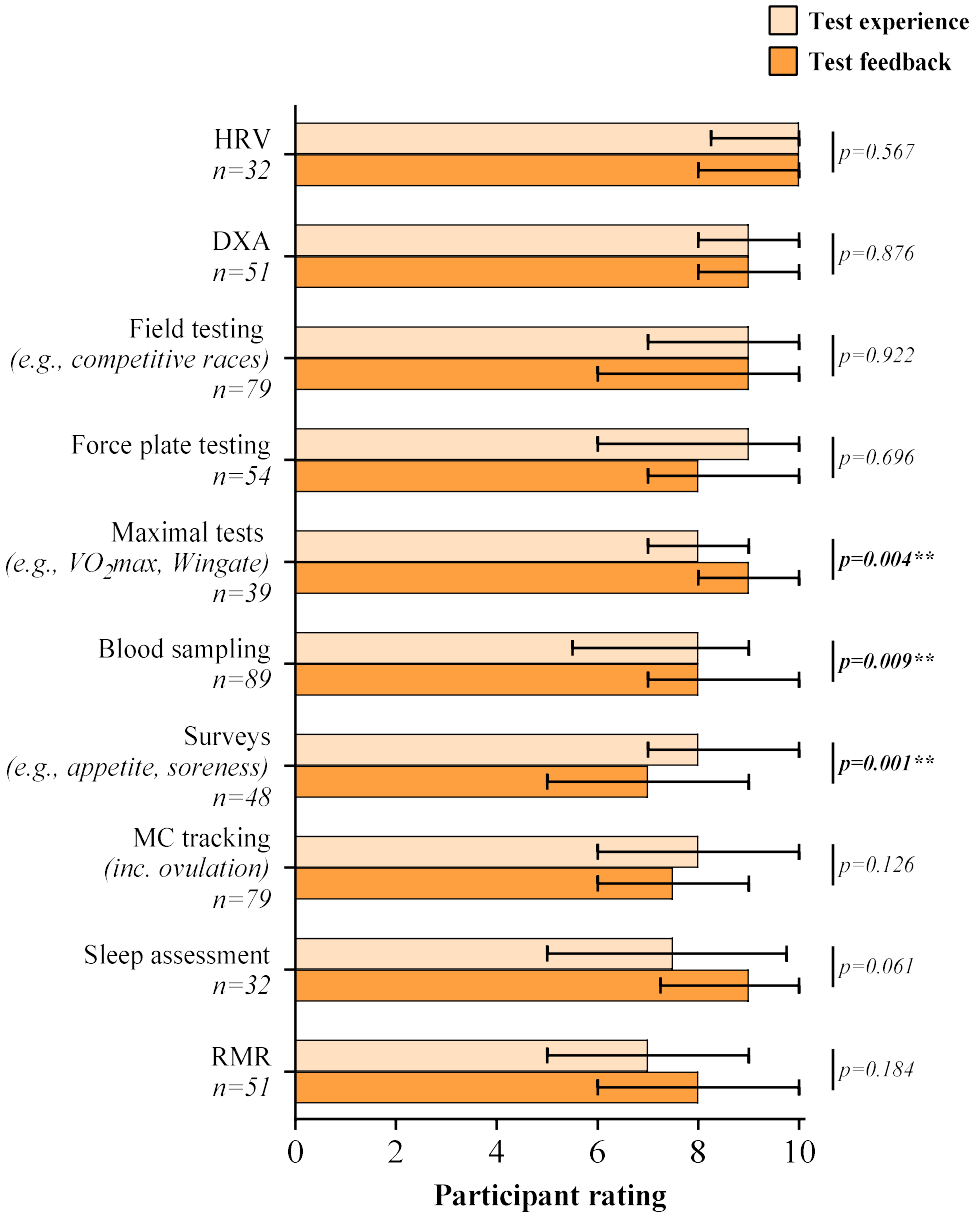


Figure 6.2. Participant rating of each test/procedure completed, whereby "1" represented "the worst experience" and "10" was "the best experience". The "test experience" refers to the experience of the test itself (e.g., the actual process of undertaking a blood test or using the at-home device to measure sleep), while "test feedback" refers to the individual feedback that participants received post-study (e.g., blood test/ sleep summary results). Data presented as the median, with error bars denoting the interquartile range. Data is presented for female athletes only with the total participant number (out of a maximum of n=89) for each test provided on the y axis labels, as most tests were not present in all studies. Wilcoxon signed-rank tests compared within-participant ratings of the test experience *vs* feedback. **denotes $p<0.01$. *HRV*; heart rate variability, *DXA*; dual x-ray absorptiometry, *MC*; menstrual cycle, *RMR;* resting metabolic rate.

Theme C (2): Overall study experiences: Participants had an overall positive experience, rated as 77±16 out of 100. The majority (94%, n=84) of participants indicated a willingness to participate in future studies, with reasons cited for non-participation including time commitments, retirement from elite sport, and dislike of a particular study aspect (blood tests and questionnaires). Moreover, most participants (87%, n=77) stated that they would apply the study findings to their sports involvement, while the remaining 13% reported that they required additional information to understand the application of results. Of the 12 participants who would not apply study findings to their sport, only one stated that they would not participate in future studies.

Sub-analyses: comparisons between sexes: Five out of ten male athletes had previously participated in research, in comparison to the two out of ten female participants. Non-participation was due to a perceived lack of opportunity across both sexes. Similarly, both sexes cited the main motivations for their participation was to receive personalized results [men, 82±15 (out of 100); women 90±11] alongside an interest in the research outcome (men, 76±21; women 87±19), while both rated as the least important motivators as food provision (men, 34±21; women 35±28) and social aspects (men, 37±26; women 44±28, Figure 6.3). There was no different between sexes ($p>0.05$) for their rating of any participation reason, although there was a large effect size for the difference between how the sexes rated the importance personalized results ($p=0.187$, d=0.61) alongside the research outcome ($p=0.256$, d=0.52). The overall research experience was identical between sexes, rated as 81±15 out of 100 by male, and 82±14 by female participants ($p=0.810$, d=0.11). Female athletes reported that the research was more time demanding than anticipated, but the difference with men (where demands matched expectations) did not reach significance ($p=0.067$, d=0.87). Physical and mental demands were similar for both sexes (rated as close to what was expected). All male participants stated that they would participate in research again, while one female participant reported that time commitments would prevent future participation, and another would retire from elite sport. There were no differences between sexes in their rating of the study procedures and feedback received (all $p>0.05$). All female participants stated they would apply the study findings to their training, whereas this was the case for only seven of the male participants.

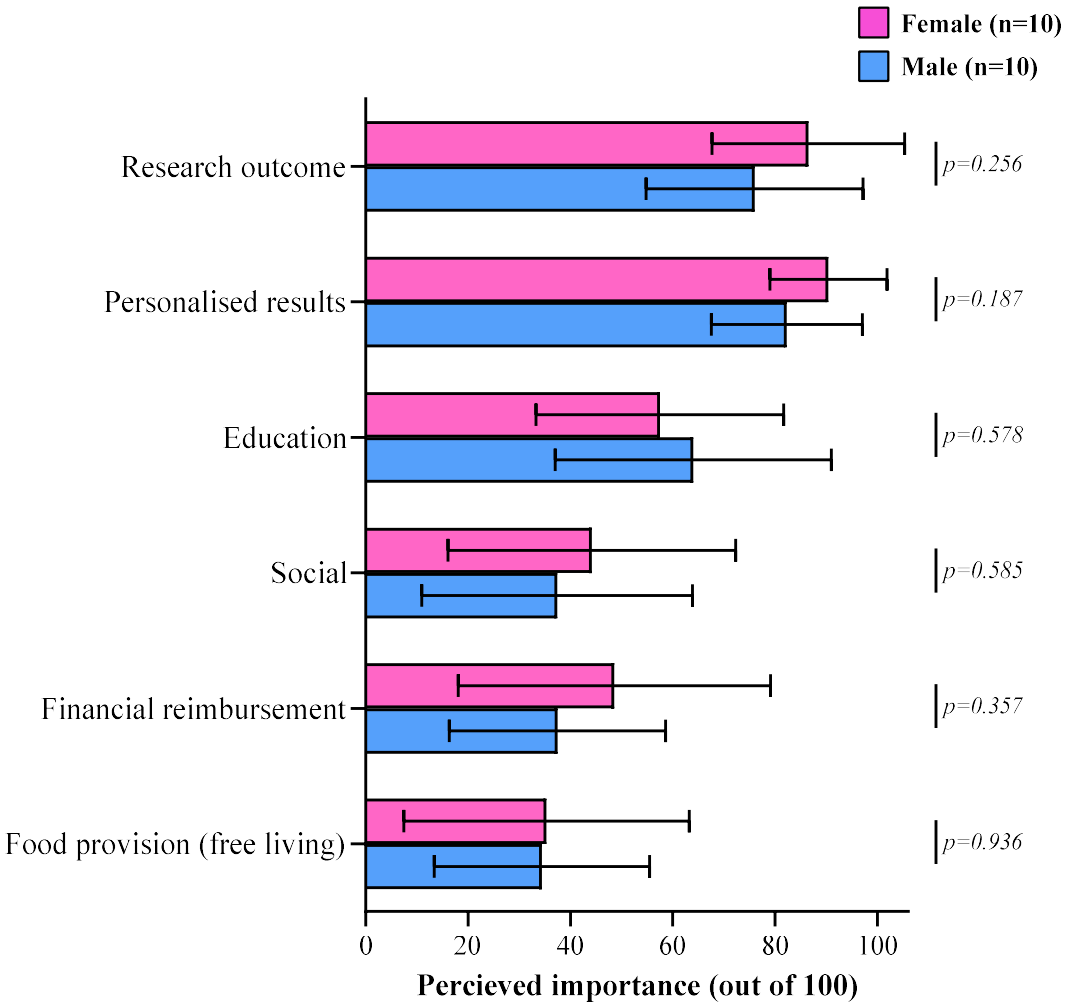


Figure 6.3. Participants' perceived importance of differing reasons for study participation, separated according to males and females in Study 1 only (n=20, 10M, 10F). Reasons for study participation were rated between 0-100 according to their perceived importance, whereby "0" represented that the reason "did not influence their decision to participate at all" and "100" was "very much influenced their decision". Data presented as mean±standard deviation. Independent t-tests compared the perceived importance of each distinct reason between males and females.

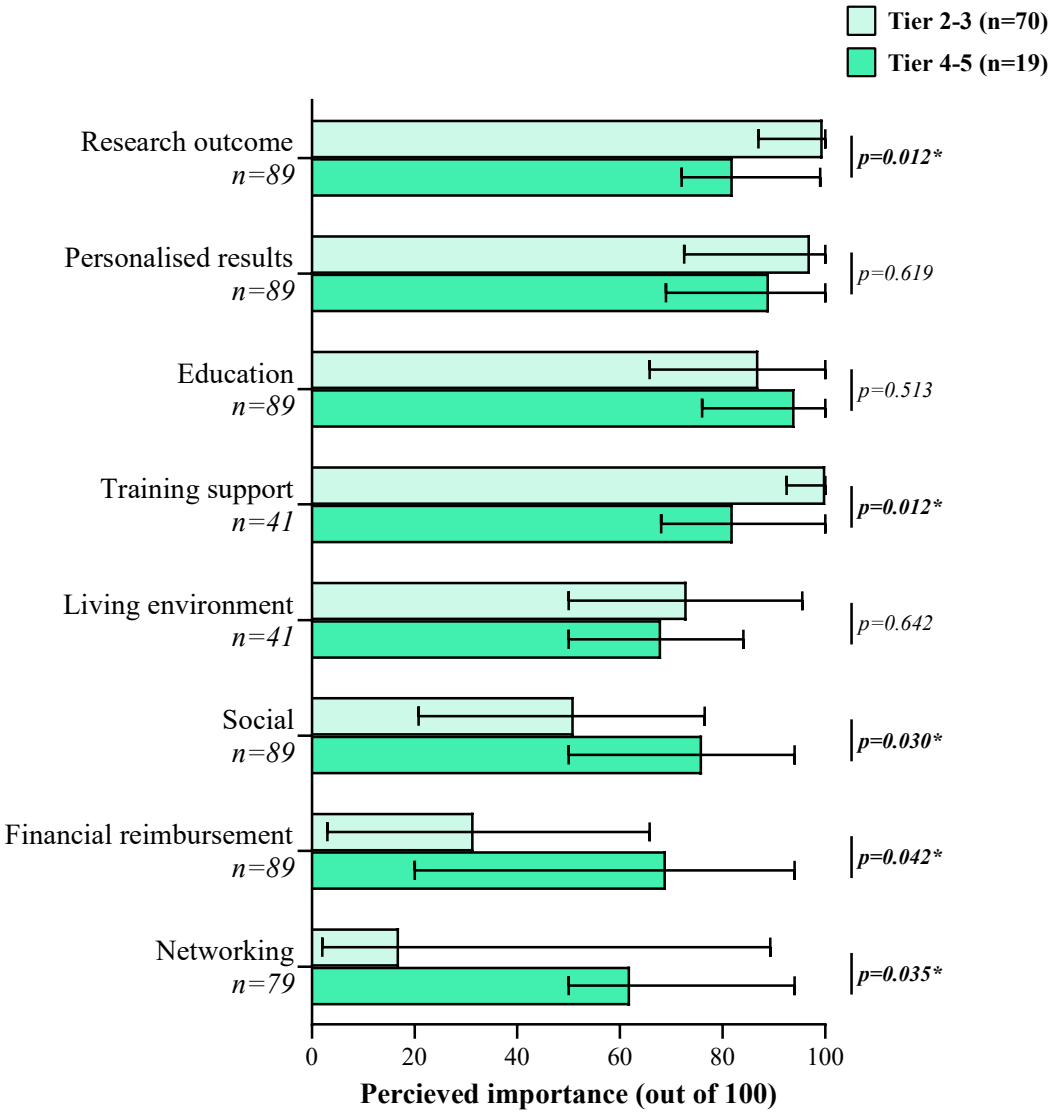


Figure 6.4. Participants' perceived importance of differing reasons for study participation, separated according to athletic tier as defined by (McKay et al., 2022b) (n=70 Tier 2-3, n=19 Tier 4-5). Reasons for study participation were rated between 0-100 according to their perceived importance across all four studies, whereby "0" represented that the reason "did not influence their decision to participate at all" and "100" was "very much influenced their decision". Data is presented for female athletes only, the total participant number for each reason is provided on the y axis labels [n=19 Tier 4-5 athletes is consistent across all reasons, while the number of Tier 2-3 athletes varied (out of a maximum of n=70) as some reasons were not applicable in all studies]. Data presented as the median, with error bars denoting the interquartile range. Mann Whitney U tests compared the perceived importance of each distinct reason between athletic tiers. *denotes $p<0.05$.

Sub-analyses: comparisons between athletic tiers: Responses of the more highly trained athletes in Study 4 (Tiers 4-5, n=19) were compared to those less highly trained across Studies 1, 2 and 3 (Tiers 2-3, n=70). A similar proportion had never previously participated in research [84% (n=16) of Tier 4-5 and 80% (n=56) of Tier 2-3], with both groups citing the primary reasons for non-participation as a perceived lack of opportunity. Regarding reasons for current study participation, Tier 2-3 rated an interest in the research outcome [100±13 (out of 100)] and training support (100±7) as their most important reasons, while for Tier 4-5 athletes these were educational opportunities (94±24) and receiving personalized results (89±28, Figure 6.4). Tier 4-5 athletes placed more importance on social aspects (76±38 vs 51±54, *P=0.030,* r_{rb}=0.33), networking opportunities (62±44 vs 17±82, *P=0.035,* r_{rb}=0.32), and financial support (69±70 vs 32±60, *P=0.042,* r_{rb}=0.31), but valued the research outcome (82±25 vs 100±13, *P=0.012,* r_{rb}=0.36) and training support (82±28 vs 100±7, *P=0.012,* r_{rb}=0.45) less highly than Tier 2-3 athletes. Participant ratings of study procedures did not differ between tiers (*p>0.05*), with the exception of maximal exercise tests for which the experience was rated higher by Tier 4-5 [9±1.5 (out of 10)] compared to Tier 2-3 athletes (7±2, *P<0.001,* r_{rb}=0.65). Overall study experience did not differ between Tier 2-3 (78±16 out of 100) and Tier 4-5 (73±18, *P=0.340,* d=0.28) athletes, while a similar proportion stated that they would apply the research findings to their sport [87% (n=61) of Tier 2-3 and 84% (n=16) of Tier 4-5] and would participate in future research [94% (n=66) of Tier 2-3 and 95% (n=18) of Tier 4-5].

6.5 Discussion

Our findings confirm the underrepresentation of women in SES research, with only 19% of participants across all four studies having previously participated in research. The experiences and reflections of the participants in these four studies highlights a *perceived* lack of opportunity for female athletes to engage in research as the primary barrier to their participation, reinforcing a sentiment well-established across the literature (Bruinvels et al., 2017; Costello et al., 2014; Cowley et al., 2021). The main motivations for female athletes to participate were an interest in the research outcomes and to receive personalized results. Encouragingly, participants had an overall positive experience of their study involvement, rating their experience 77 out of 100, with 94% stating that they would participate in future research. Taken together, our findings in this subset of female athletes suggest that, despite the barriers to participation, once involved as research participants, female athletes report positive experiences and are willing to participate in future.

The positive experiences of the female athletes in these four studies may relate to the high percentage of female researchers involved in both the investigative (on-site researchers) and authorship teams, as evidence suggests women participate more frequently in SES research when women are involved in the investigative team (James et al., 2024). Indeed, women were first authors and lead investigators across

all studies and accounted for 81±17% of the on-site investigative team and 75±8% of authorship teams across studies. Moreover, as the main motivations for participation were an interest in the research outcome and to receive personalized results, the high overall median rating of test feedback [8±4 (out of 10)] may also have contributed to a positive experience. Interestingly, there was some discord where tests rated lower for experience almost always received higher ratings for feedback. This may suggest that less pleasurable more invasive tests, such as blood sampling and maximal exercise tests, yielded the most valued results for participants. Alternatively, these invasive tests are perhaps not readily available to the athletes as part of their normal training regime and are thereby valued higher once provided during a research study.

The inclusion of male participants in Study 1 allows between-sex comparisons to be made, albeit with recognition of the small participant numbers (n=20, 10 female and 10 male athletes). Reasons for prior non-participation in research was the same between sexes: a perceived lack of opportunity. However, five out of ten male athletes had previously participated in research, in comparison to just two out of ten female participants. There was no difference between the sexes in their reasons for participation, perception of the research experience, nor how they rated the study procedures and feedback received. On the other hand, female participants reported that the research was more time-demanding than anticipated, and although the difference with men did not reach statistical significance the effect size was large (*p=0.067,* d=0.87). We speculate that this difference might be explained by the need for female participants to be tested during the active phase of their oral contraceptive pill cycle, thus extending the total study experience across a longer timeframe (due to longer washout periods) than was necessary for men, although the total number of study hours was identical between sexes.

When considering athletes of differing athletic tiers [Tiers 2-3 (trained/national level) compared to Tiers 4-5 (national/international/world class) (McKay et al., 2022b)], a similar proportion had never previously participated in research, with the primary reason for non-participation once again being a perceived lack of opportunity. Tier 2-3 athletes were similar to those of Tier 4-5 in their positive rating of the overall study experience and willingness to participate in future research. Interestingly, Tier 4-5 athletes cited their primary participation reasons as educational opportunities and to receive personalized results, while Tier 2-3 athletes valued training support and an interest in the research outcome most highly. Moreover, Tier 4-5 athletes placed more importance on social aspects, networking opportunities and financial support, but valued the training support and research outcome less highly than Tier 2-3 athletes. The higher calibre athletes also rated their experience of maximal exercise testing (such as $\dot{V}O_2$max or Wingate tests) higher than lower calibre athletes, likely reflecting the Tier 4-5 cohort's greater familiarity with maximal exercise protocols.

When comparing athletes across different tiers, differences in participant characteristics must also be considered (Table 1). Participants in Study 2 were Indigenous Women's Academy players (Smith et al., 2024c). This cohort may face unique barriers to research participation, beyond the scope of what was evaluated. Indeed, of the 43 athletes who initially consented to participate, 42% (n=18) dropped out prior to commencing the study due to lifestyle and financial challenges (McKay et al., 2024). Moreover, while female athletes are underrepresented in the literature, Aboriginal and Torres Strait Islander athlete populations are even less well studied and hence this presents an area for future research. Study 4 also included an international cohort of Tier 4-5 athletes, representing 12 different countries across 19 participants. As such, cultural and language differences in these cohorts may have resulted in different study experiences, and hence should be considered when interpreting the differences between athletic tiers.

Limitations and future research

These findings represent the experiences of a small subset of female athletes aged 18-45 years and may differ from those across different ages, while our cohort of elite (Tiers 4-5) athletes is small (n=19). Moreover, this convenience sample comprises participants already enrolled in research and therefore does not represent the views and opinions of those who do not take part in research. The questionnaire used was also novel, and although tested for clarity and understanding, was not formally piloted or validated. This initial exploratory analysis may therefore inform a subsequent more systematic and detailed exploration of the area. To reduce participant burden, the questionnaire was completed following study completion. Ideally, questions regarding prior study participation (Category A) would have been completed prior to commencing our research study to minimize any influence of their experiences on responses. However, this approach was consistent across all studies. The questionnaire was also distributed to participants after they had received their initial personalized results, but prior to the publication of full study findings. This approach was chosen to maximize the response rate, given the likelihood of a lower retention rate with a lengthy follow up time. Responses may have differed if participants received the questionnaire after the full study results were available. Separately, additional information, such as participant training age, alongside motherhood and job status, would have facilitated a more nuanced interpretation of prior research participation among our female participants.

6.6 Conclusions

Our findings suggest that of the female athletes who participate in research the majority (81%) had never previously participated, citing a perceived a lack of opportunities as the primary barrier. However, despite these barriers, once involved as research participants the subset of female athletes investigated in this study reported broadly positive experiences and the majority (94%) were willing to participate in future. The main motivations for female athletes to participate were an interest in the research

outcome and receiving personalized results. As such, it may be prudent to focus future recruitment efforts on emphasizing these study aspects, for example highlighting that participation will contribute to research and outlining what personalized results will be received. Including more detailed results and interpretation may also be important in retaining female participants for future studies. Ultimately, our findings demonstrate that a perceived lack of opportunity is the primary barrier to female athlete research participation. As such, opportunities for women to participate in high quality studies should be prioritized.

Chapter 7: Discussion and Conclusions

The call to arms to close the sex- and gender-gap in SES research has resulted in a welcomed surge in the number of studies including female athletes in recent years. However, as a consequence of the unfortunate failure to adopt a high-quality approach to the menstrual status of participants, the goal of formulating robust conclusions about the effect of menstrual status and phase on important SES issues remains unachieved. There are a number of guidelines and comprehensive methodological recommendations for the classification and control of participant menstrual status (Elliott-Sale et al., 2021; Janse de Jonge et al., 2019). Although the challenge of achieving a rigorous level of control is acknowledged, particularly in applied SES settings, the current doctoral program involved the implementation of these methodological guidelines throughout different study designs involving female athletes. This chapter discusses the practicalities of implementing such guidelines (both in the laboratory and in the field) and our applied experiences working with female athletes with varying menstrual statuses. The balance between methodological rigour, time and cost restraints, and the generalisability of findings is discussed. Secondly, resources have been produced to aid researchers in the future pursuit of high-quality research including women.

7.1 A tiering system to evaluate menstrual status classification and control in research

An adjunct activity to the work presented in this thesis involved the creation of a standardised protocol and research tools to "audit" the representation of female athletes across different sub disciplines of SES research (Smith et al., 2022b). The goal was to allow our research group, and others, to systematically examine the quantity and quality of the representation of female athletes in the literature which informs evidence-based guidelines across different areas of SES. This methodology was designed to create a gap analysis of themes in which there is little information/representation of women and/or alongside research areas with the greatest scope for development or impact. The methodological classification and control of participant menstrual status in studies including pre-menopausal female participants represented a key focus within the audit process. A tiered ranking system, based on best-practice guidelines (Elliott-Sale et al., 2021) was therefore devised to assess the quality of a study's methodological control relating to ovarian hormonal profiles, with studies ranked as either Gold/Silver/Bronze or ungraded (Smith et al., 2022b). In addition to its role in retrospectively examining the methodological quality of existing research, this tool can be used prospectively to ensure that planned research adopts an appropriate classification and control of the menstrual status of its participants. The specific methods required to achieve each tier are detailed in Table 7.1.

Table 7.1. Methodological requirements for each tier, separated for athletes using hormonal contraception (HC) and those not using HC.

Participants not using hormonal contraception (including those using a copper intrauterine device)						
	During the study				**Prior to study commencement**	
	Calendar counting	Ovulation testing	Hormonal measures	Repeated outcomes	Prior MC tracking	Prior HC use
BRONZE	✓	**X**	**X**	No (1 cycle only)	None	Not stated
SILVER	✓	✓	**X**	No (1 cycle only)	1 month	Not stated, or <3 months prior
GOLD	✓	✓*	✓	Yes (≥2 cycles)	2 months	>3 months prior

Participants using hormonal contraception						
	HC information			**Testing schedule**	**Participant group**	
	(1) Length of usage	(2) HC type: *OCP (mono, bi, or triphasic; combined or progesterone only) or other HC (implants, injections, non-copper based intrauterine devices/coils, vaginal rings, transdermal patch)*	(3) formulation *(name and concentration of exogenous hormones)*	Active/inactive days considered (OCP only)	Number of HC types [and brand (OCP only)]	
BRONZE		1 of 3 stated		**X**	>1 per group	
SILVER		2 of 3 stated		Optional (advised)	≥1 per group	
GOLD	≥3 months	stated	stated	✓	1 per group	

**gold standard can be achieved with or without ovulation testing, as blood sampling can detect the pre-ovulatory oestrogen surge. MC; menstrual cycle, HC; hormonal contraception, OCP; oral contraceptive pill.*

7.2 Implementing the tiering system

Figure 7.1 outlines a step-by-step guide to implementing the tiering system during study planning and participant recruitment to ensure the appropriate methodological classification and control of participant menstrual status. Importantly, this is implemented separately for female participants using HC and those not using HC, even within a single study.

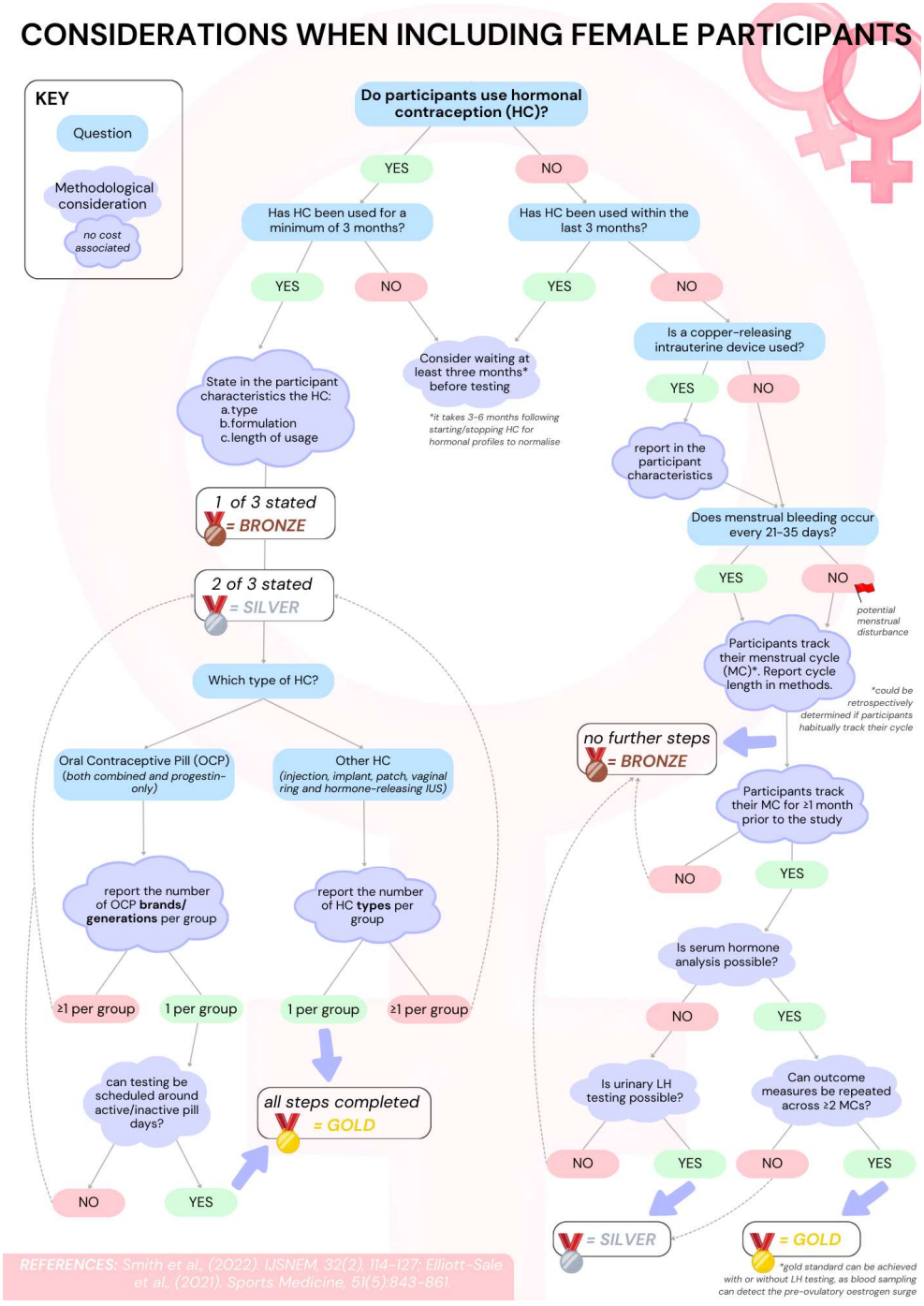


Figure 7.1. Flowchart detailing the process of implementing each tier of methodological control.

For participants not using HC, there are several available methods to achieve the three aspects of MC tracking outlined in Table 7.1. A detailed analysis of available methods is beyond the scope of this thesis but should be considered when implementing the tiering system.

1. **Calendar counting** of the MC can be achieved through a simple pen-and-paper diary, or electronically through mobile applications. Monitoring can occur in real-time whereby the participant records information daily, or retrospectively with participants recalling information about several past MCs either through a survey or interview. The issue of recall bias should however be noted when considering a retrospective approach.
2. **Ovulation testing** is most commonly achieved through proxy measures: real-time urinary LH testing, or retrospective analysis of serum oestrogen/progesterone concentrations. Other less thoroughly validated methods include the assessment of BBT, either through oral thermometers or more frequently via wearable technology such as smart watches that measure skin temperature or intra-vaginal rings. It should be noted that an ovarian ultrasound is the gold standard for ovulation detection, but this is usually unattainable in research. Therefore, the term "ovulation" will be used to describe when the pre-ovulatory LH surge has been detected, and then retrospectively confirmed with (Gold) or without (Silver) a corresponding oestradiol concentration (hence ovulation is presumed to have occurred).
3. **Hormonal measures** are currently recommended to be obtained through serum measures of oestradiol and progesterone. Analysis is usually done in batch following study completion in a retrospective and cost-effective manner but can also be done real-time for immediate verification of menstrual phase. Salivary and urinary hormone measures also exist, but current evidence is insufficient to support their accuracy comparative to blood measures.

Once the tier/methods have been implemented, it is important to align the language used within any study descriptions with appropriate interpretation of findings. This is of particular relevance when investigating participants with a natural MC (i.e., not using HC). The following section outlines the appropriate language regarding participant menstrual status, the identification of menstrual disturbances within the participant cohort, and the power of each tier to identify MC phases. Table 7.2 summarises the ability of each tier across these three areas.

Table 7.2. The ability to classify participant menstrual status in participants not using hormonal contraception, based on the tier of methodological control utilised.

	Menstrual status			Detection of menstrual disturbances				Detection of MC phases		
	Naturally menstruating	Naturally menstruating + ovulation	Eumenorrhea	Amenorrhea/ cycle length irregularity	Short luteal phase	Luteal phase deficiency	Anovulation	Bleeding and non-bleeding	Follicular and luteal	MC phases 1 to 4
BRONZE	✓	X	X	✓	X	X	X	✓	X	X
SILVER	✓	✓	X	✓	✓	X	✓	✓	✓	X
GOLD	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

MC; menstrual cycle.

Menstrual status

- Only the gold tier (or silver tier with the addition of serum oestrogen and progesterone measurements) can identify ***eumenorrhea.*** Because confirmation of eumenorrhea requires serum hormone measurements, this term cannot be used unless gold (or silver as outlined above) tier is achieved.
- Bronze tier can identify ***naturally menstruating*** individuals through determining cycle length, while silver tier is required to determine ***ovulation*** (i.e., individuals who are naturally menstruating with ovulation).
- If insufficient detail is provided to award gold, silver, or bronze, then the participants should be reported as "not using HC, with an otherwise unknown menstrual status". However, this approach is discouraged, noting that there is no additional cost to implementing the bronze tier.

Menstrual disturbances

- All tiers can identify ***amenorrhea*** or cycle-length irregularities, via calendar counting.
- ***Anovulation*** or a ***short luteal phase*** (onset of menstruation within 10 days of ovulation) can only be detected with the silver and gold tiers which measure ovulation.
- Only gold tier (or silver with the addition of serum oestrogen and progesterone measures) is able to identify a ***luteal phase deficiency*** [cycles with less than 16 $nmol \cdot L^{-1}$ of progesterone, when a single luteal phase progesterone measurement is taken (Elliott-Sale et al., 2021)], as this requires the measurement of serum hormone concentrations.

Menstrual cycle phases

- Only a gold tier method (or silver with the addition of serum oestrogen and progesterone measurements) can detect menstrual ***phases*** (1-4). This means that specific MC phases cannot be referred to in studies with silver and bronze tier methods (e.g. "testing was conducted in MC phases 1 and 4") because they have not been verified in the protocol.
- A silver approach allows the identification of ***follicular*** and ***luteal*** phases (as ovulation is identified).
- A bronze tier can only distinguish between bleeding and non-bleeding timepoints, because calendar counting alone is insufficient to determine ovulation. Although methods using this approach might state that testing was conducted on days X and Y of the cycle, assertions identifying the phase of the MC in which these have occurred cannot be made.

It is also possible to add measurement of serum oestrogen and progesterone to the silver tier, to determine eumenorrhea, identify MC phases, or detect a luteal phase deficiency. However, the achievement of gold tier requires the outcome measures to be repeated across a second MC. Separately,

although not included in the current tiering system, emerging evidence suggests that symptoms (positive or negative) associated with the MC or HC use may influence various outcome measures such as performance, independent of hormonal fluctuations (Giacomoni et al., 2000; Smith et al., 2024b). It may therefore be beneficial to also include a measure of daily symptom incidence and severity. Although this adds a small participant burden, it does not entail additional cost and can be easily incorporated into existing daily monitoring. However, the lack of validated or standardised menstrual symptom questionnaires is noted. Supplementary material S7.1 provides an example pre-screening questionnaire that can be completed by participants to provide researchers with sufficient menstrual status information to achieve the "silver" tier for participants using HC, or "bronze" for those not using HC.

7.3 Selecting the appropriate tier

The appropriate level of methodological rigour for a particular study is dependent upon numerous factors, including the research question, resources and participant recruitment. While a gold tier approach provides the highest internal validity, the resource-intensive nature of this methodology is acknowledged. Therefore, the following section provides guidance to maximise the methodological quality, irrespective of study resources. The costs and benefits of each tier are outlined in Table 7.3. For example, only a gold tier method (or silver with the addition of serum oestrogen and progesterone measurements) has the power to identify menstrual phases (1-4), and as such, if the research question examines menstrual phases as the independent variable, then this classification approach must be used. However, a bronze tier may be appropriate for studies not directly examining ovarian hormones, or where robust evidence exists that ovarian hormones do not influence the outcome measure. For example, a study examining an athletes' perception of nutritional practices is unlikely to benefit from retrospective serum hormone analysis. However, if menstrual irregularities may be confounding to the outcome measures, a higher tier may be necessary to identify these. Indeed, Noordhof et al. (2024) demonstrated that the silver tier detected only 61% of subtle menstrual disturbances such as a luteal phase deficiency.

Table 7.3. The cost-benefit of each tier of menstrual status classification and control.

	Internal Validity	Cost	Time/ resources	Participant burden
BRONZE	↓	↓	↓	↓
SILVER	↔	↔	↔	↔
GOLD	↑	↑	↑	↑

↑ *higher,* ↓*lower,* ↔ *intermediate.*

These guidelines not only apply to research examining the female sex hormones as the independent variable. We suggest that all research including women achieve a minimum of bronze tier, as this method does not entail any additional cost (i.e., reporting the HC type/formulation/usage length for participants using HC, or MC tracking via calendar counting for those not using HC) and can be achieved through completing the participant screening questionnaire (Figures S7.1 and S7.2). Even if not perceived directly relevant to the research question, or current evidence indicates that oestrogen/progesterone do not influence a particular outcome measure, this additional information may provide useful context when interpreting study findings in the future. For example, future research may uncover a currently unknown effect of oestrogen/ progesterone on a body system. Achieving a minimum of a bronze tier facilitates the appropriate future interpretation of findings in context. Furthermore, although a gold standard approach provides the highest internal validity, transparent reporting of participant menstrual status(es) is important, regardless of the tier achieved. For example, if the research question is not related to MC phases, and simply looks to include women in the cohort to improve generalisability, then the transparent reporting of participant menstrual characteristics (i.e., bronze tier) can still be considered a high-quality approach.

7.4 Selecting a participant cohort

Some consideration is required during the planning phase to decide on the appropriate population to study in terms of menstrual status.

Broadly, there are three options for study cohorts, each of which was included during this thesis:

1. Participants using HC
2. Participants not using HC
3. Mixed: a combination of those using HC and those not using HC

Considerations regarding statistical power, the number of experimental conditions, and the time and resources available to conduct the study all contribute to the decision regarding the most appropriate study cohort. For example, in scenarios where the outcome measure is sensitive, or multiple experimental conditions are planned, the goal of detecting small but real changes may be best addressed by choosing a cohort using HC to minimise the additional noise created by fluctuations in ovarian hormones. In contrast, when a larger magnitude of response is anticipated, the study may select a more heterogeneous cohort, including women not taking HC, to enhance generalisability. Some considerations for participant recruitment are outlined in Table 7.4. The prevalence of HC use in the target population may also guide this decision, both in terms of the generalisability of findings as well as the size of the potential participant pool. Here it is noted that HC-using female athletes are currently underrepresented in the SES literature (D'Souza et al., 2023), particularly those using methods other than OCPs (Flood et al., 2024).

Table 7.4. Considerations when selecting a participant cohort based on menstrual status.

Participant menstrual status	Benefits	Costs
1 – Participants using HC	A lower budget is required to achieve a gold standard as blood tests and urinary ovulation are not required. Likely quicker to conduct a study in this cohort as outcome measures do not need to be repeated across a second HC cycle.	Given the wide variety of available HC methods, achieving a homogenous group as required for a gold tier is challenging.
2 – Participants NOT using HC	Can answer research questions examining the effect of oestrogen and progesterone as the independent variable.	Costly, labour-intensive and time demanding to achieve a gold or silver standard due to the need for blood tests and/or urinary ovulation, alongside repeating outcome measures in ≥two MCs to achieve a gold tier.
3 – Mixed: both participants using HC and not using HC	The most generalisable approach, reflective of the wider female population. Can compare findings directly between individuals using HC and those not (if statistical power is achieved).	Methodological considerations for ovarian hormone concentration should be made specific to the menstrual status of each female population in the cohort and is therefore more time-intensive than using just a single cohort. This will also likely increase the required sample size to achieve statistical power if comparing between groups.

MC, menstrual cycle; *HC,* hormonal contraception; *OCP*, oral contraceptive pill.

7.4 Discussion of studies in this thesis

The studies in this thesis aimed to examine a range of methodologies and menstrual statuses in the applied setting. As such, each of the three experimental studies employed a different methodological design and recruited participants of differing menstrual statuses, to provide a broad range and depth of experience (Table 7.5). The following sections will discuss the scientific outcomes of each study, with a focus on the strengths and limitations of various methods used to control for ovarian hormones.

Study 1 – Effects of 24-hour diet- or exercise-induced energy availability manipulations on substrate utilisation and performance

Approach to the research question

Athletes commonly experience a reduction in EA (the mismatch between dietary EI and EEE, relative to FFM) to facilitate performance goals, such as meeting weigh-in targets or achieving optimal competition physique characteristics. The 2023 International Olympic Committee's consensus statement on REDs (Mountjoy et al., 2023) introduced the concepts of "adaptable" (acute and transient) *versus* "problematic" (chronic) LEA. While problematic LEA is associated with negative health and performance outcomes, that may result in REDs (Mountjoy et al., 2023), adaptable LEA may cause transient and minor metabolic adjustments and/or be associated with performance benefits. However, the implications of acute LEA on strength/power outcomes are unknown, as are the potentially divergent effects associated with the method of reducing EA (i.e., dietary restriction or increased EEE). Any moderating influence of sex on the response to acute LEA also remains unexplored. This study, therefore, sought to examine sex-based differences in substrate oxidation, metabolism, and performance in response to a 24-hour period of LEA, induced separately by decreasing EI or increasing EEE. Five different conditions of EA were investigated in a randomised order in a Latin square design, each comprising a three-day trial, totalling 15 trial days per participant. The primary outcome was to assess sex differences in PFO during fasted submaximal exercise the morning following 24 h of altered EA. It was calculated that 20 participants (10 females, 10 males) were required to detect differences in peak PFO relative to FFM between the sexes, with 90% statistical power. A rolling-recruitment strategy was employed, with 20 participants completing data collection across an eight-month period.

Participant cohort

PFO during submaximal cycling exercise was identified as the primary outcome variable of this novel research study. When selecting the appropriate female cohort, we noted that ovarian hormones may influence pathways involved in substrate oxidation (Bunt, 1990; D'Eon et al., 2002; Nicklas et al., 1989; Oosthuyse & Bosch, 2012; Ruby et al., 1997) and may therefore influence this outcome.

Table 7.5. Worked examples of the menstrual status tiering system outlined above, implemented for each study conducted during this thesis.

	Study 1	**Study 2**		**Study 3**
Menstrual status	Monophasic oral contraceptive pills	Both non-hormonal contraceptive and hormonal contraceptive users		Non-hormonal contraceptive users
		HC users	*Non-HC users*	
Tier achieved	Silver	Silver	Silver *(+serum oestradiol and progesterone measures)*	Silver *(+ serum oestradiol and progesterone measures)*
Methods implemented to achieve tier	– Stated: the length of HC usage (>3 months), HC type and formulation. – Participants were tested during active HC usage only (i.e., pill taking days, avoiding the withdrawal bleed) and instructed to take their pill at the same time of day on each testing occasion.	– Stated: the length of HC usage (>3 months), HC type and formulation. – Participants were tested during active HC usage only (i.e., pill taking days, avoiding the withdrawal bleed) and instructed to take their pill at the same time of day on each testing occasion.	– 16 weeks of MC tracking. – Urinary ovulation testing. – Blood tests to assess serum 17-β-oestradiol and progesterone. – No HC use for >3 months prior to study commencement.	– 8-10 weeks of MC tracking. – Urinary ovulation testing. – Blood tests to assess serum 17-β-oestradiol and progesterone. – No HC use for >3 months prior to study commencement.
Reason gold tier was not achieved (study limitation)	Participants were not restricted to a single brand/generation/formulation of monophasic oral contraceptive pills.	Numerous HC types within the participant cohort.	Outcome measures were only assessed across one MC (i.e., not repeated in a second cycle).	Outcome measures were only assessed across one MC (i.e., not repeated in a second cycle).
Likely impact of the study limitation	Increases between participant variability in hormone status which may influence study findings.	Increases between participant variability in hormone status which may influence study findings.	Increases the variability of the data as a consistent response between MCs cannot be confirmed.	Increases the variability of the data as a consistent response between MCs cannot be confirmed.

MC, menstrual cycle; *HC,* hormonal contraception.

Moreover, other outcome measures spanned numerous body systems with unknown responses to varying concentrations of oestrogen and progesterone. When planning the tight control of the prescribed EI and EEE for five specific conditions within this complex study design, we identified the importance of minimising fluctuations in ovarian hormones to match this intended rigour. Furthermore, we recognised the logistical challenge of recruiting and testing participants with a natural MC in a specific phase for each of the five experimental conditions, with the considerations that other aspects such as a training effect may influence outcome measures if testing was conducted across a five-month time period. Accordingly, athletes using OCPs were selected as the female cohort, noting that although it would be ideal to increase the generalisability of results to women with a natural MC, including a mixed cohort of women would have necessitated a greater sample size to achieve statistical power, which was beyond the scope and resources of this study. Participants were therefore included if they had been using a combined monophasic OCP for at least three months prior to study commencement. We initially aimed to recruit participants using a single brand/generation/formulation of OCPs (as is required to achieve a gold tier), however this proved too prohibitive in recruiting the required sample size to achieve statistical power for between-sex comparisons. Therefore, participants were recruited on the basis of using combined monophasic OCPs, but the formulation/generation/brand was not stipulated.

Menstrual status tier

Given that participants were not restricted to a single brand/generation/formulation of combined monophasic OCPs we therefore achieved a "silver" tier. However, the HC type/formulation/length of usage for each individual participant was comprehensively reported in a supplementary table (Table S3.1). We also considered active/inactive pill days when scheduling testing, with all testing occurring during active HC usage only (i.e., pill taking days, avoiding the withdrawal bleed), as is required for a gold tier, thus further increasing the robustness of our methods. Participants used calendar counting to report the dates of their withdrawal bleeds (if taken), and trial dates were scheduled accordingly.

Results, Discussion and Future Research

The findings of this rigorously controlled laboratory-based study suggest that 24 hours of LEA is not a sufficient exposure to impair strength/power, sprint capacity, or cognitive performance. However, 24 hours of exercise-induced LEA appears to influence substrate oxidation more than LEA induced by diet alone. Moreover, there were no sex-differences in response to 24 h EA manipulations across any outcome measure (performance measures of explosive power, substrate oxidation or postprandial metabolism). Overall, these findings indicate that a 24 h period of LEA can be implemented by both male and female athletes without impairments to strength/power, sprint capacity, or cognitive performance.

A lack of differential response between the sexes may be a result of studying females using OCPs, eliminating the cyclical fluctuations in endogenous oestrogen observed in naturally menstruating

women. As discussed, ovarian hormones may influence pathways involved in substrate oxidation (Bunt, 1990; D'Eon et al., 2002; Nicklas et al., 1989; Oosthuyse & Bosch, 2012; Ruby et al., 1997), and indeed sex-based differences in substrate oxidation/performance are hypothesised to be mediated by endogenous oestrogen concentrations (Devries, 2016). Further research may therefore consider examining the influence of fluctuations in oestrogen and progesterone on these findings, by focussing on naturally menstruating female athletes. Indeed, it could be hypothesised that a sex difference would be observed between males and naturally menstruating female athletes, given the potential influence of ovarian hormones on substrate oxidation. However, this remains speculative. We also note that our study involved a brief but severe LEA exposure (24 h at 15 kcal·kg^{-1} FFM·day^{-1}). Moreover, future research may consider if sex differences in these responses might be observed across longer LEA time periods, or at a higher EA (less severe LEA dose). Since women have been reported to experience more pronounced physiological effects across other body systems (Ihle & Loucks, 2004; Loucks & Thuma, 2003; Papageorgiou et al., 2017; Trexler et al., 2014), a less severe LEA threshold (20-25 kcal·kg^{-1}FFM·day^{-1}) might therefore elicit alterations in women but not men. This therefore provides an avenue for future investigation.

Study 2 – Minimal influence of the menstrual cycle or hormonal contraceptives on performance in female rugby league athletes.

Approach to the research question

The next study aimed to investigate how fluctuations in oestrogen and progesterone may influence performance in an applied setting among athletes of a higher training status [Tier 3, national level (McKay et al., 2022b)]. Here, we know that the cyclical fluctuations in oestrogen and progesterone across the MC have the potential to influence multiple biological systems associated with athletic performance. However, our understanding of any influence of oestrogen or progesterone on performance, through MC phases or with HC use, is inconclusive. This uncertainty is due in part to a general failure of studies to achieve the necessary methodological classification and control of ovarian hormonal profiles (Elliott-Sale et al., 2021; McNulty et al., 2020b) to support causality regarding any influence of oestrogen and progesterone on performance. Therefore, the aim of Study 2 was to examine any alterations in performance across a MC, alongside comparing performance between individuals using HC, and those not.

Study 2 employed an observational design in a training camp environment. Unlike the rolling-recruitment model used in Study 1 where testing was scheduled in advance, the residential nature and five-week duration of this second study facilitated access to participants and the capacity for testing at short notice. We were able to test athletes in MC phases 1, 2 and 4, via the use of daily questionnaires and urinary ovulation measures to track the progress of each MC and quick-response scheduling of

laboratory testing, often less than 24 hours in advance. On each of the three testing occasions, serum hormone measurements of oestradiol and progesterone were collected to retrospectively verify MC phase. This protocol was of particular value in attempting to assess performance during the under-researched phase 2 of the MC, where oestrogen concentration peaks, and progesterone concentration is low. Since this phase is brief (lasting approximately 24-48 h), having athletes on-site and available for testing with less than 24 h notice facilitated testing during this phase.

Participant cohort

We recruited a mixed cohort of participants using HC and those not using HC, to assess performance across a MC as well as between the two groups. Since the participants were recruited primarily on their athletic ability and without any exclusion criteria based on menstrual status, our cohort provided an authentic experience of the heterogeneity of menstrual statuses within a real-world training squad. Indeed, our cohort included HC users (54%) and non-users (46%) in similar proportions to the reported prevalence rates of HC use among athletes (Martin et al., 2018). This aspect increased the generalisability of the findings of our study and aligned with its applied focus. The actual menstrual status for athletes not using contraception was then confirmed during the study through best-practice methods (Elliott-Sale et al., 2021): calendar counting, urinary ovulation testing and serum hormone verification.

Menstrual status tier

With a stated primary outcome of examining performance across MC phases 1, 2 and 4 among athletes not using HC, measures of serum oestradiol and progesterone concentrations were an essential component of the study protocol. Such hormone analysis enabled the: (1) identification of eumenorrheic athletes alongside subtle menstrual disturbances and (2) identification of MC phases. Participants were also required to track their MC both before and during the training camp through calendar counting and urinary LH testing (for a total of 16 weeks), and to verify that no HC use had occurred in the preceding three months. However, because the training camp duration (five-weeks) prohibited the repetition of outcome measures across a second MC, our protocol achieved a "silver" tier status for athletes not using HC.

The HC type, formulation and length of usage were reported for athletes using HC. Athletes using OCPs (n=4) were tested during active HC usage only (i.e., pill taking days, avoiding the withdrawal bleed). Due to the numerous types of HC being used within the participant cohort (hormonal implants, hormonal injections and OCPs), albeit reflective of a real-world training squad, the study protocol therefore achieved "silver" tier for athletes using HC.

Results, Discussion and Future Research

Our findings demonstrated no detectable influence of MC phase or HC use on any performance outcome assessed (jump height, peak force, sprint time, distance thrown or Stroop effect). Some small variations in kinetic and kinematic CMJ/SJ outputs were observed among athletes not using HC. Naturally menstruating athletes produced a 16.8% ($p=0.021$) greater mean concentric power in MC phase 4 than 1 during the CMJ, while during the SJ impulse at 50 ms was 4.7% ($p=0.031$) lower in MC phase 4 than 1, without differences between tests for HC users. However, it could not be determined if the observed alterations exceeded between-day variability. Further research is therefore required, potentially with a larger sample size or among more highly trained athletes [Tiers 4-5, who display superior performance consistency compared to recreational athletes (Hopkins & Hewson, 2001)], to determine if kinetic and kinematic alterations across the MC exceed between-day variability.

Retrospective analysis of serum hormone measurements revealed that a true MC phase 2 (peak oestrogen concentration) was only achieved in one out of the 11 athletes not using HC. As such, phase 2 data were excluded from analysis, and the study was only able to compare performance across phases 1 and 4. This experience demonstrates the challenge of undertaking research on female athletes. Despite our unique and resource-intensive design, we were unable to use nearly a third of the performance testing data collected during the camp. Although we tracked athletes' MCs with a supervised prospective phase-based approach according to best-practice guidelines and reacted quickly to schedule testing for the next available morning (noting that protocols required a fasted, rested presentation), the reliance on retrospective analysis of hormone concentrations in blood samples collected on the test morning meant that confirmation of the true/desired MC phase did not occur after the completion of testing. Furthermore, the five-week duration of the camp prevented the repetition of testing in a second MC when the desired MC phase was not initially achieved. Therefore, future research is needed to understand the potential influence of oestrogen and progesterone across other MC phases (two and three) on performance, including kinetic and kinematic parameters. However, the challenges with prospectively identifying phase 2 may prohibit any practical alterations in the training environment, even if performance alterations are observed. In the meantime, our study reveals that evidence for changes in testing performance across a MC, or during active HC use, is insufficient to justify "menstrual phase- or status-based" testing programs at a group or team-based level among female rugby league athletes.

Study 3 – Perceived negative menstrual cycle symptoms, but not changes in oestrogen or progesterone, are associated with impaired cycling race performance.

Approach to the research question

Experiences within Study 2 demonstrated the difficulty of capturing data from female athletes during phase 2 of the MC. Its short-lived nature represents a challenge in being able to schedule testing once it is detected, even in a study design that maximised participant availability and pro-active testing opportunities. Indeed, the protocol in Study 2 resulted in the exclusion of almost one third of the collected performance data, and a failure to identify outcomes associated with the unique hormonal profile of Phase 2. Therefore Study 3 set out to use a correlational model to explore the relationship between ovarian hormonal fluctuations across the MC and performance, without seeking to prospectively test in specific MC phases. This approach was selected in an attempt to maximise data retention. We employed a novel observational study design: virtually recruiting and managing participants and using e-cycling for performance measurement, meaning that participants were based across Australia and all data collection was completed remotely. Although not causative, by using a remote study design, a larger sample size was recruited (n=41) compared to what is typically achieved in SES research, while also enabling live competitive races to enhance ecological validity.

The study design encompassed weekly live races at a fixed time for a total of four weeks (i.e., four races in total), in which athletes raced against the other study participants in a TT format. Participants underwent a blood test within 21 h of the race. Oestrogen and progesterone concentration were then retrospectively correlated to race completion time as the primary outcome measure. As a secondary outcome we also correlated race time to perceived negative MC-related symptoms on race day, and GI symptoms immediately pre- and post-race. A four-week time period was chosen to provide enough races to capture different hormonal fluctuations across each individual, therefore providing sufficient data points across a range of both "high" and "low" oestrogen/progesterone concentrations. Even if an entire MC was not captured in all participants, it was anticipated that four races (across 22 days) would result in sufficient hormonal variation across this time period for all individuals. For example, an individual with an extended cycle length (>35 days) will likely experience an extended luteal phase, and therefore had fewer than four races occurred (e.g. three races would have resulted in a 14-day testing window) there was the possibility that all races could occur under an identical hormonal milieu in such an individual.

Participant cohort

As the study objective was to examine the relationship between ovarian hormones and performance, it was necessary to exclude women using HC who would experience suppression of endogenous hormone concentrations. Inclusion criteria stated that participants had abstained from HC for at least three months

prior to study commencement. The actual menstrual status of participants was then classified within the study. Current amenorrhea was the only menstrual dysfunction excluded, based on its obvious presentation to the participant (i.e. easily diagnosed) and resulting inability to investigate the primary outcome of interest, due to suppression of endogenous hormones. Other menstrual irregularities that do not suppress endogenous hormonal profiles as severely (e.g., PCOS, endometriosis) were included to increase generalisability, noting that the continued fluctuations in hormone concentrations would still allow examination of the relationship between ovarian hormones and race performance.

Menstrual status tier

The primary outcome measure was the relationship between oestrogen/ progesterone on cycling race performance. Blood samples for the analysis of serum oestradiol and progesterone concentrations were collected at a mean time of within 11.5 h of the race start, and their analyses were standardised by engaging with various branches of the same external (commercial) pathology lab for all but four participants. Participants also tracked their MC, through calendar counting and urinary LH testing, both during the four-week race series, and before/after, in order to capture a minimum of two complete MCs per participant. A "silver" tier was achieved as the outcome measures were only assessed across one MC; the races were not repeated in a second cycle as this was deemed to be too resource intensive and onerous for participants. However, tracking two complete MCs for each participant facilitated the investigation of intra-individual variability in MC characteristics, despite races not occurring across multiple MCs.

Participants who were detected to have menstrual irregularities were still included in the study, with comprehensive reporting and a sensitivity analysis being conducted to examine any confounding influence on the outcome measures. This increased the generalisability of findings by including participants with a range of hormonal profiles, alongside maximising data retention.

Results, Discussion and Future Research

Our study observed that fluctuations in ovarian hormone concentrations across the MC do not appear to systemically affect race performance among trained cyclists. There was also no difference in performance between races conducted during follicular and luteal phases, alongside races when participants were experiencing menstrual bleeding vs not bleeding, nor when they were ovulating vs not ovulating. However, the total number of perceived negative MC symptoms recorded on race day was positively correlated to increased race time, as was the number of GI symptoms of at least "moderate" severity before the race. Therefore, the management of negative MC and GI symptoms appears important for athletic performance enhancement or to mitigate performance decline.

This novel remote study design also increased real-world validity, through the use of live competitive cycling races and flexibility with participant recruitment. Although not a causative study design, this

correlational approach also maximised data retention. Indeed, many participants did not complete all four races (n=15 completed three races and n=3 completed two). As this research design did not aim to test in specific MC phases, data from participants who had completed at least two separate races (n=37) could still be included for analysis. However, had participants been required to complete all four races to be included in analysis this would have resulted in a 49% data loss. Moreover, as the correlations between race performance and both oestradiol (r=-0.001, *p=0.992*) and progesterone (r=-0.023, *p=0.833*) concentrations were extremely weak, it is likely that additional participant numbers would not have detected a relationship.

Overall, our findings suggest that fluctuations in ovarian hormone concentrations across the MC are not systematically associated with real-world cycling performance in trained cyclists, but perception of negative self-reported MC or GI symptoms may have a greater effect. Therefore, an individualised approach, including monitoring and managing any negative symptoms, may be better for uncovering any links to individual athlete performance or mitigating performance decline. However, while we measured the severity of GI symptoms, other perceived MC-related symptoms were reported only in terms of incidence, and hence the influence of MC symptom severity on performance was not able to be examined. However, a standardised and validated questionnaire to assess the incidence and severity of MC-related symptoms does not currently exist and hence its development should be a priority in further elucidating their influence on performance with greater granularity. Moreover, further investigation is required to fully understand the relationship between symptoms and performance, both examining if specific symptoms are driving an association, and if this relationship persists into other activities beyond cycling. Indeed, it could be hypothesised that GI symptoms may be a more influential factor in sports involving a greater degree of vertical motion, such as running, although this requires investigation.

Study 4 – Female athletes report positive experiences as research participants.

Approach to the research question

Given the underrepresentation of women in SES research, we sought to understand the experiences of female athletes currently involved in applied SES research to inform future studies and potentially increase participation rates. Indeed, there is some evidence to suggest that women may have greater volunteer bias (i.e., are less likely to participate) than their male counterparts (Costello et al., 2014; Cowley et al., 2021; Nuzzo, 2021; Smith et al., 2022b), and that this may be a contributor to the sex-bias in SES research. Study 4 therefore collated the experiences of the female athletes who participated in the three experimental studies conducted during this thesis (alongside an additional 19 elite athletes) to highlight any nuances of their participation.

Participant cohort

The female participants (n=89) across all four studies completed a voluntary questionnaire upon study completion. As each study investigated women with different menstrual statuses, this study was able to capture the experiences of a broad range of women: 28 women who used some form of HC and 61 women not using any HC (i.e., with a "natural" MC). Moreover, the studies examined a range of different participant cohorts. Study 1 examined exclusively women using OCP, Study 3 was restricted to just those not using HC, while Studies 2 and the additional included study (Kuikman et al., 2024a) examined a mixed cohort of those using HC and those not using HC.

Menstrual status tier

All four experimental studies achieved a "silver" tier of menstrual status classification and control across each of the differing cohorts examined. As such, we were able to collate the experiences of women who participated in studies with a high-quality classification and control of menstrual status, which is typically deemed to be of a greater participant burden given the additional measurements required (e.g., urinary ovulation sticks, blood tests for the assessment of serum hormones, and comprehensive cycle tracking).

Results, Discussion and Future Research

Our findings confirm the underrepresentation of women in SES research: just 19% of female athletes had participated in research before. A *perceived* lack of opportunity to engage in research was cited as the primary barrier to their participation. The main motivations for female athletes to participate were an interest in the research outcomes and to receive personalised results. As such, it may be prudent to focus future recruitment efforts on emphasising these study aspects, for example highlighting that participation will contribute to research and outlining what personalised results will be received. Including more detailed results and interpretation may also be important in retaining female participants for future studies. Despite the higher participant burden associated with a "silver" methodological tier for the classification and control of menstrual status, the research experience was rated positively at a mean of 77 out of 100. Encouragingly, the majority (94%) of participants also indicated a willingness to participate in future studies. However, these findings represent a small subset of female athletes aged 18-45 years and may differ across age groups. Moreover, this convenience sample comprises participants already enrolled in research and therefore does not represent the views and opinions of those that do not take part. This initial exploratory analysis may consequently inform a subsequent more systematic and detailed exploration of the area including additional information such as participant training age, alongside motherhood and job status, to facilitate a more nuanced interpretation of prior research participation among women. Moreover, future research may seek to understand the experiences of those who do not currently participate in research. Ultimately, these data demonstrate that female

athletes are willing and interested in research participation but perceive a lack of opportunities to do so. As such, opportunities for women to participate in high quality studies should be prioritised.

7.5 Conclusions

This thesis set out to conduct studies using a range of methodologies and differing participant menstrual statuses, culminating in the development of applied recommendations and resources to guide researchers in the pursuit of high-quality applied SES research. While the overarching goal was to encourage and support an improvement in the quality and quantity of research involving female athletes, a simultaneous result would be the accrual of information on the effect of ovarian hormones on performance. This theme was chosen because it has been demonstrated that women are substantially underrepresented in studies of sports performance. In summary, the performance findings demonstrate:

1. A 24-hour period of severe LEA did not impair strength/power, sprint capacity, or cognitive performance among cyclists, and there was no difference in this response between sexes.
2. There was no detectable influence of MC phase or HC use on overall physical and cognitive performance among rugby league athletes.
3. Cycling performance was not systemically altered with fluctuations in oestradiol or progesterone across the MC, but performance may be influence by negative menstrual/ GI symptoms.

Together, these outcomes demonstrate a lack of association between ovarian hormone fluctuations, either across a MC or with HC use, and the performance measures assessed, either in the laboratory setting as assessed during Study 2 or a real-world race as measured during Study 3. As such, an individualised approach to MC monitoring/tracking (monitoring the MC through calendar counting and associated symptomology to uncover any potential patterns) is likely to represent the current best-practice approach for athletes, given the lack of consistent response at the group level.

Meanwhile, the overarching aim of this thesis has been addressed in developing insights and resources to aid researchers in conducting high-quality applied research in female athletes. A key goal of these outputs was to balance the desire for high quality methodology regarding menstrual status classification and control with the practical challenges of conducting research in the applied setting. A range of resources have been developed to aid this pursuit, whilst the methodological decisions and practical learnings from each of the experimental studies conducted during this thesis are detailed to inform future high-quality research, and ultimately facilitate a correction of the sex-bias in SES research.

Chapter 8: References

Achten, J., Gleeson, M., & Jeukendrup, A. E. (2002). Determination of the exercise intensity that elicits maximal fat oxidation. *Medicine and Science in Sports and Exercise, 34*(1), 92-97.

Alghannam, A. F., Gonzalez, J. T., & Betts, J. A. (2018). Restoration of muscle glycogen and functional capacity: Role of post-exercise carbohydrate and protein co-ingestion. *Nutrients, 10*(2), 253.

Ansdell, P., Brownstein, C. G., Skarabot, J., Hicks, K. M., Simoes, D. C. M., Thomas, K., Howatson, G., Hunter, S. K., & Goodall, S. (2019). Menstrual cycle-associated modulations in neuromuscular function and fatigability of the knee extensors in eumenorrheic women. *Journal of Applied Physiology, 126*(6), 1701-1712.

Arkinstall, M. J., Bruce, C. R., Clark, S. A., Rickards, C. A., Burke, L. M., & Hawley, J. A. (2004). Regulation of fuel metabolism by preexercise muscle glycogen content and exercise intensity. *Journal of Applied Physiology, 97*(6), 2275-2283.

Armour, M., Parry, K. A., Steel, K., & Smith, C. A. (2020). Australian female athlete perceptions of the challenges associated with training and competing when menstrual symptoms are present. *International Journal of Sports Science and Coaching, 15*(3), 316-323.

Bailey, S. P., Zacher, C. M., & Mittleman, K. D. (2000). Effect of menstrual cycle phase on carbohydrate supplementation during prolonged exercise to fatigue. *Journal of Applied Physiology, 88*(2), 690-697.

Bale, P., & Nelson, G. (1985). The effects of menstruation on performance of swimmers. *Australian Journal of Science and Medicine in Sport, 17*(1), 19-22.

Baltgalvis, K. A., Greising, S. M., Warren, G. L., & Lowe, D. A. (2010). Estrogen regulates estrogen receptors and antioxidant gene expression in mouse skeletal muscle. *PloS One, 5*(4), e10164.

Barel, E., Krispil, M., & Yaari, I. (2019). Cognitive performance across the menstrual cycle. *Journal of Cognitive Psychology, 4*(47), 41-47.

Barrack, M. T., Gibbs, J. C., De Souza, M. J., Williams, N. I., Nichols, J. F., Rauh, M. J., & Nattiv, A. (2014). Higher incidence of bone stress injuries with increasing female athlete triad-related risk factors: A prospective multisite study of exercising girls and women. *American Journal of Sports Medicine, 42*(4), 949-958.

Bbc Sport. (2020). Bbc elite british sportswomen's survey results. Accessed in 2024.

Beals, K. A., & Manore, M. M. (2002). Disorders of the female athlete triad among collegiate athletes. *International Journal of Sport Nutrition and Exercise Metabolism, 12*(3), 281-293.

Beidleman, B. A., Rock, P. B., Muza, S. R., Fulco, C. S., Forte Jr, V. A., & Cymerman, A. (1999). Exercise v̇e and physical performance at altitude are not affected by menstrual cycle phase. *Journal of Applied Physiology, 86*(5), 1519-1526.

Bemben, D. A., Boileau, R. A., Bahr, J. M., Nelson, R. A., & Misner, J. E. (1992). Effects of oral contraceptives on hormonal and metabolic responses during exercise. *Medicine and Science in Sports and Exercise, 24*(4), 434-441.

Bemben, D. A., Salm, P. C., & Salm, A. J. (1995). Ventilatory and blood lactate responses to maximal treadmill exercise during the menstrual cycle. *Journal of Sports Medicine and Physical Fitness, 35*(4), 257-262.

Bird, S. R., & Hawley, J. A. (2016). Update on the effects of physical activity on insulin sensitivity in humans. *BMJ Open Sport and Exercise Medicine, 2*(1).

Bonen, A., Haynes, F. W., & Graham, T. E. (1991). Substrate and hormonal responses to exercise in women using oral-contraceptives. *Journal of Applied Physiology, 70*(5), 1917-1927.

Brinton, R. D., Thompson, R. F., Foy, M. R., Baudry, M., Wang, J., Finch, C. E., Morgan, T. E., Pike, C. J., Mack, W. J., & Stanczyk, F. Z. (2008). Progesterone receptors: Form and function in brain. *Frontiers in Neuroendocrinology, 29*(2), 313-339.

Brook, E. M., Tenforde, A. S., Broad, E. M., Matzkin, E. G., Yang, H. Y., Collins, J. E., & Blauwet, C. A. (2019). Low energy availability, menstrual dysfunction, and impaired bone health: A survey of elite para athletes. *Scandinavian Journal of Medicine and Science in Sports, 29*(5), 678-685.

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? *British Journal of Sports Medicine, 51*(6), 487-488.

Bryner, R. W., Toffle, R. C., Ullrich, I. H., & Yeater, R. A. (1996). Effect of low dose oral contraceptives on exercise performance. *British Journal of Sports Medicine, 30*(1), 36-40.

Brynhildsen, J., Lennartsson, H., Klemetz, M., Dahlquist, P., Hedin, B., & Hammar, M. (1997). Oral contraceptive use among female elite athletes and age-matched controls and its relation to low back pain. *Acta Obstetricia et Gynecologica Scandinavica, 76*(9), 873-878.

Bunt, J. C. (1990). Metabolic actions of estradiol: Significance for acute and chronic exercise responses. *Medicine and Science in Sports and Exercise, 22*(3), 286-290.

Burke, L. M. (2021). Ketogenic low-cho, high-fat diet: The future of elite endurance sport? *The Journal of Physiology, 599*(3), 819-843.

Burke, L. M., Slater, G. J., Matthews, J. J., Langan-Evans, C., & Horswill, C. A. (2021). Acsm expert consensus statement on weight loss in weight-category sports. *Current Sports Medicine Reports, 20*(4), 199-217.

Burke, L. M., Van Loon, L. J., & Hawley, J. A. (2017). Postexercise muscle glycogen resynthesis in humans. *Journal of Applied Physiology, 122*(5), 1055-1067.

Burke, L. M., Whitfield, J., Ross, M. L., Tee, N., Sharma, A. P., King, A. J., Heikura, I. A., Morabito, A., & Mckay, A. K. (2023). Short severe energy restriction with refueling reduces body mass without altering training-associated performance improvement. *Medicine and Science in Sports and Exercise, 55*(8), 1487-1498.

Burrows, M., & Bird, S. R. (2005). Velocity at vo2(max) and peak treadmill velocity are not influenced within or across the phases of the menstrual cycle. *European Journal of Applied Physiology, 93*(5-6), 575-580.

Burrows, M., Nevill, A. M., Bird, S., & Simpson, D. (2003). Physiological factors associated with low bone mineral density in female endurance runners. *British Journal of Sports Medicine, 37*(1), 67-71.

Bushman, B., Masterson, G., & Nelsen, J. (2006). Anaerobic power performance and the menstrual cycle: Eumenorrheic and oral contraceptive users. *Journal of Sports Medicine and Physical Fitness, 46*(1), 132-137.

Campbell, S. E., Angus, D. J., & Febbraio, M. A. (2001). Glucose kinetics and exercise performance during phases of the menstrual cycle: Effect of glucose ingestion. *American Journal of Physiology-Endocrinology and Metabolism, 281*(4), E817-825.

Carmichael, M. A., Thomson, R. L., Moran, L. J., & Wycherley, T. P. (2021). The impact of menstrual cycle phase on ahletes' performance: A narrative review. *International Journal of Environmental Research and Public Health, 18*(4), 1667.

Carpenter, A. J., & Nunneley, S. A. (1988). Endogenous hormones subtly alter women's response to heat stress. *Journal of Applied Physiology, 65*(5), 2313-2317.

Casazza, G. A., Suh, S. H., Miller, B. F., Navazio, F. M., & Brooks, G. A. (2002). Effects of oral contraceptives on peak exercise capacity. *Journal of Applied Physiology, 93*(5), 1698-1702.

Chidi-Ogbolu, N., & Baar, K. (2018). Effect of estrogen on musculoskeletal performance and injury risk. *Frontiers in Physiology, 9*, 1834.

Chrzanowski-Smith, O. J., Edinburgh, R. M., Thomas, M. P., Haralabidis, N., Williams, S., Betts, J. A., & Gonzalez, J. T. (2020). The day-to-day reliability of peak fat oxidation and fatmax. *European Journal of Applied Physiology, 120*(8), 1745-1759.

Chrzanowski-Smith, O. J., Edinburgh, R. M., Smith, E., Thomas, M. P., Walhin, J. P., Koumanov, F., Williams, S., Betts, J. A., & Gonzalez, J. T. (2021). Resting skeletal muscle pnpla2 (atgl) and cpt1b are associated with peak fat oxidation rates in men and women but do not explain observed sex differences. *Experimental Physiology, 106*(5), 1208-1223.

Clarke, A. C., Bruinvels, G., Julian, R., Inge, P., Pedlar, C. R., & Govus, A. D. (2021). Hormonal contraceptive use in football codes in australia. *Frontiers in Sports and Active Living, 3*, 16.

Claus, L. (2020). Why gender equal pay is so hard to achieve in sport. Accessed in 2021. Retrieved from https://www.mgmt.ucl.ac.uk/news/gender-equality-sports-%E2%80%93-why-gender-equal-pay-so-hard-achieve-sport

Comfort, P., Dos' Santos, T., Beckham, G. K., Stone, M. H., Guppy, S. N., & Haff, G. G. (2019). Standardization and methodological considerations for the isometric midthigh pull. *Strength and Conditioning Journal, 41*(2), 57-79.

Constantini, N. W., Dubnov, G., & Lebrun, C. M. (2005). The menstrual cycle and sport performance. *Clinical Journal of Sport Medicine, 24*(2), e51-82, xiii-xiv.

Cook, C. J., Kilduff, L. P., & Crewther, B. T. (2018). Basal and stress-induced salivary testosterone variation across the menstrual cycle and linkage to motivation and muscle power. *Scandinavian Journal of Medicine and Science in Sports, 28*(4), 1345-1353.

Coste, O., Paris, F., Galtier, F., Letois, F., Maimoun, L., & Sultan, C. (2011). Polycystic ovary-like syndrome in adolescent competitive swimmers. *Fertility and Sterility, 96*(4), 1037-1042.

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.

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, 29*(2), 146-151.

Crewther, B. T., Hamilton, D., Kilduff, L. P., Drawer, S., & Cook, C. J. (2018). The effect of oral contraceptive use on salivary testosterone concentrations and athlete performance during international field hockey matches. *Journal of Science and Medicine in Sport, 21*(5), 453-456.

D'eon, T. M., Sharoff, C., Chipkin, S. R., Grow, D., Ruby, B. C., & Braun, B. (2002). Regulation of exercise carbohydrate metabolism by estrogen and progesterone in women. *American Journal of Physiology-Endocrinology and Metabolism, 283*(5), E1046-1055.

D'souza, A. C., Wageh, M., Williams, J. S., Colenso-Semple, L. M., Mccarthy, D. G., Mckay, A. K., Elliott-Sale, K. J., Burke, L. M., Parise, G., & Macdonald, M. J. (2023). Menstrual cycle hormones and oral contraceptives: A multimethod systems physiology-based review of their impact on key aspects of female physiology. *Journal of Applied Physiology, 135*(6), 1284-1299.

Dadgostar, H., Razi, M., Aleyasin, A., Alenabi, T., & Dahaghin, S. (2009). The relation between athletic sports and prevalence of amenorrhea and oligomenorrhea in iranian female athletes. *BMC Sports Science Medicine and Rehabilitation, 1*(1), 1-7.

Daggett, A., Davies, B., & Boobis, L. (1983). Physiological and biochemical responses to exercise following oral-contraceptive use. *Medicine and Science in Sports and Exercise, 15*(2), 174-174.

Dam, T. V., Dalgaard, L. B., Sevdalis, V., Bibby, B. M., De Jonge, X. J., Gravholt, C. H., & Hansen, M. (2022). Muscle performance during the menstrual cycle correlates with psychological well-

being, but not fluctuations in sex hormones. *Medicine and Science in Sports and Exercise, 54*(10), 1678.

Davies, B. N., Elford, J. C. C., & Jamieson, K. F. (1991). Variations in performance in simple muscle tests at different phases of the menstrual-cycle. *Journal of Sports Medicine and Physical Fitness, 31*(4), 532-537.

De Jonge, X. J., Boot, C., Thom, J., Ruell, P., & Thompson, M. (2001). The influence of menstrual cycle phase on skeletal muscle contractile characteristics in humans. *The Journal of Physiology, 530*(Pt 1), 161.

De Souza, M. J., Maguire, M. S., Rubin, K. R., & Maresh, C. M. (1990). Effects of menstrual phase and amenorrhea on exercise performance in runners. *Medicine and Science in Sports and Exercise, 22*(5), 575-580.

De Souza, M. J., Toombs, R., Scheid, J., O'donnell, E., West, S. L., & Williams, N. I. (2010). High prevalence of subtle and severe menstrual disturbances in exercising women: Confirmation using daily hormone measures. *Human Reproduction, 25*(2), 491-503.

Dean, T. M., Perreault, L., Mazzeo, R. S., & Horton, T. J. (2003). No effect of menstrual cycle phase on lactate threshold. *Journal of Applied Physiology, 95*(6), 2537-2543.

Debruynprevost, P., Masset, C., & Sturbois, X. (1984). Physiological-response from 18-25 years women to aerobic and anaerobic physical-fitness tests at different periods during the menstrual-cycle. *Journal of Sports Medicine and Physical Fitness, 24*(2), 144-148.

Deeks, J. J., Higgins, J. P., Altman, D. G., & Group, C. S. M. (2019). Analysing data and undertaking meta-analyses. *Cochrane Handbook for Systematic Reviews of Interventions*, 241-284.

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. *Experimental Physiology, 101*(2), 243-249.

Devries, M. C., Hamadeh, M. J., Phillips, S. M., & Tarnopolsky, M. A. (2006). Menstrual cycle phase and sex influence muscle glycogen utilization and glucose turnover during moderate-intensity endurance exercise. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology, 291*(4), R1120-1128.

Diabetesaustralia. (2023). Blood glucose level range. Accessed in 2024. Retrieved from https://www.diabetesaustralia.com.au/managing-diabetes/blood-glucose-range/

Dibrezzo, R., Fort, I. L., & Brown, B. (1991). Relationships among strength, endurance, weight and body fat during three phases of the menstrual cycle. *Journal of Sports Medicine and Physical Fitness, 31*(1), 89-94.

Dokumaci, B., & Hazir, T. (2019). Effects of the menstrual cycle on running economy: Oxygen cost versus caloric cost. *Research Quarterly for Exercise and Sport, 90*(3), 318-326.

Dombovy, M. L., Bonekat, H. W., Williams, T. J., & Staats, B. A. (1987). Exercise performance and ventilatory response in the menstrual cycle. *Medicine and Science in Sports and Exercise, 19*(2), 111-117.

Douglas, L. (2018). The rise of women's sports. Accessed in 2021. Retrieved from https://www.nielsen.com/wp-content/uploads/sites/3/2019/04/the-rise-of-womens-sports.pdf

Dragutinovic, B., Moser, F., Notbohm, H. L., Ihalainen, J. K., Bloch, W., & Schumann, M. (2024). Influence of menstrual cycle and oral contraceptive phases on strength performance, neuromuscular fatigue and perceived exertion. *Journal of Applied Physiology*.

Dusek, T. (2001). Influence of high intensity training on menstrual cycle disorders in athletes. *Croatian Medical Journal, 42*(1), 79-82.

Ekenros, L., Hirschberg, A. L., Heijne, A., & Fridén, C. (2013). Oral contraceptives do not affect muscle strength and hop performance in active women. *Clinical Journal of Sport Medicine, 23*(3), 202-207.

Elliott-Sale, K., Ross, E., Burden, R., & Hicks, K. (2020a). The bases expert statement on conducting and implementing female athlete-based research. *The Sport and Exercise Scientist*.

Elliott-Sale, K. J., & Hicks, K. M. (2018). Hormonal-based contraception and the exercising female. In *The Exercising Female* (pp. 30-43): Routledge.

Elliott-Sale, K. J., Mcnulty, K. L., Ansdell, P., Goodall, S., Hicks, K. M., Thomas, K., Swinton, P. A., & Dolan, E. (2020b). The effects of oral contraceptives on exercise performance in women: A systematic review and meta-analysis. *Sports Medicine, 50*(10), 1785-1812.

Elliott-Sale, K. J., Mcnulty, K. L., Ansdell, P., Goodall, S., Hicks, K. M., Thomas, K., Swinton, P. A., & Dolan, E. (2020c). The effects of oral contraceptives on exercise performance in women: A systematic review and meta-analysis. *Sports Medicine*, 1-28.

Elliott-Sale, K. J., Minahan, C. L., De Jonge, X., Ackerman, K. E., Sipila, S., Constantini, N. W., Lebrun, C. M., & 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, 51*(5), 843-861.

Emmonds, S., Heyward, O., & Jones, B. (2019). The challenge of applying and undertaking research in female sport. *Sports Medicine Open, 5*(1), 51.

Eshre, T. R., & Group, A.-S. P. C. W. (2004). Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome. *Fertility and Sterility, 81*(1), 19-25.

Fensham, N. C., Heikura, I. A., Mckay, A. K., Tee, N., Ackerman, K. E., & Burke, L. M. (2022). Short-term carbohydrate restriction impairs bone formation at rest and during prolonged exercise to a greater degree than low energy availability. *Journal of Bone and Mineral Research, 37*(10), 1915-1925.

Findlay, R. J., Macrae, E. H. R., Whyte, I. Y., Easton, C., & Forrest, L. J. (2020). How the menstrual cycle and menstruation affect sporting performance: Experiences and perceptions of elite female rugby players. *British Journal of Sports Medicine, 54*(18), 1108-+.

Flood, T. R., Clausen, E., Kuikman, M. A., Smith, E., Mckay, A. K., Burke, L. M., & Elliott-Sale, K. J. (2024). Auditing the representation of hormonal contraceptives in studies assessing exercise performance in women. *Journal of Sports Sciences, 42*(9), 825-839.

Fomin, S., Pivovarova, V., & Voronova, V. (1989). Changes in the special working capacity and mental stability of well-trained woman skiers at various phases of the biological cycle. *Research in Sports Medicine: An International Journal, 1*(2), 89-92.

Frandsen, J., Pistoljevic, N., Quesada, J. P., Amaro-Gahete, F. J., Ritz, C., Larsen, S., Dela, F., & Helge, J. W. (2020). Menstrual cycle phase does not affect whole body peak fat oxidation rate during a graded exercise test. *Journal of Applied Physiology, 128*(3), 681-687.

Frankovich, R. J., & Lebrun, C. M. (2000). Menstrual cycle, contraception, and performance. *Clinical Journal of Sport Medicine, 19*(2), 251-271.

Fruzzetti, F., Lello, S., Lazzarini, V., Fratta, S., Orru, M., Sorge, R., Minerba, L., Ricci, C., Genazzani, A. R., & Melis, G. B. (2007). The oral contraceptive containing 30 µg of ethinylestradiol plus 3 mg of drospirenone is able to antagonize the increase of extracellular water occurring in healthy young women during the luteal phase of the menstrual cycle: An observational study. *Contraception, 75*(3), 199-203.

Giacomoni, M., Bernard, T., Gavarry, O., Altare, S., & Falgairette, G. (2000). Influence of the menstrual cycle phase and menstrual symptoms on maximal anaerobic performance. *Medicine and Science in Sports and Exercise, 32*(2), 486-492.

Goldsmith, E., & Glaister, M. (2020). The effect of the menstrual cycle on running economy. *Journal of Sports Medicine and Physical Fitness, 60*(4), 610-617.

Gordon, D., Hughes, F., Young, K., Scruton, A., Keiller, D., Caddy, O., Baker, J., & Barnes, R. (2013). The effects of menstrual cycle phase on the development of peak torque under isokinetic conditions. *Isokinetics and Exercise Science, 21*(4), 285-291.

Graja, A., Kacem, M., Hammouda, O., Borji, R., Bouzid, M. A., Souissi, N., & Rebai, H. (2020). Physical, biochemical, and neuromuscular responses to repeated sprint exercise in eumenorrheic female handball players: Effect of menstrual cycle phases. *Journal of Strength and Conditioning Research*.

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. *Experimental Physiology, 101*(2), 230-242.

Greenhall, M., Taipale, R. S., Ihalainen, J. K., & Hackney, A. C. (2021). Influence of the menstrual cycle phase on marathon performance in recreational runners. *International Journal of Sports Physiology and Performance, 16*(4), 601-604.

Guo, H., Dong, X., & Zhang, J. (2005). Hormone levels at different menstrual cycle phases and changes of athletic performance in female athletes. *Chinese Journal of Tissue Engineering Research*, 188-190.

Hackney, A. C. (1999). Influence of oestrogen on muscle glycogen utilization during exercise. *Acta Physiologica Scandinavica, 167*(3), 273-274.

Hackney, A. C., Mccracken-Compton, M. A., & Ainsworth, B. (1994). Substrate responses to submaximal exercise in the midfollicular and midluteal phases of the menstrual cycle. *International Journal of Sport Nutrition and Exercise Metabolism, 4*(3), 299-308.

Hagmar, M., Berglund, B., Brismar, K., & Hirschberg, A. L. (2009). Hyperandrogenism may explain reproductive dysfunction in olympic athletes. *Medicine and Science in Sports and Exercise, 41*(6), 1241-1248.

Hampson, E. (1990). Variations in sex-related cognitive abilities across the menstrual cycle. *Brain and Cognition, 14*(1), 26-43.

Hara, Y., Waters, E. M., Mcewen, B. S., & Morrison, J. H. (2015). Estrogen effects on cognitive and synaptic health over the lifecourse. *Physiological Reviews, 95*(3), 785-807.

Harris, P. A., Taylor, R., Minor, B. L., Elliott, V., Fernandez, M., O'neal, L., Mcleod, L., Delacqua, G., Delacqua, F., & Kirby, J. (2019). The redcap consortium: Building an international community of software platform partners. *Journal of Biomedical Informatics, 95*, 103208.

Harris, P. A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., & Conde, J. G. (2009). Research electronic data capture (redcap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics, 42*(2), 377-381.

Hearris, M. A., Hammond, K. M., Fell, J. M., & Morton, J. P. (2018). Regulation of muscle glycogen metabolism during exercise: Implications for endurance performance and training adaptations. *Nutrients, 10*(3), 298.

Heikura, I. A., Burke, L. M., Hawley, J. A., Ross, M. L., Garvican-Lewis, L., Sharma, A. P., Mckay, A. K., Leckey, J. J., Welvaert, M., & Mccall, L. (2020). A short-term ketogenic diet impairs markers of bone health in response to exercise. *Frontiers In Endocrinology, 10*, 880.

Heikura, I. A., Uusitalo, A. L. T., Stellingwerff, T., Bergland, D., Mero, A. A., & Burke, L. M. (2018). Low energy availability is difficult to assess but outcomes have large impact on bone injury rates in elite distance athletes. *International Journal of Sport Nutrition and Exercise Metabolism, 28*(4), 403-411.

Hertel, J., Williams, N. I., Olmsted-Kramer, L. C., Leidy, H. J., & Putukian, M. (2006). Neuromuscular performance and knee laxity do not change across the menstrual cycle in female athletes. *Knee Surgery, Sports Traumatology, Arthroscopy, 14*(9), 817-822.

Hogwood, A. C., Ortiz De Zevallos, J., Kruse, K. E., De Guzman, J., Buckley, M., Weltman, A., & Allen, J. D. (2023). The effects of inorganic nitrate supplementation on exercise economy and endurance capacity across the menstrual cycle. *Journal of Applied Physiology, 135*(5), 1167-1175.

Hopkins, W. G. (2015). Individual responses made easy. *Journal of Applied Physiology, 118*(12), 1444-1446.

Hopkins, W. G., & Hewson, D. J. (2001). Variability of competitive performance of distance runners. *Medicine and Science in Sports and Exercise, 33*(9), 1588-1592.

Horne, J. A., & Östberg, O. (1976). A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms. *International Journal of Chronobiology*.

Horton, T. J., Miller, E. K., Glueck, D., & Tench, K. (2002). No effect of menstrual cycle phase on glucose kinetics and fuel oxidation during moderate-intensity exercise. *American Journal of Physiology-Endocrinology and Metabolism, 282*(4), E752-E762.

Horton, T. J., Pagliassotti, M. J., Hobbs, K., & Hill, J. O. (1998). Fuel metabolism in men and women during and after long-duration exercise. *Journal of Applied Physiology, 85*(5), 1823-1832.

Howell, D. C., Rogier, M., Yzerbyt, V., & Bestgen, Y. (1998). *Statistical Methods in Human Sciences* (Vol. 721): New York: Wadsworth.

Huber, J. (1998). Pharmacokinetics of implanon®: An integrated analysis. *Contraception, 58*(6), 85S-90S.

Hunter, S. K. (2016). Sex differences in fatigability of dynamic contractions. *Experimental Physiology, 101*(2), 250-255.

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 Medicine Open, 7*(1), 43.

Ihle, R., & Loucks, A. B. (2004). Dose-response relationships between energy availability and bone turnover in young exercising women. *Journal of Bone and Mineral Research, 19*(8), 1231-1240.

Impellizzeri, F. M., & Maffiuletti, N. A. (2007). Convergent evidence for construct validity of a 7-point likert scale of lower limb muscle soreness. *Clinical Journal of Sport Medicine, 17*(6), 494-496.

International Olympic Committee. (2019). Olympic charter. In: Force As From June.

International Olympic Committee. (2021). Tokyo 2020 first ever gender-balanced olympic games in history, record number of female competitors at paralympic games. Accessed in 2021. Retrieved from https://olympics.com/ioc/news/tokyo-2020-first-ever-gender-balanced-olympic-games-in-history-record-number-of-female-competitors-at-paralympic-games

International Olympic Committee. (2024). #genderequalolympics: Celebrating full gender parity on the field of play at paris 2024. Accessed in 2024. Retrieved from https://olympics.com/ioc/news/genderequalolympics-celebrating-full-gender-parity-on-the-field-of-play-at-paris-2024

International Paralympic Committee. (2024). Paris 2024: Record number of delegations and females to compete. Accessed in 2024. Retrieved from https://www.paralympic.org/paris-2024/news/paris-2024-record-number-delegations-and-females-compete#:~:text=Competing%20across%20549%20medal%20events,lined%20up%20at%20Sydney%202000.

Isacco, L., & Boisseau, N. (2017). Sex hormones and substrate metabolism during endurance exercise. In *Sex Hormones, Exercise and Women* (pp. 35-58): Springer.

Isacco, L., Thivel, D., Meddahi-Pelle, A., Lemoine-Morel, S., Duclos, M., & Boisseau, N. (2014). Exercise per se masks oral contraceptive-induced postprandial lipid mobilization. *Applied Physiology Nutrition and Metabolism, 39*(11), 1222-1229.

Isacco, L., Thivel, D., Pelle, A. M., Zouhal, H., Duclos, M., Duche, P., & Boisseau, N. (2012). Oral contraception and energy intake in women: Impact on substrate oxidation during exercise. *Applied Physiology Nutrition and Metabolism, 37*(4), 646-656.

Ishibashi, A., Kojima, C., Tanabe, Y., Iwayama, K., Hiroyama, T., Tsuji, T., Kamei, A., Goto, K., & Takahashi, H. (2020). Effect of low energy availability during three consecutive days of endurance training on iron metabolism in male long distance runners. *Physiological Reports, 8*(12), e14494.

James, J. J., Klevenow, E. A., Atkinson, M. A., Vosters, E. E., Bueckers, E. P., Quinn, M. E., Kindy, S. L., Mason, A. P., Nelson, S. K., & Rainwater, K. A. (2024). Reply to nuzzo and deaner:"Investigator bias" is a potent influence on the underrepresentation of women research participants in biomedical research. *Journal of Applied Physiology, 136*(1), 56-57.

James, L. P., Talpey, S. W., Young, W. B., Geneau, M. C., Newton, R. U., & Gastin, P. B. (2023). Strength classification and diagnosis: Not all strength is created equal. *Strength and Conditioning Journal, 45*(3), 333-341.

Janse De Jonge, X., Thompson, B., & Han, A. (2019). Methodological recommendations for menstrual cycle research in sports and exercise. *Medicine and Science in Sports and Exercise, 51*(12), 2610-2617.

Janse De Jonge, X. A. (2003). Effects of the menstrual cycle on exercise performance. *Sports Medicine, 33*(11), 833-851.

Jeukendrup, A., & Wallis, G. (2005). Measurement of substrate oxidation during exercise by means of gas exchange measurements. *International journal of Sports Medicine, 26*(S 1), S28-S37.

Joyce, S., Sabapathy, S., Bulmer, A., & Minahan, C. (2013). Effect of long-term oral contraceptive use on determinants of endurance performance. *Journal of Strength and Conditioning Research, 27*(7), 1891-1896.

Julian, R., Hecksteden, A., Fullagar, H. H., & Meyer, T. (2017). The effects of menstrual cycle phase on physical performance in female soccer players. *PloS One, 12*(3), e0173951.

Julian, R., Skorski, S., Hecksteden, A., Pfeifer, C., Bradley, P. S., Schulze, E., & Meyer, T. (2021). Menstrual cycle phase and elite female soccer match-play: Influence on various physical performance outputs. *Science and Medicine in Football, 5*(2), 97-104.

Jurkowski, J. E., Jones, N. L., Toews, C. J., & Sutton, J. R. (1981). Effects of menstrual cycle on blood lactate, o2 delivery, and performance during exercise. *Journal of Applied Physiology: Respiratory, Environmental and Exercise Physiology, 51*(6), 1493-1499.

Jurov, I., Keay, N., & Rauter, S. (2022a). Reducing energy availability in male endurance athletes: A randomized trial with a three-step energy reduction. *Journal of the International Society of Sports Nutrition, 19*(1), 179-195.

Jurov, I., Keay, N., Spudic, D., & Rauter, S. (2022b). Inducing low energy availability in trained endurance male athletes results in poorer explosive power. *European Journal of Applied Physiology, 122*(2), 503-513.

Kanaley, J. A., Boileau, R. A., Bahr, J. M., Misner, J. E., & Nelson, R. A. (1992). Cortisol levels during prolonged exercise: The influence of menstrual phase and menstrual status. *International Journal of Sport Nutrition and Exercise Metabolism, 13*(4), 332-336.

Kelly, M. K., Smith, E. S., Brown, H. A., Jardine, W. T., Convit, L., Bowe, S. J., Condo, D., Guy, J. H., Burke, L. M., Periard, J. D., Snipe, R. M. J., Snow, R. J., & Carr, A. J. (2024). Auditing the representation of females versus males in heat adaptation research. *International Journal of Sport Nutrition and Exercise Metabolism, 1(aop)*, 1-11.

Kishali, N. F., Imamoglu, O., Atan, T., Akyol, P., Burmaoglu, G., Bas, M., & Yildirim, K. (2004). Some performance parameter changes during menstrual cycle periods of athletes and non-athletes. *Annals of the Rheumatic Diseases, 63*, 490-490.

Klentrou, P., & Plyley, M. (2003). Onset of puberty, menstrual frequency, and body fat in elite rhythmic gymnasts compared with normal controls. *British Journal of Sports Medicine, 37*(6), 490-494.

Knowles, O. E., Aisbett, B., Main, L. C., Drinkwater, E. J., Orellana, L., & Lamon, S. (2019). Resistance training and skeletal muscle protein metabolism in eumenorrheic females: Implications for researchers and practitioners. *Sports Medicine, 49*(11), 1637-1650.

Koehler, K., Hoerner, N. R., Gibbs, J. C., Zinner, C., Braun, H., De Souza, M. J., & Schaenzer, W. (2016). Low energy availability in exercising men is associated with reduced leptin and insulin but not with changes in other metabolic hormones. *Journal of Sports Sciences, 34*(20), 1921-1929.

Kojima, C., Ishibashi, A., Tanabe, Y., Iwayama, K., Kamei, A., Takahashi, H., & Goto, K. (2020). Muscle glycogen content during endurance training under low energy availability. *Medicine and Science in Sports and Exercise, 52*(1), 187-195.

Kovalchik, S. A., & Reid, M. (2017). Comparing matchplay characteristics and physical demands of junior and professional tennis athletes in the era of big data. *Journal of Sports Science and Medicine, 16*(4), 489-497.

Kozaki, T., & Yasukouchi, A. (2009). Sex differences on components of mental rotation at different menstrual phases. *International Journal of Neuroscience, 119*(1), 59-67.

Krauss, R. M., & Burkman, R. T., Jr. (1992). The metabolic impact of oral contraceptives. *American Journal of Obstetrics and Gynecology, 167*(4 Pt 2), 1177-1184.

Kuikman, M. A., Mckay, A. K., Smith, E. S., Ackerman, K. E., Harris, R., Elliott-Sale, K. J., Stellingwerff, T., & Burke, L. M. (2023a). Female athlete representation and dietary control methods among studies assessing chronic carbohydrate approaches to support training. *International Journal of Sport Nutrition and Exercise Metabolism, 33*(4), 198-208.

Kuikman, M. A., Mckay, A. K. A., Mccormick, R., Tee, N., Vallance, B., Ackerman, K. E., Harris, R., Elliott-Sale, K. J., Stellingwerff, T., & Burke, L. M. (2024a). The temporal effects of altitude and low energy availability manipulation on resting metabolic rate in female athletes. *Medicine and Science in Sports and Exercise, aop*.

Kuikman, M. A., Smith, E. S., Mckay, A. K., Ackerman, K. E., Harris, R., Elliott-Sale, K. J., Stellingwerff, T., & Burke, L. M. (2023b). Fueling the female athlete: Auditing her representation in studies of acute carbohydrate intake for exercise. *Medicine and Science in Sports and Exercise, 55*(3), 569.

Kuikman, M. A., Smith, E. S., Mckay, A. K. A., Mccormick, R., Ackerman, K. E., Harris, R., Elliott-Sale, K. J., Stellingwerff, T., & Burke, L. M. (2024b). Impact of acute dietary and exercise manipulation on rmr measurements and dxa body composition estimates. *Medicine and Science in Sports and Exercise, aop*.

Lane, A. R., O'leary, C. B., & Hackney, A. C. (2015). Menstrual cycle phase effects free testosterone responses to prolonged aerobic exercise. *Acta Physiologica Hungarica, 102*(3), 336-341.

Langan-Evans, C., Germaine, M., Artukovic, M., Oxborough, D. L., Areta, J. L., Close, G. L., & Morton, J. P. (2020). The psychological and physiological consequences of low energy availability in a male combat sport athlete. *Medicine and Science in Sports and Exercise, 53*(4), 673-683.

Lara, B., Gutierrez-Hellin, J., Garcia-Bataller, A., Rodriguez-Fernandez, P., Romero-Moraleda, B., & Del Coso, J. (2020a). Ergogenic effects of caffeine on peak aerobic cycling power during the menstrual cycle. *European Journal of Nutrition, 59*(6), 2525-2534.

Lara, B., Gutiérrez Hellín, J., Ruíz-Moreno, C., Romero-Moraleda, B., & Del Coso, J. (2020b). Acute caffeine intake increases performance in the 15-s wingate test during the menstrual cycle. *British Journal of Clinical Pharmacology, 86*(4), 745-752.

Larsen, B., Morris, K., Quinn, K., Osborne, M., & Minahan, C. (2020). Practice does not make perfect: A brief view of athletes' knowledge on the menstrual cycle and oral contraceptives. *Journal of Science and Medicine in Sport, 23*(8), 690-694.

Lavoie, J. M., Dionne, N., Helie, R., & Brisson, G. R. (1987). Menstrual cycle phase dissociation of blood glucose homeostasis during exercise. *Journal of Applied Physiology, 62*(3), 1084-1089.

Lebrun, C. M. (1993). Effect of the different phases of the menstrual cycle and oral contraceptives on athletic performance. *Sports Medicine, 16*(6), 400-430.

Lebrun, C. M., Mckenzie, D. C., Prior, J. C., & Taunton, J. E. (1995). Effects of menstrual cycle phase on athletic performance. *Medicine and Science in Sports and Exercise, 27*(3), 437-444.

Lebrun, C. M., Petit, M. A., Mckenzie, D. C., Taunton, J. E., & Prior, J. C. (2003). Decreased maximal aerobic capacity with use of a triphasic oral contraceptive in highly active women: A randomised controlled trial. *British Journal of Sports Medicine, 37*(4), 315-320.

Lebrun, C. M. T. (1991). *Effects of the menstrual cycle and oral contraceptives on athletic performance.* University of British Columbia,

Lee, S. J., Sim, M., Van Rens, F. E., & Peiffer, J. J. (2024). Fatigue resistance is altered during the high-hormone phase of eumenorrheic females but not oral contraceptive users. *Medicine and Science in Sports and Exercise, 56*(1), 92-102.

Leung, K. C., Johannsson, G., Leong, G. M., & Ho, K. K. (2004). Estrogen regulation of growth hormone action. *Endocrine Reviews, 25*(5), 693-721.

Lewis, C. A., Kimmig, A. S., Zsido, R. G., Jank, A., Derntl, B., & Sacher, J. (2019). Effects of hormonal contraceptives on mood: A focus on emotion recognition and reactivity, reward processing, and stress response. *Current Psychiatry Reports, 21*(11), 115.

Lieberman, H. R., Bukhari, A. S., Caldwell, J. A., Wilson, M. A., Mahoney, C. R., Pasiakos, S. M., Mcclung, J. P., & Smith, T. J. (2017). Two days of calorie deprivation induced by underfeeding and aerobic exercise degrades mood and lowers interstitial glucose but does not impair cognitive function in young adults. *Journal of Nutrition, 147*(1), 110-116.

Lieberman, H. R., Caruso, C. M., Niro, P. J., Adam, G. E., Kellogg, M. D., Nindl, B. C., & Kramer, F. M. (2008). A double-blind, placebo-controlled test of 2 d of calorie deprivation: Effects on cognition, activity, sleep, and interstitial glucose concentrations. *American Journal of Clinical Nutrition, 88*(3), 667-676.

Loucks, A., Verdun, M., & Heath, E. (1998). Low energy availability, not stress of exercise, alters lh pulsatility in exercising women. *Journal of Applied Physiology, 84*(1), 37-46.

Loucks, A. B. (2004). Energy balance and body composition in sports and exercise. *Journal of Sports Sciences, 22*(1), 1-14.

Loucks, A. B. (2006). The response of luteinizing hormone pulsatility to 5 days of low energy availability disappears by 14 years of gynecological age. *The Journal of Clinical Endocrinology and Metabolism, 91*(8), 3158-3164.

Loucks, A. B. (2007). Low energy availability in the marathon and other endurance sports. *Sports Medicine, 37*(4-5), 348-352.

Loucks, A. B. (2013). Energy balance and energy availability. *The encyclopaedia of Sports Medicine: An IOC medical commission publication, 19*, 72-87.

Loucks, A. B., & Heath, E. (1994). Dietary restriction reduces luteinizing hormone (lh) pulse frequency during waking hours and increases lh pulse amplitude during sleep in young menstruating women. *The Journal of Clinical Endocrinology and Metabolism, 78*(4), 910-915.

Loucks, A. B., Kiens, B., & Wright, H. H. (2013). Energy availability in athletes. *Food, Nutrition and Sports Performance III*, 15-24.

Loucks, A. B., & Thuma, J. R. (2003). Luteinizing hormone pulsatility is disrupted at a threshold of energy availability in regularly menstruating women. *The Journal of Clinical Endocrinology and Metabolism, 88*(1), 297-311.

Lowe, D. A., Baltgalvis, K. A., & Greising, S. M. (2010). Mechanisms behind estrogen's beneficial effect on muscle strength in females. *Exercise and Sport Sciences Reviews, 38*(2), 61-67.

Lynch, N. J., & Nimmo, M. A. (1998). Effects of menstrual cycle phase and oral contraceptive use on intermittent exercise. *European Journal of Applied Physiology and Occupational Physiology, 78*(6), 565-572.

Maciejczyk, M., Wiecek, M., Szymura, J., Szygula, Z., Wiecha, S., & Cempla, J. (2014). The influence of increased body fat or lean body mass on aerobic performance. *PloS One, 9*(4), e95797.

Martin, D., Papageorgiou, M., Colgan, H., Bandelow, S., Greeves, J. P., Tang, J. C. Y., Fraser, W. D., Cooper, S. B., Sale, C., & Elliott-Sale, K. J. (2021). The effects of short-term low energy availability, achieved through diet or exercise, on cognitive function in oral contraceptive users and eumenorrheic women. *Applied Physiology Nutrition and Metabolism, 46*(7), 781-789.

Martin, D., Sale, C., Cooper, S. B., & Elliott-Sale, K. J. (2018). Period prevalence and perceived side effects of hormonal contraceptive use and the menstrual cycle in elite athletes. *International Journal of Sports Physiology and Performance, 13*(7), 926-932.

Martin, J. G., & Buono, M. J. (1997). Oral contraceptives elevate core temperature and heart rate during exercise in the heat. *Clinical Physiology, 17*(4), 401-408.

Mccallum, S. (2022). Jessica ennis-hill: Why i now train around my period. Accessed in 2022. Retrieved from https://www.bbc.com/news/technology-61591417

Mckay, A. K., Minahan, C., Harris, R., Mccormick, R., Skinner, J., Ackerman, K. E., & Burke, L. M. (2024). Female athlete research camp: A unique model for conducting research in high performance female athletes. *Medicine and Science in Sports and Exercise, 56*, 706-716.

Mckay, A. K., Peeling, P., Pyne, D. B., Tee, N., Whitfield, J., Sharma, A. P., Heikura, I. A., & Burke, L. M. (2022a). Six days of low carbohydrate, not energy availability, alters the iron and immune response to exercise in elite athletes. *Medicine and Science in Sports and Exercise, 54*, 377-387.

Mckay, A. K. A., Stellingwerff, T., Smith, E. S., Martin, D. T., Mujika, I., Goosey-Tolfrey, V. L., Sheppard, J., & Burke, L. M. (2022b). Defining training and performance caliber: A participant classification framework. *International Journal of Sports Physiology and Performance, 17*(2), 317-331.

Mclay, R. T., Thomson, C. D., Williams, S. M., & Rehrer, N. J. (2007). Carbohydrate loading and female endurance athletes: Effect of menstrual-cycle phase. *International Journal of Sport Nutrition and Exercise Metabolism, 17*(2), 189-205.

Mcnamara, A., Harris, R., & Minahan, C. (2022). 'That time of the month'… for the biggest event of your career! Perception of menstrual cycle on performance of australian athletes training for the 2020 olympic and paralympic games. *BMJ Open Sport and Exercise Medicine, 8*(2), e001300.

Mcnulty, K. L., Elliott-Sale, K. J., Dolan, E., Swinton, P. A., Ansdell, P., Goodall, S., Thomas, K., & Hicks, K. M. (2020a). The effects of menstrual cycle phase on exercise performance in eumenorrheic women: A systematic review and meta-analysis. *Sports Medicine*, 1-15.

Mcnulty, K. L., Elliott-Sale, K. J., Dolan, E., Swinton, P. A., Ansdell, P., Goodall, S., Thomas, K., & Hicks, K. M. (2020b). The effects of menstrual cycle phase on exercise performance in eumenorrheic women: A systematic review and meta-analysis. *Sports Medicine, 50*(10), 1813-1827.

Meallogger. Meallogger [Mobile application software]. Retrieved from https://www.meallogger.com/

Melanson, E. L., Maclean, P. S., & Hill, J. O. (2009). Exercise improves fat metabolism in muscle but does not increase 24-h fat oxidation. *Exercise and Sport Sciences Reviews, 37*(2), 93-101.

Meyer, C., Dostou, J. M., Welle, S. L., & Gerich, J. E. (2002). Role of human liver, kidney, and skeletal muscle in postprandial glucose homeostasis. *American Journal of Physiology-Endocrinology and Metabolism, 282*(2), E419-E427.

Miskec, C. M., Potteiger, J. A., Nau, K. L., & Zebas, C. J. (1997). Do varying environmental and menstrual cycle conditions affect anaerobic power output in female athletes? *Journal of Strength and Conditioning Research, 11*(4), 219-223.

Montes, A., Lally, D., & Hale, R. W. (1983). The effects of oral contraceptives on respiration. *Fertility and Sterility, 39*(4), 515-519.

Mountjoy, M., Ackerman, K., Bailey, D., Burke, L., Constantini, N., Hackney, A., Heikura, I., Melin, A., Pensgaard, A., & Stellingwerff, T. (2023). 2023 international olympic committee's (ioc) consensus statement on relative energy deficiency in sport (reds). *British Journal of Sports Medicine, 57*(17).

Myburgh, K. H., Hutchins, J., Fataar, A. B., Hough, S. F., & Noakes, T. D. (1990). Low bone density is an etiologic factor for stress fractures in athletes. *Annals of Internal Medicine, 113*(10), 754-759.

Narang, B. J., Atkinson, G., Gonzalez, J. T., & Betts, J. A. (2020). A tool to explore discrete-time data: The time series response analyser. *International Journal of Sport Nutrition and Exercise Metabolism, 30*(5), 374-381.

Nicklas, B. J., Hackney, A. C., & Sharp, R. L. (1989). The menstrual cycle and exercise: Performance, muscle glycogen, and substrate responses. *International Journal of Sport Nutrition and Exercise Metabolism, 10*(4), 264-269.

Noordhof, D. A., Taylor, M. Y., Topranin, V. D. M., Engseth, T. P., Sandbakk, Ø., & Osborne, J. O. (2024). Agreement between the 2-and 3-step methods for identifying subtle menstrual disturbances. *International Journal of Sports Physiology and Performance, 1*(aop), 1-5.

Nose-Ogura, S., Yoshino, O., Dohi, M., Kigawa, M., Harada, M., Hiraike, O., Onda, T., Osuga, Y., Fujii, T., & Saito, S. (2019). Risk factors of stress fractures due to the female athlete triad: Differences in teens and twenties. *Scandinavian Journal of Medicine and Science in Sports, 29*(10), 1501-1510.

Notelovitz, M., Zauner, C., Mckenzie, L., Suggs, Y., Fields, C., & Kitchens, C. (1987). The effect of low-dose oral contraceptives on cardiorespiratory function, coagulation, and lipids in exercising young women: A preliminary report. *American Journal of Obstetrics and Gynecology, 156*(3), 591-598.

Nuzzo, J. (2021). Volunteer bias and female participation in exercise and sports science research. *Quest, 73*(1), 82-101.

Oester, C., Norris, D., Scott, D., Pedlar, C., Bruinvels, G., & Lovell, R. (2024). Inconsistencies in the perceived impact of the menstrual cycle on sports performance and in the prevalence of menstrual cycle symptoms: A scoping review of the literature. *Journal of Science and Medicine in Sport*.

Oosthuyse, T., & Bosch, A. N. (2010). The effect of the menstrual cycle on exercise metabolism: Implications for exercise performance in eumenorrhoeic women. *Sports Medicine, 40*(3), 207-227.

Oosthuyse, T., & Bosch, A. N. (2012). Oestrogen's regulation of fat metabolism during exercise and gender specific effects. *Current Opinion in Pharmacology, 12*(3), 363-371.

Oosthuyse, T., Bosch, A. N., & Jackson, S. (2005). Cycling time trial performance during different phases of the menstrual cycle. *European Journal of Applied Physiology, 94*(3), 268-276.

Otaka, M., Chen, S.-M., Zhu, Y., Tsai, Y.-S., Tseng, C.-Y., Fogt, D. L., Lim, B.-H., Huang, C.-Y., & Kuo, C.-H. (2018). Does ovulation affect performance in tennis players? *BMJ Open Sport and Exercise Medicine, 4*(1), e000305.

Owen, C., Till, K., Weakley, J., & Jones, B. (2020). Testing methods and physical qualities of male age grade rugby union players: A systematic review. *PloS One, 15*(6), e0233796.

Owen, N. J., Watkins, J., Kilduff, L. P., Bevan, H. R., & Bennett, M. A. (2014). Development of a criterion method to determine peak mechanical power output in a countermovement jump. *Journal of Strength and Conditioning Research, 28*(6), 1552-1558.

Oxfeldt, M., Dalgaard, L. B., Jorgensen, A. A., & Hansen, M. (2020). Hormonal contraceptive use, menstrual dysfunctions, and self-reported side effects in elite athletes in denmark. *International Journal of Sports Physiology and Performance, 15*(10), 1377-1384.

Pallavi, L. C., Uj, D. S., & Shivaprakash, G. (2017). Assessment of musculoskeletal strength and levels of fatigue during different phases of menstrual cycle in young adults. *Journal of Clinical and Diagnostic Research, 11*(2), CC11-CC13.

Papageorgiou, M., Dolan, E., Elliott-Sale, K. J., & Sale, C. (2018a). Reduced energy availability: Implications for bone health in physically active populations. *European Journal of Nutrition, 57*(3), 847-859.

Papageorgiou, M., Elliott-Sale, K. J., Parsons, A., Tang, J. C., Greeves, J. P., Fraser, W. D., & Sale, C. (2017). Effects of reduced energy availability on bone metabolism in women and men. *Bone, 105*, 191-199.

Papageorgiou, M., Martin, D., Colgan, H., Cooper, S., Greeves, J. P., Tang, J. C., Fraser, W. D., Elliott-Sale, K. J., & Sale, C. (2018b). Bone metabolic responses to low energy availability achieved by diet or exercise in active eumenorrheic women. *Bone, 114*, 181-188.

Peronnet, F., & Massicotte, D. (1991). Table of nonprotein respiratory quotient: An update. *Canadian Journal of Sport Sciences, 16*(1), 23-29.

Pessali-Marques, B., Burden, A. M., Morse, C. I., & Onambélé-Pearson, G. L. (2024). Musculoskeletal morphology and joint flexibility-associated functional characteristics across three time points during the menstrual cycle in female contemporary dancers. *Journal of Functional Morphology and Kinesiology, 9*(1), 38.

Peters, C., & Burrows, M. (2006). Androgenicity of the progestin in oral contraceptives does not affect maximal leg strength. *Contraception, 74*(6), 487-491.

Petrofsky, J. S., Ledonne, D. M., Rinehart, J. S., & Lind, A. R. (1976). Isometric strength and endurance during the menstrual cycle. *European Journal of Applied Physiology, 35*(1), 1-10.

Phillips, S. K., Sanderson, A. G., Birch, K., Bruce, S. A., & Woledge, R. C. (1996). Changes in maximal voluntary force of human adductor pollicis muscle during the menstrual cycle. *The Journal of Physiology, 496 (Pt 2)*(2), 551-557.

Pugh, J. N., Sparks, A. S., Doran, D. A., Fleming, S. C., Langan-Evans, C., Kirk, B., Fearn, R., Morton, J. P., & Close, G. L. (2019). Four weeks of probiotic supplementation reduces gi symptoms during a marathon race. *European Journal of Applied Physiology, 119*, 1491-1501.

Quadagno, D., Faquin, L., Lim, G.-N., Kuminka, W., & Moffatt, R. (1991). The menstrual cycle: Does it affect athletic performance? *The Physician and Sportsmedicine, 19*(3), 121-124.

Rael, B., Romero-Parra, N., Alfaro-Magallanes, V. M., Barba-Moreno, L., Cupeiro, R., Janse De Jonge, X., Peinado, A. B., & Iron, F. S. G. (2021). Body composition over the menstrual and oral contraceptive cycle in trained females. *International Journal of Sports Physiology and Performance, 16*(3), 375-381.

Randell, R. K., Rollo, I., Roberts, T. J., Dalrymple, K. J., Jeukendrup, A. E., & Carter, J. M. (2017). Maximal fat oxidation rates in an athletic population. *Medicine and Science in Sports and Exercise, 49*(1), 133-140.

Rechichi, C., & Dawson, B. (2009). Effect of oral contraceptive cycle phase on performance in team sport players. *Journal of Science and Medicine in Sport, 12*(1), 190-195.

Rechichi, C., & Dawson, B. (2012). Oral contraceptive cycle phase does not affect 200-m swim time trial performance. *Journal of Strength and Conditioning Research, 26*(4), 961-967.

Rechichi, C., Dawson, B., & Goodman, C. (2009). Athletic performance and the oral contraceptive. *International Journal of Sports Physiology and Performance, 4*(2), 151-162.

Redman, L. M., & Loucks, A. B. (2005). Menstrual disorders in athletes. *Sports Medicine, 35*(9), 747-755.

Redman, L. M., & Weatherby, R. P. (2004). Measuring performance during the menstrual cycle: A model using oral contraceptives. *Medicine and Science in Sports and Exercise, 36*(1), 130-136.

Reed, B., & Carr, B. (2000). The normal menstrual cycle and the control of ovulation. *Inc.: South Dartmouth, MA, USA*.

Rickenlund, A., Carlstrom, K., Ekblom, B., Brismar, T. B., Von Schoultz, B., & Hirschberg, A. L. (2004). Effects of oral contraceptives on body composition and physical performance in female athletes. *Journal of Clinical Endocrinology & Metabolism, 89*(9), 4364-4370.

Rodrigues, P., De Azevedo Correia, M., & Wharton, L. (2019). Effect of menstrual cycle on muscle strength. *Journal of Exercise Physiology Online, 22*(5), 89-96.

Rogers, S., & Baker, M. (1996). Thermoregulation during exercise in women who are taking oral contraceptives. *European Journal of Applied Physiology, 75*(1), 34-38.

Romero-Moraleda, B., Coso, J. D., Gutierrez-Hellin, J., Ruiz-Moreno, C., Grgic, J., & Lara, B. (2019). The influence of the menstrual cycle on muscle strength and power performance. *Journal of Human Kinetics, 68*, 123-133.

Ruby, B. C., Robergs, R. A., Waters, D. L., Burge, M., Mermier, C., & Stolarczyk, L. (1997). Effects of estradiol on substrate turnover during exercise in amenorrheic females. *Medicine and Science in Sports and Exercise, 29*(9), 1160-1169.

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.

Sarwar, R., Niclos, B. B., & Rutherford, O. M. (1996). Changes in muscle strength, relaxation rate and fatiguability during the human menstrual cycle. *The Journal of Physiology, 493 (Pt 1)*(1), 267-272.

Saunders, P. U., Telford, R. D., Pyne, D. B., Cunningham, R., Gore, C. J., Hahn, A. G., & Hawley, J. A. (2004). Improved running economy in elite runners after 20 days of simulated moderate-altitude exposure. *Journal of Applied Physiology, 96*(3), 931-937.

Schoene, R. B., Robertson, H. T., Pierson, D. J., & Peterson, A. P. (1981). Respiratory drives and exercise in menstrual cycles of athletic and nonathletic women. *Journal of Applied Physiology: Respiratory, Environmental and Exercise Physiology, 50*(6), 1300-1305.

Seebauer, M., Fruhwirth, M., & Moser, M. (2002). Changes of respiratory sinus arrhythmia during the menstrual cycle depend on average heart rate. *European Journal of Applied Physiology, 87*(4-5), 309-314.

Shakhlina, L., Roda, O., Kalytka, S., Romaniuk, O., Matskevych, N., & Zakhozhyi, V. (2016). Physical performance during the menstrual cycle of female athletes who specialize in 800 m and 1500 m running. *Journal of Physical Education and Sport, 16*(4), 1345.

Sheel, A. W., Dominelli, P. B., & Molgat-Seon, Y. (2016). Revisiting dysanapsis: Sex-based differences in airways and the mechanics of breathing during exercise. *Experimental Physiology, 101*(2), 213-218.

Šimić, N., & Santini, M. (2012). Verbal and spatial functions during different phases of the menstrual cycle. *Psychiatria Danubina, 24*(1.), 73-79.

Sims, S. T., & Heather, A. K. (2018). Myths and methodologies: Reducing scientific design ambiguity in studies comparing sexes and/or menstrual cycle phases. *Experimental Physiology, 103*(10), 1309-1317.

Slater, G., Townsend, N., Morabito, A., Burke, L., Schultz, C., Brookes, D., Farley, A., Meerkin, J., & Ducker, K. (2022). Practitioner best practice guidelines for dxa assessment of body composition.

Smekal, G., Von Duvillard, S. P., Frigo, P., Tegelhofer, T., Pokan, R., Hofmann, P., Tschan, H., Baron, R., Wonisch, M., Renezeder, K., & Bachl, N. (2007). Menstrual cycle: No effect on exercise cardiorespiratory variables or blood lactate concentration. *Medicine and Science in Sports and Exercise, 39*(7), 1098-1106.

Smith, E. S., Kuikman, M. A., Weakley, J., Tee, N., Mccormick, R., Ackerman, K. E., Elliott-Sale, K. J., Stellingwerff, T., Harris, R., Mckay, A. K. A., & Burke, L. M. (2024a). Effects of 24-hour diet- or exercise-induced energy availability manipulations on substrate utilization and performance. *Medicine and Science in Sports and Exercise, (online ahead of print)*.

Smith, E. S., Mccormick, R., Mckay, A. K. A., Ackerman, K. E., Elliott-Sale, K. J., Stellingwerff, T., Harris, R., & Burke, L. M. (2024b). Perceived negative menstrual cycle symptoms, but not changes in oestrogen or progesterone, are associated with impaired cycling race performance. *Medicine and Science in Sports and Exercise, (online ahead of print)*.

Smith, E. S., Mckay, A. K., Kuikman, M., Ackerman, K. E., Harris, R., Elliott-Sale, K. J., Stellingwerff, T., & Burke, L. M. (2022a). Managing female athlete health: Auditing the representation of female versus male participants among research in supplements to manage diagnosed micronutrient issues. *Nutrients, 14*(16), 3372.

Smith, E. S., Mckay, A. K. A., Ackerman, K. E., Harris, R., Elliott-Sale, K. J., Stellingwerff, T., & Burke, L. M. (2022b). 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), 114-127.

Smith, E. S., Mckay, A. K. A., Kuikman, M., Ackerman, K. E., Harris, R., Elliott-Sale, K. J., Stellingwerff, T., & Burke, L. M. (2022c). Auditing the representation of female versus male athletes in sports science and Sports Medicine research: Evidence-based performance supplements. *Nutrients, 14*(5), 953.

Smith, E. S., Weakley, J., Mckay, A. K. A., Mccormick, R., Tee, N., Kuikman, M. A., Harris, R., Minahan, C., Buxton, S., Skinner, J., Ackerman, K. E., Elliott-Sale, K. J., Stellingwerff, T., & Burke, L. M. (2024c). Minimal influence of the menstrual cycle or hormonal contraceptives on performance in female rugby league athletes. *European Journal of Sport Science, 24*(8), 1067–1078.

Smith, M., Dyson, R., Hale, T., Hamilton, M., Kelly, J., & Wellington, P. (2001). The effects of restricted energy and fluid intake on simulated amateur boxing performance. *International Journal of Sport Nutrition and Exercise Metabolism, 11*(2), 238-247.

Smith, M. J., Adams, L. F., Schmidt, P. J., Rubinow, D. R., & Wassermann, E. M. (2002). Effects of ovarian hormones on human cortical excitability. *Annals of Neurology, 51*(5), 599-603.

Somboonwong, J., Chutimakul, L., & Sanguanrungsirikul, S. (2015). Core temperature changes and sprint performance of elite female soccer players after a 15-minute warm-up in a hot-humid environment. *Journal of Strength and Conditioning Research, 29*(1), 262-269.

Stachon, A. J. (2016). Menstrual changes in body composition of female athletes. *Collegium Antropologicum, 40*(2), 111-122.

Štefanovský, M., Péterová, A., Vanderka, M., & Lengvarský, L. (2016). Influence of selected phases of the menstrual cycle on performance in special judo fitness test and wingate test. *Acta Gymnica, 46*(3), 136-142.

Stellingwerff, T. (2018). Case study: Body composition periodization in an olympic-level female middle-distance runner over a 9-year career. *International Journal of Sport Nutrition and Exercise Metabolism, 28*(4), 428-433.

Stephenson, L. A., Kolka, M. A., & Wilkerson, J. E. (1982). Metabolic and thermoregulatory responses to exercise during the human menstrual cycle. *Medicine and Science in Sports and Exercise, 14*(4), 270-275.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*(6), 643-662.

Suh, S. H., Casazza, G. A., Horning, M. A., Miller, B. F., & Brooks, G. A. (2002). Luteal and follicular glucose fluxes during rest and exercise in 3-h postabsorptive women. *Journal of Applied Physiology, 93*(1), 42-50.

Suh, S. H., Casazza, G. A., Horning, M. A., Miller, B. F., & Brooks, G. A. (2003). Effects of oral contraceptives on glucose flux and substrate oxidation rates during rest and exercise. *Journal of Applied Physiology, 94*(1), 285-294.

Sunderland, C., & Nevill, M. (2003). Effect of the menstrual cycle on performance of intermittent, high-intensity shuttle running in a hot environment. *European Journal of Applied Physiology, 88*(4-5), 345-352.

Sutresna, N. (2016). *Women Athletes' Endurance And Menstruation Cycle; Pre-Menstruation, 2nd Day Of Menstruation And 5th Day Of Menstruation.* Paper presented at the Proceedings of the 6th International Conference on Educational, Management, Administration and Leadership.

Swinton, P. A., Hemingway, B. S., Saunders, B., Gualano, B., & Dolan, E. (2018). A statistical framework to interpret individual response to intervention: Paving the way for personalized nutrition and exercise prescription. *Frontiers in Nutrition, 5*, 41.

Taim, B. C., Lye, J., Suppiah, H. T., Chan, T. W., Chia, M., & Clarke, A. (2024). Menstrual cycle characteristics, perceived impact on performance, and barriers to communication: Perspectives of high-performance adolescent athletes in singapore. *Scandinavian Journal of Medicine and Science in Sports, 34*(1), e14488.

Taim, B. C., Ó Catháin, C., Renard, M., Elliott-Sale, K. J., Madigan, S., & Ní Chéilleachair, N. (2023). The prevalence of menstrual cycle disorders and menstrual cycle-related symptoms in female athletes: A systematic literature review. *Sports Medicine, 53*(10), 1963-1984.

Tarnopolsky, L. J., Macdougall, J. D., Atkinson, S. A., Tarnopolsky, M. A., & Sutton, J. R. (1990). Gender differences in substrate for endurance exercise. *Journal of Applied Physiology, 68*(1), 302-308.

Tasmektepligil, M. Y., Agaoglu, S. A., Turkmen, L., & Turkmen, M. (2010). The motor performance and some physical characteristics of the sportswomen and sedentary lifestyle women during menstrual cycle. *Archives of Budo, 6*(4), 195-203.

Tenan, M. S. (2017). Sex hormone effects on the nervous system and their impact on muscle strength and motor performance in women. In *Sex Hormones, Exercise and Women* (pp. 59-70): Springer.

Tenan, M. S., Hackney, A. C., & Griffin, L. (2016). Maximal force and tremor changes across the menstrual cycle. *European Journal of Applied Physiology, 116*(1), 153-160.

Thein-Nissenbaum, J. M., Rauh, M. J., Carr, K. E., Loud, K. J., & Mcguine, T. A. (2012). Menstrual irregularity and musculoskeletal injury in female high school athletes. *Journal of Athletic Training, 47*(1), 74-82.

Thompson, B. M., Drover, K. B., Stellmaker, R. J., Sculley, D. V., & Janse De Jonge, X. A. (2021). The effect of the menstrual cycle and oral contraceptive cycle on muscle performance and perceptual measures. *International Journal of Environmental Research and Public Health, 18*(20), 10565.

Tounsi, M., Jaafar, H., Aloui, A., & Souissi, N. (2018). Soccer-related performance in eumenorrheic tunisian high-level soccer players: Effects of menstrual cycle phase and moment of day. *The Journal of Sports Medicine and Physical Fitness, 58*(4), 497-502.

Townes, C. (2019). A decade of fighting for equal pay in sports. Accessed in 2021. Retrieved from https://www.forbes.com/sites/ceceliatownes/2020/12/31/a-decade-of-fighting-for-equal-pay/?sh=6557a29a1376

Trexler, E. T., Smith-Ryan, A. E., & Norton, L. E. (2014). Metabolic adaptation to weight loss: Implications for the athlete. *Journal of the International Society of Sports Nutrition, 11*(1), 7.

Tsampoukos, A., Peckham, E. A., James, R., & Nevill, M. E. (2010). Effect of menstrual cycle phase on sprinting performance. *European Journal of Applied Physiology, 109*(4), 659-667.

Vaiksaar, S., Jürimäe, J., Mäestu, J., Purge, P., Kalytka, S., Shakhlina, L., & Jürimäe, T. (2011). No effect of menstrual cycle phase and oral contraceptive use on endurance performance in rowers. *Journal of Strength and Conditioning Research, 25*(6), 1571-1578.

Venables, M. C., Achten, J., & Jeukendrup, A. E. (2005). Determinants of fat oxidation during exercise in healthy men and women: A cross-sectional study. *Journal of Applied Physiology, 98*(1), 160-167.

Vermeulen, A., Verdonck, L., & Kaufman, J. M. (1999). A critical evaluation of simple methods for the estimation of free testosterone in serum. *The Journal of Clinical Endocrinology and Metabolism, 84*(10), 3666-3672.

Weakley, J., Black, G., Mclaren, S., Scantlebury, S., Suchomel, T. J., Mcmahon, E., Watts, D., & Read, D. B. (2022). Testing and profiling athletes: Recommendations for test selection, implementation, and maximizing information. *Strength and Conditioning Journal*, 10.1519.

Weakley, J., Mccosker, C., Chalkley, D., Johnston, R., Munteanu, G., & Morrison, M. (2023). Comparison of sprint timing methods on performance, and displacement and velocity at timing initiation. *Journal of Strength and Conditioning Research, 37*(1), 234-238.

Wiecek, M., Szymura, J., Maciejczyk, M., Cempla, J., & Szygula, Z. (2016). Effect of sex and menstrual cycle in women on starting speed, anaerobic endurance and muscle power. *Acta Physiologica Hungarica, 103*(1), 127-132.

Wirth, J. C., & Lohman, T. G. (1982). The relationship of static muscle function to use of oral-contraceptives. *Medicine and Science in Sports and Exercise, 14*(1), 16-20.

Zderic, T. W., Coggan, A. R., & Ruby, B. C. (2001). Glucose kinetics and substrate oxidation during exercise in the follicular and luteal phases. *Journal of Applied Physiology, 90*(2), 447-453.

Zhang, H., Huizenga, C., Arens, E., & Wang, D. (2004). Thermal sensation and comfort in transient non-uniform thermal environments. *European Journal of Applied Physiology, 92*, 728-733.

Zwift. (2024). Category enforcement faq. Accessed in 2021. Retrieved from https://support.zwift.com/en_us/category-enforcement-faq-rJ5CQrpvc?fbclid=IwAR3SOjO4ESxyZy2D1Da2bLBYlb9SwlUtt4gSBtUjQBdvaqbsguE4bWOsGT8

Chapter 9: Research Portfolio Appendix

9.1 Statement of contributions

Study 1

Smith, E. S., Kuikman, M. A., Weakley, J., Tee, N., McCormick, R., Ackerman, K. E., Elliott-Sale, K. J., Stellingwerff, T., Harris, R., McKay, A.K.A., Burke, L. M. Effects of 24-hour diet- or exercise-induced energy availability manipulations on substrate utilization and performance. *Medicine and Science in Sports and Exercise*, (aop). https://doi.org/10.1249/ MSS.0000000000003608.

Contribution statement: ES contributed to the conception and design, collection and assembly of data, data analysis and interpretation, drafting, revising and approval of the final manuscript. MK contributed to the conception and design, collection and assembly of data, revising and approval of the final manuscript. JW contributed to the collection and assembly of data, data analysis and interpretation, revising and approval of the final manuscript. NT contributed to the collection and assembly of data, revising and approval of the final manuscript. RM contributed to the collection and assembly of data, revising and approval of the final manuscript. KA contributed to the data interpretation and revising and approval of the final manuscript. KES contributed to the data interpretation and revising and approval of the final manuscript. TS contributed to the data interpretation and revising and approval of the final manuscript. RH contributed to the data interpretation and revising and approval of the final manuscript. AM contributed to the conception and design, collection and assembly of data, data analysis and interpretation, revising and approval of the final manuscript. LB contributed to the conception and design, collection and assembly of data, data analysis and interpretation, revising and approval of the final manuscript.

Approximate percentage contributions: E.S. Smith 60%; M.A. Kuikman 10%; J. Weakley 5%; N. Tee 2.5%; R. McCormick 2.5%; K.E. Ackerman 2.5%; K.J. Elliott-Sale 2.5%; T. Stellingwerff 2.5%; R. Harris 2.5%; A.K.A. McKay 5%; L.M. Burke 5%.

I acknowledge that my contribution to the above paper is above 50%

E.S. Smith 16/09/2024

As principal supervisor of this project, I certify that the above contributions are true and correct:

L.M. Burke 16/09/2024

Co-authors:

16/09/2024

M.A. Kuikman

16/09/2024

J. Weakley

18/09/2024

N. Tee

17/09/2024

R. McCormick

16/09/2024

K.E. Ackerman

16/09/2024

K.J. Elliott-Sale

15/09/2024

T. Stellingwerff

17/09/2024

R. Harris

16/09/2024

A.K.A. McKay

Study 2

Smith, E. S., Weakley, J., McKay, A. K. A., McCormick, R., Tee, N., Kuikman, M. A., Harris, R., Minahan, C., Buxton, S., Skinner, J., Ackerman, K. E., Elliott-Sale, K. J., Stellingwerff, T., Burke, L. M. (2024). Minimal influence of the menstrual cycle or hormonal contraceptives on performance in female rugby league athletes. *European Journal of Sport Science, 24*(8), 1067–1078.

Contribution statement: ES contributed to the conception and design, collection and assembly of data, data analysis and interpretation, drafting, revising and approval of the final manuscript. JW contributed to the collection and assembly of data, data analysis and interpretation, revising and approval of the final manuscript. AM contributed to the conception and design, collection and assembly of data, data analysis and interpretation, revising and approval of the final manuscript. RM contributed to the collection and assembly of data, revising and approval of the final manuscript. NT contributed to the collection and assembly of data, revising and approval of the final manuscript. MK contributed to the collection and assembly of data, revising and approval of the final manuscript. RH contributed to the conception and design, data interpretation and revising and approval of the final manuscript. CM contributed to the conception and design, data interpretation and revising and approval of the final manuscript. SB contributed to revising and approval of the final manuscript. JS contributed to revising and approval of the final manuscript. KA contributed to the data interpretation and revising and approval of the final manuscript. KES contributed to the data interpretation and revising and approval of the final manuscript. TS contributed to the data interpretation and revising and approval of the final manuscript. LB contributed to the conception and design, collection and assembly of data, data analysis and interpretation, revising and approval of the final manuscript.

Approximate percentage contributions: E.S. Smith 60%; J. Weakley 5%; A.K.A. McKay 5%; R. McCormick 2.5%; N. Tee 2.5%; M.A. Kuikman 2.5%; R. Harris 2.5%; C. Minahan 2.5%; S. Buxton 2.5%; J. Skinner 2.5%; K.E. Ackerman 2.5%; K.J. Elliott-Sale 2.5%; T. Stellingwerff 2.5%; L.M. Burke 5%.

I acknowledge that my contribution to the above paper is above 50%

E.S. Smith 16/09/2024

As principal supervisor of this project, I certify that the above contributions are true and correct:

L.M. Burke 16/09/2024

Co-authors:

16/09/2024

J. Weakley

16/09/2024

A.K.A. McKay

17/09/2024

R. McCormick

18/09/2024

N. Tee

16/09/2024

M.A. Kuikman

17/09/2024

R. Harris

16/09/2024

C. Minahan

17/09/2024

S. Buxton

16/09/2024

J. Skinner

16/09/2024

K.E. Ackerman

16/09/2024

K.J. Elliott-Sale

15/09/2024

T. Stellingwerff

Study 3

Smith, E. S., McCormick, R., McKay, A. K. A., Ackerman, K. E., Elliott-Sale, K. J., Stellingwerff, T., Harris, R., Burke, L. M.. Perceived negative menstrual cycle symptoms, but not changes in oestrogen or progesterone, are associated with impaired cycling race performance. *Medicine and Science in Sports and Exercise*, (aop). https://doi.org/10.1249/ MSS.0000000000003587.

Contribution statement: ES contributed to the conception and design, collection and assembly of data, data analysis and interpretation, drafting, revising and approval of the final manuscript. RM contributed to the conception and design, collection and assembly of data, revising and approval of the final manuscript. AM contributed to the conception and design, data analysis and interpretation, revising and approval of the final manuscript. KA contributed to the data interpretation and revising and approval of the final manuscript. KES contributed to the data interpretation and revising and approval of the final manuscript. TS contributed to the data interpretation and revising and approval of the final manuscript. RH contributed to the data interpretation and revising and approval of the final manuscript. LB contributed to the conception and design, data analysis and interpretation, revising and approval of the final manuscript.

Approximate percentage contributions: E.S. Smith 75%; R. McCormick 5%; A.K.A. McKay 5%; K.E. Ackerman 2.5%; K.J. Elliott-Sale 2.5%; T. Stellingwerff 2.5%; R. Harris 2.5%; L.M. Burke 5%.

I acknowledge that my contribution to the above paper is above 50%

E.S. Smith 16/09/2024

As principal supervisor of this project, I certify that the above contributions are true and correct:

L.M. Burke 16/09/2024

Co-authors:

17/09/2024

R. McCormick

16/09/2024

A.K.A. McKay

16/09/2024

K.E. Ackerman

16/09/2024

K.J. Elliott-Sale

15/09/2024

T. Stellingwerff

17/09/2024

R. Harris

Study 4

Smith, E. S., McKay, A. K. A., Ackerman, K. E., Elliott-Sale, K. J., Stellingwerff, T., Harris, R., Burke, L. M. Original investigation: Female athletes report positive experiences as research participants. *International journal of sports nutrition and exercise metabolism, (in review)*.

Contribution statement: ES contributed to the conception and design, collection and assembly of data, data analysis and interpretation, drafting, revising and approval of the final manuscript. AM contributed to the conception and design, data analysis and interpretation, revising and approval of the final manuscript. KA contributed to the data interpretation and revising and approval of the final manuscript. KES contributed to the data interpretation and revising and approval of the final manuscript. TS contributed to the data interpretation and revising and approval of the final manuscript. RH contributed to the data interpretation and revising and approval of the final manuscript. LB contributed to the conception and design, data analysis and interpretation, revising and approval of the final manuscript.

Approximate percentage contributions: E.S. Smith 80%; A.K.A. McKay 5%; K.E. Ackerman 2.5%; K.J. Elliott-Sale 2.5%; T. Stellingwerff 2.5%; R. Harris 2.5%; L.M. Burke 5%.

I acknowledge that my contribution to the above paper is above 50%

E.S. Smith 16/09/2024

As principal supervisor of this project, I certify that the above contributions are true and correct:

L.M. Burke 16/09/2024

Co-authors:

16/09/2024

A.K.A. McKay

16/09/2024

K.E. Ackerman

16/09/2024

K.J. Elliott-Sale

15/09/2024

T. Stellingwerff

17/09/2024

R. Harris

9.2 Appendices to publications

Study 1

Table S3.1. The type, formulation, brand, and length of usage of oral contraceptive pill across n=10 female participants.

Pill brand	Pill type (mono-/bi-/tri-phasic)	Pill formulation	Length of usage (years)
Estelle 35 ED	Monophasic	35 µg ethinyloestradiol, 2 mg cyproterone acetate	6
Evelyn ED 150/30	Monophasic	30 µg ethinyloestradiol, 150 µg levonorgestrel	24
Femme-Tab 20/100	Monophasic	20 µg ethinyloestradiol, 100 µg levonorgestrel	1
Isabelle	Monophasic	30 µg ethinyloestradiol, 3mg drospirenone	0.3
Levlen ED 150/30	Monophasic	30 µg ethinyloestradiol, 150 µg levonorgestrel	10
Levlen ED 150/30	Monophasic	30 µg ethinyloestradiol, 150 µg levonorgestrel	5
Levlen ED 150/30	Monophasic	30 µg ethinyloestradiol, 150 µg levonorgestrel	3
Micronelle 30 ED	Monophasic	30 µg ethinyloestradiol, 150 µg levonorgestrel	20
Yang 20/3	Monophasic	20 µg ethinyloestradiol, 3mg drospirenone	0.3
Yasmin	Monophasic	30 µg ethinyloestradiol, 3mg drospirenone	1.5
		Median	**4.0**
		Min	**0.3**
		Max	**24.0**

Length of usage taken from the time of commencing the study.

Table S3.2. The 7-point Likert scale used to assess lower limb muscle soreness (Impellizzeri & Maffiuletti, 2007)

Please indicate how your leg muscles feel at this present moment.

0	A complete absence of soreness
1	A light soreness in the muscle felt only when touched/ a vague ache
2	A moderate soreness/ pain felt only when touched/ a slight persistent ache/ pain
3	A light muscle soreness/ pain when walking up or down stairs
4	A light muscle soreness, pain, stiffness, or weakness when walking on a flat surface
5	A moderate muscle soreness, pain, stiffness, or weakness when walking
6	A severe muscle soreness, stiffness or weakness that limits my ability to move

Table S3.3. Participant fat free mass as measured by DXA scan (participants fasted), and body mass as measured on the force plate by ForceDecks software (participants fed), across all five conditions of energy availability. Results are averaged across sexes.

Condition	Fat free mass (kg)	*p value (vs LEA_{REST})*	*p value (vs LEA_{EX})*	Body mass (kg)	*p value (vs LEA_{REST})*
LEA_{REST}	54.6±11.0	-	*0.733*	73.3±13.7	-
LEA_{EX}	54.8±11.1	*0.733*	-	73.7±14.0	*0.447*
HEA_{REST}	55.2±11.0	***0.022****	*0.339*	74.0±14.1	***0.030****
HEA_{EX}	55.5±11.2	***0.0004*****	***0.021****	73.9±13.8	*0.056*
GEA	55.5±11.3	***0.0002*****	***0.011****	74.2±14.1	***0.0006*****

*Values displayed as mean±standard deviation. LEA, low energy availability; HEA, high energy availability; GEA, high energy availability for mass gain/growth. *denotes significances $p<0.05$, ***denotes significance $p<0.001$.*

Table S3.4. Energy expenditure, carbohydrate oxidation and test duration of the FATMAX test across all five conditions of energy availability. Results are averaged across sexes.

Condition	Energy expenditure (kcal)	Carbohydrate oxidation ($g \cdot min^{-1}$)	Test duration (min: seconds)
LEA_{REST}	200±90 [c]	29.2±13.7	22:12 [c e]
LEA_{EX}	217±96 [c e]	28.9±14.0	22:48 [c e]
HEA_{REST}	178±96 [a b d]	28.2±14.6	20:30 [a b]
HEA_{EX}	205±101 [c]	29.5±13.7	22:00 [e]
GEA	182±89 [b]	31.7±15.0	20:12 [a b d]

Values displayed as mean±standard deviation. LEA, low energy availability; HEA, high energy availability; GEA, high energy availability for mass gain/growth. [a] significantly different from LEA_{REST} condition, [b] significantly different from LEA_{EX} condition, [c] significantly different from HEA_{REST} condition, [d] significantly different from HEA_{EX} condition, [e] significantly different from GEA condition.

Table S3.5. Outcome measures assessed during the countermovement jump across all five conditions of energy availability and between sexes.

		LEA_{REST}	LEA_{EX}	HEA_{REST}	HEA_{EX}	GEA	*p value (main effect of condition)*	*p value (main effect of sex)*	*p value (sex*condition interaction)*
Jump height (cm)							*0.158*	***0.002****	*0.503*
	Females	21.1±4.3	21.8±4.9	21.3±5.0	21.3±5.2	20.8±4.9			
	Males	29.9±5.9	29.4±5.6	29.6±5.4	29.5±5.2	28.8±5.6			
	Average	25.5±6.7	25.4±6.6	25.4±6.6	25.4±6.6	24.8±6.6			
Velocity at take-off ($m \cdot s^{-1}$)							*0.217*	***0.003****	*0.567*
	Females	2.03±0.22	2.06±0.25	2.05±0.25	2.03±0.28	2.01±0.24			
	Males	2.42±0.24	2.40±0.23	2.40±0.23	2.41±0.21	2.38±0.23			
	Average	2.22±0.30	2.23±0.29	2.23±0.30	2.22±0.31	2.20±0.30			
Peak velocity ($m \cdot s^{-1}$)							*0.069*	***0.001****	*0.530*
	Females	2.18±0.19	2.20±0.22	2.20±0.21	2.18±0.23	2.16±0.21			
	Males	2.56±0.21	2.53±0.21	2.54±0.20	2.54±0.19	2.51±0.21			
	Average	2.37±0.27	2.36±0.27	2.37±0.27	2.36±0.28	2.33±0.27			
Mean velocity ($m \cdot s^{-1}$)							***0.029****	***<0.0004****	*0.602*
	Females	1.17±0.12	1.12±0.09	1.12±0.09	1.14±0.16	1.13±0.16			
	Males	1.37±0.10	1.37±0.07	1.37±0.07	1.37±0.08	1.32±0.08			
	Average	**1.27±0.15^**	1.25±0.15	1.25±0.15	1.25±0.17	**1.23±0.15^**			
Relative peak power ($W \cdot kg^{-1}$)							***0.033***$^{@}$	***0.014****	*0.780*
	Females	36.6±6.5	36.5±5.8	35.6±5.4	35.6±5.6	35.4±5.7			
	Males	44.2±6.7	43.9±6.7	43.4±6.4	43.5±6.7	43.2±6.7			

	Average	40.4±7.5	40.2±7.2	39.7±7.0	39.5±7.3	39.5±7.3			
Relative mean power ($W \cdot kg^{-1}$)							*0.109*	***0.009*****	*0.984*
	Females	2.03±0.80	1.83±0.60	1.58±0.58	1.82±0.70	1.66±0.78			
	Males	2.68±0.74	2.54±0.63	2.38±0.53	2.55±0.79	2.45±0.74			
	Average	2.35±0.82	2.19±0.70	1.98±0.68	2.18±0.82	2.06±0.84			
Relative peak force ($N \cdot kg^{-1}$)							*0.715*	*0.075*	*0.676*
	Females	19.6±1.3	20.0±2.0	19.3±1.8	19.4±2.1	19.9±2.1			
	Males	21.0±1.7	21.3±1.6	21.4±1.3	21.0±1.1	20.8±1.9			
	Average	20.3±1.7	20.7±1.9	20.4±1.8	20.2±1.8	20.4±2.0			
Relative mean force ($N \cdot kg^{-1}$)							*0.136*	*0.383*	*0.377*
	Females	12.4±0.7	12.2±0.5	12.1±0.4	12.1±0.5	12.2±0.4			
	Males	12.5±0.6	12.3±0.6	12.3±0.5	12.5±0.6	12.3±0.6			
	Average	12.4±0.7	12.2±0.5	12.2±0.4	12.3±0.6	12.2±0.5			
Rate of force development at 50 ms ($N \cdot s^{-1}$)							*0.265*	*0.798*	*0.962*
	Females	3590±1804	3056±2234	3370±1997	3635±1959	4477±3247			
	Males	5086±3589	3494±2465	2892±1412	4127±2282	3385±1838			
	Average	4377±2912	3275±2301	3131±1701	3881±2086	3931±2628			
Rate of force development at 100 ms ($N \cdot s^{-1}$)							*0.435*	*0.274*	*0.647*
	Females	3143±1178	2952±1839	2854±1326	3038±1567	3746±2438			
	Males	4608±2509	3547±1961	3332±1438	3975±1771	3550±1864			
	Average	3914±2081	3249±1875	3093±1369	3507±1697	3648±2114			
Rate of force development at 150 ms ($N \cdot s^{-1}$)							*0.847*	*0.050*	*0.861*
	Females	3359±1667	2706±1480	2749±1180	2842±1296	3149±1693			

	Males	4003±1605	4042±2108	3619±1450	4116±1417	3733±1749			
	Average	3681±1627	3374±1901	3184±1362	3479±1474	3441±1702			
Rate of force development at 200 ms ($N \cdot s^{-1}$)							*0.896*	***0.0003****	*0.929*
	Females	2519±746	2324±979	2281±942	2476±878	2414±1126			
	Males	3685±957	3854±1825	3648±1010	4005±988	3457±1597			
	Average	3102±1027	3089±1627	2965±1181	3240±1202	2936±1448			
Impulse at 50 ms (N·s)							*0.403*	***0.0007****	*0.397*
	Females	64.0±17.9	57.2±7.5	59.7±10.4	58.6±10.6	59.2±10.8			
	Males	79.7±11.5	80.5±16.9	80.6±13.2	80.5±12.5	78.6±11.0			
	Average	71.8±16.8	68.9±17.5	70.1±15.7	69.6±15.9	68.9±14.5			
Impulse at 100 ms (N·s)							*0.324*	***0.0006****	*0.186*
	Females	126.7±30.1	114.4±14.2	117.3±19.7	116.6±19.6	119.1±20.9			
	Males	158.7±25.2	160.5±32.7	159.8±26.0	159.8±25.6	157.5±23.3			
	Average	142.7±31.6	137.4±34.1	138.5±31.3	138.2±31.4	138.3±29.2			
Impulse at 150 ms (N·s)							*0.418*	***0.0007****	*0.483*
	Females	182.8±33.1	171.8±22.5	173.6±28.6	172.8±27.0	177.4±31.1			
	Males	234.2±38.0	234.7±45.8	233.1±38.6	234.5±38.7	233.0±35.3			
	Average	208.5±43.6	203.2±47.7	203.3±45.0	203.6±45.3	205.2±43.2			
Impulse at 200 ms (N·s)							*0.421*	***0.001****	*0.766*
	Females	234.2±36.8	231.6±34.0	230.4±38.5	229.7±34.7	237.9±42.5			
	Males	306.9±50.3	303.6±49.9	303.5±50.2	306.4±50.7	306.4±46.8			
	Average	272.5±57.1	267.6±55.6	266.9±57.5	268.1±57.8	272.2±55.9			
Total impulse (N·s)							*0.081*	***0.002****	*0.608*

	Females	320.3±77.1	334.8±72.8	341.0±68.4	343.3±77.6	325.4±57.9			
	Males	448.6±91.5	452.8±102.1	461.1±91.8	454.8±81.4	453.5±97.5			
	Average	384.4±105.4	393.8±105.4	401.1±100.0	399.0±96.3	389.5±102.0			
Flight time: contraction time (s)							*0.097*	*0.493*	*0.293*
	Females	0.55±0.14	0.50±0.10	0.49±0.09	0.48±0.11	0.51±0.09			
	Males	0.56±0.12	0.52±0.12	0.52±0.10	0.56±0.12	0.53±0.12			
	Average	0.55±0.13	0.51±0.11	0.50±0.09	0.52±0.12	0.52±0.11			
Contraction time (s)							*0.208*	*0.188*	*0.336*
	Females	0.82±0.13	0.89±0.16	0.89±0.11	0.92±0.12	0.85±0.09			
	Males	0.93±0.17	0.94±0.19	1.00±0.17	0.91±0.14	0.93±0.16			
	Average	0.87±0.16	0.91±0.17	0.94±0.15	0.92±0.13	0.89±0.13			
Concentric time (s)							*0.209*	*0.118*	*0.203*
	Females	0.53±0.08	0.58±0.13	0.58±0.07	0.60±0.09	0.56±0.08			
	Males	0.61±0.11	0.67±0.20	0.66±0.13	0.59±0.08	0.67±0.19			
	Average	0.57±0.10	0.63±0.17	0.62±0.11	0.59±0.09	0.61±0.15			
Eccentric time (s)							*0.098*	*0.430*	*0.675*
	Females	0.29±0.06	0.31±0.05	0.32±0.05	0.32±0.05	0.30±0.03			
	Males	0.32±0.06	0.32±0.07	0.33±0.07	0.33±0.06	0.32±0.07			
	Average	0.30±0.06	0.32±0.06	0.32±0.06	0.32±0.05	0.31±0.06			
Centre of mass displacement (cm)							***0.003***	***0.018***	*0.933*
	Females	-29.1±4.6	-28.3±3.6	-30.6±4.5	-29.9±3.8	-27.7±2.9			
	Males	-36.1±7.4	-35.1±9.0	-38.2±9.5	-38.1±9.1	-34.5±8.5			
	Average	-32.8±7.1	**-31.7±7.6**$^{\#}$	**-34.4±8.2**$^{\wedge\#}$	**-34.0±8.0**$^{\lambda}$	**-31.1±7.1**$^{\wedge\lambda}$			

Eccentric utilisation ratio							*0.205*	*0.232*	*0.452*
	Females	1.06±0.07	1.08±0.13	1.06±0.07	1.01±0.07	1.06±0.07			
	Males	1.06±0.09	1.09±0.09	1.12±0.14	1.09±0.07	1.10±0.07			
	Average	1.06±0.08	1.09±0.11	1.09±0.11	1.04±0.08	1.08±0.07			
Reactive strength index							*0.171*	***0.045****	*0.691*
	Females	0.010±0.01	0.012±0.02	0.012±0.01	0.003±0.02	0.011±0.01			
	Males	0.019±0.03	0.024±0.03	0.032±0.03	0.021±0.02	0.028±0.02			
	Average	0.015±0.02	0.018±0.02	0.022±0.03	0.012±0.02	0.020±0.02			
Dynamic strength index							*0.634*	*0.504*	*0.318*
	Females	0.73±0.19	0.70±0.14	0.70±0.11	0.68±0.11	0.70±0.12			
	Males	0.65±0.15	0.65±0.15	0.67±0.17	0.66±0.13	0.69±0.17			
	Average	0.69±0.17	0.67±0.14	0.69±0.14	0.67±0.12	0.70±0.14			

*Values displayed as mean±standard deviation. LEA, low energy availability; HEA, high energy availability; GEA, high energy availability for mass gain/growth. *denotes significance between the sexes/conditions p<0.05, **denotes significant difference between the sexes/conditions p<0.01, ***denotes significant difference between the sexes/conditions p<0.001. [@]denotes no significant differences between conditions in post-hoc testing. ^/[#]/[λ] significantly different from other condition with a matching symbol.*

Table S3.6. Outcome measures assessed during the squat jump across all five conditions of energy availability and between sexes.

		LEA_{REST}	LEA_{EX}	HEA_{REST}	HEA_{EX}	GEA	*p value (main effect of condition)*	*p value (main effect of sex)*	*p value (sex*condition interaction)*
Jump height (cm)							***0.008****	***0.004****	*0.282*
	Females	19.8±4.7	21.0±6.0	20.1±4.4	21.0±4.8	19.7±5.0			
	Males	28.0±4.2	27.0±4.5	26.4±4.3	27.4±4.3	26.0±4.1			
	Average	**24.1±6.0^**	24.2±6.0	23.2±5.3	**24.2±5.5$^{\#}$**	**22.9±5.5^$^{\#}$**			
Velocity at take-off ($m \cdot s^{-1}$)							***0.039****	***0.006****	*0.541*
	Females	1.97±0.25	2.01±0.32	1.98±0.23	2.02±0.27	1.96±0.25			
	Males	2.35±0.17	2.30±0.19	2.28±0.18	2.32±0.18	2.27±0.17			
	Average	2.17±0.28	2.17±0.29	2.13±0.25	**2.17±0.27^**	**2.11±0.26^**			
Peak velocity ($m \cdot s^{-1}$)							***0.016****	***0.002****	*0.354*
	Females	2.12±0.20	2.18±0.28	2.14±0.20	2.18±0.21	2.12±0.21			
	Males	2.49±0.15	2.44±0.18	2.43±10.7	2.46±0.16	2.41±0.15			
	Average	**2.32±0.26^**	2.32±0.26	2.29±0.23	**2.32±0.23$^{\#}$**	**2.26±0.23^$^{\#}$**			
Mean velocity ($m \cdot s^{-1}$)							*0.400*	***0.0006*****	*0.575*
	Females	0.76±0.10	0.78±0.20	0.80±0.10	0.83±0.15	0.75±0.12			
	Males	0.97±0.09	0.95±0.06	0.92±0.13	0.97±0.11	0.93±0.06			
	Average	0.87±0.14	0.87±0.16	0.86±0.13	0.90±0.15	0.84±0.13			
Relative peak power ($W \cdot kg^{-1}$)							*0.051*	***0.040****	*0.216*
	Females	35.3±5.9	36.4±6.7	35.7±5.8	36.4±5.8	35.0±5.7			
	Males	42.3±5.3	41.4±5.8	41.0±4.8	40.8±5.0	40.6±5.7			

	Average	39.0±6.5	39.0±6.6	38.4±5.9	38.6±5.7	37.8±6.3			
Relative mean power ($W \cdot kg^{-1}$)							*0.269*	***0.004****	*0.571*
	Females	10.8±1.8	11.3±3.9	11.5±2.0	12.0±3.2	10.5±2.3			
	Males	14.3±1.5	14.0±1.2	13.4±2.4	14.3±2.0	13.6±1.1			
	Average	12.7±2.4	12.7±3.1	12.5±2.4	13.2±2.8	12.1±2.4			
Relative peak force ($N \cdot kg^{-1}$)							*0.619*	*0.624*	*0.583*
	Females	19.6±2.4	19.8±1.7	19.7±2.5	19.7±2.6	19.3±1.8			
	Males	19.4±1.7	19.3±1.6	19.2±1.3	18.9±1.4	19.2±1.8			
	Average	19.5±2.0	19.5±1.6	19.4±2.0	19.3±2.1	19.2±1.8			
Relative mean force ($N \cdot kg^{-1}$)							*0.406*	*0.259*	*0.379*
	Females	14.0±1.2	14.0±1.6	14.2±1.3	14.3±1.7	13.7±1.2			
	Males	14.7±0.8	14.6±0.7	14.4±0.8	14.6±0.7	14.5±0.7			
	Average	14.3±1.0	14.3±1.2	14.3±1.1	14.4±1.3	14.1±1.0			
Rate of force development at 50 ms ($N \cdot s^{-1}$)							*0.913*	*0.198*	*0.233*
	Females	1351±1188	1604±1747	1552±1013	1942±1631	1109±867			
	Males	1930±1302	2070±1313	1304±885	1836±1686	2108±1539			
	Average	1655±1250	1849±1509	1434±936	1892±1612	1609±1319			
Rate of force development at 100 ms ($N \cdot s^{-1}$)							*0.854*	*0.058*	*0.328*
	Females	1666±1374	1868±1857	1954±1017	2337±1852	1331±1043			
	Males	2709±1464	2696±1490	2307±1592	2776±1766	2692±1570			
	Average	2215±1482	2304±1680	2131±1313	2556±1776	2011±1473			
Rate of force development at 150 ms ($N \cdot s^{-1}$)							*0.692*	***0.044****	*0.449*
	Females	1638±1060	1667±1334	1988±960	2151±1592	1490±982			

	Males	2773±1069	2639±997	2356±1135	2662±1087	2567±921			
	Average	2236±1187	2179±1240	2172±1040	2406±1352	2028±1079			
Rate of force development at 200 ms ($N \cdot s^{-1}$)							*0.919*	***0.009****	*0.448*
	Females	1357±631	1461±936	1685±834	1661±973	1448±873			
	Males	2305±768	2343±564	2076±734	2318±550	2324±565			
	Average	1856±842	1925±868	1881±790	1990±840	1886±845			
Impulse at 50 ms (N·s)							*0.281*	***0.005****	*0.836*
	Females	33.8±4.0	33.3±5.4	33.5±4.8	34.2±4.8	33.3±4.7			
	Males	40.8±6.1	41.9±6.6	41.8±6.9	41.9±7.0	41.8±6.9			
	Average	37.5±6.2	37.8±7.4	37.7±7.2	38.1±7.1	37.6±7.2			
Impulse at 100 ms (N·s)							*0.525*	***0.003****	*0.521*
	Females	73.0±7.0	71.9±13.6	72.9±7.8	75.2±10.3	70.6±9.0			
	Males	90.1±14.0	91.5±14.9	89.7±16.2	92.1±16.7	91.9±16.0			
	Average	82.0±14.0	82.2±17.2	81.3±15.1	83.7±16.0	81.2±16.7			
Impulse at 150 ms (N·s)							*0.604*	***0.002****	*0.628*
	Females	116.3±11.1	115.2±24.3	117.8±10.9	122.0±18.3	112.2±14.6			
	Males	147.3±23.7	158.7±25.7	145.1±26.8	149.3±27.0	148.7±26.8			
	Average	132.6±24.3	132.8±29.8	131.4±24.4	135.7±26.5	130.5±28.1			
Impulse at 200 ms (N·s)							*0.654*	***0.001****	*0.519*
	Females	163.1±17.3	160.5±35.0	166.6±17.1	171.6±28.3	157.7±21.7			
	Males	209.5±33.7	210.9±36.3	204.4±36.9	211.5±36.0	210.7±36.3			
	Average	187.6±35.6	187.0±43.2	185.5±34.0	191.5±37.6	184.2±39.8			
Total impulse (N·s)							*0.934*	***0.029****	*0.205*

	Females	463.6±114.1	464.4±143.5	443.9±108.4	458.3±124.1	470.5±125.9			
	Males	581.3±122.7	572.3±109.2	591.0±120.3	577.8±108.0	579.2±103.6			
	Average	525.5±130.3	521.2±134.8	517.4±134.6	518.0±128.8	524.8±125.3			
Flight time: contraction time (s)							*0.547*	*0.185*	*0.256*
	Females	0.86±0.23	0.87±0.32	0.91±0.27	0.92±0.33	0.83±0.24			
	Males	0.99±0.14	0.99±0.14	0.94±0.16	0.98±0.13	0.97±0.14			
	Average	0.93±0.20	0.93±0.24	0.93±0.21	0.95±0.25	0.90±0.21			
Contraction time (s)							*0.803*	*0.813*	*0.174*
	Females	0.50±0.10	0.52±0.14	0.47±0.10	0.49±0.13	0.52±0.10			
	Males	0.49±0.08	0.48±0.07	0.50±0.07	0.48±0.06	0.49±0.07			
	Average	0.49±0.08	0.50±0.11	0.49±0.09	0.49±0.10	0.50±0.08			

*Values displayed as mean±standard deviation. LEA, low energy availability; HEA, high energy availability; GEA, high energy availability for mass gain/growth. *denotes significance between the sexes/conditions p<0.05, **denotes significant difference between the sexes/conditions p<0.01, ***denotes significant difference between the sexes/conditions p<0.001. [@]denotes no significant differences between conditions in post-hoc testing. ^/ [#] significantly different from other condition with a matching symbol.*

Table S3.7. Outcome measures assessed during the isometric mid-thigh pull across all five conditions of energy availability and between sexes.

		LEA_{REST}	LEA_{EX}	HEA_{REST}	HEA_{EX}	GEA	*p value (main effect of condition)*	*p value (main effect of sex)*	*p value (sex*condition interaction)*
Relative peak force ($N \cdot kg^{-1}$)							*0.874*	*0.346*	*0.731*
	Females	25.4±3.3	25.2±4.2	24.7±3.5	25.7±3.7	25.3±4.5			
	Males	28.1±7.6	27.7±6.6	28.5±7.3	27.9±6.1	27.4±7.3			
	Average	26.7±5.9	26.4±5.5	26.6±5.9	26.8±5.0	26.3±6.0			
Time to peak force (s)							*0.969*	*0.850*	*0.213*
	Females	2.56±1.44	2.31±1.12	2.22±1.39	2.46±1.15	2.01±1.15			
	Males	2.22±1.00	2.54±1.16	2.49±1.38	2.01±1.33	2.69±0.09			
	Average	2.39±1.22	2.42±1.11	2.36±1.35	2.24±1.23	2.35±1.06			
Rate of force development at 50 ms ($N \cdot s^{-1}$)							*0.919*	*0.439*	*0.702*
	Females	2813±2600	2646±2323	2650±2008	2801±2810	3750±2819			
	Males	4208±3434	3991±2815	3635±3350	4677±3079	3499±2345			
	Average	3511±3050	3318±2605	3117±2694	3739±3026	3624±2527			
Rate of force development at 100 ms ($N \cdot s^{-1}$)							*0.755*	*0.064*	*0.876*
	Females	2983±2394	2547±1964	3477±2049	2691±2387	3572±2002			
	Males	4326±3116	3898±2554	4982±3426	4612±2922	3933±2274			
	Average	3654±2791	3222±2323	4229±2854	3652±2778	3753±2093			
Rate of force development at 150 ms ($N \cdot s^{-1}$)							*0.887*	***0.030****	*0.873*
	Females	2833±2129	2566±1721	3167±1544	2603±2010	3287±1660			
	Males	4146±2821	3951±2486	4710±2904	4717±2837	3966±2303			

	Average	3489±2524	3259±2199	3938±2398	3660±2627	3627±1985			
Rate of force development at 200 ms (N·s^{-1})							*0.947*	***0.018****	*0.924*
	Females	2666±1883	2716±1444	3059±1432	2677±1721	3113±1446			
	Males	4044±2500	3913±2220	4412±2445	4568±2502	3905±2130			
	Average	3355±2267	3314±1923	3736±2070	3622±2304	3509±1818			
Rate of force development at 250 ms (N·s^{-1})							*0.942*	***0.024****	*0.985*
	Females	2439±1567	2601±1243	2867±1072	2512±1469	2751±1195			
	Males	3608±2088	3472±1874	3936±2005	3784±2065	3518±1761			
	Average	3024±1894	3036±1611	3401±1658	3115±1844	3134±1517			
Impulse at 50 ms (N·s)							*0.421*	***0.008****	*0.735*
	Females	41.3±8.1	41.4±7.0	40.2±5.5	41.1±6.7	43.2±7.9			
	Males	54.4±12.1	55.4±13.5	54.7±13.3	53.2±11.3	55.1±13.2			
	Average	47.9±12.1	48.4±12.7	47.4±12.4	47.1±11.0	49.2±12.2			
Impulse at 100 ms (N·s)							*0.800*	***0.004****	*0.756*
	Females	91.0±20.7	90.1±15.4	89.7±11.8	89.5±17.0	96.4±16.7			
	Males	121.0±26.8	121.3±29.0	123.1±31.5	120.1±25.3	121.1±28.2			
	Average	106.0±27.9	105.7±27.7	106.4±28.8	104.8±26.2	108.7±25.9			
Impulse at 150 ms (N·s)							*0.819*	***0.003****	*0.809*
	Females	147.1±36.3	144.2±25.5	147.6±21.3	143.5±30.4	156.5±27.2			
	Males	197.4±44.8	196.7±47.9	202.6±50.0	197.0±42.6	197.2±47.9			
	Average	172.3±47.4	170.4±46.0	175.1±46.8	170.3±45.3	176.9±43.3			
Impulse at 200 ms (N·s)							*0.862*	***0.002****	*0.805*
	Females	195.6±47.5	193.2±33.8	195.6±26.4	191.7±39.7	208.4±36.6			

	Males	262.5±58.6	262.5±62.0	269.1±64.7	263.4±56.5	262.6±62.9			
	Average	229.1±62.2	227.8±60.2	232.4±61.1	227.5±60.1	235.5±57.3			
Impulse at 250 ms (N·s)							*0.957*	***0.0007******	*0.942*
	Females	276.1±74.9	274.4±50.8	282.6±42.8	271.7±64.5	295.6±54.8			
	Males	376.0±88.4	374.2±90.3	387.8±88.8	383.3±89.5	376.0±95.8			
	Average	326.0±94.8	324.3±87.8	335.2±86.7	327.5±95.1	335.8±86.4			

*Values displayed as mean±standard deviation. LEA, low energy availability; HEA, high energy availability; GEA, high energy availability for mass gain/growth. *denotes significance between the sexes/conditions p<0.05, **denotes significant difference between the sexes/conditions p<0.01, ***denotes significant difference between the sexes/conditions p<0.001.*

Table S3.8. Outcome measures assessed during the Stroop Colour and Word Test across all five conditions of energy availability and between sexes.

		LEA_{REST}	LEA_{EX}	HEA_{REST}	HEA_{EX}	GEA	*p value (main effect of condition)*	*p value (main effect of sex)*	*p value (sex*condition interaction)*
Stroop effect, accuracy (%)							*0.520*	*0.115*	*0.836*
	Females	0.2±2.0	-0.7±2.7	-1.5±2.9	-0.2±1.7	-0.2±1.7			
	Males	-1.2±2.2	-1.1±1.7	-2.0±2.6	-1.3±3.4	-2.2±3.9			
	Average	-0.5±2.2	-0.9±2.2	-1.7±2.7	-0.7±2.7	-1.2±3.1			
Stroop effect, reaction time (ms)							*0.913*	*0.798*	*0.452*
	Females	84±48	75±44	63±38	76±41	78±45			
	Males	69±50	95±62	90±61	78±43	70±45			
	Average	76±49	85±53	76±51	77±41	74±44			

Values displayed as mean±standard deviation. LEA, low energy availability; HEA, high energy availability; GEA, high energy availability for mass gain/growth.

Figure S4.1.

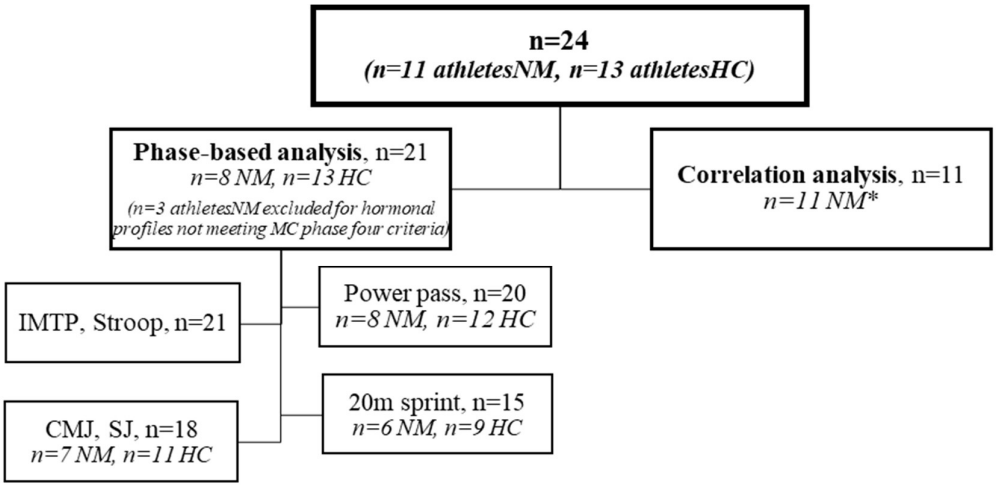


Figure S4.1. Flow chart illustrating the participant number reported for each outcome measure for both phase-based and correlational analysis approaches. Athletes excluded for each variable were due to sickness/injury. *a single progesterone value was excluded from correlational analysis because it was >2.5 standard deviations above the mean. NM; naturally menstruating, HC; hormonal contraception.

Figure S4.2.

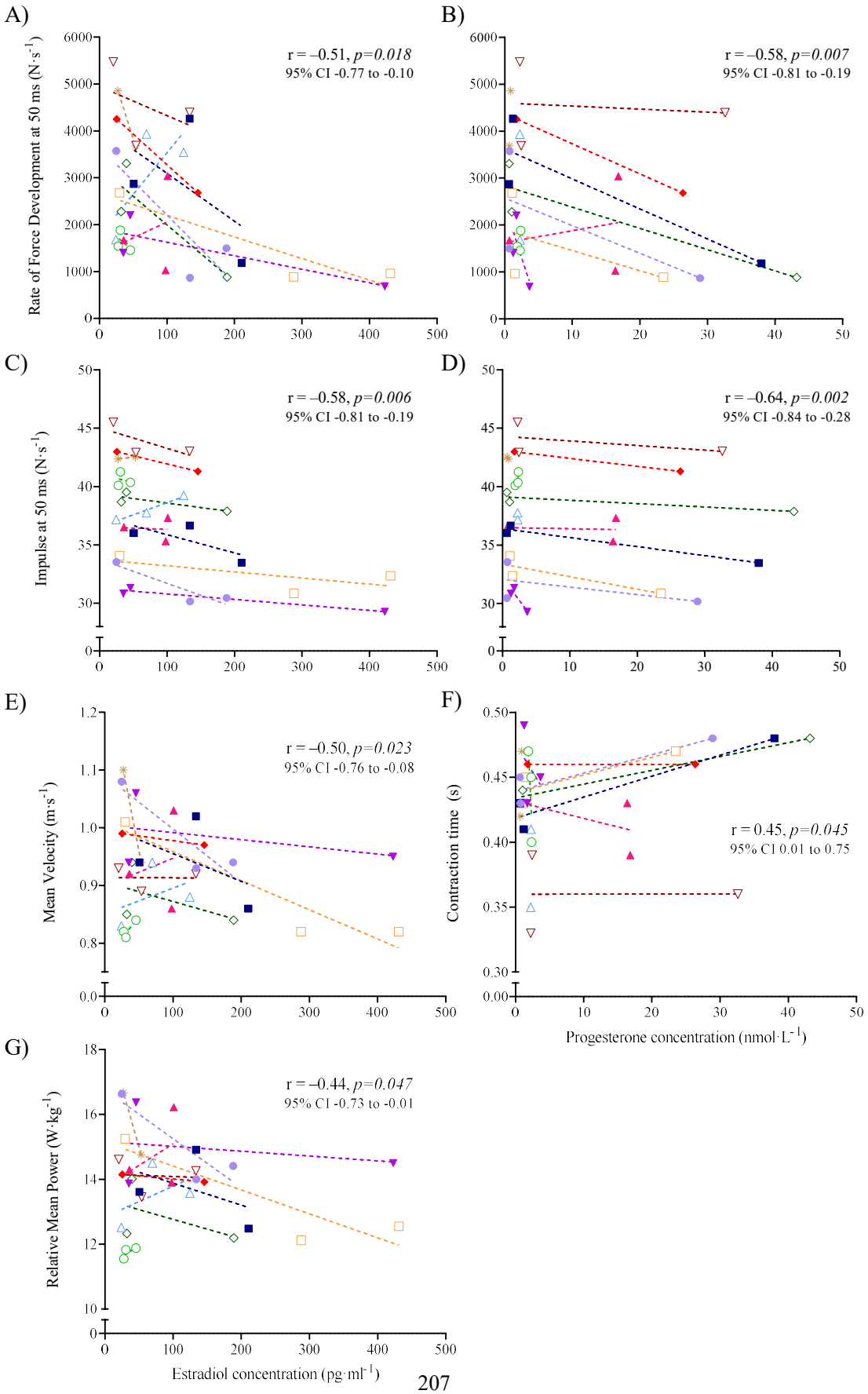

A)
Rate of Force Development at 50 ms (N·s-1)
r = –0.51, p=0.018
95% CI -0.77 to -0.10
B)
r = –0.58, p=0.007
95% CI -0.81 to -0.19
C)
Impulse at 50 ms (N·s-1)
r = –0.58, p=0.006
95% CI -0.81 to -0.19
D)
r = –0.64, p=0.002
95% CI -0.84 to -0.28
E)
Mean Velocity (m·s-1)
r = –0.50, p=0.023
95% CI -0.76 to -0.08
F)
Contraction time (s)
r = 0.45, p=0.045
95% CI 0.01 to 0.75
Progesterone concentration (nmol·L-1)
G)
Relative Mean Power (W·kg-1)
r = –0.44, p=0.047
95% CI -0.73 to -0.01
Estradiol concentration (pg·ml-1)

Figure S4.2. Repeated measures correlations during the squat jump between rate of force development at 50 ms and (A) oestradiol and (B) progesterone, between impulse at 50 ms and (C) oestradiol and (D) progesterone, between (E) mean velocity and oestradiol, (F) contraction time and progesterone, (G) relative mean power and oestradiol. Correlations among naturally menstruating athletes (n=11). The different colour lines and symbols represent different participants.

Figure S4.3.

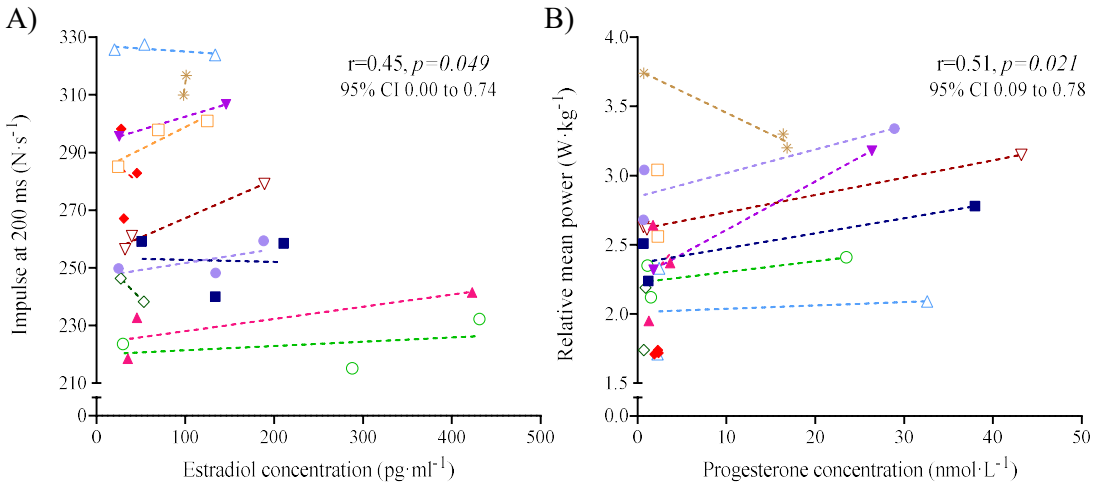


Figure S3. Repeated measures correlations during the countermovement jump between (A) impulse at 200 ms and oestradiol, and (B) relative mean power and progesterone. Correlations among naturally menstruating athletes (n=11). The different colour lines and symbols represent different participants.

Table S4.1. Outcome measures assessed across all performance tests among naturally menstruating athletes and athletes using hormonal contraception.

Performance Test	Outcome measure	Naturally menstruating athletes (n=11)		Athletes using hormonal contraception (n=13)	
		Phase 1	Phase 4	Test one	Test three
Countermovement Jump	Jump height (m)	0.26±0.04	0.27±0.04	0.27±0.05	0.27±0.04
	Velocity at take-off ($m \cdot s^{-1}$)	2.27±0.16	2.31±0.17	2.32±0.18	2.31±0.21
	Relative peak force ($N \cdot kg^{-1}$)	21.9±1.36	22.8±1.49	21.8±1.69	22.3±2.27
	Relative mean force ($N \cdot kg^{-1}$)	12.6±0.3	12.8±0.3	12.6±0.6	12.7±0.6
	Relative peak power ($W \cdot kg^{-1}$)	42.2±4.7	43.3±5.2	43.5±5.1	43.6±5.4
	Relative mean power ($W \cdot kg^{-1}$)	**2.4±0.4**	**2.9±0.5***	2.6±0.7	2.6±0.7
	Peak velocity ($m \cdot s^{-1}$)	2.39±0.15	2.43±0.16	2.45±0.16	2.44±0.19
	Mean velocity ($m \cdot s^{-1}$)	1.27±0.04	1.30±0.07	1.31±0.09	1.31±0.11
	RFD at 50 ms ($N \cdot s^{-1}$)	-1517±1517	-1515±2219	-1150±1241	-615±2235
	RFD at 100 ms ($N \cdot s^{-1}$)	-1674±1479	-1577±1812	-1444±1153	-1113±1949
	RFD at 150 ms ($N \cdot s^{-1}$)	-1152±1278	-1269±1176	-1057±961	-1043±1460
	RFD at 200 ms ($N \cdot s^{-1}$)	-1162±1472	-1998±2681	-1359±1846	-1690±2399
	RFD minimum-maximum ($N \cdot s^{-1}$)	3974±1513	4373±1866	3933±1588	4212±1643
	Impulse at 50 ms (N·s)	71.5±10.3	73.8±11.1	77.8±11.8	78.8±12.9
	Impulse at 100 ms (N·s)	139.7±19.0	144.1±21.2	153.2±23.9	155.7±22.7
	Impulse at 150 ms (N·s)	205.4±28.5	212.1±33.4	225.7±34.1	229.5±30.3
	Impulse at 200 ms (N·s)	270.8±33.7	276.1±38.0	297.2±42.8	299.3±37.1
	Total impulse (N·s)	356.2±54.7	348.2±46.4	390.7±74.5	384.1±75.5

	Flight time: contraction time (s)	0.60±0.08	0.65±0.06	0.53±0.19	0.61±0.12
	Contraction time (s)	0.83±0.10	0.76±0.07	0.84±0.15	0.82±0.17
	Concentric time (s)	0.54±0.08	0.49±0.05	0.56±0.12	0.54±0.14
	Eccentric time (s)	0.28±0.04	0.27±0.04	0.29±0.05	0.28±0.04
	Centre of mass displacement (m)	-0.30±0.05	-0.29±0.05	-0.31±0.06	-0.30±0.07
Squat Jump	Jump height (m)	0.24±0.04	0.24±0.02	0.24±0.04	0.25±0.05
	Velocity at take-off ($m \cdot s^{-1}$)	2.19±0.15	2.18±0.09	2.17±0.19	2.21±0.23
	Relative peak force ($N \cdot kg^{-1}$)	19.9±1.4	19.9±1.5	20.4±2.3	20.2±1.9
	Relative mean force ($N \cdot kg^{-1}$)	15.0±0.6	14.7±0.6	14.6±0.9	14.7±1.0
	Relative peak power ($W \cdot kg^{-1}$)	39.8±3.9	39.8±3.3	40.4±5.6	41.1±6.2
	Relative mean power ($W \cdot kg^{-1}$)	14.2±1.5	13.2±0.9	12.7±2.3	13.3±2.7
	Peak velocity ($m \cdot s^{-1}$)	2.32±0.14	2.31±0.09	2.31±0.17	2.36±0.21
	Mean velocity ($m \cdot s^{-1}$)	0.95±0.09	0.89±0.05	0.86±0.14	0.89±0.15
	RFD at 50 ms ($N \cdot s^{-1}$)	3258±1287	2063±1471	2237±1811	2176±1389
	RFD at 100 ms ($N \cdot s^{-1}$)	3479±901	2576±1352	2635±1836	2855±1681
	RFD at 150 ms ($N \cdot s^{-1}$)	2912±920	2440±634	2270±1233	2552±1257
	RFD at 200 ms ($N \cdot s^{-1}$)	2213±953	2071±666	1896±826	2084±915
	RFD minimum-maximum ($N \cdot s^{-1}$)	2213±935	1984±634	2169±681	2129±629
	Impulse at 50 ms (N·s)	**38.3±4.5**	**36.6±5.1***	40.0±7.5	40.0±7.0
	Impulse at 100 ms (N·s)	86.8±11.7	80.8±14.3	88.1±20.4	88.1±17.7
	Impulse at 150 ms (N·s)	141.9±20.0	131.3±23.6	141.7±34.7	143.8±31.6
	Impulse at 200 ms (N·s)	199.7±31.1	186.2±32.4	198.2±48.6	201.7±44.7

	Total impulse (N·s)	443.5±50.9	457.9±52.7	508.5±71.9	510.8±72.9
	Flight time: contraction time (s)	1.11±0.11	0.92±0.38	1.00±0.18	1.01±0.19
	Contraction time (s)	0.42±0.04	0.44±0.05	0.46±0.06	0.46±0.08
Isometric Mid-Thigh Pull	Relative peak force ($N \cdot kg^{-1}$)	24.2±2.7	26.1±2.3	24.9±3.8	24.2±3.6
	Time to peak force (s)	2.7±1.5	3.6±1.1	3.0±1.2	2.4±0.8
	RFD at 50 ms ($N \cdot s^{-1}$)	2909±2127	3179±2643	1349±2255	2549±2995
	RFD at 100 ms ($N \cdot s^{-1}$)	3013±2261	3336±2357	1307±2239	2563±2842
	RFD at 150 ms ($N \cdot s^{-1}$)	3206±2458	3471±2268	1196±1185	2615±2448
	RFD at 200 ms ($N \cdot s^{-1}$)	3032±2297	3481±1941	1129±1600	2527±1970
	RFD at 250 ms ($N \cdot s^{-1}$)	2491±2125	3205±1521	1226±1357	2260±1655
	Impulse at 50 ms (N·s)	57.9±17.2	49.3±11.3	50.6±9.8	59.9±18.7
	Impulse at 100 ms (N·s)	123.1±34.0	107.8±27.8	104.5±20.2	126.0±38.5
	Impulse at 150 ms (N·s)	197.0±50.3	174.8±47.8	161.1±33.5	198.9±59.3
	Impulse at 200 ms (N·s)	262.4±66.3	233.4±63.8	214.8±43.5	264.6±77.8
	Impulse at 250 ms (N·s)	363.9±81.8	334.4±96.2	283.0±63.2	361.2±99.4
Calculated metrics	Eccentric Utilisation Ratio	1.13±0.06	1.16±0.08	1.16±0.13	1.12±0.08
	Reactive Strength Index	3.57±1.65	4.27±2.02	3.61±2.20	2.85±1.60
	Dynamic Strength Index	0.58±0.10	0.58±0.05	0.59±0.06	0.56±0.07

Power Pass	Distance thrown (m)	5.49±0.45	5.56±0.50	5.86±0.52	5.72±0.48
20 m Sprint	Time to 5m (s)	1.18±0.04	1.17±0.03	1.25±0.07	1.23±0.07
	Time to 10m (s)	2.02±0.07	2.00±0.05	2.11±0.13	2.08±0.10
	Time to 20m (s)	3.50±0.12	3.46±0.11	3.66±0.27	3.58±0.20
Stroop Colour and	Stroop effect accuracy (%)	-1.3±3.1	-0.4±1.7	-1.3±4.0	-3.4±4.3
Word Test	Stroop effect reaction time (ms)	93.2±24.3	107±63.5	151±93.0	141±93.0

*Results presented as mean±standard deviation. *significantly different from menstrual cycle phase 1 ($p<0.05$). RFD; rate of force development.*

Table S4.2. Mean total training load across the five-week training camp.

	Naturally menstruating athletes (n=11)	Athletes using hormonal contraception (n=13)	*p value*
Training duration (hours)	32.3±4.4	32.1±5.1	*0.944*
RPE x duration (AU)	10,782±2,144	10,572±3,210	*0.856*
Distance covered in field sessions (km)	30.6±13.4	34.6±7.3	*0.387*
Distance covered in field sessions per minute played ($m \cdot min^{-1}$)	346±128	385±80	*0.390*
Gym volume load (AU)	38,519±9,388	40,122±7,436	*0.645*

Results presented as mean±standard deviation.

Table S4.3. Mean intra-phase coefficient of variation for naturally menstruating athletes and inter-test coefficient of variation for athletes using HC for performance outcome measures and kinematic outputs.

Performance Test	Outcome measure	Naturally menstruating athletes (n=11)	Athletes using hormonal contraception (n=13)
		Intra-phase CV (%)	Inter-test CV (%)
Countermovement Jump	Jump height (m)	4.0	16.6
	Velocity at take-off ($m \cdot s^{-1}$)	1.9	8.5
	Relative peak force ($N \cdot kg^{-1}$)	3.2	9.6
	Relative mean force ($N \cdot kg^{-1}$)	1.3	4.6
	Relative peak power ($W \cdot kg^{-1}$)	2.2	13.0
	Relative mean power ($W \cdot kg^{-1}$)	9.4	28.1
	Peak velocity ($m \cdot s^{-1}$)	1.4	7.2
	Mean velocity ($m \cdot s^{-1}$)	2.5	8.1
	RFD at 50 ms ($N \cdot s^{-1}$)	338.2	213.9
	RFD at 100 ms ($N \cdot s^{-1}$)	4.0	131.1
	RFD at 150 ms ($N \cdot s^{-1}$)	54.6	119.7
	RFD at 200 ms ($N \cdot s^{-1}$)	34.9	126.6
	RFD minimum-maximum ($N \cdot s^{-1}$)	17.5	37.3
	Impulse at 50 ms (N·s)	3.6	15.3
	Impulse at 100 ms (N·s)	3.0	14.1
	Impulse at 150 ms (N·s)	2.6	13.0
	Impulse at 200 ms (N·s)	2.2	12.4

	Total impulse (N·s)	2.4	19.4
	Flight time: contraction time (s)	6.1	36.3
	Contraction time (s)	5.1	17.8
	Concentric time (s)	7.0	21.1
	Eccentric time (s)	4.3	16.1
	Centre of mass displacement (m)	5.6	19.2
Squat Jump	Jump height (m)	5.3	20.7
	Velocity at take-off ($m \cdot s^{-1}$)	2.7	10.4
	Relative peak force ($N \cdot kg^{-1}$)	1.7	10.5
	Relative mean force ($N \cdot kg^{-1}$)	2.4	6.4
	Relative peak power ($W \cdot kg^{-1}$)	2.9	15.1
	Relative mean power ($W \cdot kg^{-1}$)	8.4	19.6
	Peak velocity ($m \cdot s^{-1}$)	2.3	8.6
	Mean velocity ($m \cdot s^{-1}$)	6.9	15.8
	RFD at 50 ms ($N \cdot s^{-1}$)	43.5	71.8
	RFD at 100 ms ($N \cdot s^{-1}$)	35.4	62.1
	RFD at 150 ms ($N \cdot s^{-1}$)	24.5	49.2
	RFD at 200 ms ($N \cdot s^{-1}$)	17.7	40.7
	RFD minimum-maximum ($N \cdot s^{-1}$)	7.7	29.8
	Impulse at 50 ms (N·s)	2.8	17.2
	Impulse at 100 ms (N·s)	4.8	19.7
	Impulse at 150 ms (N·s)	5.7	20.8

	Impulse at 200 ms (N·s)	5.6	20.8
	Total impulse (N·s)	3.5	16.3
	Flight time: contraction time (s)	10.8	25.6
	Contraction time (s)	5.7	14.3
Isometric Mid-Thigh Pull	Relative peak force ($N \cdot kg^{-1}$)	4.7	15.2
	Time to peak force (s)	39.6	38.1
	RFD at 50 ms ($N \cdot s^{-1}$)	66.8	123.4
	RFD at 100 ms ($N \cdot s^{-1}$)	52.1	111.1
	RFD at 150 ms ($N \cdot s^{-1}$)	51.2	99.4
	RFD at 200 ms ($N \cdot s^{-1}$)	51.3	86.7
	RFD at 250 ms ($N \cdot s^{-1}$)	55.3	78.6
	Impulse at 50 ms (N·s)	10.3	24.8
	Impulse at 100 ms (N·s)	11.6	24.8
	Impulse at 150 ms (N·s)	12.4	25.1
	Impulse at 200 ms (N·s)	12.2	24.7
	Impulse at 250 ms (N·s)	13.0	25.2
Power Pass	Distance thrown (m)	3.0	8.1
20 m Sprint	Time to 5m (s)	1.3	6.3
	Time to 10m (s)	0.8	6.0
	Time to 20m (s)	0.7	6.7

Stroop Colour and Word Test	Stroop effect accuracy (%)	N/A	3.1
	Stroop effect reaction time (ms)	N/A	75.8

Results presented as mean±standard deviation. Intra-phase CV for the Stroop Test is not applicable for athletesNM as only one test was completed per phase (i.e., no repeats). CV; coefficient of variation.

Study 3

Figure S5.1. Daily menstrual questionnaire.

1. In the last 24 hours, did you have period blood flow?

   Yes No

2. Are you within the first 24 hours of this period?

   Yes No

3. How would you describe your period flow the past 24 hours?

   Light flow

   Medium flow

   Heavy flow

4. Did you use an ovulation stick this morning?

   Yes No

5. What was the ovulation stick result?

   No smiley face

   Flashing smiley face

   Static smiley face

6. Please identify any of the following you have experienced in the past 24 hours (please select all that apply).

   Abdominal cramp

   Acne

   Appetite changes

   Bladder incontinence

   Bloating

   Breast pain

   Constipation

   Diarrhoea

Fatigue

Headache

Lower back pain

Mood changes

Nausea

Night sweats

Pelvic pain

Sleep changes

Other _____

Nothing identified

7. How much did these symptoms affect your training today?

Not at all affected training ---------- Completely

(Place a mark on the scale above)

8. Please identify any medications you have taken in the last 24 hours.

No medications

Paracetamol (e.g. panadol)

Ibuprofen (e.g. nurofen)

Anti-histamines

Cold and flu

Anti-depressants

Other _____

9. How many training sessions did you complete yesterday? (please enter "0" on rest days)

_____

10. What type of training sessions you completed yesterday?

    Zwift cycling session

    Outdoor cycling session

    Gym session

    Other

_____

11. How many minutes was the training session?

_____

12. How many kilometres was the training session? Please write "N/A" if necessary.

_____

13. On the scale below, please rate the intensity of yesterday's first training session.

    0 - Rest

    1 - Very , very easy

    2 - Easy

    3 - Moderate

    4 - Somewhat Hard

    5 - Hard

    6

    7 - Very Hard

    8

    9

    10 - Maximal

14. Please detail any other training sessions completed yesterday (type, minutes, kilometres and RPE)

_____

15. Are you currently injured?

    Yes No

Please provide detail on this injury

(body part, diagnosis etc.)

_____

16. Did this injury prevent you from training/competing?

    Yes No

17. Use the following scale to rate your current fatigue levels

    1 - Always tired

    1.5

    2 - More tired than normal

    2.5

    3 - Normal

    3.5

    4 - Fresh

    4.5

    5 - Very fresh

18. Use the following scale to rate your sleep last night

    1 - Insomnia

    1.5

    2 - Restless sleep

2.5

3 - Difficulty falling asleep

3.5

4 - Good

4.5

5 - Very restful

19. Use the following scale to rate your current levels of general muscle soreness

1 - Very sore

1.5

2 - Increase in soreness/tightness

2.5

3 - Normal

3.5

4 - Feeling good

4.5

5 - Feeling great

20. Use the following scale to rate your current stress levels

1 - Highly stressed

1.5

2 - Feeling stressed

2.5

3 - Normal

3.5

4 - Relaxed

4.5

5 - Very relaxed

21. Use the following scale to rate your current mood

1 - Highly annoyed, irritable or down

1.5

2 - Snappiness at team-mates, family or co-workers

2.5

3 - Less interested in others &/or activities than usual

3.5

4 - A generally good mood

4.5

5 - Very positive mood

Table S5.1. Likert scale for gastrointestinal symptoms.

Please rate your current gastrointestinal symptoms (1 = no symptoms, 10 = extreme symptoms)

	1	2	3	4	5	6	7	8	9	10
Nausea										
Urge to vomit										
Vomiting										
Belching										
Bloating										
Stomach pain/cramps										
Gastric acidosis										
Constipation										
Diarrhoea										
Urge to defecate										
Gas										

Table S5.2. Likert scale for thermal sensation.

Please rate your current thermal sensation

0	very hot
1	hot
2	warm
3	slightly warm
4	neutral
5	slightly cool
6	cool
7	cold
8	very cold

Table S5.3. Likert scale for thermal comfort.

Please rate your current thermal comfort

0	very comfortable
1	comfortable
2	just comfortable
3	just uncomfortable
4	uncomfortable
5	very uncomfortable

Table S5.4. Individual participant menstrual cycle information, as confirmed through retrospective classification following two months of menstrual cycle monitoring. Classifications as defined by Elliott-Sale (22): eumenorrhea (menstrual cycle length 21-35 days, confirmed urinary luteinising hormone surge, serum progesterone concentration >16 nmol·L^{-1}), naturally menstruating (cycle length 21-35 days without confirmed ovulation or hormonal profiles), oligomenorrhea (cycle length >35 days), polymenorrhea (cycle length <21 days), anovulatory (negative urinary luteinising hormone surge testing for two consecutive cycles). "Unknown" is reported if the participant did not record the characteristic.

Athlete	**Cycle 1**			**Cycle 2**			**Progesterone >16 nmol·L^{-1} detected?**	**Classification (and any diagnoses)**
	Cycle length (days)	Bleeding days	Ovulation day	Cycle length (days)	Bleeding days	Ovulation day		
1	28	9	17	27	9	16	Yes	Eumenorrhea
2	14	2	not detected	25	6	not detected	No	Polymenorrhea* with suspected anovulation
3	37	5	17	29	7	12	Yes	Eumenorrhea*
4	36	8	20	35	6	20	No	Oligomenorrhea
5	26	5	15	26	3	13	Yes	Eumenorrhea
6	40	6	12	31	6	10	Yes	Oligomenorrhea* (Endometriosis)
7	25	6	not detected	24	6	11	Yes	Naturally menstruating
8	28	7	16	28	7	15	Yes	Eumenorrhea
9	29	5	15	28	5	14	Yes	Eumenorrhea
10	37	6	21	33	5	16	Yes	Eumenorrhea*
11	27	6	12	28	6	13	No	Naturally menstruating with ovulation
12	27	6	not detected	27	4	12	Yes	Naturally menstruating
13	29	3	14	28	6	20	Yes	Eumenorrhea
14	24	4	11	26	5	11	Yes	Eumenorrhea (PCOS)

15	26	5	11	27	5	11	Yes	Eumenorrhea
16	27	4	15	27	4	14	Yes	Eumenorrhea (PCOS)
17	26	4	13	24	5	13	Yes	Eumenorrhea
18	29	5	not detected	32	5	not detected	Yes	Naturally menstruating with suspected anovulation
19	31	5	13	30	7	15	Yes	Eumenorrhea
20	31	6	17	28	5	15	Yes	Eumenorrhea
21	28	5	13	29	6	14	Yes	Eumenorrhea
22	27	5	11	29	6	12	Yes	Eumenorrhea
23	32	5	17	30	6	16	Yes	Eumenorrhea
24	26	5	14	24	5	16	No	Naturally menstruating with ovulation
25	27	5	13	22	4	19	Yes	Eumenorrhea
26	32	7	20	27	6	not detected	No	Naturally menstruating
27	28	5	24	28	5	18	Yes	Eumenorrhea
28	28	6	13	28	5	not detected	Yes	Naturally menstruating
29	28	6	15	34	5	21	No	Naturally menstruating with ovulation
30	26	4	12	24	5	18	No	Naturally menstruating with ovulation
31	24	5	13	41	unknown	unknown	No	Naturally menstruating with ovulation* (PCOS)
32	24	6	13	25	6	13	Yes	Eumenorrhea
33	30	5	17	31	6	not detected	Yes	Naturally menstruating
34	25	4	12	27	4	13	Yes	Eumenorrhea
35	23	4	12	26	4	13	No	Naturally menstruating with ovulation (Endo)
36	23	7	12	27	8	16	No	Naturally menstruating with ovulation
37	17	7	16	24	8	14	No	Naturally menstruating with ovulation*

*Blood hormone concentrations were measured during one menstrual cycle per athlete. *denotes some irregularity in menstrual cycle length. Endo, endometriosis; PCOS, polycystic ovary syndrome.*

Study 4

Figure S6.1. Survey completed by participants upon study completion.

Questions in grey font only appeared to participants on the online survey if a relevant answer to a prior question was selected.

A. Please complete the following questions regarding your prior participation in other research studies before youparticipated in our study

Have you participated in a research study before?	◯ Yes (1 study only) ◯ Yes (more than 1 study) ◯ Started, but dropped out before completion ◯ No
Why did you drop out from this study?	_____ (please provide as much detail as possible)
How many studies have you previously participated in?	_____
What type of project was it?	◯ survey ◯ lab trial (one-off visit)lab ◯ trial (1-3 visits) lab trial ◯ (>3 visits) ◯ research/training camp
How much of your time in TOTAL did this project require? (i.e., how long from your first lab visit to your last, for example 3 x 3 h trials over 1 month would total 1 month)	◯ < 30 min ◯ 30 min - 1 hour ◯ 1-4 hours ◯ 4-8 hours ◯ 8-24 hours ◯ 1-2 days ◯ >2 days ◯ >1 week ◯ >2 weeks ◯ >1 month
Was any financial support for provided your participation?	◯ Yes - direct compensation (e.g., funding, prize money) ◯ Yes - reimbursement of some costs (e.g. food, transport) ◯ Yes - some free gifts and resources but no money ◯ No - no direct income or compensation (either direct compensation/reimbursement for yourtime or contribution towards costs such as transport)
Please provide details of this financial support	_____ (including amount provided and rationale (e.g., $500 prize money))
Please provide separate details for each of thestudies you have participated in	_____ *(e.g. 1x survey (< 30 min, no financial support), 1x lab study (1-4 hours, 1 trial visit, $50 Amazonvoucher))*

Why have you not participated in a study before?(please select all that apply)

- Never had the opportunity ☐
- Couldn't afford it (lack of financial support) ☐
- Time demand was too great ☐
- Too young ☐
- Not interested in research Unaware of research opportunities ☐
- Concerns regarding lab testing (needles, maximalexercise etc.) ☐
- Concerns around exercise demands ☐
- Dietary concerns (e.g., having to adhere to a standardised diet) ☐
- Other _____ ☐

Please rank the reasons you identified from 1 (main reason for not previously participating), onwards. If you only identified one reason, please just mark it as the main reason.

	Main reason	2	3	4	5
Never had the opportunity	○	○	○	○	○
Couldn't afford it (lack of financial support)	○	○	○	○	○
Time demand was too great	○	○	○	○	○
Too young	○	○	○	○	○
Not interested in research	○	○	○	○	○
Unaware of research opportunities	○	○	○	○	○
Concerns regarding lab testing (needles, maximal exercise etc.)	○	○	○	○	○
Concerns around exercise demands	○	○	○	○	○
Dietary controls (e.g. having to adhere to a standardised diet)	○	○	○	○	○
Other	○	○	○	○	○

B. Please complete the following questions regarding why you decided to participate in this study

To what degree did the following influence your decision to participate in OUR research project?

Interest in the research outcome (e.g., how the results could be applied to my training/ the wider community)

not at all — somewhat — very much so

(Place a mark on the scale above)

Educational opportunities

not at all — somewhat — very much so

(Place a mark on the scale above)

Opportunity to receive personalised results (e.g., the influence of your menstrual cycle on performance)

not at all — somewhat — very much so

(Place a mark on the scale above)

Financial support (directly for your participation)	not at all — somewhat — very much so *(Place a mark on the scale above)*
Social aspect/peer support	not at all — somewhat — very much so *(Place a mark on the scale above)*
Networking (visibility for my sporting ability (talent scouts etc.))	not at all — somewhat — very much so *(Place a mark on the scale above)*
Other reasons listed were study specific	not at all — somewhat — very much so *(Place a mark on the scale above)*
Other (please specify)	_____

To what extent do you agree with the following statements?

I understood the goals of this research study prior to my participation	completely disagree — somewhat agree — completely agree *(Place a mark on the scale above)*
The information given to me before I joined the trial was everything I wanted to know (e.g., procedures, time commitment, who to contact with questions)	completely disagree — somewhat agree — completely agree *(Place a mark on the scale above)*
The information given to me before I joined the trial was easy to understand (e.g., procedures, time commitment, who to contact with questions)	completely disagree — somewhat agree — completely agree *(Place a mark on the scale above)*
I felt comfortable that I could ask any questions before and during the study	completely disagree — somewhat agree — completely agree *(Place a mark on the scale above)*
How did you hear about this study?	◯ Online/ social media ◯ Through my university ◯ Through my sports club/coach ◯ Word of mouth ◯ Saw a flyer ◯ Other

C. Please complete the following questions regarding your experience during this study

Please rate your **experiences** of each part of the study (1 = the worst experience, 10 = the best experience)										
	1	2	3	4	5	6	7	8	9	10
Tests listed were specific to the study	○	○	○	○	○	○	○	○	○	○
Tests listed were specific to the study	○	○	○	○	○	○	○	○	○	○
	○	○	○	○	○	○	○	○	○	○
	○	○	○	○	○	○	○	○	○	○
	○	○	○	○	○	○	○	○	○	○
	○	○	○	○	○	○	○	○	○	○
	○	○	○	○	○	○	○	○	○	○

Please rate the **feedback** received for each of the following parts of the study - how valuableor successful was this? (1 = the worst experience, 10 = the best experience)										
	1	2	3	4	5	6	7	8	9	10
Tests listed were specific to the study	○	○	○	○	○	○	○	○	○	○
Tests listed were specific to the study	○	○	○	○	○	○	○	○	○	○
	○	○	○	○	○	○	○	○	○	○
	○	○	○	○	○	○	○	○	○	○
	○	○	○	○	○	○	○	○	○	○
	○	○	○	○	○	○	○	○	○	○
	○	○	○	○	○	○	○	○	○	○

Please answer the following questions regarding the credibility of the research/ future use of the research

Please rate the credibility/professionalism of the researchers	not at all credible — somewhat credible — extremely credible *(Place a mark on the scale above)*
How important are the general findings of this research project likely to be in your sport and your sporting career?	not at all valuable — somewhat valuable — extremely valuable *(Place a mark on the scale above)*
Will you apply anything you've learnt in this study toyour training?	◯ Yes ◯ No
Please provide details	_____
Knowing how the research was undertaken, if these outcomes are applied to the top athletes in your sport, to what extent would you recommend/support this?	would not recommend at all — would somewhat recommend — would completely recommend *(Place a mark on the scale above)*
Knowing how the research was undertaken, how would yourate any publications or publicity about its findings?	Highly unbelievable (project didn't reflect real world or was conducted in a "loose manner") — Somewhat believable — Highly believable (real-world issues, strict research control) *(Place a mark on the scale above)*

Please answer the following questions regarding your overall study experiences

Was your overall research experience similar to what you had anticipated?	completely different — somewhat the same — exactly the same *(Place a mark on the scale above)*

How did the following aspects compare to your expectations?

Time commitment	much less demanding — as expected — much more demanding *(Place a mark on the scale above)*
Physical demands	much less demanding — as expected — much more demanding *(Place a mark on the scale above)*
Mental demands	much less demanding — as expected — much more demanding *(Place a mark on the scale above)*
What were your favourite aspects/highlights of participating in the study and why?	(please provide as much detail as possible)
What aspects of the study would you have liked to be different/ what could have been better to improve your experience and why?	(please provide as much detail as possible)
I felt as if the study's research outcomes and my own training experiences were equally prioritized	completely disagree — somewhat agree — completely agree *(Place a mark on the scale above)*
Please rate your overall experience of participating in this research study	Likely to be the worst study ever conducted — Likely to be the best study ever conducted *(Place a mark on the scale above)*
Following this experience, would you participate in another research study?	◯ Yes ◯ No
Why?	

Figure S6.2.

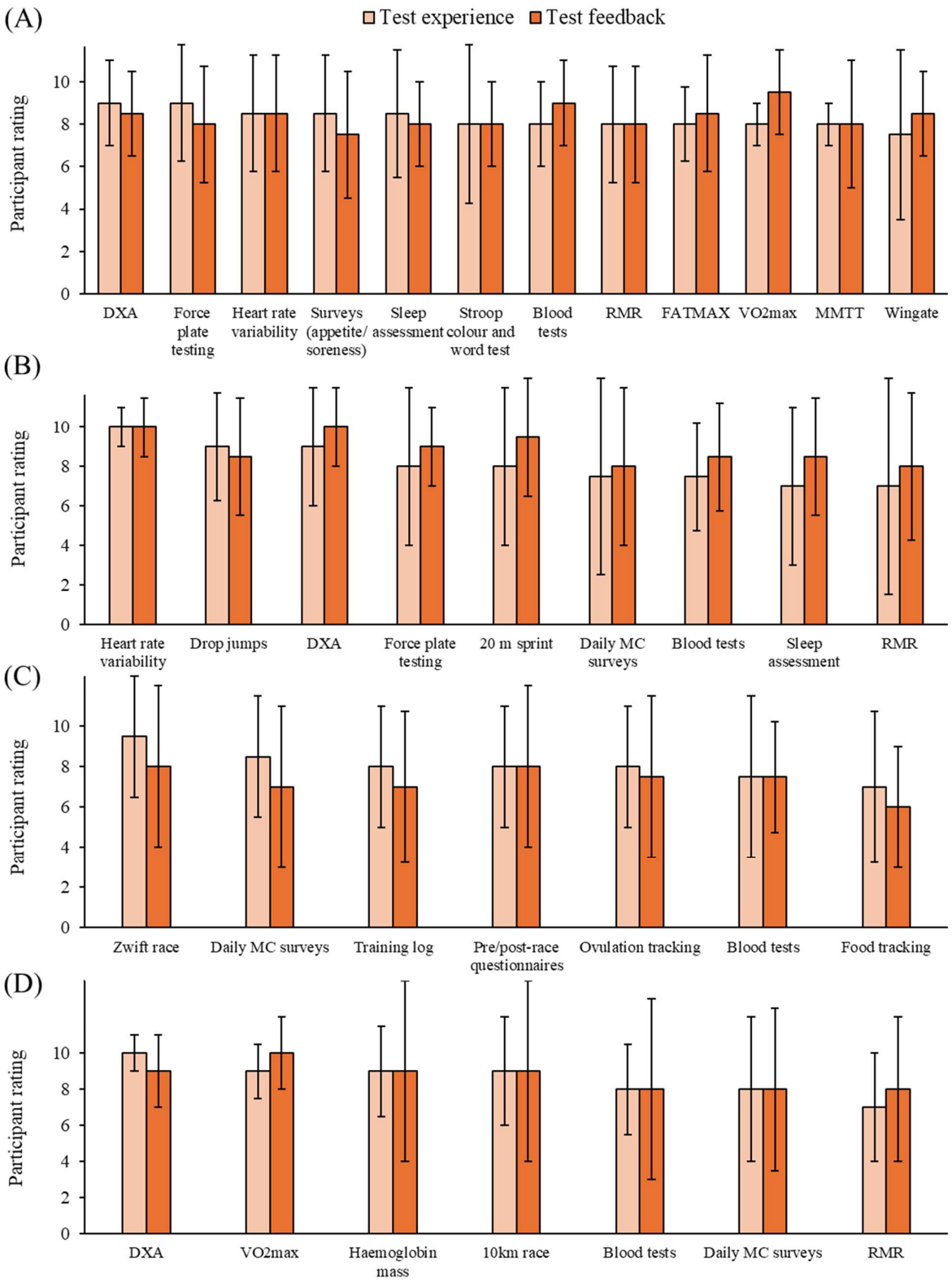

(A)
Test experience
Test feedback
Participant rating
0
2
4
6
8
10
DXA
Force plate testing
Heart rate variability
Surveys (appetite/ soreness)
Sleep assessment
Stroop colour and word test
Blood tests
RMR
FATMAX
VO2max
MMTT
Wingate
(B)
Participant rating
Heart rate variability
Drop jumps
DXA
Force plate testing
20 m sprint
Daily MC surveys
Blood tests
Sleep assessment
RMR
(C)
Participant rating
Zwift race
Daily MC surveys
Training log
Pre/post-race questionnaires
Ovulation tracking
Blood tests
Food tracking
(D)
Participant rating
DXA
VO2max
Haemoglobin mass
10km race
Blood tests
Daily MC surveys
RMR

Figure S6.2. Participant rating of each test completed, displayed separately across the four studies: (A) Study 1 (n=10), (B) Study 2 (n=22), (C) Study 3 (n=38), and (D) Study 4 (n=19), whereby "1" represented "the worst experience" and "10" was "the best experience". The "test experience" refers to the experience of the test itself (e.g., the actual process of undertaking a blood test or using the at-home device to measure sleep), while "test feedback" refers to the individual feedback that participants received post-study (e.g., blood test/ sleep summary results). Data presented as the median, with error bars denoting the interquartile range. Tests are listed in order of participant experience of the test itself (left to right, from best to worst). *DXA*; dual x-ray absorptiometry, *RMR*; resting metabolic rate, *FATMAX*; the intensity eliciting peak fat oxidation, *MMTT*; mixed meal tolerance test, *MC*; menstrual cycle.

Figure S7.1 Proposed pre-screening questionnaire for female participants in research studies.

Pre-screening questionnaire for women in research studies: Menstrual function & contraceptives

1. Do you use any kind of hormonal contraceptives?

☐ Yes *(go to Q11)*
☐ No

2. Have you used any form of hormonal contraception in the last 3 months?

☐ Yes
☐ No *(go to Q4)*

3. Which contraceptive did you previously use?

☐ Oral contraceptive pill
☐ Intrauterine system (IUS)
☐ Implant
☐ Hormonal injection (e.g., Depo)
☐ Vaginal ring (e.g., NuvaRing)

4. Do you use the copper IUD (non-hormonal contraceptive)?

☐ Yes *(go to Q5)*
☐ No *(go to Q6)*

5. When was the copper IUD inserted?

_____ (DD/MM/YYYY)

6. How many periods have you had during the last year?

☐ More than 17
☐ 13-16
☐ 10-12
☐ 7-9
☐ 4-6
☐ 0-3

7. When was your last period?

☐ 0-4 weeks ago
☐ 1-2 months ago
☐ 3-4 months ago
☐ 5 months ago or more

8. Are your periods regular? (i.e. typically close to the same number of days apart)

- ☐ Yes, most of the time
- ☐ No, mostly not

9. On average, how many days apart are your periods? (i.e. days separating day 1 of bleeding between consecutive periods)

_____

10. How many days do you normally bleed during your period?

- ☐ 1-2 days
- ☐ 3-4 days
- ☐ 5-6 days
- ☐ 7-8 days
- ☐ 9 or more days

(go to Q17)

11. Which type of hormonal contraception do you use?

- ☐ Oral contraceptive pill *(go to Q12)*
- ☐ Intrauterine system (IUS) *(go to Q15)*
- ☐ Implant *(go to Q15)*
- ☐ Hormonal injection (e.g., Depo) *(go to Q16)*
- ☐ Hormonal patch *(go to Q16)*
- ☐ Vaginal ring (e.g., NuvaRing) *(go to Q17)*

12. Please list the brand and hormonal dosage of your oral contraceptive pill: (this information will be available on the packet)

_____

13. How long have you been taking this oral contraceptive pill?

_____

14. Do you typically take the 7-day break in the 28-day oral contraceptive pill cycle? (i.e., do you have a withdrawal bleed once per month?)

- ☐ Yes
- ☐ No

(go to Q17)

15. When was this IUS/Implant inserted?

_____ (DD/MM/YYYY)

(go to Q17)

16. When was your last injection/ patch placement?

_____ (DD/MM/YYYY)

17. Have your periods ever stopped for 3 consecutive months or longer (besides pregnancy or hormonal contraception)?

- ☐ Yes, that's the situation now
- ☐ Yes, it has happened before
- ☐ No, never

18. Have you ever been diagnosed with any of the following conditions? (tick all that apply)

- ☐ Endometriosis
- ☐ Polycystic ovary syndrome (PCOS)
- ☐ Primary amenorrhoea
- ☐ Secondary amenorrhoea
- ☐ Dysmenorrhoea
- ☐ Pre-menstrual syndrome
- ☐ Menorrhagia
- ☐ Polymenorrhea
- ☐ Oligomenorrhea
- ☐ None of the above

19. How old were you when you had your first period?

_____

20. Have you ever had problems with heavy bleeding?

- ☐ Yes
- ☐ No

Figure S7.2. A guide for researchers to interpret participant responses to the proposed pre-screening questionnaire for female participants in research studies

USER GUIDE: Pre-screening Questionnaire for women in research studies: Menstrual function & contraceptives

This questionnaire will provide sufficient detail on your participants to achieve a minimum of "bronze" standard of methodological classification and control.

The first question identifies if participants use HC, methodological control is then completed separately for participants using HC vs those not using HC.

Questions that may signify action for the researcher:

2. If participants have used HC in the last 3 months – consider waiting at least three months before testing this participant as it takes 3-6 months following the cessation of HC use for hormonal profiles to normalise

4. as the copper IUD is a non-hormonal form of contraception, the same methodological steps should be carried out as for naturally menstruating participants, but the presence of the copper IUD should be noted, alongside date of insertion (if known, Q5)

6-9. these questions in combination seek to identify irregularities in cycle length, or amenorrhea.

- >17 periods in the last year (Q6), combined with <21 days separating periods (Q9) may suggest polymenorrhea.
- <10 periods in the past year (Q6), combined with a last period 1-2 or 3-4 months ago (Q7), and >35 days separating periods (Q9) may suggest oligomenorrhea.
- 0-3 periods in the last year (Q6), combined with a last period 3-4 months ago or >5 months ago (Q7) may suggest amenorrhea. Check in combination with Q17.

Check if the participant has reported any diagnosed menstrual conditions in question 18 and follow up with the participant, regardless. Importantly, identification of MC disturbances does not automatically necessitate exclusion from the research study, this should be carefully considered in line with your research question, as there are many ways women with MC irregularities can be included.

9. The self-reported MC length can be reported in your participant characteristics section, as long as it is stated that this is retrospectively determined/self-reported and not tracked as part of your study.

11-16. these questions provide details around HC usage, each participants HC use (type, formulation and length of usage) should be reported. This can be in the supplementary material and does not need to be in the main text.

17. If "yes" is answered, follow up with participant as this may suggest amenorrhea (check Q6,7 as aforementioned).

18. consider if excluding individuals with these diagnoses is necessary, this will be context dependent. If included report in your participant characteristics.