International Journal of Kinesiology & Sports Science
ISSN: 2202-946X
www.ijkss.aiac.org.au

Performance of a Modified, Three-Step Menstrual Cycle Tracking Method in Naturally Cycling Females

Marissa L. Doroshuk[1], Patricia K. Doyle-Baker[1,2]

[1]*Human Performance Lab, Faculty of Kinesiology, University of Calgary, 2500 University Dr. NW Alberta, Canada.*

[2]*Alberta Children's Hospital Research Institute, University of Calgary, 2500 University Dr. NW Alberta, Canada*

Corresponding Author: Patricia Doyle-Baker, DrPH/PhD, E-mail: pdoyleba@ucalgary.ca

ARTICLE INFO

Article history
Received: February 01, 2025
Revised: March 21, 2025
Accepted: April 11, 2025
Published: April 30, 2025
Volume: 13 Issue: 2

Conflicts of interest: The authors declare they have no known conflicts of interest.
Funding: The funding for this study was supported by the Doyle-Baker Lab.

ABSTRACT

Background: An objective method of menstrual cycle tracking while minimizing participant burden and cost for field-based research is needed. A modification was proposed to a well-known three-step (m3-step) method to improve accessibility for participants and athletes with difficult travel schedules. **Methods:** A longitudinal design was employed, and the m3-step method (calendar counting, urinary ovulation, and salivary hormones) was completed over three consecutive cycles to assess performance while classifying menstrual variability. Naturally cycling females (N=28; age 18-36 years) from across Canada were recruited prospectively. Participants shipped their hormone samples to the lab where they were classified as "high" or "low" hormone based on ovulation status and a progesterone/estradiol (P4/E2) ratio of 100 pg/mL. Cycle length (mean, ±; SD) was self-reported (28.9 ± 4.16 days) and salivary testing occurred on cycle day 22.5 ± 3.26. **Results:** The average luteinizing hormone surge for those with a positive test occurred on cycle day 14.2 ± 2.27 (22/28). Average cycle length ($t(24.1) = 2.44$, $p = 0.02$), progesterone ($t(21.1) = -4.72$, $p < 0.01$) and P4/E2 ratios ($t(18.9) = -7.74$, $p < 0.01$) were statistically significant between high (12/28) and low (16/28) hormone groups. A logistic regression explored the relationship of progesterone to the hormone classification criteria using a crudes odd ratio (1.98 (95% CI 1.24 – 3.17, $p < 0.01$)). **Conclusion:** The m3-step method yielded a sensitivity of 65% and specificity of 91% using the P4/E2 ratio of 100 pg/mL. Limitations included self-reported naturally cycling, the day of the testing and the P4/E2 value used. In summary, this study examined the feasibility of a m3-step menstrual cycle tracking method to classify hormones as high or low in naturally cycling females for potential implementation in a field-based setting.

Key words: Menstrual Cycle, Ovulation, Salivary Hormones, Luteinizing Hormone, Estradiol, Progesterone

INTRODUCTION

Sport scientists and researchers are increasingly recognizing the importance of reducing methodological limitations by distinguishing between "normal" menstrual cycles (i.e., eumenorrheic) and those that do not fit this criterion (Meignié et al., 2021). A eumenorrheic cycle is defined as lasting between 21 and 35 days, with confirmed ovulation and healthy levels of the hormones estradiol and progesterone (Elliott-Sale et al., 2021). These criteria are important because shifting hormones can mask changes in the cycle related to both physical (muscle, bone, endurance, energy level) and mental stress (attention) and subsequently mislabel the cycle as "normal" (Fehring et al., 2006; Nillni et al., 2021). Identifying eumenorrheic cycles can help researchers interpret results by reducing potential confounding variables in the data and lead to a greater understanding of evaluating potential effects on physical performance (Meignié et al., 2021).

High-quality methodological approaches to identify menstrual cycle phases, including cost and participant burden, are challenges that are surmountable in lab-based research (Allen et al., 2016; Janse De Jonge et al., 2019; McNulty et al., 2020; Tivis et al., 2005). These challenges, however, are exacerbated in field-based research, where sports scientists and athletes may have limited access to gold standards such as serum hormone tests to verify estradiol and progesterone or transvaginal ultrasound for ovulation detection due to demanding schedules and excess travel (Emmonds et al., 2019; Rossiter et al., 2023). While urinary luteinizing hormone (LH) tests are readily available for field-based research and shown to be accurate predictors of ovulation (Johnson et al., 2011; Johnson et al., 2015), they cannot identify hormone levels of estradiol and progesterone. Identifying correct hormone levels post-ovulation detection can provide insight into menstruating females' hormonal health

Published by Australian International Academic Centre PTY.LTD.

http://dx.doi.org/10.7575/aiac.ijkss.v.13n.2p.12

(Janse De Jong et al., 2019). For menstruating athletes, the ability to identify and classify menstrual hormone profiles can be an important indicator of overall training load and athlete health (Cabre et al., 2022).

Researchers have found that a combination of methods, known as the three-step verification method (Schaumberg et al., 2017) of calendar counting/menstrual cycle mapping, urinary LH tests, and serum hormone verification measuring estradiol and progesterone, is better than serum testing alone. This method is needed to reliably classify cycle length, menstruation, precursor to ovulation, and the rise of the mid-luteal phase hormones – estradiol and progesterone – in eumenorrheic menstrual cycles (Allen et al., 2016; Elliott-Sale et al., 2021; Janse De Jonge et al., 2019; Schaumberg et al., 2017; Schmalenberger et al., 2021). This method is doable in a variety of settings, yet a gap remains in field-based research, such as remote settings where serum collection and testing cannot be completed (i.e., safely or too costly).

A modification from the well-known three-step verification method in 2017 by Schaumberg et al. was proposed as the rationale for this study to help improve menstrual cycle research methodologies for sports scientists. This modification involved a multistep approach beginning with calendar counting in the 1st month, followed by 2nd month of urinary LH tests, and 3rd month of salivary hormone verification (Allen et al., 2016). To the best of our knowledge, the performance of this modified three-step (m3-step) method has not been tested in the literature yet may be an accessible and reliable menstrual cycle tracking method in field-based research (Colenso-Semple et al., 2023; Elliott-Sale et al., 2021; Wideman et al., 2013). Therefore, the purpose of this study was to describe and assess the performance of the m3-step method and its ability to classify menstrual cycle variability in naturally cycling females.

MATERIALS AND METHODS

Study Design

A prospective study design was employed, and participants were recruited across Canada primarily through word of mouth, posters, snowball sampling, and social media. An informed consent prior to the study start, which began October 2022 and ended in November 2023 was completed. The duration of the protocol was over four to six months with one month for each measurement method and up to three months for missing data. Once enrolled, participants began on their next menstrual cycle with weekly reminders to complete the protocol until completion. This study was approved by the Conjoint Health Research Ethics Board (CHREB) at the University of Calgary (REB21-0610) and adhered to the Declaration of Helsinki guidelines.

Population

Healthy, naturally cycling (Elliott-Sale et al., 2021), female-sex assigned at birth individuals, aged 18 to 36 years were eligible. Participants had to reside in Canada and access the internet and FedEx shipping accessibility to receive the urinary and salivary testing kits to complete the protocol. Exclusion criteria included a) hormonal contraceptives, hormonal therapies, non-hormonal intrauterine device implantation, pregnancy, or lactation within six months, b) history of reproductive surgery or cigarette smoking, and c) presence of a chronic disease(s) and/or prescribed drug use that may interfere with gonadotrophin or steroid hormone levels and metabolism.

Study Protocol

Questionnaires & follow-up

Data was collected and managed using a secure, web-based electronic database tool (REDCap) hosted at the University of Calgary (Harris et al., 2013; Harris et al., 2009). Participants completed a baseline questionnaire on demographics, menstrual cycle history, and physical activity sent via REDCap. Follow-up questionnaires were completed at the end of each cycle to record menstrual dates, period length, symptoms, physical activity, record the LH surge (precursor to ovulation) and salivary hormone testing completion. Reminders to complete questionnaires were sent every three days for a maximum of three times or until the survey was complete. If after the third reminder the participant still did not complete the survey, the research coordinator would send a follow-up email or phone call if consent for contact by phone was provided.

Calendar counting – menstrual cycle 1 to 3

Menstrual cycle start dates and menstruation length were self-reported and recorded in REDCap. These variables were self-reported for each cycle in the protocol, resulting in three recorded menstrual cycle lengths per participant.

Ovulation testing – menstrual cycle 2

A sign of ovulation was determined in cycle two by measuring the urinary LH surge with Clearblue Digital Ovulation Test Kits (SPD Swiss Precision Diagnostics GmbH, Geneva, Switzerland). The Clearblue Digital Ovulation Test Kit contains an automated optical reader optimizing LH concentration readings through computerized interpretation of the results with either an empty circle, indicating a negative LH surge test, or a smiling face, indicating a positive test (Johnson et al., 2015).

Ten strips were provided to participants, and testing began on cycle day eight (Janse De Jonge et al., 2019; Schaumberg et al., 2017; Schmalenberger et al., 2021). Instructions were given to complete testing with first-morning urine, following the manufacturer's instructions. Participants continued testing until either 1) a positive LH surge result was detected or 2) all test strips had been used, indicating negative results (i.e., no LH surge). Those who had a positive result were asked to confirm by uploading a photo of the positive test and the date of testing into REDCap.

Personalized baselines of urine LH were established in the assay, and changes from baseline were detected at a

sensitivity of 40mIU/mL (SPD Swiss Precision Diagnostics GmbH, Geneva, Switzerland). The sensitivity value was retrieved from Clearblue Digital Ovulation Test Kit (SPD Swiss Precision Diagnostics GmbH, Geneva, Switzerland).

Salivary hormone testing – menstrual cycle 3

Saliva tests were completed by participants during the third menstrual cycle of the study using kits from Labrix by Doctor's Data Inc (DDI; Doctor's Data in St. Charles, IL, USA). The assigned testing day aimed to capture the mid-luteal hormone peak. Day of collection was assigned by the research team using the recorded LH surge day from cycle two and adding eight days (Schaumberg et al., 2017; Schmalenberger et al., 2021). If no LH surge was detected, the testing day was calculated by subtracting 13 days from the longest recorded cycle length (Hampson, 2020). Four unstimulated samples were taken across the day – 30 minutes post-awakening, noon, evening, and before bed. The time of collection and date were recorded for each sample. Participants were directed to follow the instructions provided in the kit and avoid brushing their teeth, eating, or drinking one hour prior to sampling. Samples were frozen and asked to be shipped for analysis within two days of collection.

Salivary hormone tests were analyzed by DDI using MALDI Biotyper® matrix-assisted laser desorption/ionization with time-of-flight (MALDI-TOF) mass spectrometry (Doctor's Data in St. Charles, IL, USA). Each sample was tested for estradiol and progesterone, and values were averaged at the four collection timepoints. Assay quality parameters were retrieved from DDI and the sensitivity for estradiol was 0.1 pg/mL with an intra-assay CV of 8.1% at 3.8 pg/mL, 6.3% at 7.2 pg/mL, and 7.0% at 20.3 pg/mL. The sensitivity for progesterone was 5.0 pg/mL with an intra-assay CV of 8.4% at 39.2 pg/mL and 4.0% at 884.6 pg/mL.

Statistical Analysis

Analysis was conducted using STATA v. 17.0 (StataCorp. 2021. *Stata Statistical Software: Release 17*. College Station, TX: StataCorp LLC). Missing data on menstrual cycle or length was collected by email follow-up and used for analysis. Data were assessed for normality using histograms. Results were expressed as mean (±, SD) and ranges. Our sample size for differences between groups was calculated using Welch's t test with the Welch-Satterthwaite equation**.** This estimated a N=18 with 9 per group (Satterthwaite, 1946).

Participants were classified as either high or low hormone based on LH surge status (i.e., yes/no), and a salivary progesterone/estradiol ratio (P4/E2) > 100 pg/mL. The P4/E2 ratio cut-off was determined through grouping of participants by their estradiol and progesterone hormone values. Group one included those whose hormones were within the reference intervals of estradiol (0.6 – 4.5 pg/mL) and progesterone (127 – 446 pg/mL) provided by DDI. Group two included participants whose estradiol or progesterone values were outside of these reference intervals. The mean mid-luteal P4/E2 ratio was taken for each group (within reference intervals: 137.6 ± 48.86 pg/mL; outside reference intervals: 64.3 ± 39.49 pg/mL) and the mean P4/E2 ratio between the two groups became the cut-off value (100 pg/mL). This calculation is based on the reference intervals provided by DDI and may differ between assay methods (Hillebrand et al., 2021).

Differences in menstrual cycle characteristics between high and low hormone groups were identified using Welch's t-test (Table 1), and significance levels were set to p-value < 0.05. Different menstrual characteristics between hormone groups were used to inform the logistic regression model which explored associations between salivary progesterone and hormone group classification. Performance of the m3-step method was assessed with sensitivity (true positive/(true positive + false negative)), specificity (true negative/(true negative + false positive)), positive predictive value (PPV = true positive/(true positive + false positive)), and negative predictive value (NPV = true negative/(true negative + false negative)) using the P4/E2 ratio of 100 pg/mL and the estimated mid-luteal phase day range. True positives included participants whose mid-luteal phase was captured and had a P4/E2 ratio > 100 pg/mL. True negatives were defined as those whose mid-luteal phase were not captured and had a P4/E2 ratio < 100 pg/mL.

RESULTS

Demographics

Data was collected from 28 healthy, young (mean age 29.0 ± 4.25 yrs; range 18.00 – 36.00 yrs), physically active female-sexed individuals from across Canada (Western Canada 78.6%, Eastern Canada 21.4%). Most participants identified as North American (57.1%) or European (25.0%), with other ethnicities including Latin, Central, or South American, and Asian (17.9%). All participants had some university-level education, with 85.7% meeting the recommended average hours of sleep (7 – 9 hours) per night for adults identified in the Canadian 24-Hour Movement Guidelines (Ross et al., 2020). Calculated body mass index based on self-reported height and weight at enrolment was 24.0 ± 4.64 kg/m^2 (18.54 – 36.14 kg/m^2).

Menstrual Cycle Characteristics

Menstrual cycle characteristics and hormone values are summarized in Table 1. The average self-reported cycle length for all participants' three menstrual cycles (n_{cycles} = 84) was 28.9 ± 4.92 days with menstruation lasting 4.9 ± 1.05 days. Salivary testing occurred on average, cycle day 22.5 ± 3.26 (n = 28). Mid-luteal estradiol was 1.2 ± 0.34 pg/mL and progesterone 117.1 ± 72.07 pg/mL for the total sample. Mean follicular (13.3 ± 2.27 days) and luteal phase (12.4 ± 2.06 days) lengths were calculated for those with a positive LH surge (n = 22). For those who recorded a positive LH surge test, the mean LH surge day was cycle day 14.3 ± 2.27 (n = 22). Significant differences were found in 1) average cycle lengths with the high hormone group having shorter cycle lengths than the low hormone group which had

Table 1. Participants' menstrual cycle characteristics (mean; ± SD; range) by hormone groups (high and low)

Menstrual Cycle Characteristics	All N=28	High Hormone n=12	Low Hormone n=16	Significance p<0.05
*Cycle Length days	28.9±4.92 (23.00 – 57.00)	27.0±2.61 (23.00 – 34.00)	30.3±5.73 (24.00 – 57.00)	*t* (70.5) = 3.54 p=0.02
*Menstruation days	4.9±1.05 (2.00 – 8.00)	4.9±1.03 (2.00 – 7.00)	4.9±1.07 (3.00 – 8.00)	*t* (76.7) = 0.16 p=0.87
**Follicular Phase days	13.3±2.27 (9.00 – 18.00)	12.8±2.77 (9.00 – 18.00)	13.9±1.37 (12.00 – 16.00)	*t* (17.8) = 1.27 p=0.22
**Luteal Phase days	12.4±2.06 (7.00 – 17.00)	13.0±1.91 (9.00 – 17.00)	11.7±2.11 (7.00 – 14.00)	*t* (20.3) = -1.50 p=0.15
***Cycle* Day -LH Surge	14.3±2.27 (10.00 – 19.00)	13.8±2.77 (10.00 – 19.00)	14.9±1.37 (13.00 – 17.00)	*t* (17.8) = 1.27 p=0.22
Testing Day of ML Peak	22.5±3.26 (15.00 – 31.00)	21.8±2.66 (19.00 – 27.00)	23.0±3.65 (15.00 – 31.00)	*t* (28.0) = 0.98 p=0.54
Estradiol (pg/mL)	1.2±0.34 (0.60 – 1.90)	1.2±0.38 (0.60 – 1.80)	1.3±0.30 (0.80 – 1.90)	*t* (21.8) = 1.26 p=0.22
Progesterone (pg/mL)	117.1±72.07 (21.00 – 302.00)	173.9±61.17 (95.00 – 302.00)	74.6±45.76 (21.00 – 168.00)	*t* (21.1) = -4.72 p<0.01
P4/E2 Ratio (pg/mL)	97.3±59.67 (17.50 – 223.00)	154.8±39.25 (106.00 – 223.00)	54.2±25.47 (17.50 – 98.80)	*t* (18.9) = -7.74 p<0.01

LH luteinizing hormone; *ML* mid luteal; *P4/E2* progesterone/estradiol ratio. *Mean of 3 self-reported cycles for each participant ($n_{cycle\ length}$=84 cycles; n_{menses}=82 cycles). **LH surge day, follicular, luteal phase calculations only available for those with LH surge confirmation (n=22 total, n=12 high hormone, n=10 low hormone). Significance set at alpha <0.05

longer cycle lengths (high: 27.0 ± 2.61; low: 30.3 ± 5.73; t (70.5) = 3.54, p = 0.00), 2) salivary progesterone levels (high: 173.9 ± 61.17; low: 74.6 ± 45.76; t (21.1) = -4.72, p = 0.00) and 3) the P4/E2 ratio (high: 154.8 ± 39.25; low: 54.2 ± 25.47; t (18.9) = -7.74, p = 0.00) between hormone groups. No significant differences were found in the remaining menstrual characteristics between hormone groups.

Modified Three-Step Method Performance

The m3-step method classified 42.9% of participants as high hormone while 57.1% were classified as low hormone. A P4/E2 ratio cut-off of 100 pg/mL yielded a sensitivity of 64.7% and specificity of 90.9% with a PPV of 91.7% and NPV of 62.5%. An exploratory analysis was performed using a P4/E2 ratio of 90 pg/mL and 110 pg/mL to further assess the precision of the P4/E2 ratio of 100 pg/mL used in the m3-step method. By decreasing the cut-off value to 90 pg/mL an inverse relationship was observed where the m3-step method classified more participants (13/28; 46.4%) as high hormone. Increasing the cut-off value to 110 pg/mL had an opposing effect, resulting in less participants (11/28; 39.3%) classified as high hormone.

Association of Salivary Progesterone to Classify Menstrual Variability

Salivary progesterone was selected as the primary predictor as the main outcome of the m3-step in capturing the mid-luteal peak and its statistical significance between groups. Progesterone values were grouped by 20 pg/mL unit increases to capture the lowest hormone range in the sample with an aim to provide clinical relevance. The crude odds ratio (OR) of the logistic regression using salivary progesterone as the primary predictor for the high and low hormone groups was 1.98 (95% CI 1.24 – 3.17; p < 0.01). This result was statistically significant, meaning every 20 pg/mL increase in salivary progesterone, the odds of being classified as high hormone was 1.98 times the odds of being classified as low hormone (p < 0.01).

DISCUSSION

The m3-step method proposed in this study was successfully able to identify the mid-luteal phase in 43% (11/28) of participants using the P4/E2 ratio cut-off > 100 pg/mL with a sensitivity of 65%. The results from this study differed from Schaumberg et al. (2017), whose method was 70% successful in identifying the mid-luteal phase in naturally cycling individuals when luteal phase deficiency was not considered. This may be due to variable prevalence of luteal phase deficiency between study populations which can range from 15% to 30% in naturally cycling, physically active females (Bull et al., 2019; Janse De Jonge et al., 2019). To further improve the sensitivity of the methods used in this study, the urinary ovulation and salivary hormone testing could be completed in the same cycle to increase confidence in identifying the mid-luteal phase (Bull et al., 2019; Janse De Jonge et al., 2019), and serial testing of multiple cycles could be performed to increase biomarker accuracy (Shea & Vitzthum, 2020; Vitzthum 2021). Furthermore, completing the measurement of three or more time points in a single cycle to detect the change of menstrual hormones across the cycle would increase the ability to understand individual hormone variability (Janse De Jonge et al., 2019; Westerlind & Williams, 2007). Future research, therefore, may consid-

er measuring hormones during 1) menstruation to capture individual hormone threshold levels (approximately day 3 of cycle), 2) the late-follicular phase capturing peak estradiol and LH surge, and 3) the mid-luteal phase with aims to capture the progesterone and estradiol peak (Janse De Jonge et al., 2019).

Using a P4/E2 ratio cut-off value of 100 pg/mL, we had a 65% ability to correctly classify the proportion of individuals whose mid-luteal phase was captured. The specificity indicates that the P4/E2 ratio cut-off value of 100 pg/mL had a 91% ability to correctly classify the proportion of individuals whose mid-luteal phase was not captured. These results are similar to those found in 2022 by Yu et al., where specificity values were higher, and sensitivity was lower for predicting menstruation and the fertile window based on measured physiological parameters.

The high PPV (92%) and moderately high NPV (63%) demonstrate how likely a participant was (PPV) or was not (NPV) in their mid-luteal phase at the time of salivary hormone collection based on the P4/E2 ratio cut-off. However, future studies may consider the indication of a higher predictive value of P4/E2 cut-off as a potential tool for clinical classification of naturally cycling females as eumenorrheic (Ferrer et al., 2024; Parikh et al., 2008).

The regression model demonstrated that for every 20 pg/mL increase in salivary progesterone, the odds of being classified as high hormone was 2 times the odds of being classified as low hormone ($p < 0.01$). This OR suggests that salivary progesterone may be meaningful in classifying menstrual cycles as high—indicating eumenorrhea—or low hormone in healthy females (Elliott-Sale et al., 2021). The wide confidence interval is suggestive of high individual variability and/or the influence of a small sample size (Nemes et al., 2009). Progesterone was observed as the primary influence on the P4/E2 ratio since mean estradiol levels were not significantly different between hormone groups.

These findings align with current recommendations in menstrual cycle research to use robust methods of hormone verification to identify menstrual phases and to classify cycles as eumenorrheic (Colenso-Semple et al., 2023; Elliott-Sale et al., 2021; Janse De Jonge et al., 2019). However, the development of salivary reference intervals for estradiol, progesterone, and the P4/E2 ratio for menstrual cycle classification is also needed. Additionally, measuring progesterone and estradiol on three consecutive days in the mid-luteal phase may improve the certainty of the hormonal rise and plateau (Westerlind & Williams, 2007). These measurement techniques have been previously done in serum samples but have not been established using saliva (Janse De Jonge et al., 2019; Schuamberg et al., 2017; Wideman et al., 2013). Salivary measures taken during these timepoints may help develop reference intervals for given cycle phases and could potentially be used in menstrual cycle phase identification (Vitzthum, 2021).

Strengths and Limitations

This study illustrated the challenges and feasibility associated with implementing the m3-step method in field-based research where no lab visits could occur. Misinterpretation of the results by participants was unlikely with the digital ovulation tests. These tests have been reported to minimize user interpretation errors and increase certainty in results due to the simplicity of a digital display (Johnson et al., 2011). Johnson et al. (2015) also reported that this digital test was reliable in identifying the LH surge at a mean of 0.81 days prior to ovulation, as verified by transvaginal ultrasound. Forward counting from the identified LH surge day is the current recommendation when aiming to prospectively identify the mid-luteal phase (Janse De Jong et al., 2019; Schmalenberger et al., 2021).

Assay parameters of sensitivity and CVs for salivary and urinary methods used were reported where available for a more complete understanding of hormone results. The parameters reported are based on the CVs less than 10% and low limits of detection suggesting high sensitivity (Schultheiss & Stanton, 2009). Salivary hormone results from this study may contribute to the literature to help develop reference intervals of the mid-luteal phase in premenopausal females.

Identification of anovulatory cycles is challenging, and many of our participants self-reported they thought had a eumenorrheic cycle at the time of recruitment. Anovulation, however, is not uncommon given the prevalence can be as high as 37% (Prior et al., 2015) among regularly menstruating women. In hindsight, we should have included a pre-enrollment with a positive LH surge using the for confirmation (Elliott-Sale et al., 2021; Janse De Jonge et al., 2019). It is likely that the m3-step method would have performed better and classified more participants as high hormone, if a eumenorrheic menstrual status was verified prior to starting the protocol.

We also employed the use of backward counting based on cycle length, which has been noted to be less reliable than forward counting from ovulation confirmation (Schmalenberger et al., 2021). However, when no ovulation data is available, backward counting has been recommended to capture the mid-luteal phase (Schmalenberger et al., 2021). Lastly the P4/E2 ratio cut-off was determined using exploratory methods as there were no known guidelines in the literature for using salivary hormone values in relation to menstrual cycle phases and may potentially contribute to misclassification.

Our results from the logistic regression, PPV, and NPV suggest that salivary progesterone and subsequently, the P4/E2 ratio, may be useful predictors for menstrual cycle hormone classification and identification of the mid-luteal phase. This may provide insight into the overall hormone health of naturally cycling females but warrants the development of standardized salivary progesterone reference intervals for menstrual cycle phases in naturally cycling females. To develop standardized salivary progesterone intervals and capture menstrual cycle variability, serial tracking of several menstrual cycles with urinary ovulation and salivary hormone testing, and comparison to serum are recommended (Janse De Jonge et al., 2021; Schmalenberger et al., 2021; Tourville et al., 2016; Vitzthum 2021). Integrating these methods as standardized protocols may increase access to high-quality and feasible measures that can identify men-

strual cycle phases, improving field-based menstrual cycle research quality and providing insights into rural participants and female athletes health (Ferrer et al., 2024; Janse De Jonge et al., 2021).

CONCLUSION

Our results were pragmatic and showed promise in the effectiveness of implementing high-quality, field-based menstrual tracking methods in sports science research. Alterations to the method that may further improve its sensitivity include pre-enrollment urinary ovulation testing as perhaps the "surest sign of ovulation" – not just a regular cycle of 21 to 35 days (Dunne, 2018)—and/or repeated measures of ovulation and salivary hormone verification in the same cycle (Schmalenberger et al., 2021). Serial testing of these measures for multiple cycles and measuring hormones on three consecutive days in the mid-luteal phase to identify if progesterone is rising, declining, or plateaued would likely provide greater confidence for mid-luteal phase identification (Schmalenberger et al., 2021; Westerlind & Williams 2007). Therefore, our study examined the feasibility of a 3-step menstrual cycle tracking method to classify naturally cycling females as high or low hormones. The sequencing of urinary ovulation testing in cycle two and the use of salivary hormones in cycle three allowed the method to be employed where field-based testing is often needed for individuals or athletes who are unable to come to a lab for testing, thereby reducing the participant burden of extra travel time and the economic burden of serum lab costs (Allen et al., 2016). This is of particular interest with the uptake in menstrual cycle tracking of athletes when training, competition, and travel dominate their schedules (Antero et al., 2023; De Paula Oliveira et al., 2021; Winegarden et al., 2021).

DECLARATIONS

Ethical Approval

This study was approved by the Conjoint Health Research Ethics Board (CHREB) at the University of Calgary (REB21-0610) and adhered to the Declaration of Helsinki guidelines.

Data Availability Statement

The data is not publicly available due to the conditions of ethical approval for this study.

Acknowledgments

We acknowledge the participants for their contribution to data collection. We thank Dr. Constance Lebrun for contributing as the study physician and Carla van den Berg for her statistical mentorship with MLD.

Author Contributions

Conceptualization, MLD and PKDB; Methodology and Analysis, MLD, and PKDB.; Project administration, PKDB; Writing—original draft, MLD and PKDB.; Writing—review and editing, PKDB. All authors have read and agreed to the published version of the manuscript.

REFERENCES

Allen, A. M., McRae-Clark, A. L., Carlson, S., Saladin, M. E., Gray, K. M., Wetherington, C. L., McKee, S. A., & Allen, S. S. (2016). Determining menstrual phase in human biobehavioral research: A review with recommendations. *Experimental and Clinical Psychopharmacology*, *24*(1), 1–11. https://doi.org/10.1037/pha0000057

Antero, J., Golovkine, S., Niffoi, L., Meignié, A., Chassard, T., Delarochelambert, Q., Duclos, M., Maitre, C., Maciejewski, H., Diry, A., & Toussaint, J.-F. (2023). Menstrual cycle and hormonal contraceptive phases' effect on elite rowers' training, performance and wellness. *Frontiers in Physiology*, *14*, 1110526. https://doi.org/10.3389/fphys.2023.1110526

Bull, J. R., Rowland, S. P., Scherwitzl, E. B., Scherwitzl, R., Danielsson, K. G., & Harper, J. (2019). Real-world menstrual cycle characteristics of more than 600,000 menstrual cycles. *Npj Digital Medicine*, *2*(1), 83. https://doi.org/10.1038/s41746-019-0152-7

Cabre, H., Moore, S., Smith-Ryan, A., & Hackney, A. (2022). Relative energy deficiency in sport (RED-S): scientific, clinical, andpractical implications for the female athlete. *Deutsche Zeitschrift Für Sportmedizin/German Journal of Sports Medicine*, *73*(7), 225–234. https://doi.org/10.5960/dzsm.2022.546

Colenso-Semple, L. M., D'Souza, A. C., Elliott-Sale, K. J., & Phillips, S. M. (2023). Current evidence shows no influence of women's menstrual cycle phase on acute strength performance or adaptations to resistance exercise training. *Frontiers in Sports and Active Living*, *5*, 1054542. https://doi.org/10.3389/fspor.2023.1054542

De Paula Oliveira, T., Bruinvels, G., Pedlar, C. R., Moore, B., & Newell, J. (2021). Modelling menstrual cycle length in athletes using state-space models. *Scientific Reports*, *11*(1), 16972. https://doi.org/10.1038/s41598-021-95960-1

Dunne, C. (2018). Infertility testing and diagnosis for the community physician. *BCMJ*, 60(4): 203-209.

Elliott-Sale, K. J., Minahan, C. L., de Jonge, X. A. K. J., Ackerman, K. E., Sipilä, 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. https://doi.org/10.1007/s40279-021-01435-8

Emmonds, S., Heyward, O., & Jones, B. (2019). The Challenge of Applying and Undertaking Research in Female Sport. *Sports Medicine - Open*, *5*(1), 51. https://doi.org/10.1186/s40798-019-0224-x

Fehring, R. J., Schneider, M., & Raviele, K. (2006). Variability in the phases of the menstrual cycle. *Journal of Obstetric, Gynecologic, and Neonatal Nursing : JOGNN*, *35*(3), 376–384.

Ferrer, E, Rodas, G, Casals, G, Trilla, A, Balagué-Dobon, L, González, J. R., Ridley, K., White, R., & Burden, R. J. (2024). The use of saliva and blood progesterone to profile the menstrual cycles of youth professional football players. *Frontiers in Sports and Active Living*, 6: 1430158. doi: 10.3389/fspor.2024.1430158

Hampson, E. (2020). A brief guide to the menstrual cycle and oral contraceptive use for researchers in behavioral endocrinology. *Hormones and Behavior*, *119*, 104655. https://doi.org/10.1016/j.yhbeh.2019.104655

Harris, P. A., Taylor, R., Minor, B. L., Elliott, V., Fernandez, M., O'Neal, L., McLeod, L., Delacqua, G., Delacqua, F., Kirby, J., & Duda, S. N. (2019). The REDCap consortium: Building an international community of software platform partners. *Journal of Biomedical Informatics*, *95*, 103208. https://doi.org/10.1016/j.jbi.2019.103208

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. https://doi.org/10.1016/j.jbi.2008.08.010

Janse De Jonge X., Thompson, B., & Han, A. (2019). Methodological Recommendations for Menstrual Cycle Research in Sports and Exercise. *Medicine & Science in Sports & Exercise*, *51*(12), 2610–2617. https://doi.org/10.1249/MSS.0000000000002073

Johnson S., Ellis J., Godbert S., Ali S., & Zinaman M. (2011). Comparison of a digital ovulation test with three popular line ovulation tests to investigate user accuracy and certainty. *Expert Opinion on Medical Diagnostics*, *5*(6), 467–473. https://doi.org/10.1517/17530059.2011.617737

Johnson, S., Weddell, S., Godbert, S., Freundl, G., Roos, J., & Gnoth, C. (2015). Development of the first urinary reproductive hormone ranges referenced to independently determined ovulation day. *Clinical Chemistry and Laboratory Medicine*, *53*(7), 1099–1108. https://doi.org/10.1515/cclm-2014-1087

McNulty, K. L., Elliott-Sale, K. J., Dolan, E., Swinton, P. A., Ansdell, P., Goodall, S., Thomas, K., & Hicks, K. M. (2020). The Effects of Menstrual Cycle Phase on Exercise Performance in Eumenorrheic Women: A Systematic Review and Meta-Analysis. *Sports Medicine*, *50*(10), 1813–1827. https://doi.org/10.1007/s40279-020-01319-3

Meignié, A., Duclos, M., Carling, C., Orhant, E., Provost, P., Toussaint, J.-F., & Antero, J. (2021). The Effects of Menstrual Cycle Phase on Elite Athlete Performance: A Critical and Systematic Review. *Frontiers in Physiology*, *12*, 654585. https://doi.org/10.3389/fphys.2021.654585

Nemes, S., Jonasson, J. M., Genell, A., & Steineck, G. (2009). Bias in odds ratios by logistic regression modelling and sample size. *BMC Medical Research Methodology*, *9*(1), 56. https://doi.org/10.1186/1471-2288-9-56

Nillni, Y. I., Rasmusson, A. M., Paul, E. L., & Pineles, S. L. (2021). The Impact of the Menstrual Cycle and Underlying Hormones in Anxiety and PTSD: What Do We Know and Where Do We Go From Here? *Current Psychiatry Reports*, *23*(2), 8. https://doi.org/10.1007/s11920-020-01221-9

Parikh, R., Mathai, A., Parikh, S., Sekhar, G. C., & Thomas, R. (2008). Understanding and using sensitivity, speciÞcity and predictive values. *Indian Journal of Ophthalmology*, *56*(1).

Prior, J. C., Naess, M., Langhammer, A., & Forsmo, S. (2015). Ovulation Prevalence in Women with Spontaneous Normal-Length Menstrual Cycles - A Population-Based Cohort from HUNT3, Norway. *PloS one*, *10*(8), e0134473. https://doi.org/10.1371/journal.pone.0134473

Rossiter, A., Comyns, T. M., Sherwin, I., Nevill, A. M., & Warrington, G. D. (2023). Elite athletes' and support staffs' experiences and perceptions of long-haul travel, and the self-management strategies they use. *International Journal of Sports Science & Coaching*, 174795412311700. https://doi.org/10.1177/17479541231170080

Ross, R., Chaput, J.-P., Giangregorio, L. M., Janssen, I., Saunders, T. J., Kho, M. E., Poitras, V. J., Tomasone, J. R., El-Kotob, R., McLaughlin, E. C., Duggan, M., Carrier, J., Carson, V., Chastin, S. F., Latimer-Cheung, A. E., Chulak-Bozzer, T., Faulkner, G., Flood, S. M., Gazendam, M. K., … Tremblay, M. S. (2020). Canadian 24-Hour Movement Guidelines for Adults aged 18–64 years and Adults aged 65 years or older: an integration of physical activity, sedentary behaviour, and sleep. *Applied Physiology, Nutrition, and Metabolism*, *45*(10 (Suppl. 2)), S57–S102. https://doi.org/10.1139/apnm-2020-0467

Schaumberg, M. A., Jenkins, D. G., Janse de Jonge, X. A. K., Emmerton, L. M., & Skinner, T. L. (2017). Three-step method for menstrual and oral contraceptive cycle verification. *Journal of Science and Medicine in Sport*, *20*(11), 965–969. https://doi.org/10.1016/j.jsams.2016.08.013

Schmalenberger, K. M., Tauseef, H. A., Barone, J. C., Owens, S. A., Lieberman, L., Jarczok, M. N., Girdler, S. S., Kiesner, J., Ditzen, B., & Eisenlohr-Moul, T. A. (2021). How to study the menstrual cycle: Practical tools and recommendations. *Psychoneuroendocrinology*, *123*, 104895. https://doi.org/10.1016/j.psyneuen.2020.104895

Schultheiss, O.C., & Stanton, S.J. (2009). Assessment of salivary hormones, In: Harmon-Jones E, & Beer JS Eds., Methods in Social Neuroscience, 17–44. Guilford Press.

Shea, A. A., & Vitzthum, V. J. (2020). The extent and causes of natural variation in menstrual cycles: Integrating empirically-based models of ovarian cycling into research on women's health. *Drug Discovery Today: Disease Models*, *32*, 41–49. https://doi.org/10.1016/j.ddmod.2020.11.002

Satterthwaite, F. E. (1946). An Approximate Distribution of Estimates of Variance Components. *Biometrics Bulletin, 2*(6), 110-114.

Tivis, L. J., Richardson, M. D., Peddi, E., & Arjmandi, B. (2005). Saliva versus serum estradiol: Implications for research studies using postmenopausal women. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *29*(5), 727–732. https://doi.org/10.1016/j.pnpbp.2005.04.029

Tourville, T. W., Shultz, S. J., Vacek, P. M., Knudsen, E. J., Bernstein, I. M., Tourville, K. J., Hardy, D. M., Johnson, R. J., Slauterbeck, J. R., & Beynnon, B. D. (2016). Evaluation of an Algorithm to Predict Menstrual-Cycle Phase at the Time of Injury. *Journal of Athletic Training*, *51*(1), 47–56. https://doi.org/10.4085/1062-6050-51.3.01

Vitzthum, V. J., Thornburg, J., Spielvogel, H., & Deschner, T. (2021). Recognizing normal reproductive biology: A comparative analysis of variability in menstrual cycle biomarkers in German and Bolivian women. *American Journal of Human Biology*, *33*(5), e23663. https://doi.org/10.1002/ajhb.23663

Westerlind, K. C., & Williams, N. I. (2007). Effect of Energy Deficiency on Estrogen Metabolism in Premenopausal Women. *Medicine & Science in Sports & Exercise*, *39*(7), 1090–1097. https://doi.org/10.1097/mss.0b013e3180485727

Wideman, L., Montgomery, M. M., Levine, B. J., Beynnon, B. D., & Shultz, S. J. (2013). Accuracy of calendar-based methods for assigning menstrual cycle phase in women. *Sports Health*, *5*(2), 143–149. https://doi.org/10.1177/1941738112469930

Winegarden, A., Lebrun, C., Passfield, L., & Doyle-Baker, P. (2021). Three km track time trial performance changes after HIIT in competitive cross-country skiers. *International Journal of Kinesiology and Sports Science.*, *9*(1), 15–23. https://doi.org/10.7575/AIAC.IJKSS.V.9N.1P.15

Yu, J.-L., Su, Y.-F., Zhang, C., Jin, L., Lin, X.-H., Chen, L.-T., Huang, H.-F., & Wu, Y.-T. (2022). Tracking of menstrual cycles and prediction of the fertile window via measurements of basal body temperature and heart rate as well as machine-learning algorithms. *Reproductive Biology and Endocrinology*, *20*(1), 118. https://doi.org/10.1186/s12958-022-00993-4