Importance of Daily Sex Hormone Measurements Within the Menstrual Cycle for Fertility Estimates in Cyclical Shifts Studies

Urszula M. Marcinkowska[1,2] ⓘ

Abstract

Recent discussions have highlighted the importance of fertility measurements for the study of peri-ovulatory shifts in women's mating psychology and mating-related behaviors. Participants in such studies typically attend at least two test sessions, one of which is, at least in theory, scheduled to occur during the high-fertility, peri-ovulatory phase of the menstrual cycle. A crucial part of this debate is whether luteinizing hormone (LH) tests alone are sufficient to accurately assign test sessions to the peri-ovulatory phase. This article adds to this ongoing debate by presenting analyses of a detailed database of daily estradiol levels and LH tests for 102 menstrual cycles. Based on more than 4,000 hormonal measurements, it is clear that individual differences in length of the cycle, length of the luteal phase and, perhaps most importantly, the discrepancy between the timing of the LH surge and the drop in estradiol that follows it are pronounced. Less than 40% of analyzed cycles followed the *textbook* pattern commonly assumed to occur in fertility-based research, in which the LH surge is assumed to occur not more than 48 hr before the estradiol drop. These results suggest that LH tests alone are not sufficient to assign test sessions to the peri-ovulatory phase and that analyses of sex hormones are essential to identify whether the participant was tested during the peri-ovulatory phase.

Whether women's mate preferences and mating-related behaviors change during the high-fertility (i.e., peri-ovulatory) phase of the menstrual cycle so that women direct their attention and mating effort toward men displaying cues of good genes (Gangestad & Thornhill, 1998; Penton-Voak et al., 1999) has been extensively debated over the last 5 years. When investigating this peri-ovulatory shifts hypothesis, the authors aim to compare women's preferences and behaviors during the fertile phase of the menstrual cycle and low-fertility phases. After many years of research, there is still considerably little agreement about the robustness of claimed cyclic shifts, with some researchers arguing that cyclic changes in factors such as mate preferences are reliable, while other researchers have demonstrated they are not (Gangestad, 2016; Gildersleeve et al., 2014; B. C. Jones et al., 2018, 2019; Marcinkowska et al., 2017; Wood et al., 2014). There are two crucial issues in the discussion on cyclical shifts: (1) methodological concerns related to identifying the fertile window within the cycle

and (2) underestimating the complexity of determinants of women's preferences (i.e., not considering women's relationship status, self-judged attractiveness, or sexual openness).

In a typical ovulatory menstrual cycle, during the initial period of about 2 weeks (follicular phase), the follicle develops in the ovary. After maturation, high doses of estradiol are secreted from the ovary, the follicle ruptures, and an egg is released (ovulation occurs). Luteinizing hormone (LH) is a pituitary hormone secreted in response to increased levels of

[1] Faculty of Health Sciences, Jagiellonian University Medical College, Kraków, Poland
[2] Department of Anthropology, Yale University, New Haven, CT, USA

Corresponding Author:
Urszula M. Marcinkowska, Faculty of Health Sciences, Jagiellonian University Medical College, Grzegórzecka 20, 31-531 Kraków, Poland.
Email: ummarcinkowska@gmail.com

bloodstream estradiol. The LH surge usually initiates 35–44 hr before the ovulation (Cahill et al., 1998) and, together with the midcycle abrupt decline in estradiol levels (Lipson & Ellison, 1996), can be used as a reliable physiological estimate of the fertile window within a menstrual cycle (Dunson et al., 2001). After ovulation, the dominant follicle transforms into corpus luteum and begins releasing high doses of progesterone until the next menstrual bleeding occurs (R. Jones & Lopez, 2014).

Ever since the first mentions of the fertile window being around the 14th day of the menstrual cycle (Ogino, 1930), its duration has been established to be 48 hr on average, based on the basal body temperature (BBT; Siegler, 1944). Studies based on BBT observed increased fecundity starting from 5 days before the peak temperature, reaching its maximum (.3 probability of pregnancy per sexual intercourse) 2 days after the peak temperature (Barrett & Marshall, 1969). Two other methods that can be efficiently conducted outside a laboratory setting are monitoring mucus characteristics and salivary ferning, yielding 48% and 37% accuracy, respectively (Guida et al., 1999). A more recent study showed that combining mucus observation with simultaneous monitoring of LH levels provides best possible accuracy, that is, 97–99%, with peak fertility occurring 24 hr from the first positive result of the LH test (Leiva et al., 2017). Another marker of ovulation is an abrupt decline in the ratio of urinary estradiol and progesterone metabolites (Baird et al., 1991). The hormonal ratio was found to indicate the exact day of ovulation with .65 probability (Dunson et al., 2001). The only accessible direct measure of releasing the follicle from the ovary is ultrasound scanning (even though it was initially suggested to potentially induce ovulation itself; Dunson et al., 2001). This most accurate method is also the most invasive one and virtually impossible to implement in large-scale, nonlaboratory setting studies (which describes most fertility-based studies of sexual preferences).

Following recent research aimed at establishing the most accurate methodology (Blake et al., 2016) and underlining the importance of replication and validation of previously established paradigms (Marcinkowska et al., 2017), I would like to offer additional insight into fertility-based studies of human sexual preferences based on the detailed database of menstrual cycles I gathered in south Poland between 2004 and 2008. I hope to complement and provide further support for previously published discussion (Bachofner & Lobmaier, 2018; Blake, 2018; Lobmaier & Bachofner, 2018; Roney, 2018; Wallen, 2018) with more examples of good practices and recommendations for future tests of cyclic shifts in mate preferences and mating-related behaviors.

The database is arguably unique due to the robustness of hormonal measurements conducted per cycle: 15 daily measurements of two sex hormones. All daily samples were assayed in duplicates to increase the robustness of hormone measures. Moreover, recruitment was not limited solely to university students (mean age = 28 years, SD = 4.6 years), which means the data and conclusions are arguably more generalizable to the population level. Creation of this database of hormonal measurements has been costly in terms of both time and finance; hence, it is worth sharing with a broader audience in the hope that it will prove useful. Multiple articles based on this data set have been published to date (Dixson et al., 2018; Marcinkowska et al., 2017, 2018, 2019; Reynolds et al., 2018; Richards et al., 2018). This data set provides an important insight into the hormonal underpinnings of any possible cyclical fluctuations among reproductive age, contemporary women from an industrialized country.

Method

Participating women (n = 102) were between 21 and 37 years of age (mean = 28.8, SD = 4.6), did not have any medical conditions that could affect levels of reproductive steroid hormones, had been neither pregnant nor breastfeeding in the 3 months before recruitment, and had regular menstrual cycles (i.e., a difference in length between consecutive cycles of 5 days or less). Participants were asked to collect saliva samples each morning starting from the first day of menstrual bleeding, until the end of the menstrual cycle (i.e., the day before the onset of the next menstrual bleeding). Women were verbally instructed by the principal investigator concerning collecting and storing the saliva and were given a set of 2-ml centrifuge tubes with the minimum amount of required saliva marked on a tube together with written instructions. All participants received LH kits. Each kit consisted of sterilized urine cups and 10 LH tests. Women were instructed to conduct the tests from the 10th until the 20th day of the cycle or until a positive result was obtained.

Saliva samples were collected in the morning, preferably before eating, drinking, or smoking, or not earlier than 30 min after eating, drinking, or smoking. Immediately after collection, samples were frozen in the participant's home freezer. After the end of the cycle, samples were transported in portable freezers to the laboratory where hormonal assays for measurements of 17-β-estradiol (E2) and 17-α-hydroxy-progesterone (P) were conducted. Hormonal measurements were conducted using commercially available hormonal assays of DRG International, Inc. ELISA plates SLV4188 (sensitivity: 0.4 pg/ml, standard range: 1–100 pg/ml) for E2 and SLV3140 for P (sensitivity: 2.5 pg/ml, standard range: 10–5,000 pg/ml). All samples were assayed in duplicates. The quality of hormonal measurements was monitored for each plate separately by including samples of known concentrations with low and high E2 and P levels. Inter-assay variability was 10.1%, intra-assay variability was 7.5% for E2 and 14.1% and 4.9% for P, which is within acceptable values for steroid sex hormones assays (Schultheiss & Stanton, 2009).

Results and Discussion

Cycle Description

The mean duration of the cycle was 28.25 days (Min = 21, Max = 36, SD = 2.95). Peak fertility was assessed by two independent methods: the result of LH tests and daily measurements of
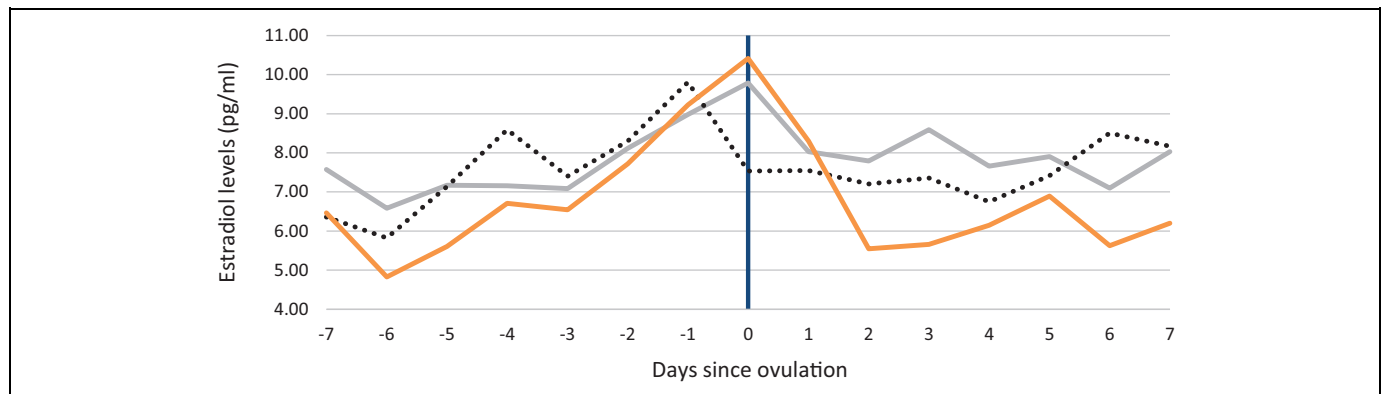
Figure 1. Averaged levels of estradiol (pg/ml) for cycles with positive luteinizing hormone (LH) test result (solid gray line), cycles with a positive LH test result and estradiol drop observed 48 hr after (solid orange line), and cycles without a positive result of LH test (dotted line, centered around −14 day of the cycle).

E2. Positive LH test results were obtained in 75 of the cycles, and midcycle E2 drops were observed in 91 cycles. LH surge occurred, on average, 14.4 days before the onset of the next menses (Min = 8, Max = 23, SD = 2.43) and E2 drop occurred, on average, 13.2 days before the onset of the next menses (Min = 6, Max = 19, SD = 2.67).

LH Surge Verification

Although LH test producers claim that once the intensity of the test line reaches the control line ovulation is most likely to occur, a significant number of participants in our sample did not obtain a positive LH test result throughout the entire cycle. LH tests depend on average (i.e., "normal") hormonal ranges, which may not correspond with those of individual participants (Direito et al., 2013). Of the 102 participants, 75 confirmed obtaining a positive LH result. Eleven of them reported that the test line appeared but never reached the intensity of the control line. Of that group E2 drop was not recorded for only one, meaning that for the remaining 10 cycles, although the results of LH tests were "negative" (defined by the producer), there may still have been ovulation, as signaled by the E2 drop. For further analyses, a total number of 75 women have been included in the "LH surge present" group (including all women for whom a pale control line appeared, Electronic Supplementary Material [ESM] 1).

There are two reasons why an LH test would not detect ovulation. The usual cutoff point (and one used in this study) for obtaining a positive result of the LH test is 25 mIU/ml. Sensitivity (true positive result) and specificity (true negative result) for a 25 mIU/ml test are 0.54 (0.29–0.77) and 0.97 (0.95–0.98), respectively (Leiva et al., 2017). Another cause of misdetection is variety in LH fluctuation patterns. Based on 43 regularly cycling women,

> the onset of urinary LH surge was categorized into rapid-onset type (within 1 day, 42.9%) and gradual-onset type (over 2–6 days, 57.1%). Configurations of LH surge can be categorized into three types: (a) spiking (41.9%); (b) biphasic (44.2%); and

(c) plateau (13.9%), and two (4.3%) women demonstrated LH surge without ovulation. (Su et al., 2017).

It is hard to predict how over-the-counter ovulation tests will react for each type of onset (and to my knowledge, such a study has not yet been conducted).

LH Surge versus Estradiol-Based Fertility Peak

The? Highest midcycle drop in E2 has been suggested as a reasonable estimate of the exact timing of ovulation (Lipson & Ellison, 1996). The E2 drop appeared, on average, 1 day after the LH surge (Min. = −4 days, Max. = 6 days, SD = 2.77; Figure 1). Sixty-nine women experienced both LH surge and E2 drop, and 54 of them experienced E2 drop after the LH surge. For 32 of them, E2 drop occurred not later than 48 hr after the LH surge, following the *textbook*. For seven women, although LH levels surged, no E2 drop was recorded (hence, their cycle was probably nonovulatory despite registering an LH surge).

Interestingly, for 22 women, although LH surge was not recorded, E2 drop was observed during the cycle. This occurred in 81% of cycles from the group with negative results of the LH tests. It is possible that these women did experience ovulation, but based on results of the LH tests, their cycles could have been erroneously classified as nonovulatory. If a woman's physiologically typical LH levels were lower than average, LH tests attuned to populational levels would not detect them: The amount of LH present in urine would not pass the threshold for color change on the test strip. Ovulation tests convert a continuous change in LH levels into a binary result (positive or negative), they do not track cyclical variation in LH, and a relatively large change in LH in an ovulatory cycle with overall low LH levels can be undetected. Such an explanation is further supported by the changes in E2 levels, averaged for all cycles where no positive LH test result was obtained, manifesting in the highest E2 drop 1 day after the −14 day of the cycle (Figure 1).
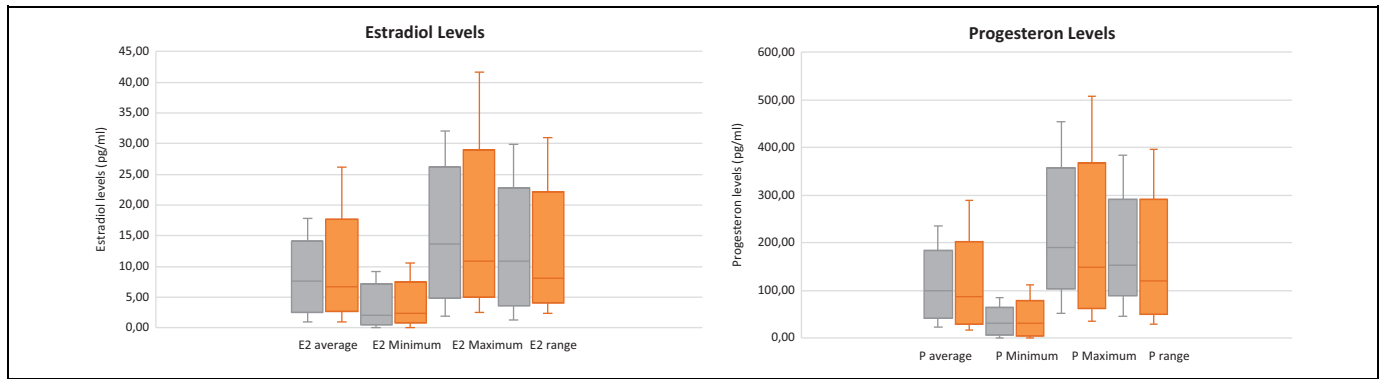
Figure 2. Hormonal characteristics averaged for *textbook* (orange) and nonovulatory (gray) cycles, the middle line of the box represents the median, the bottom and top line of the box represents the median of first quartile and third quartile, and the whiskers show minimum and maximum value.

These findings are of importance in the light of studies that use only LH tests to determine the day of ovulation and schedule research procedures based solely on LH results. From the total sample of 102 healthy, regularly menstruating, reproductive age women, only about 32 experienced the E2 drop in the vicinity of the positive result of the LH test, that is, not more than 48 hr after. The mean time difference between a positive LH test result and E2 drop was 24 hr, as expected; however, the variation here was noteworthy (from 4 days before till 6 days after), further supporting the weaker than expected time relation between LH tests results and ovulation (Bachofner & Lobmaier, 2018).

Importance of the Estradiol Measurements

As described above, measuring the LH surge alone was not sufficient to determine the occurrence and timing of the peak fertility with high accuracy. Based on a large data set of daily collected saliva samples, it is possible to narrow post hoc analyses (as done in Marcinkowska et al., 2017) based on the participants' segregation to various groups (e.g., women who did experience both LH surge and E2 drop in the right order for ovulation to occur). Without such post hoc extra examination, some participants may be misclassified as in the fertile phase when in reality they are not or may be excluded from the high-fertility group, while actually at their peak fertility.

Roney (2018) argues that by testing participants even 24 hr after the positive LH test result, one could in fact test participants in their low-fertility phase and that could lead to null findings due to phase misclassification. Author presented E/P ratios centered around ovulation day as defined by the highest E2 drop within the cycle (our measurements of mean E2 levels reflect a similar pattern; Figure 1). However, it is no surprise that E2/P ratio will echo E2 levels, and hence E2/P drop will mirror E2 drop. To show that studies might be misclassifying women as fertile when testing them 24 or 48 hr after the possible ovulation, one needs to present daily hormonal levels centered around ovulation as defined, for example, by the

positive LH test result (which also seems to be the prevalent method of pinpointing ovulation).

Differences in Sex Hormones Between Ovulatory and Nonovulatory Cycles

As overall differences of sex hormones were suggested to be a marker of overall fertility (Blackwell et al., 2018; Jasienska & Jasienski, 2008), levels of estradiol and progesterone between nonovulatory and ovulatory (based on the LH test result) cycles were compared. In a medical setting, the timing of the ovulation has often been gauged by the occurrence of an abrupt decline in the ratio of urinary metabolites of E2 and P (Baird et al., 1995). Progesterone peak during the luteal phase is often presented as the best marker for the detection of the ovulation (reaching 97% of detection rate); however, it is not as precise for ascertaining the timing of the ovulation (O'Connor et al., 2006).

In this sample, ovulatory cycles did not differ significantly from *textbook* ones in any of the following hormonal measurements: average cycle level, minimum daily level, maximum daily level, and range (within-cycle-level changes; see Figure 2 and ESM 1). Lack of differences can be caused by the aforementioned interparticipant variation being significantly larger than the intraparticipant ones (for another example of the extent of intra- and interwomen variation, see Stricker et al., 2006).

Conclusions

Although hormonal assays are demanding (from both a time and a financial point of view), based on this data set, I suggest that they are the only way that future studies can define (and validate) fertile windows of participants more accurately (see also B. C. Jones et al., 2019). As demonstrated, solely using LH tests to determine ovulation leads to both false positives (where participants who did not experience E2 drop are classified as being in their fertile window) and false negatives (where participants who might have experienced ovulation,
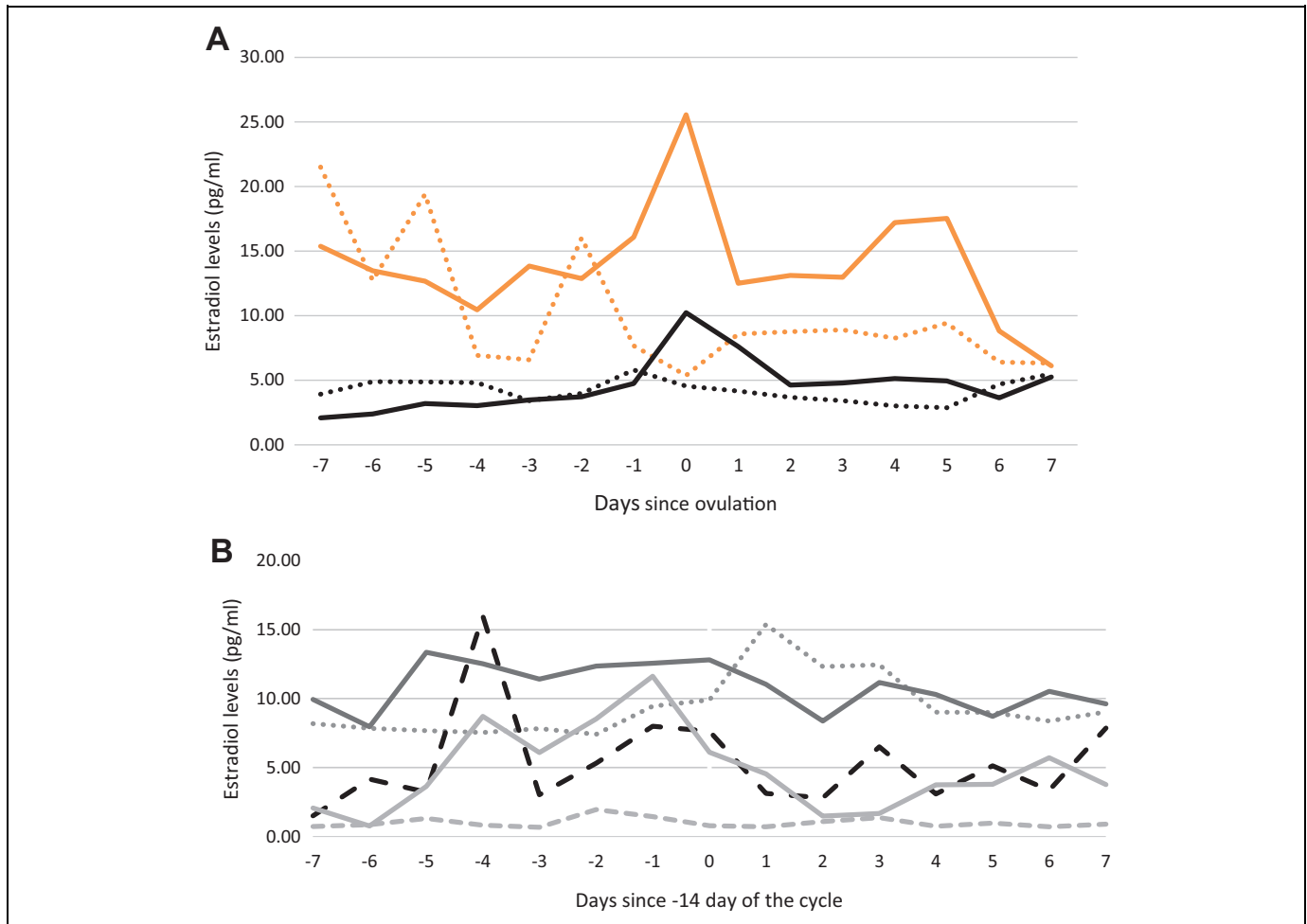
Figure 3. Daily levels of estradiol (pg/ml) from chosen cycles. (A) Ovulatory cycles centered around the day when a positive LH test result was obtained ("Day 0"). Although all cycles were "ovulatory" according to the luteinizing hormone (LH) test result, there is an array of E2 fluctuation patterns: midcycle E2 drop not present (dotted black line), E2 drop was present but relatively low (solid black line), *textbook* example of the E2 fluctuations (solid orange), and E2 drop present before the positive LH test result (dotted orange line). (B) Nonovulatory cycles with varying levels of estradiol centered around Day −14 of the cycle.

but did not obtain a positive LH test result, are excluded from the analyses).

Importantly, as there is significant interspecific variation in levels of sex hormones (Kol & Homburg, 2008; Figure 3 and ESM 2), the relative levels and the general, within-cycle trends, are of greater importance than any interspecific comparisons of daily hormonal levels (Bachofner & Lobmaier, 2018). This considerable variability can stem from the differences in genetic makeup and developmental and living conditions (Jasienska & Ellison, 2004; Jasienska & Jasienski, 2008; Jasienska, Kapiszewska, Ellison, et al., 2006; Jasienska, Ziomkiewcz, Lipson, Thune, & Ellison, 2006; Jasienska, Ziomkiewicz, Thune, Lipson, & Ellison, 2006). Conducting multiple measurements of sex hormones within one cycle seems to be crucial for possible post hoc analyses and gauging peak fertility. Reducing false negatives is essential to maximize the return on costly data collection by avoiding the unnecessary exclusion of subjects. Taking a broader perspective, it should be emphasized that the interindividual variations in hormonal levels are vast, and what we perceive as a

physiological norm should be more comprehensive. I do not want to claim that no averages can be presented to the broader public, but this should always be accompanied by an example of naturally occurring variation from a contemporary population.

Another recommendation is to increase sample sizes in the peri-ovulatory shift hypotheses studies. As we have demonstrated, fewer than 40% of participants of the initial sample could have attended a meeting in the high-fertility window (i.e., experienced E2 drop not more than 48 hr after obtaining a positive result of the LH test). If we included a stringent criterion of meeting participants not more than 24 hr after the positive LH test results, as recently suggested by Lobmaier et al. (Blake, 2018; Lobmaier & Bachofner, 2018), this percentage would have decreased even more due to difficulties arranging the meetings. If we want to test the peri-ovulatory shift hypotheses reliably, both E2 drop validation and a possibility for post hoc participant exclusion from analyses have to be in place (see also Blake, et al., 2017, for an example of post hoc sensitivity analysis). This is even more important due to

numerous recent reports of null findings in peri-ovulatory shift studies (B. C. Jones et al., 2018).

Additionally, increased sample sizes would allow for controlling of possible confounding variables (as done in Marcinkowska et al., 2017), as it is also possible that around-ovulation shifts appear only in certain environments (e.g., in short-term or long-term mating-oriented women, partnered or single, of high or low self-judged attractiveness; as also suggested by B. C. Jones et al., 2019). What is more, studies on this topic should not be conducted solely on students, as the number of nonovulatory cycles can be elevated in women younger than 20 years old (Collett et al., 1954). Daily sex hormone measurements and large sample sizes should be a gold standard in fertility-based studies of mating psychology, which examine these arguments in the future.

While designing fertility-based studies of human behavior, posing research questions, and stating hypotheses, we should be aware of the fact that we are searching for endocrine bases common in two communication streams: brain-behavior and brain-reproductive system axis. Following Roney (2018) and Bachofner and Lobmaier (2018), I underline that we should not treat conception probability solely as a dichotomous variable, but rather focus on the endocrine signaling, which provides exchange of information between the reproductive system and the brain.

Acknowledgments

Declaration of Conflicting Interests

Ethical Statement

The study was conducted with the understanding and written, informed consent from each subject, with the approval of Jagiellonian University Medical College Ethics Board, and in compliance with national legislation and the Code of Ethical Principles for Medical Research Involving Human Subjects of the World Medical Association (Declaration of Helsinki).

Funding

ORCID iD

Urszula M. Marcinkowska 🆔 https://orcid.org/0000-0002-6851-4513

Supplemental Material

Supplemental material for this article is available online.

References

Bachofner, L. M., & Lobmaier, J. S. (2018). Towards a more holistic view of fertility: The need to consider biological underpinnings rather than only data. *Hormones and Behavior*, *106*, A10–A11. https://doi.org/10.1016/j.yhbeh.2018.07.004

Baird, D. D., McConnaughey, D. R., Weinberg, C. R., Musey, P. I., Collins, D. C., Kesner, J. S., Knecht, E. A., & Wilcox, A. J. (1995). Application of a method for estimating day of ovulation using urinary estrogen and progesterone metabolites. *Epidemiology*, *6*, 547–550. https://doi.org/10.1097/00001648-199509000-00015

Baird, D. D., Weinberg, C. R., Wilcox, A. J., & McConnaughey, D. R. (1991). Using the ratio of urinary estrogen to progesterone metabolites to estimate day of ovulation. *Statistics in Medicine*, *10*, 255–266. https://doi.org/10.1002/sim.4780100209

Barrett, J. C., & Marshall, J. (1969). Risk of conception on different days of the menstrual cycle. *Population Studies—A Journal of Demography*, *23*, 455–461. https://doi.org/10.2307/2172882

Blackwell, L. F., Cooke, D. G., & Brown, S. (2018). The use of estrone-3-glucuronide and pregnanediol-3-glucuronide excretion rates to navigate the continuum of ovarian activity. *Frontiers in Public Health*, *6*. https://doi.org/10.3389/fpubh.2018.00153

Blake, K. R. (2018). Resolving speculations of methodological inadequacies in the standardized protocol for characterizing women's fertility: Comment on Lobmaier and Bachofner (2018). *Hormones and Behavior*, *106*, A4–A6. https://doi.org/10.1016/j.yhbeh.2018.05.022

Blake, K. R., Bastian, B., O'Dean, S. M., & Denson, T. F. (2017). High estradiol and low progesterone are associated with high assertiveness in women. *Psychoneuroendocrinology*, *75*, 91–99. https://doi.org/10.1016/j.psyneuen.2016.10.008

Blake, K. R., Dixson, B. J. W., O'Dean, S. M., & Denson, T. F. (2016). Standardized protocols for characterizing women's fertility: A data-driven approach. *Hormones and Behavior*, *81*, 74–83. https://doi.org/10.1016/j.yhbeh.2016.03.004

Cahill, D. J., Wardle, P. G., Harlow, C. R., & Hull, M. G. (1998). Onset of the preovulatory luteinizing hormone surge: Diurnal timing and critical follicular prerequisites. *Fertility and Sterility*, *70*, 56–59. https://doi.org/10.1016/s0015-0282(98)00113-7

Collett, M. E., Wertenberger, G. E., & Fiske, V. M. (1954). The effect of age upon the pattern of the menstrual cycle. *Fertility and Sterility*, *5*, 437–448.

Direito, A., Bailly, S., Mariani, A., & Ecochard, R. (2013). Relationships between the luteinizing hormone surge and other characteristics of the menstrual cycle in normally ovulating women. *Fertility and Sterility*, *99*, 279–U338. https://doi.org/10.1016/j.fertnstert.2012.08.047

Dixson, B. J. W., Lee, A. J., Blake, K. R., Jasienska, G., & Marcinkowska, U. M. (2018). Women's preferences for men's beards show no relation to their ovarian cycle phase and sex hormone levels. *Hormones and Behavior*, *97*, 137–144. https://doi.org/10.1016/j.yhbeh.2017.11.006

Dunson, D. B., Weinberg, C. R., Baird, D. D., Kesner, J. S., & Wilcox, A. J. (2001). Assessing human fertility using several markers of ovulation. *Statistics in Medicine*, *20*, 965–978. https://doi.org/10.1002/sim.716

Gangestad, S. W. (2016). Comment: Wood et al.'s (2014) speculations of inappropriate research practices in ovulatory cycle studies. *Emotion Review*, *8*, 87–90.

Gangestad, S. W., & Thornhill, R. (1998). Menstrual cycle variation in women's preferences for the scent of symmetrical men. *Proceedings of the Royal Society B—Biological Sciences*, *265*, 927–933.

Gildersleeve, K., Haselton, M. G., & Fales, M. R. (2014). Do women's mate preferences change across the ovulatory cycle? A meta-analytic review. *Psychological Bulletin*, *140*, 1205–1259. https://doi.org/10.1037/a0035438

Guida, N., Tommaselli, G. A., Palomba, S., Pellicano, M., Moccia, G., Di Carlo, C., & Nappi, C. (1999). Efficacy of methods for determining ovulation in a natural family planning program. *Fertility and Sterility*, *72*, 900–904. https://doi.org/10.1016/s0015-0282(99)00365-9

Jasienska, G., & Ellison, P. T. (2004). Energetic factors and seasonal changes in ovarian function in women from rural Poland. *American Journal of Human Biology*, *16*, 563–580. https://doi.org/10.1002/ajhb.20063

Jasienska, G., & Jasienski, M. (2008). Interpopulation, interindividual, intercycle, and intracycle natural variation in progesterone levels: A quantitative assessment and implications for population studies. *American Journal of Human Biology*, *20*, 35–42. https://doi.org/10.1002/ajhb.20686

Jasienska, G., Kapiszewska, M., Ellison, P. T., Kalemba-Drozdz, M., Nenko, I., Thune, I., & Ziomkiewicz, A. (2006). CYP17 genotypes differ in salivary 17-beta estradiol levels: A study based on hormonal profiles from entire menstrual cycles. *Cancer Epidemiology Biomarkers & Prevention*, *15*, 2131–2135. https://doi.org/10.1158/1055-9965.epi-06-0450

Jasienska, G., Ziomkiewcz, A., Lipson, S. F., Thune, I., & Ellison, P. T. (2006). High Ponderal index at birth predicts high estradiol levels in adult women. *American Journal of Human Biology*, *18*, 133–140. https://doi.org/10.1002/ajhb.20462

Jasienska, G., Ziomkiewicz, A., Thune, I., Lipson, S. F., & Ellison, P. T. (2006). Habitual physical activity and estradiol levels in women of reproductive age. *European Journal of Cancer Prevention*, *15*, 439–445. https://doi.org/10.1097/00008469-200610000-00009

Jones, B. C., Hahn, A. C., & DeBruine, L. M. (2019). Ovulation, sex hormones, and women's matins psychology. *Trends in Cognitive Sciences*, *23*, 51–62. https://doi.org/10.1016/j.tics.2018.10.008

Jones, B. C., Hahn, A. C., Fisher, C. I., Wang, H., Kandrik, M., Han, C., Fasolt, V., Morrison, D., Lee, A. J., Holzleitner, I. J., O'Shea, K. J., Roberts, S. C., Little, A. C., & DeBruine, L. M. (2018). No compelling evidence that preferences for facial masculinity track changes in women's hormonal status. *Psychological Science*, *29*. https://doi.org/10.1177/0956797618760197

Jones, R., & Lopez, K. H. (2014). *Human reproductive biology*. Elsevier.

Kol, S., & Homburg, R. (2008). Change, change, change: Hormonal actions depend on changes in blood levels. *Human Reproduction*, *23*, 1004–1006. https://doi.org/10.1093/humrep/den061

Leiva, R. A., Bouchard, T. P., Abdullah, S. H., & Ecochard, R. (2017). Urinary luteinizing hormone tests: Which concentration threshold best predicts ovulation? *Frontiers in Public Health*, *5*, 320–320. https://doi.org/10.3389/fpubh.2017.00320

Lipson, S. F., & Ellison, P. T. (1996). Comparison of salivary steroid profiles in naturally occurring conception and non-conception cycles. *Human Reproduction*, *11*, 2090–2096.

Lobmaier, J. S., & Bachofner, L. M. (2018). Timing is crucial: Some critical thoughts on using LH tests to determine women's current fertility. *Hormones and Behavior*, *106*, A2–A3. https://doi.org/10.1016/j.yhbeh.2018.07.005

Marcinkowska, U. M., Galbarczyk, A., & Jasienska, G. (2017). La donna e mobile? Lack of cyclical shifts in facial symmetry, and facial and body masculinity preferences—A hormone based study. *Psychoneuroendocrinology*, *88*, 47–53. https://doi.org/10.1016/j.psyneuen.2017.11.007

Marcinkowska, U. M., Helle, S., Jones, B. C., & Jasienska, G. (2019). Does testosterone predict women's preference for facial masculinity? *PLoS One*, *14*. https://doi.org/10.1371/journal.pone.0210636

Marcinkowska, U. M., Kaminski, G., Little, A. C., & Jasienska, G. (2018). Average ovarian hormone levels, rather than daily values and their fluctuations, are related to facial preferences among women. *Hormones and Behavior*, *102*, 114–119. https://doi.org/10.1016/j.yhbeh.2018.05.013

O'Connor, K. A., Brindle, E., Brindle, R. C., Shofer, J. B., Ferrell, R. J., Klein, N. A., Soules, M. R., Holman, D. J., Mansfield, P. K., & Wood, J. W. (2006). Ovulation detection methods for urinary hormones: Precision, daily and intermittent sampling and a combined hierarchical method. *Human Reproduction*, *21*, 1442–1452. https://doi.org/10.1093/humrep/dei497

Ogino, K. (1930). Ovulationstermin und kon-zeptionstermin [Ovulation time and conception time]. *Zentralbl Gynakol*, *54*, 464–478.

Penton-Voak, I. S., Perrett, D. I., Castles, D. L., Kobayashi, T., Burt, D. M., Murray, L. K., & Minamisawa, R. (1999). Menstrual cycle alters face preference. *Nature*, *399*, 741–742. https://doi.org/10.1038/21557

Reynolds, T. A., Makhanova, A., Marcinkowska, U. M., Jasienska, G., McNulty, J. K., Eckel, L. A., Nikonova, L., & Maner, J. K. (2018). Progesterone and women's anxiety across the menstrual cycle. *Hormones and Behavior*, *102*, 34–40. https://doi.org/10.1016/j.yhbeh.2018.04.008

Richards, G., Klimek, M., Jasienska, G., & Marcinkowska, U. M. (2018). Digit ratio (2D:4D) and circulating testosterone, oestradiol, and progesterone levels across the menstrual cycle. *Early Human Development*, *117*, 68–73. https://doi.org/10.1016/j.earlhumdev.2017.12.006

Roney, J. R. (2018). Hormonal mechanisms and the optimal use of luteinizing hormone tests in human menstrual cycle research. *Hormones and Behavior*, *106*, A7–A9. https://doi.org/10.1016/j.yhbeh.2018.05.021

Schultheiss, O. C., & Stanton, S. J. (2009). *Assessment of salivary hormones*. In E. Harmon-Jones & J. S. Beer (Eds.), *Methods in social neuroscience*. Guilford Press.

Siegler, S. L. (1944). *Fertility in women*. Lippincott J. B.

Stricker, R., Eberhart, R., Chevailler, M. C., Quinn, F. A., Bischof, P., & Stricker, R. (2006). Establishment of detailed reference values

for luteinizing hormone, follicle stimulating hormone, estradiol, and progesterone during different phases of the menstrual cycle on the Abbott Architect (R) analyzer. *Clinical Chemistry and Laboratory Medicine*, *44*, 883–887. https://doi.org/10.1515/cclm.2006.160

Su, H.-W., Yi, Y.-C., Wei, T.-Y., Chang, T.-C., & Cheng, C.-M. (2017). Detection of ovulation, a review of currently available methods. *Bioengineering & Translational Medicine*, *2*, 238–246. https://doi.org/10.1002/btm2.10058

Wallen, K. (2018). LH and fertility: A discussion of identifying the period of maximal fertility in naturally cycling women. *Hormones and Behavior*, *106*, A1. https://doi.org/10.1016/j.yhbeh.2018.08.001

Wood, W., Kressel, L., Joshi, P. D., & Louie, B. (2014). Meta-analysis of menstrual cycle effects on women's mate preferences. *Emotion Review*, *6*, 229–249. https://doi.org/10.1177/1754073914523073