Published in final edited form as:

Fertil Steril. 2014 August ; 102(2): 511–518.e2. doi:10.1016/j.fertnstert.2014.04.035.

Assessment of Anovulation in Eumenorrheic Women: Comparison of Ovulation Detection Algorithms

Kristine E. Lynch, Ph.D.^{a,b}, Sunni L. Mumford, Ph.D.^a, Karen C. Schliep, Ph.D.^a, Brian W. Whitcomb, Ph.D.^c, Shvetha M. Zarek, M.D.^{a,d}, Anna Z Pollack, Ph.D.^{a,e}, Elizabeth R. Bertone-Johnson, Sc.D.^c, Michelle Danaher, Ph.D.^a, Jean Wactawski-Wende, Ph.D.^f, Audrey J. Gaskins, B.S.E.^{a,g}, and Enrique F. Schisterman, Ph.D.^a

^aEpidemiology Branch, Division of Intramural Population Health Research, Eunice Kennedy Shriver National Institute of Child Health & Human Development, NIH, DHHS, Bethesda, Maryland ^bDepartment of Public Health and Nursing, Westminster College, Salt Lake City, Utah ^cDivision of Biostatistics and Epidemiology, University of Massachusetts School of Public Health and Health Sciences, Amherst, Massachusetts ^dProgram of Reproductive and Adult Endocrinology, Eunice Kennedy Shriver National Institute of Child Health & Human Development, NIH, DHHS, Bethesda, Maryland ^eDepartment of Global and Community Health, George Mason University, Fairfax, Virginia ^fDepartment of Social and Preventive Medicine, University at Buffalo, Buffalo, New York ^gDepartments of Nutrition and Epidemiology, Harvard School of Public Health, Boston, Massachusetts

Abstract

Objective—To compare previously used algorithms to identify anovulatory menstrual cycles in women self-reporting regular menses.

Design—Prospective cohort study

Setting—Western New York

Study participants—259 healthy, regularly menstruating women followed for one (n=9) or two (n=250) menstrual cycles (2005–2007).

Intervention(s)—None.

Main Outcome Measure(s)—Prevalence of sporadic anovulatory cycles identified using eleven previously defined algorithms that utilize estradiol, progesterone, and luteinizing hormone (LH) concentrations.

Result(s)—Algorithms based on serum LH, estradiol, and progesterone levels detected a prevalence of anovulation across the study period of 5.5% to 12.8% (concordant classification for

Corresponding Author: Enrique F. Schisterman, *Eunice Kennedy Shriver* National Institute of Child Health and Human Development 6100 Executive Blvd., 7B03; Rockville, MD 20852; Phone (301) 435-6893; schistee@mail.nih.gov.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

91.7% to 97.4% of cycles). The prevalence of anovulatory cycles varied from 3.4% to 18.6% using algorithms based on urinary LH alone or with the primary estradiol metabolite, estrone-3-glucuronide (E3G), levels.

Conclusion(s)—The prevalence of anovulatory cycles among healthy women varied by algorithm. Mid-cycle LH surge urine-based algorithms used in over-the-counter fertility monitors tended to classify a higher proportion of anovulatory cycles compared to luteal phase progesterone serum-based algorithms. Our study demonstrates that algorithms based on the LH surge, or in conjunction with E3G, potentially estimate a higher percentage of anovulatory episodes. Addition of measurements of post-ovulatory serum progesterone or urine pregnanediol may aid in detecting ovulation.

Keywords

ovulation; menstrual cycles; progesterone; luteinizing hormone; estradiol

Introduction

Chronic anovulation has been associated with increased risk of infertility (1), decreased bone mineral density (2), endometrial cancer (3), and when associated with irregular menses, is one of several diagnostic criteria for polycystic ovary syndrome (4). The relationships between dietary (5–7), behavioral (8), environmental factors (9), and anovulation represent modifiable factors to improve fertility and chronic health conditions. Valid estimates of these relationships, however, depend on accurately identifying anovulatory menstrual cycles.

Identification of anovulatory cycles is challenging. Transvaginal ultrasound, the gold-standard (10), involves daily, mid-cycle ultrasounds, which is resource intensive and thus impractical in epidemiological studies. In the absence of gold standard methods, daily or multiple well-timed measurements of reproductive hormone concentrations are commonly used to identify ovulatory status in research settings. However, given previous research showing anovulation prevalence ranging from 3.7% (11) to 23% (12) among regularly menstruating women using hormone assessment, the best strategy to identify ovulation remains under debate (6, 13–28). It is not clear whether differences in prevalence estimates are due to ovulation assessment method or differences in study design, length of follow-up, or study population.

Algorithms to identify anovulation differ in the hormones assessed, thresholds used, and the cycle phase. A comparison of algorithms used to assess ovulatory status has not been performed. Therefore, the purpose of the present study was to compare the prevalence of anovulatory cycles across two menstrual cycles using previously established algorithms among a cohort of healthy, premenopausal women.

MATERIALS AND METHODS

Study Population

The BioCycle Study enrolled 259 regularly menstruating women for one (n=9) or two (n=250) cycles as previously described (29). Participants were female volunteers aged 18 to 44 years from the Western New York region. Women reporting at least three regular menstrual periods in the past three months, no cycle less than 21 days or more than 35 days in the past six months, and no history of gynecological disorders or chronic disease were eligible. Women were excluded if they had used hormonal contraceptives in the past three months (12 months if long-acting), or were currently pregnant or breastfeeding. The University at Buffalo Health Sciences Institutional Review Board (IRB) approved the study and served as the IRB designated by the National Institutes of Health under a reliance agreement. All participants provided written informed consent.

Hormone Assessment

Participants attended up to eight clinic visits per cycle. The visits were scheduled to occur on approximately the second day of menstruation, one visit in the mid follicular phase, three visits during the periovulatory phase, and one visit each in the early, mid, and late luteal phases. Clinic visits were initially scheduled using an algorithm accounting for each woman's self-reported cycle length, with mid-cycle visits adjusted based on fertility monitor indication of days of peak fertility (Clearblue® Easy Fertility Monitor, Inverness Medical, Waltham, MA, USA) (30). Participants synchronized the monitor to their cycle and then checked the monitor daily for a prompt from the monitor to submit a sample. Thus, between the sixth and the ninth days of a woman's cycle (depending on her cycle length history), the monitor begins to request daily tests for 10 days. On test days, the woman briefly submerges a test stick in her first-morning urine and then inserts the test stick into the monitor. The test stick has a nitrocellulose strip with an anti-LH antibody zone and an estrone-3-glucuronide (E3G) conjugate zone. The monitor optically reads the level of E3G and LH in the urine by the intensity of the lines in the corresponding zones. Each day, the monitor assigns the woman to low, high, or peak fertility on the basis of her E3G and LH levels (peak corresponds with approximately 30 IU/L for LH). Thus, unlike home LH test sticks, the monitor provides information to help anticipate the LH surge. If the monitor indicated 'peak fertility' on a day without a scheduled visit, the participant was asked to come in that morning and the following two mornings. Using the fertility monitor, menstrual cycle phase could be defined more accurately (31). Women were highly adherent to the study protocol, with 94% of women completing 7 or 8 visits per cycle.

Monitors were adapted to capture and store hormone levels on an internal computer chip that was available for download for this study. Data card readers were used to download daily urine E3G and LH values for each participant at every clinic visit. Fasting morning blood samples were collected at each of the eight cycle visits and processed according to standardized protocols. Samples were frozen at -80°C and sent as complete participant cycle batches to the Kaleida Health Center for Laboratory Medicine (Buffalo, NY) for analysis of hormone concentrations. Estradiol, progesterone, LH, and follicle stimulating hormone (FSH) were measured using IMMUNLITE® 2000 Solid phase competitive

Chemiluminescent Enzymatic Immunoassay by Specialty Laboratories, Inc. (Valencia, CA) on the DPC Immulite 2000 analyzer (Siemens Medical Solutions Diagnostics, Deerfield, IL).

Algorithm Descriptions

We compared eleven ovulation detection algorithms, nine of which have been previously described (11, 14, 17–19, 22, 32, 34, 36), and two that were specifically adapted for the BioCycle Study (6, 33). Detailed descriptions of the algorithms are given below. We applied algorithms designed for urinary hormone measures (18) to serum samples, since urinary and serum steroid hormone levels have high concordance (1). The algorithms used to classify cycles fall into three categories: 1) luteal phase progesterone activity (algorithms 1–5); 2) luteal day transition (algorithm 6); and 3) mid-cycle LH surge (algorithms 7–11).

Algorithms

Luteal Phase Progesterone Activity Algorithms (Algorithms 1–5)—After ovulation, the corpus luteum produces elevated amounts of progesterone to prepare the endometrium for implantation (1).

Algorithm 1, Progesterone Ratio (P-R), Relative threshold of progesterone—Designed for daily progesterone measurements, the relative threshold algorithm compares daily progesterone levels to baseline levels. The baseline level is usually calculated by averaging daily progesterone concentrations taken on days six to ten after the start of menses. The baseline level is used as the denominator and the daily progesterone level is the numerator. A cycle is classified as anovulatory if the ratio of daily to baseline progesterone does not exceed 3.0 for three or more days. This threshold is thought to reflect steady progesterone production by the corpus luteum in the luteal phase (17).

Algorithm 2, Absolute threshold of progesterone 5 ng/mL (P5)—Progesterone above 5.0 ng/mL (1 ng/mL=3.2 nmol/L) on at least one luteal phase sample day from multiple measurements to constitute an ovulatory cycle (11).

Algorithm 3, Absolute threshold of progesterone 3 ng/mL (P3)—Progesterone above 3.0 ng/mL on at least one luteal phase sample day from multiple measurements to constitute an ovulatory cycle (32).

Algorithm 4, BioCycle Study specific absolute threshold of progesterone and timing of LH peak (Bio-P5-LH)—In the BioCycle Study, the Bio-P5-LH algorithm was operationalized as follows: a cycle was defined as anovulatory if peak progesterone concentrations were ≤ 5.0 ng/mL and there was no observed serum LH peak on the mid or late luteal phase visits. For application to other studies with measures of LH peak and progesterone, the timing of the LH peak in relation to the peak progesterone measurements can inform whether an adequate number of luteal serum draws were obtained to assess progesterone rise (6).

Algorithm 5, BioCycle Study specific absolute threshold of progesterone and timing of LH peak (Bio-P3-LH)—We applied the same algorithm described above (Bio-P5-LH) using 3.0 ng/mL as the threshold progesterone level. Cycles were defined as anovulatory if peak progesterone concentrations were ≤ 3 ng/mL and no serum LH peak was observed on the mid or late luteal phase visits (33).

Luteal Day Transition (LDT) Algorithm (Algorithm 6)—LDT recognizes the hormonal changes that transpire during the transition from ovulation to the luteal phase, which include decreased follicular estrogen production and increased progesterone production from the corpus luteum (20). The peak ratio of urinary E3G and pregnanediol-3-glucuronide, metabolites of estrogen and progesterone respectively, serves as a proxy for ovulation because of its strong concordance with the LH peak. The LDT algorithm compares five-day sequences of this ratio throughout the cycle and identifies a set of five days in which the ratio value is highest at day one of the sequence and is 40% less than day one at days four and five. A cycle that does not have such a sequence is classified as anovulatory, while the reverse is considered ovulatory.

Algorithm 6, LDT—In this study, the LDT algorithm was applied using serum estradiol and progesterone levels, with linear interpolation utilized to estimate daily hormone levels between measured levels (details described in the statistical analysis section) (18).

Mid-cycle LH Surge (Algorithms 7–11)—The abrupt increase in mid-cycle LH levels is a positive feedback response to the late follicular phase increase in estrogen levels that occur as ovarian follicles mature (1). As LH is essential for complete maturation of the ovarian follicle, initiation of follicular rupture, and the expulsion of the ovum from the fully matured follicle, a mid-cycle LH surge is commonly used as a marker of ovulation. The majority of ovulation prediction kits or fertility monitors use the LH surge to identify estimated day of ovulation.

Algorithm 7, LH Fertility Monitor (LH-FM), internal algorithm of the fertility monitor to detect LH surge (34)—The Clearblue ® Easy Fertility Monitor used for the BioCycle Study identified low, high, and peak fertility based on the manufacturer's proprietary algorithm. A reading of "high fertility" indicates urinary E3G concentrations (typically between 20 and 30 ng/mL) that correlate with the estrogen rise typically observed in the follicular phase (34). The monitor continues to provide a high reading until a threshold level of urinary LH is detected, at approximately >30 IU/I, indicating peak fertility. Usually the monitor indicates at least two days of high fertility followed by typically two days of peak fertility. Rarely, the day of the estrogen threshold coincides with the day of the LH surge, and the monitor will register peak fertility immediately following low fertility. Some women only see low and high signals, particularly if they miss test days or if the cycle is anovulatory. We classified cycles with a 'peak fertility' reading as ovulatory and anovulatory otherwise. While fertility monitors are used by clinicians primarily to identify the timing of ovulation rather than the presence of ovulation, fertility monitors have been used to identify ovulation in epidemiologic studies (35).

Algorithm 8, LH-S1—LH surge defined as a 2.5-fold increase in LH from that of baseline LH levels (14).

Algorithm 9, LH-S2—LH surge defined as a 4.0-fold increase in LH from that of baseline LH levels (19).

Algorithm 10, LH-S3—LH surge defined as a LH level > 180% above the average of the preceding four LH values (22).

Algorithm 11, LH-S4—LH surge defined as the first LH value exceeding +3 SD above the averages of both the immediately preceding and following five days that is also of 2 days duration (36).

For the above algorithms (LH-S1 to LH-S4), we classified cycles as anovulatory if no LH surge was observed and as ovulatory otherwise.

Statistical Analysis

Descriptive statistics were calculated for the study population. We calculated the prevalence of anovulatory cycles observed across the study period (two cycles) according to each algorithm. We further compared pairs of algorithms using percent agreement and Cohen's κ statistics, along with calculating percent of cycles that were considered anovulatory by all serum and/or urine algorithms. Linear interpolation was utilized for two algorithms requiring daily measurements to estimate hormone levels on days where a serum sample was not collected (P-R, LDT). The five LH algorithms utilized available daily urine LH and/or E3G data measured by the Clearblue ® Easy Fertility Monitor. Algorithms using serum data were applied to the full dataset (n=509 cycles; n=3903 (95.8%) total visits) whereas cycles that had complete fertility monitor data (n=445 (87.4%)) (30) were used for the urine algorithm analyses. A sensitivity analysis on the fertility monitor data was conducted to determine if excluding cycles that were thought to have failed to reach a peak for non-hormonal reasons (e.g., missed test, monitor data available only before peak, and test stick reading errors) affected our study conclusions (30).

We evaluated the hormonal patterns between women with cycles classified as ovulatory or anovulatory by all algorithms, and present geometric mean hormone levels across the cycle for two commonly used algorithms, P5 and LH-FM, to represent patterns utilizing both serum and urine algorithms. All analyses were carried out using SAS version 9.3 (SAS Institute, Cary, NC).

RESULTS

Participants were relatively young (mean age=27.3 $\pm$ 8.2), of white race (n=154 [59.5%]), and of normal BMI (24.1 $\pm$ 3.9) (Table 1). The prevalence of anovulatory cycles varied from 3.4% to 18.6% overall, with a range of 5.5% to 12.8% using algorithms based on serum hormone concentrations and from 3.4% to 18.6% using algorithms applied to urinary LH concentrations from the fertility monitor (Table 2). There were 12 (2.4%) specific cycles considered anovulatory by all serum-based algorithms, though only 4 of these were also

considered anovulatory by at least one of the urine-based algorithms (LH-FM, LH-S1, LH-S4). In contrast, no specific cycles were considered anovulatory by all of the urine-based algorithms. Among the serum hormone algorithms, the absolute progesterone level ≥ 5 ng/mL algorithm (P5) identified 12.8% of cycles, compared to Bio-P3-LH, which identified 5.5%. For the algorithms that utilized daily urine measurements, the proportion of anovulatory cycles depended on the LH surge definition, with the lowest proportion for LH $\geq 180\%$ of the mean plus 2 SDs (3.4%, LH-S3), and the highest proportion for LH values not exceeding the mean plus 3 SDs (18.6%, LH-S4). We found similar results when we excluded the 33 cycles that failed to reach peak for non-hormonal reasons (range for fertility monitor algorithms was 2.7% to 17.4%, data not shown).

The six serum-based algorithms provided concordant classification for on average 94.8% of the cycles (range: 91.7% to 97.4%) (Table 3). As expected, similar algorithms (i.e., P-R, P5, P3, Bio-P5-LH, and Bio-P3-LH) had the highest magnitude of agreement (κ statistic mean: 0.66, range: 0.53 to 0.84) compared to serum algorithms using more dissimilar criteria (i.e., LDT, versus Bio-P5-LH, or Bio-P3-LH), which had the lowest magnitude of agreement (0.39 and 0.40, respectively). Overall pairwise concordant classification for the five urine-based algorithms averaged 80.1% (range: 73.0% to 86.0%), though κ statistics were much lower than for the serum based algorithms (range: -0.11 to 0.49), which can partially be attributed to the unbalanced nature of the monitor data (37). Cross classification within the serum algorithm group and within the urinary algorithm group is provided in Supplementary Tables 1 and 2.

Geometric mean hormone concentrations for cycles classified as ovulatory and anovulatory by P5 and LH-FM are shown in Figure 1. Cycles classified as ovulatory using either algorithm tended to display classical hormone profiles, including a clear estrogen peak preceding the mid-cycle LH surge. These cycles also displayed visual evidence of a luteal phase with a late cycle progesterone rise and subsequent decline prior to the start of the next cycle. However, cycles classified as anovulatory using P5 had consistently lower geometric mean hormone concentrations across the cycle than anovulatory cycles based on LH-FM using the fertility monitor. In addition, anovulatory cycles based on LH-FM demonstrated rises in estrogen, LH, and progesterone, though at lower concentrations than the ovulatory cycles.

DISCUSSION

The prevalence of anovulation varied greatly among this cohort of healthy women. In general, algorithms based on serum LH and luteal progesterone tended to estimate a lower proportion of anovulatory cycles than algorithms based solely on mid-cycle urinary LH measurements. These findings are particularly relevant for research on risk factors for anovulation, as the prevalence of anovulation can vary from 3.4–18.6% among healthy, eumenorrheic menstruating women depending on the algorithm used, and would likely vary more among a less healthy population. Our study importantly demonstrates that algorithms based on the LH surge alone, or in conjunction with E3G, potentially estimate a higher percentage of anovulatory episodes. Further research is needed to describe which hormone-

based algorithms most closely correspond to the gold standard ultrasound measurement, and how population characteristics influence algorithm accuracy.

Studies assessing luteal phase activity algorithms in eumenorrheic women reported prevalences of 6.4 to 10% for the P-R algorithm (13, 17), and 3.7% for the P5 algorithm (11), whereas comparatively, our study population demonstrated anovulation prevalences of 8.5% and 12.8% respectively. For the P5 algorithm, these differences may be due to the strict definition of eumenorrhea, which restricted to cycles that varied ± 1 day (11), as compared to the self-reported cycle range of 21–35 days in the present study. The prevalence of anovulation was similar to studies using the LDT algorithm, which reported a prevalence of anovulation of 10–13% (18) compared to 6.3% found in our study.

Defining anovulation by fertility monitor LH surge resulted in more than four-fold variability. Specifically, in previous studies prevalences varied from 10.0% for LH-FM (34), 6.5% for LH-S1 (14), 17.0% for LH-S2 (19), and 29.0% for LH-S3 (22), whereas anovulation prevalences for these algorithms in our study population were 17.4%, 15.6%, 11.0% and 3.4%, respectively. The marked difference for LH-S3, which identified 29.0% anovulatory cycles (22) compared to 3.4% in our study, was most likely due to differences in study population and sample size, as Testart et al. followed only 20 participants undergoing ovulation induction (22). Apart from this exception, anovulation prevalence was fairly stable using the same algorithms when comparing anovulation among different populations of healthy, eumenorrheic women. Validation studies comparing an ovulation algorithm to the gold standard of ultrasound guided ovulation detection would be ideal as the one available ultrasound study was limited to 53 participants and only identified one cycle as anovulatory (34).

Our results suggest that the proportion of cycles with subtle hormonal fluctuations is common and explains the algorithm disagreement. Algorithm disagreement occurs where there are subtle deviations from the classic hormonal profiles, such as less defined rises in LH and estrogen mid-cycle. In the present study, the serum based non-fertility monitor algorithms disagreed on ovulatory status in 17.3% of the cycles. Algorithms perform similarly when cycles display either the classical menstrual cycle pattern or an obvious deviation from that pattern (i.e., low, flat levels of all hormones). Our data suggest that algorithm choice is important when there are modest changes from typical ovulatory or anovulatory cycles.

We observed, on average, a higher proportion of anovulatory cycles and lower percent agreement based upon the LH surge with daily urine measurements, compared to serum progesterone alone or in combination with serum LH and/or estrogen. Our results are not surprising given the variability in timing and magnitude of an LH surge. Direito et al. observed extreme variability in LH surge duration, amplitude, and configuration among a cohort of regularly menstruating women 18–45 years of age (38). Using urinary samples, LH levels 2.1 to 61 times baseline levels were observed (38). New monitors that include urinary measurements of progesterone metabolites (24, 25) should help reduce the variability in assessment of ovulatory status; however, further research is needed to compare monitor performance to transvaginal ultrasound among larger cohorts of women (34).

Each clinic visit in the present study was scheduled based on a calendar and aided by the use of a daily home fertility monitor. While fertility monitors reliably detect the LH surge (30, 34), their ability to do so decreases with short or long cycles (23). Therefore algorithms that incorporate timing of the LH surge, such as Bio-P5-LH and Bio-P3-LH, may be more appropriate in studies with sampling designs that are sensitive to varying cycle lengths.

The present study had some limitations. In particular, participants were healthy, regularly menstruating women, and results may not be generalizable to populations at increased risk for chronic anovulatory cycles. The inclusion/exclusion criteria were selected to represent those with no known gynecologic disorders or menstrual irregularities, and accordingly these results are relevant to regularly cycling women. Many of the algorithms we considered are based on daily hormone urine samples, and our serum sampling scheme was limited to eight measurements per cycle. Therefore, missing data could have limited our ability to apply such algorithms, though our study had much more frequent measures than most. We used linear interpolation to generate hormone values for missing values between clinic visits. The days for bio-specimen collection were intended as the most dynamic cycle days, such that linear interpolation should be a close approximation to the true curvilinear functions. Though the current study used fertility monitors to help time the clinic visits, it is possible that visits were mistimed, which would affect the interpolation and prevalence estimates. Finally, our inability to compare algorithms with gold-standard transvaginal ultrasound is a limitation and future research is needed to distinguish true anovulatory cycles from those with more subtle disorders such as luteal phase deficiency (39).

Our study had several strengths. Previous research evaluating anovulatory episodes has been limited to small cohorts of women (12, 15, 34, 35) who were seeking care (35) or had fertility problems (15, 22). In contrast, our work is broadly applicable to healthy, premenopausal women. Additionally, ours is the first study to utilize both urinary daily fertility monitor data and serum hormone measurements timed at periods of key hormonal variability among a relatively large sample of women.

The proportion of anovulatory cycles observed among healthy, regularly menstruating women varied by algorithm, with mid-cycle LH surge algorithms tending to identify a higher proportion of anovulatory cycles (average 13.2%), and luteal progesterone algorithms identifying a lower proportion (average 8.4%). These results suggest that addition of measurements of post-ovulatory progesterone or urine pregnanediol may aid in detecting ovulation. Our findings advance knowledge in the prevalence of sporadic anovulatory cycles among regularly menstruating women, but future research is warranted to determine which to utilize when ultrasound is unavailable or not practical, such as in large epidemiologic studies or randomized clinical trials. Practical considerations including available hormone assays and timing and frequency of biospecimen collection may also play a role in determining the optimal algorithm for clinical or research use. A future validation study that compares algorithms to a gold standard method such as ultrasound is essential, particularly in determining the true prevalence of anovulatory cycles among premenopausal women, risk factors for anovulation, long-term implications of anovulation, and for clinical recommendations.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was supported by the Intramural Research Program of the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD), National Institutes of Health, Bethesda, Maryland (Contract # HHSN275200403394C). Audrey Gaskins is supported by NIH training grants T32DK007703-16 and T32HD060454.

REFERENCES

1. Fauser, BCJM. Medial Approaches to Ovarian Stimulation in Infertility. In: Strauss, JF., III; Robert, L.; Barbieri, RL., editors. *Yen & Jaffe's Reproductive Endocrinology: Physiology, Pathophysiology, and Clinical Management*. Philadelphia: Elsevier Saunders; 2013. p. 701-733.
2. Drinkwater BL, Nilson K, Chesnut CH 3rd, Bremner WJ, Shainholtz S, Southworth MB. Bone mineral content of amenorrheic and eumenorrheic athletes. *N Engl J Med*. 1984; 311:277–281. [PubMed: 6738640]
3. Kaaks R, Lukanova A, Kurzer MS. Obesity, endogenous hormones, and endometrial cancer risk: a synthetic review. *Cancer Epidemiol Biomarkers Prev*. 2002; 11:1531–1543. [PubMed: 12496040]
4. Franks S. Polycystic ovary syndrome. *N Engl J Med*. 1995; 333:853–861. [PubMed: 7651477]
5. Fenster L, Quale C, Waller K, Windham GC, Elkin EP, Benowitz N, et al. Caffeine consumption and menstrual function. *Am J Epidemiol*. 1999; 149:550–557. [PubMed: 10084244]
6. Gaskins AJ, Mumford SL, Zhang C, Wactawski-Wende J, Hovey KM, Whitcomb BW, et al. Effect of daily fiber intake on reproductive function: the BioCycle Study. *Am J Clin Nutr*. 2009; 90:1061–1069. [PubMed: 19692496]
7. Schliep KC, Schisterman EF, Mumford SL, Pollack AZ, Zhang C, Ye A, et al. Caffeinated beverage intake and reproductive hormones among premenopausal women in the BioCycle Study. *Am J Clin Nutr*. 2012; 95:488–497. [PubMed: 22237060]
8. De Souza MJ, Van Heest J, Demers LM, Lasley BL. Luteal phase deficiency in recreational runners: evidence for a hypometabolic state. *J Clin Endocrinol Metab*. 2003; 88:337–346. [PubMed: 12519874]
9. Bretveld RW, Thomas CM, Scheepers PT, Zielhuis GA, Roeleveld N. Pesticide exposure: the hormonal function of the female reproductive system disrupted? *Reprod Biol Endocrinol*. 2006; 4:30. [PubMed: 16737536]
10. Ecochard R, Boehringer H, Rabilloud M, Marret H. Chronological aspects of ultrasonic, hormonal, and other indirect indices of ovulation. *Bjog*. 2001; 108:822–829. [PubMed: 11510707]
11. Malcolm CE, Cumming DC. Does anovulation exist in eumenorrheic women? *Obstet Gynecol*. 2003; 102:317–318. [PubMed: 12907106]
12. Wathen NC, Perry L, Lilford RJ, Chard T. Interpretation of single progesterone measurement in diagnosis of anovulation and defective luteal phase: observations on analysis of the normal range. *Br Med J (Clin Res Ed)*. 1984; 288:7–9.
13. Waller K, Swan SH, Windham GC, Fenster L, Elkin EP, Lasley BL. Use of urine biomarkers to evaluate menstrual function in healthy premenopausal women. *Am J Epidemiol*. 1998; 147:1071–1080. [PubMed: 9620051]
14. Park SJ, Goldsmith LT, Skurnick JH, Wojtczuk A, Weiss G. Characteristics of the urinary luteinizing hormone surge in young ovulatory women. *Fertil Steril*. 2007; 88:684–690. [PubMed: 17434509]
15. McConnell HJ, O'Connor KA, Brindle E, Williams NI. Validity of methods for analyzing urinary steroid data to detect ovulation in athletes. *Med Sci Sports Exerc*. 2002; 34:1836–1844. [PubMed: 12439091]

16. O'Connor KA, Brindle E, Miller RC, Shofer JB, Ferrell RJ, Klein NA, et al. Ovulation detection methods for urinary hormones: precision, daily and intermittent sampling and a combined hierarchical method. *Hum Reprod.* 2006; 21:1442–1452. [PubMed: 16439502]
17. Kassam A, Ovestreet JW, Snow-Harter C, De Souza MJ, Gold EB, Lasley BL. Identification of anovulation and transient luteal function using a urinary pregnanediol-3-glucuronide ratio algorithm. *Environ Health Perspect.* 1996; 104:408–413. [PubMed: 8732951]
18. Baird DD, Weinberg CR, Wilcox AJ, McConaughy DR, Musey PI. Using the ratio of urinary oestrogen and progesterone metabolites to estimate day of ovulation. *Stat Med.* 1991; 10:255–266. [PubMed: 2052803]
19. Johansson ED, Wide L, Gemzell C. Luteinizing hormone (LH) and progesterone in plasma and LH and oestrogens in urine during 42 normal menstrual cycles. *Acta Endocrinol (Copenh).* 1971; 68:502–512. [PubMed: 5171366]
20. Hoff JD, Quigley ME, Yen SS. Hormonal dynamics at midcycle: a reevaluation. *J Clin Endocrinol Metab.* 1983; 57:792–796. [PubMed: 6411753]
21. Brown J, Skurnick J, Sharma N, Adel T, Santoro N. Frequent intermittent ovarian function in women with premature menopause: a longitudinal study. *Endocr J.* 1993; 1:467–474.
22. Testart J, Frydman R, Feinstein MC, Thebault A, Roger M, Scholler R. Interpretation of plasma luteinizing hormone assay for the collection of mature oocytes from women: definition of a luteinizing hormone surge-initiating rise. *Fertil Steril.* 1981; 36:50–54. [PubMed: 7250407]
23. Ecochard R, Leiva R, Bouchard T, Boehringer H, Direito A, Mariani A, et al. Use of urinary pregnanediol 3-glucuronide to confirm ovulation. *Steroids.* 2013; 78:1035–1040. [PubMed: 23831784]
24. Blackwell LF, Vigil P, Gross B, d'Arcanges C, Cooke DG, Brown JB. Monitoring of ovarian activity by measurement of urinary excretion rates of estrone glucuronide and pregnanediol glucuronide using the Ovarian Monitor, Part II: reliability of home testing. *Hum Reprod.* 2012; 27:550–557. [PubMed: 22131389]
25. Blackwell LF, Vigil P, Cooke DG, d'Arcanges C, Brown JB. Monitoring of ovarian activity by daily measurement of urinary excretion rates of oestrone glucuronide and pregnanediol glucuronide using the Ovarian Monitor, Part III: Variability of normal menstrual cycle profiles. *Hum Reprod.* 2013; 28:3306–3315. [PubMed: 24170744]
26. McGovern PG, Myers ER, Silva S, Coutifaris C, Carson SA, Legro RS, et al. Absence of secretory endometrium after false-positive home urine luteinizing hormone testing. *Fertil Steril.* 2004; 82:1273–1277. [PubMed: 15533341]
27. Santoro N, Crawford SL, Allsworth JE, Gold EB, Greendale GA, Korenman S, et al. Assessing menstrual cycles with urinary hormone assays. *Am J Physiol Endocrinol Metab.* 2003; 284:E521–E530. [PubMed: 12441312]
28. Nielsen MS, Barton SD, Hatasaka HH, Stanford JB. Comparison of several one-step home urinary luteinizing hormone detection test kits to Ovunque. *Fertil Steril.* 2001; 76:384–387. [PubMed: 11476792]
29. Wactawski-Wende J, Schisterman EF, Hovey KM, Howards PP, Browne RW, Hediger M, et al. BioCycle study: design of the longitudinal study of the oxidative stress and hormone variation during the menstrual cycle. *Paediatr Perinat Epidemiol.* 2009; 23:171–184. [PubMed: 19159403]
30. Howards PP, Schisterman EF, Wactawski-Wende J, Reschke JE, Frazer AA, Hovey KM. Timing clinic visits to phases of the menstrual cycle by using a fertility monitor: the BioCycle Study. *Am J Epidemiol.* 2009; 169:105–112. [PubMed: 18974081]
31. Mumford SL, Schisterman EF, Gaskins AJ, Pollack AZ, Perkins NJ, Whitcomb BW, et al. Realignment and multiple imputation of longitudinal data: an application to menstrual cycle data. *Paediatr Perinat Epidemiol.* 2011; 25:448–459. [PubMed: 21819426]
32. Abraham GE, Maroulis GB, Marshall JR. Evaluation of ovulation and corpus luteum function using measurements of plasma progesterone. *Obstet Gynecol.* 1974; 44:522–525. [PubMed: 4413313]
33. Hambridge HL, Mumford SL, Mattison DR, Ye A, Pollack AZ, Bloom MS, Mendola P, Lynch KL, Wactawski-Wende J, Schisterman EF. The influence of sporadic anovulation on hormone levels in ovulatory cycles. *Hum Reprod.* 2013; 28:1687–1694. [PubMed: 23589536]

34. Behre HM, Kuhlage J, Gassner C, Sonntag B, Schem C, Schneider HP, et al. Prediction of ovulation by urinary hormone measurements with the home use ClearPlan Fertility Monitor: comparison with transvaginal ultrasound scans and serum hormone measurements. *Hum Reprod.* 2000; 15:2478–2482. [PubMed: 11098014]
35. MacGregor EA, Frith A, Ellis J, Aspinall L. Predicting menstrual migraine with a home-use fertility monitor. *Neurology.* 2005; 64:561–563. [PubMed: 15699399]
36. Brown JB. Timing of ovulation. *Med J Aust.* 1977; 2:780–783. [PubMed: 611378]
37. Feinstein AR, Cicchetti DV. High agreement but low kappa: I. The problems of two paradoxes. *J Clin Epidemiol.* 1990; 43:543–549. [PubMed: 2348207]
38. Direito A, Bailly S, Mariani A, Ecochard R. Relationships between the luteinizing hormone surge and other characteristics of the menstrual cycle in normally ovulating women. *Fertil Steril.* 2013; 99:279–285. [PubMed: 22999798]
39. Schliep KC, Mumford SL, Hammoud AO, Stanford JB, Kissell KA, Sjaarda LA, Perkins NJ, Ahrens KA, Wactaswki Wende J, Mendola P, Schisterman EF. Luteal phase deficiency in regularly menstruating women: prevalence and overlap in identification based on clinical and biochemical diagnostic criteria. *J Clin Endocrinol Metab.* In press.

Figure 1.

Geometric mean hormone concentrations across the menstrual cycle for cycles classified as ovulatory and anovulatory using luteal progesterone threshold algorithm (P5) and the fertility monitor algorithm (LH-FM). Refer to Table 2 for full description of P5 and LH-FM algorithms.

Table 1Selected characteristics^a of BioCycle Study (2005–2007) participants

Participants [<i>n</i> (%)]	Total 259
Age (years) (mean ± SD)	27.3 ± 8.2
Race [<i>n</i> (%)]	
White	154 (59)
Black	51 (20)
Asian	37 (14)
Other	17 (7)
Body Mass Index (mean ± SD)	24.1 ± 3.9
Education, post-secondary [<i>n</i> (%)]	
High School	33 (12.7)
Post-secondary	226 (87.3)
Marital status [<i>n</i> (%)]	
Single/divorced	193 (74.5)
Married/living as married	66 (25.5)
Hormonal contraception ever [<i>n</i> (%)]	140 (55)
Nulligravid [<i>n</i> (%)]	177 (69)
Nulliparous [<i>n</i> (%)]	189 (74)
Sexual activity [<i>n</i> (%)]	
Previous, not current	64 (25)
Sexually active (<1/week)	69 (27)
Sexually active (≥1/week)	63 (25)
Physical activity [<i>n</i> (%)]	
Low	25 (9.7)
Moderate	92 (35.5)
High	142 (54.8)
Current smoker [<i>n</i> (%)]	42 (16.2)
Alcohol consumption in past 12 months [<i>n</i> (%)]	
< 12 drinks	85 (32.8)
12 drinks	172 (66.4)

^c Baseline characteristics of study population were obtained using standard questionnaires.

Description of algorithms for classifying anovulatory menstrual cycles and proportion observed in the BioCycle Study, 2005–2007

Table 2

Algorithm Number	Algorithm Name	Hormones	Linear Interpolation used?	Anovulatory classification criteria	Anovulatory Cycles n (%)
Luteal phase progesterone activity algorithms¹					
1	P-R	Serum P	Yes	P ratio (daily:baseline) does not exceed 3 for 3 days	43 (8.5)
2	P5	Serum P	No	Absolute P levels 5 ng/mL	65 (12.8)
3	P3	Serum P	No	Absolute P levels 3 ng/mL	46 (9.0)
4	Bio-P5-LH	Serum P, LH	No	P 5 ng/mL and no late LH peak is observed	42 (8.3)
5	Bio-P3-LH	Serum P, LH	No	P 3 ng/mL and no late LH peak is observed	28 (5.5)
Luteal day transition algorithm¹					
6	LDT	Serum E2, P	Yes	Absence of a 5 day sequence of E2/P where day 4 and day 5 are < day 1 and where day 1 is the highest in the sequence	32 (6.3)
Mid-cycle LH surge algorithms (using fertility monitor urine measurements)²					
7	LH-FM	Urinary E3G, LH	No	“Peak” reading on monitor not achieved	78 (17.4)
8	LH-S 1	Urinary LH	No	<2.5-fold increase in LH from that of baseline LH levels	70 (15.6)
9	LH-S 2	Urinary LH	No	<4.0 fold increase in LH from any day taken in the follicular phase	47 (11.0)
10	LH-S 3	Urinary LH	No	No LH value 180% of the mean plus 2 SD of the 4 preceding days	15 (3.4)
11	LH-S 4	Urinary LH	No	No LH value exceeds the mean plus 3 SD of the 5 preceding days	83 (18.6)

E3G, estrone-3-glucuronide; E2, estradiol; LDT, luteal day transition; LH, luteinizing hormone; P, progesterone; P-R, relative progesterone; P5, P3, absolute progesterone; Bio-P5-LH, Bio-P3-LH, absolute progesterone and timing of LH peak; 1–4, urinary monitor algorithms; LH-FM, LH fertility monitor; SD, standard deviation.

¹ Based on 509 cycles;

² Based on 445 cycles with complete visit and Clear Plan Fertility Monitor data

Table 3

Cross Classification of serum and urine algorithms[/] and corresponding kappa statistics among BioCycle Study participants (n=509 cycles for serum; n=445 cycles for urine)

κ (95% CI) (% agreement)	P-R	P-5	P-3	Bio-P5-LH	Bio-P3-LH	LDT
P-R	1					
P-5	0.75 (0.66, 0.85) (96.3)	1				
P-3	0.84 (0.75, 0.93) (97.4)	0.81 (0.73, 0.89) (96.3)	1			
Bio-P5-LH	0.53 (0.39, 0.66) (92.7)	0.76 (0.67, 0.85) (95.4)	0.60 (0.48, 0.73) (93.7)	1		
Bio-P3-LH	0.53 (0.39, 0.68) (93.9)	0.57 (0.45, 0.69) (92.7)	0.74 (0.62, 0.85) (96.4)	0.79 (0.68, 0.89) (97.2)	1	
LDT	0.73 (0.61, 0.84) (96.3)	0.58 (0.47, 0.70) (92.7)	0.67 (0.54, 0.79) (95.3)	0.39 (0.24, 0.54) (91.7)	0.40 (0.23, 0.56) (93.3)	1

κ (95% CI) (% agreement)	LH-FM	LH-S1	LH-S2	LH-S3	LH-S4
LH-FM	1				
LH-S1	0.34 (0.23, 0.45) (81.8)	1			
LH-S2	-0.11 (-0.17, -0.06) (73.0)	0.03 (-0.07, 0.13) (77.8)	1		
LH-S3	0.08 (-0.01, 0.17) (82.0)	-0.01 (-0.07, 0.05) (81.8)	-0.05 (-0.08, -0.03) (86.0)	1	
LH-S4	0.34 (0.24, 0.45) (80.7)	0.49 (0.38, 0.59) (85.4)	0.00 (-0.08, 0.09) (74.8)	-0.06 (-0.09, -0.03) (78.0)	1

^K Kappa statistic; CI, confidence interval

[/] Refer to Table 2 for full description of algorithms.