



Practice of Epidemiology

Timing Clinic Visits to Phases of the Menstrual Cycle by Using a Fertility Monitor: The BioCycle Study

Penelope P. Howards, Enrique F. Schisterman, Jean Wactawski-Wende, Jennifer E. Reschke, Andrea A. Frazer, and Kathleen M. Hovey

Initially submitted April 1, 2008; accepted for publication August 15, 2008.

Planning study visits during specific menstrual cycle phases is important if the exposure or outcome is influenced by hormonal variation. However, hormone profiles differ across cycles and across women. The value of using fertility monitors to time clinic visits was evaluated in the BioCycle Study (2005–2007). Women aged 18–44 years (mean, 27.4) with self-reported menstrual cycle lengths of 21–35 days were recruited in Buffalo, New York, for 2 cycles ($n = 250$). Participants were provided with home fertility monitors that measured urinary estrone-3-glucuronide and luteinizing hormone (LH). The women were instructed to visit the clinic for a blood draw when the monitor indicated an LH surge. The monitor recorded a surge during 76% of the first cycles and 78% of the second cycles. Scheduling visits by using set cycle days or algorithms based on cycle length, such as a midcycle window or a window determined by assuming a fixed luteal phase length, would be simpler. However, even with perfect attendance in a 3-day window, these methods would have performed poorly, capturing the monitor-detected LH surge only 37%–57% of the time. Fertility monitors appear to be useful in timing clinic visits in a compliant population with flexible schedules.

hormones; menstrual cycle; ovulation detection; ovulation prediction; research design

Abbreviations: E3G, estrone-3-glucuronide; LH, luteinizing hormone; SD, standard deviation.

Hormone levels in premenopausal women may be of interest as exposures, outcomes, or confounders in many areas of research related to women's health (e.g., fertility, cancer, cardiovascular disease). However, measuring hormone levels is complicated by the fact that they change across the menstrual cycle (Figure 1). In addition, hormone levels and the timing of key phases of the menstrual cycle vary across both women and a given woman's cycles (1). Therefore, it is difficult to anticipate when to measure hormone levels in order to capture key events such as the estrogen rise during the follicular phase, the luteinizing hormone (LH) surge, or the increase in progesterone during the luteal phase.

Researchers have used a variety of approaches in an attempt to measure hormone levels in a meaningful way given the inherent variability of these markers. Nevertheless, many studies rely on a single serum sample drawn on an arbitrary day of each woman's menstrual cycle, which

seems inadequate, even with adjustment for cycle day or for cycle phase status (determined by the sample's progesterone level) (2–4). Other studies have attempted to time specimen collection to key hormonal events based on self-reported cycle length or cycle day or the assumption that the luteal phase is 14 days (5–7), but these methods assume that women can reliably report their cycle length or that the timing of hormonal fluctuations is the same across women, both of which are unlikely to be true (1, 8–11). The “gold standard” is daily collection of first-morning urine specimens, which ensures that critical hormone windows are captured if compliance is adequate, but the protocol is burdensome, which affects recruitment and compliance. Ideally, it would be possible to characterize hormone levels across the menstrual cycle by collecting fewer, well-timed measurements, but it is unclear what algorithm should be applied to prospectively assess when the samples should be collected.

Correspondence to Dr. Penelope P. Howards, Epidemiology Department, Emory University, 1518 Clifton Road, NE, Room 450, Atlanta, GA 30322 (e-mail: penelope.howards@emory.edu).

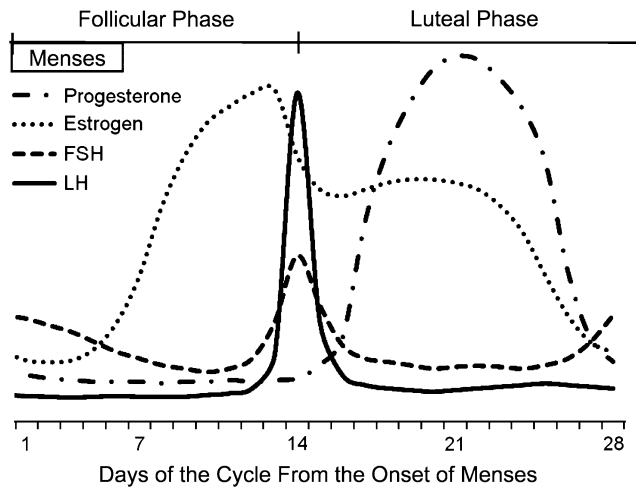


Figure 1. Estrogen, luteinizing hormone (LH), follicle-stimulating hormone (FSH), and progesterone levels during an idealized 28-day menstrual cycle with the LH surge around day 14.

In this paper, we examine the utility of providing study participants with a home fertility monitor that can detect an LH surge and compare the timing of the LH surge detected by the monitor with the timing assumed by algorithms based on cycle length.

MATERIALS AND METHODS

The BioCycle Study was carried out at the University at Buffalo under an Intramural Research Program contract from the Eunice Kennedy Shriver National Institute of Child Health and Human Development. The study is described in greater detail elsewhere (12). In brief, 259 healthy women were recruited from the community and enrolled for 2 menstrual cycles. Nine women contributed only 1 cycle. Women had to be aged 18–44 years, to have a self-reported cycle length from 21 to 35 days, and to have no known conditions that might affect their menstrual cycle function (e.g., being underweight, current use of hormonal contraception, current breastfeeding). Table 1 contains the complete list of exclusion criteria (12).

Eligible women who consented to participate in the study were asked to come to the clinic 8 times per cycle with 3 visits planned around the time of the expected LH surge. At each cycle visit, the women provided fasting blood and spot urine specimens. Estradiol, follicle-stimulating hormone, LH, and progesterone were measured in all available serum samples by using IMMULITE 2000 chemiluminescent enzymatic immunoassays (Kaleida Health, Buffalo, New York). Participants were also provided with and trained to use the Clearblue Easy Fertility Monitor (Inverness Medical Innovations, Inc., Waltham, Massachusetts) (www.clearblueeasy.com). In addition, they were asked to complete a series of questionnaires and brief daily diaries at home.

The fertility monitor was originally developed to assist women in becoming pregnant by helping them to identify

their fertile window through measurement of both estrone-3-glucuronide (E3G) and LH in urine as described in greater detail elsewhere (13). Briefly, the monitor is synchronized to a woman's cycle and then is checked daily to see whether a test is requested. 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 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 (13).

Each day, the monitor assigns the woman to low, high, or peak fertility on the basis of her E3G and LH levels. Thus, unlike home LH test sticks, the monitor provides information to help anticipate the LH surge. High fertility is determined by the level of E3G and correlates with the rise in estrogen during the follicular phase (Figure 1). Peak fertility is assessed by the level of LH and correlates with the LH surge prior to ovulation. The monitor first requests 10 consecutive days of testing for all women. If the woman does not reach peak in those first 10 days (whether from poor compliance with testing or low levels of LH), the monitor requests an additional 10 tests for a total of 20 days. The monitor initially determines high and peak fertility levels on the basis of predetermined cutpoints (peak corresponds with approximately 30 IU/L for LH), but if the woman does not reach those cutpoints, the monitor adjusts the cutpoint criteria according to the woman's specific hormone levels.

Although monitor users see only their fertility status and the day of their cycle in the display screen, more detailed data are stored by the monitor and can be downloaded, which is another advantage over home LH test sticks. The data include the following: the date and time the monitor was turned on, the day of the cycle, and whether a test was performed. For days when a test was completed, the fertility level (low, high, peak) and the actual levels of E3G and LH (in percentage transmission units) are also available.

In this paper, we examine summary data from the monitor to see whether the monitor was useful in identifying the LH surge and in helping to time the clinic visits. In addition, we compare 3 alternative methods of prospectively anticipating the timing of the LH surge to see whether they actually capture the surge according to the monitor. The first alternative (fixed cycle days method) schedules all women for a blood draw on cycle days 13–15, which should capture the LH surge in an idealized 28-day cycle (Figure 1). The second method (the luteal-phase method) is based on the assumption that variability in the cycle length is predominately due to variability in the follicular phase and that the luteal phase is more stable across women, averaging approximately 14 days. For this method, a 3-day window around the estimated LH surge is created by subtracting 15 days and 13 days from the woman's self-reported usual cycle length. For example, for a 30-day cycle, this would correspond to cycle days 15–17. The final method (the midpoint method) assumes that the LH surge occurs around the midpoint of the cycle, and therefore, the 3-day window is created around the midpoint of the self-reported cycle length. For a 30-day

Table 1. Exclusion Criteria of the BioCycle Study, Buffalo, New York, 2005–2007

Oral contraceptive or other hormone supplement use in the past 3 months
Depo-Provera, ^a Norplant, ^a or IUD use in the past 12 months
Planning to attempt to conceive in the next 3 months
Actively trying to conceive in the past 6 months
Pregnant in the last 6 months or currently trying to conceive
Breastfeeding in the last 6 months
Abnormal Papanicolaou smear in the past 6 months with no subsequent normal results
Laparoscopy-confirmed endometriosis
Current uterine fibroids or removal of a fibroid in the last 12 months
History of polycystic ovary disease
History of <i>Chlamydia</i> infection or positive IgG screen at screening
Untreated gynecologic infection or any genitourinary infection in the past 6 months
Gynecologic surgery in the past year
Treatment for infertility, ever (excluding male factor issues)
History or clinical signs of gynecologic problems
Infectious disease treated by a physician in the past 6 months
Treatment for allergies with chronic medication (at least once a week in the past 3 months)
Liver or kidney disease requiring treatment in the past year
Psychiatric condition requiring medical therapy in the past year
Body mass index of less than 18 or greater than 35 kg/m ²
Plan to restrict diet for weight loss or medical reasons in next 3 months
Gastrointestinal conditions associated with malabsorption (e.g., Crohn's disease)
Unwilling to stop regular intake of vitamin, mineral, or herbal antioxidant supplements during study
Chronic use of certain medications including lipid-lowering drugs, antihypertensive medications, aspirin, and others
Antibiotic use in the past 3 months
History of chronic diseases such as heart disease, diabetes mellitus, cancer, inflammatory diseases, autoimmune disease, liver disease, kidney disease, thyroid diseases, or any endocrine dysfunction
Current treatment for anemia or anemia associated with pregnancy in the past 12 months
History of alcohol abuse or dependency disorder or other substance abuse in past 30 days
Self-reported current regular illicit drug use in the past 30 days
Diet high in phytoestrogen content (e.g., soy-based diet)

Abbreviations: IgG, immunoglobulin G; IUD, intrauterine device.

^a Depo-Provera (Pfizer Incorporated, New York, NY); Norplant (The Population Council, New York, NY).

cycle, this would correspond to days 14–16. We also examined the latter 2 methods using the women's actual cycle lengths even though they would not be known prospectively and therefore could not be used to schedule visits in an actual study.

RESULTS

Participants in the BioCycle Study tended to be young, non-Hispanic white, highly educated, not living as married, and to never have been pregnant (Table 2). A range of income levels was represented in this study. Of the 259 participants who contributed at least 1 study cycle, 249 (96%) brought their monitors in to have the data downloaded at least once during the first cycle, and 232 (93%) of the 250 women who participated for 2 cycles had data downloaded

at least once for the second cycle. We considered adherence to testing for the first 10 tests that are requested of all women (Table 3). Of the women who brought in their monitors at least once in a cycle, 82% missed 2 tests or fewer in cycle 1, and 75% missed 2 tests or fewer in cycle 2. Approximately 85% of the women reached high fertility, and over three-quarters reached peak fertility during each cycle. At least three-quarters of the women who reached peak came into the clinic for a blood test on the day the monitor read peak.

Among the women who did not reach peak fertility, 72% in cycle 1 and 64% in cycle 2 missed 2 tests or fewer out of the first 10 requested tests. An increase in progesterone during the luteal phase can be considered a marker for ovulation (Figure 1), so we examined the maximum detected level of serum progesterone. Among the women who did not reach peak on the monitor, 65% in cycle 1 and 85% in cycle

Table 2. Summary Demographic Data on the 249 Participants With Any Fertility Monitor Data in the BioCycle Study, Buffalo, New York, 2005–2007

Variable	No.	%
Study cycles completed		
1	8	3.2
2	241 ^a	96.8
Age, years		
18–24	127	51.0
25–29	39	15.7
30–34	19	7.6
35–39	32	12.9
40–44	32	12.9
Race/ethnicity		
White, non-Hispanic	148	59.4
Black, non-Hispanic	46	18.5
Other	55	22.1
Education		
High school degree or less	30	12.1
Some college or associate's degree	117	47.0
College degree or higher	102	41.0
Student		
No	87	35.1
Full time	142	57.3
Part time	19	7.7
Missing	1	
Marital status		
Married or living as married	64	25.7
Neither married nor living as married	185	74.3
Income		
≤\$19,999	51	20.7
\$20,000–\$39,999	60	24.3
\$40,000–\$74,999	69	27.9
\$75,000–\$99,999	43	17.4
≥\$100,000	24	9.7
Missing	2	
Gravidity		
0	168	69.1
1	16	6.6
≥2	59	24.3
Missing	6	

^a A total of 241 women with any monitor data participated in the study for 2 cycles; they did not necessarily have monitor data for both cycles.

2 had at least 1 serum progesterone measurement that was greater than or equal to 5 ng/mL, which suggests that they did ovulate (Table 3).

In order to further evaluate why women did not reach peak on the monitor, we examined the serum hormone data in conjunction with the detailed monitor hormone data for

each woman who did not reach peak. Classification of the reasons for not reaching peak was subjective but is intended to help characterize the most likely challenges to using the monitor in a study. The most common reasons for not reaching peak varied by cycle (Table 3). Failure to comply with the monitor protocol, whether due to missed tests or failing to bring the monitor in to have the data downloaded, was a primary reason for not detecting peak in both cycles. Peak was also not observed when there was a monitor error in reading the test stick, which can occur if the stick is too wet or too dry. In addition, the monitor does not recognize LH surges prior to day 9 as peak and requests no more than 20 tests regardless of cycle length. Some women did not reach peak because they had an early or late surge in LH (not necessarily followed by an increase in progesterone) or a long cycle. Approximately one-fifth of the women who did not reach peak had evidence of a small LH surge followed by elevated progesterone, but many of their cycles looked atypical (e.g., estrogen remained flat throughout the cycle). Finally, some women truly appeared not to have evidence of an LH surge.

For both cycles, the mean day of peak was cycle day 15 (standard deviation (SD), 3.3) among women whose monitor reached peak (Table 4). As expected, the mean day of peak was earlier for women with shorter cycles and later for women with longer cycles. In addition, the mean day of peak decreased with increasing age among women who reached peak. The mean day of peak occurred 13 days (SD, 2.9) before the end of the cycle on average for both cycles. Peak was closer to the end of the cycle for shorter cycles but did not change across age groups. The mean number of days between reaching high fertility (E3G surge) and peak fertility was 5 days (SD, 3.2 for cycle 1 and 2.9 for cycle 2) among women whose monitor reading captured both. Of the women who reached peak, 27 (14%) in cycle 1 and 31 (17%) in cycle 2 went to peak without reaching high first.

Among women who had a monitor peak, the peak day fell between days 13 and 15 (fixed-cycle-day method), 41% of the time in cycle 1 and 37% in cycle 2 (Table 5). The luteal-phase and midpoint methods performed similarly when self-reported cycle length was used to determine the windows, but they improved slightly when the actual cycle length was used with the luteal-phase method performing the best.

Cycle length self-reported at baseline differed by 3 or more days compared with observed study cycle length over 40% of the time (Table 6). In addition, the actual study cycle length for the first cycle differed from that of the second cycle by 3 days or more for almost half the participants who participated for 2 cycles. When peak fertility was observed in both cycles, the cycle day of occurrence differed from one cycle to the next by 3 or more days 37% of the time. However, only one-quarter of the women had the day of peak fertility differ by 3 or more days when the timing of the peak day was assessed from the actual end of each cycle.

DISCUSSION

For many studies, the timing of biospecimen collection within a woman's menstrual cycle is important because the exposure or outcome is affected by hormone levels that

Table 3. Summary Data on Adherence to Use of the Fertility Monitor and Fertility Level Reached Among the 249 Women Who Had Any Fertility Monitor Data in the BioCycle Study, Buffalo, New York, 2005–2007

	Cycle 1		Cycle 2	
	No.	%	No.	%
Incomplete monitor data ^a	14	5.6	37	16.0
No. of missed tests out of the first 10 ^b				
0	120	48.2	98	42.2
1–2	83	33.3	77	33.2
3–4	23	9.2	27	11.6
5–10	22	8.8	18	7.8
Unknown ^c	1	0.4	12	5.2
Estrone-3-glucuronide level ^d				
High	215	86.4	195	84.1
No high	30	12.1	33	14.2
Unknown ^e	4	1.6	4	1.7
Luteinizing hormone level				
Peak	189	75.9	180	77.6
No peak	49	19.7	29	12.5
Unknown ^e	11	4.4	23	9.9
Clinic visit on peak day among women who reached peak				
Yes	142	75.1	148	82.2
No	47	24.9	32	17.8
No. of missed tests out of first 10 ^b among women with no peak ^f				
0	24	40.0	15	35.7
1–2	19	31.7	12	28.6
3–4	5	8.3	5	11.9
5–10	12	20.0	10	23.8
Unknown ^c	0		10	

Table continues

change across the menstrual cycle. However, it is difficult to prospectively anticipate the timing of critical phases of the cycle without daily hormonal measurements. In the BioCycle Study, the fertility monitor was useful in scheduling clinic visits around the time of the LH surge.

Participants in the BioCycle Study did not provide daily serum or urine specimens at the clinic, so it is not possible to validate the LH peak as detected by the monitor in this study. Nevertheless, we felt confident that, with proper usage, the fertility monitor would be able to detect the LH surge on the basis of the findings of Behre et al. (13). They compared fertility monitor data with serum hormone levels and transvaginal ultrasound scans in a study of 53 women. In that study, the monitor recorded peak fertility in 135 of 149 ovulatory cycles as determined by ultrasound (13). Of the 149 ovulatory cycles, a serum LH surge was detected in only 139 cycles, which means that the monitor detected an LH surge when there was a detectable surge 97% of the time (13). In addition, the monitor did not reach peak in the 1 anovulatory cycle (13).

Table 3. Continued

	Cycle 1		Cycle 2	
	No.	%	No.	%
Maximum serum progesterone below 5.0 ng/mL among women with no peak ^f				
<5.0 ng/mL	21	35.0	8	15.4
≥5.0 ng/mL	39	65.0	44	84.6
Most likely reason for not reaching peak among women with no peak ^f				
Missed test	10	16.7	12	23.1
Monitor data available only before peak	5	8.3	12	23.1
Test stick reading error	7	11.7	10	19.2
Long cycle	8	13.3	1	1.9
Normal cycle length but late luteinizing hormone peak	2	3.3	1	1.9
Early luteinizing hormone surge	4	6.7	1	1.9
Possible low luteinizing hormone surge, serum progesterone ≥5.0 ng/mL	13	21.7	11	21.2
Unlikely ever peaked	11	18.3	4	7.7

^a Data are incomplete if the woman did not bring her monitor in to have the data downloaded at the end of the testing period. Some of these women brought in their monitors after 10 tests were requested and reached high or peak fertility even though their monitor data were incomplete.

^b Ten tests were requested of all women (except one woman in cycle 1 whose next period started before 10 tests could be requested).

^c “Unknown” includes women whose monitor data are incomplete, and therefore it is unknown how many of the first 10 tests were missed.

^d Only includes women who achieved high fertility prior to the peak (the monitor automatically assigns a value of high fertility the day after the last peak day regardless of the estrone-3-glucuronide level; these women were not counted as high in this analysis).

^e “Unknown” includes women who did not reach high or peak fertility in the recorded data but who had incomplete monitor data downloaded. It does not include women with complete recorded monitor data who missed tests.

^f Includes 11 women for cycle 1 and 13 women for cycle 2 who had complete data for the first 10 tests and no observed peak but who had incomplete data for the final 10 requested tests.

Although the work of Behre et al. (13) suggests that the monitor can reliably detect the LH surge, the monitor is helpful only as a research tool if the women use it properly. In the BioCycle Study, the participants had high but not perfect adherence to the monitor-testing protocol. Because the LH surge occurs during a very small window, even missing a few tests could be problematic, depending on when in the cycle the tests were missed. Nevertheless, peak was observed for over 75% of the cycles, and high fertility, which is not confined to a single day and therefore is less dependent on consistent testing, was observed in approximately 85% of the cycles.

There are several factors that contribute to not observing peak fertility. The logistic factors include incomplete data

Table 4. Mean Day of Peak Fertility and Mean Day From the End of the Cycle for Peak Fertility Among Women Who Reached Peak, by Observed Cycle Length and Age, in the BioCycle Study, Buffalo, New York, 2005–2007

	Cycle 1		Cycle 2	
	No.	Mean (SD)	No.	Mean (SD)
Cycle day for all	189	15 (3.3)	180	15 (3.3)
Cycle day by cycle length, days				
<26	27	13 (2.3)	29	13 (3.0)
26–31	138	15 (2.8)	103	15 (2.5)
>31	24	19 (3.9)	33	18 (3.7)
Unknown ^a	0		15	15 (3.2)
Cycle day by age, years				
18–24	92	16 (3.7)	82	16 (3.5)
25–34	44	15 (2.5)	50	15 (3.3)
35–44	53	14 (2.7)	48	14 (2.4)
Days from cycle end for all	189	13 (2.9)	165 ^a	13 (2.9)
Days from cycle end by cycle length, days				
<26	27	12 (2.4)	29	11 (2.6)
26–31	138	13 (2.5)	103	14 (2.3)
>31	24	16 (4.0)	33	16 (3.2)
Unknown ^a	0		15	
Days from cycle end by age, years				
18–24	92	13 (3.6)	73	14 (3.5)
25–34	44	14 (2.1)	47	13 (2.3)
35–44	53	13 (1.9)	45	13 (2.4)
Unknown ^a	0		15	
Days between high and peak ^b	162	5 (3.2)	149	5 (2.9)

Abbreviation: SD, standard deviation.

^a Cycle length is unknown for women who did not call in and report the end of their second cycle.

^b Includes only women who reached high on the monitor before they reached peak.

because the woman did not bring her monitor in to have the data downloaded, poor compliance with testing, good compliance with testing combined with an unfortunately timed missed test, and test-stick reading errors. In some cases though, the lack of observed peak fertility may be due to anovulatory cycles or atypical hormone profiles that may or may not have resulted in ovulation. Our results are similar to those of Robinson et al. (14), who observed peak in approximately 80% of their study cycles, which also reflects a combination of underlying cycle characteristics and compliance.

Timing blood draws at the clinic around the LH surge depended on both the monitor results and the ability of the participants to come to the clinic at the appropriate time. The study protocol instructed participants whose monitors read peak to go to the clinic that morning or to call to schedule a visit as soon as possible. Over three-quarters of the women who reached peak fertility on the monitor were able to attend the clinic for a blood draw on that day. In general, the BioCycle Study participants were an adherent group with schedules that allowed them to come to the clinic on short notice. The need for flexibility in attending the clinic was emphasized at enrollment to increase the likelihood of compliant participants. Once enrolled, participants were given individualized calendars with a projected clinic visit schedule that was revised as needed. These calendars were designed to help the women to anticipate their visit schedule as much as possible. To further encourage adherence to the visit schedule, study staff were sensitive to the time commitment of the participants and followed procedures to minimize the duration of each visit. This flexibility might not be possible in all study settings.

The standardized approaches to scheduling clinic visits (fixed-cycle days, luteal-phase method, and midpoint method) would be easier to implement logistically but would also be less likely to capture the day of peak fertility. The fact that women with longer cycles reached peak fertility later than did women with shorter cycles on average supports the concept behind the standardized methods that cycle length is correlated with the timing of ovulation. In addition, the mean cycle day of peak fertility changed across

Table 5. Percentage of Cycles That Reached Peak Fertility on a Cycle Day Within Standardized Windows Based on Cycle Length Out of All Cycles That Reached Peak Fertility According to the Fertility Monitor Used in the BioCycle Study, Buffalo, New York, 2005–2007

	Cycle 1 (n = 189)		Cycle 2 (n = 180)	
	No.	%	No.	%
Fixed cycle days method (days 13–15)	77	40.7	66	36.7
Luteal-phase method (15–13 days from end of cycle)				
Self-reported cycle length	73	38.6	66	36.7
Actual cycle length ^a	108	57.1	83	50.3
Midpoint method (midpoint of cycle \pm 1 day)				
Self-reported cycle length	76	40.2	67	37.2
Actual cycle length ^a	94	49.7	73	44.2

^a Cycle 2 calculations were based on the 165 women whose actual cycle length was known.

Table 6. Comparison of Self-reported Cycle Length With Actual Cycle Length, of Actual Cycle Length in Cycles 1 and 2, and of the Timing of the Monitor Peak in Cycle 1 Versus Cycle 2 Among Participants in the BioCycle Study, Buffalo, New York, 2005–2007

No. of Days' Difference	Self-reported Cycle Length vs. Actual Cycle Length				Actual Cycle Length (Cycle 1 vs. Cycle 2)		Day of Monitor Peak (Cycle 1 vs. Cycle 2) ^a		Days From End of Cycle of Monitor Peak (Cycle 1 vs. Cycle 2) ^a	
	Cycle 1		Cycle 2		No.	%	No.	%	No.	%
	No.	%	No.	%						
0	33	13.8	36	17.2	21	10.1	30	20.3	37	27.0
1	49	20.4	52	24.9	55	26.3	43	29.1	45	32.9
2	54	22.5	38	18.2	32	15.3	21	14.2	20	14.6
≥3	104	43.3	83	39.7	101	48.3	54	36.5	35	25.6
Unknown ^b	9		23		23				11	

^a Among 148 women who reached peak fertility in cycles 1 and 2.

^b Among women with any monitor data, the actual cycle length was unknown for 9 women for cycle 1 and 23 women for cycle 2; 11 women who peaked in both cycles were missing cycle length data.

age categories, but the mean day from the end of the cycle of peak fertility did not, which provides some support specifically for the luteal-phase method. However, the day of peak fertility was farther from the end of the cycle for longer cycles and closer to the end for shorter cycles on average. Despite the compatibility of the mean peak fertility day with general expectations, use of any of the standardized methods would have missed the peak fertility day for the majority of women, even assuming perfect attendance at the clinic during the predetermined 3-day window. The luteal-phase and midpoint methods are based on cycle length, but only self-reported cycle length can be known prospectively, and it performs poorly for both algorithms. Even if the actual cycle length could be used in a prospective study, it was only marginally better than self-reported cycle length. The luteal-phase method slightly improved on the midpoint method, but only if the actual cycle length was used.

Several factors may have contributed to the poor performance of the standardized methods. Natural biologic variability in individual women is complicated by the fact that the event (LH surge) occurs in a small window (approximately 1 day), so that even modest errors in the timing of observation can miss the event completely. In addition, both we and others (9, 11) have found that women's self-reported cycle lengths, which are used to schedule visits in the standardized methods, do not correlate well with their actual cycle lengths. This may be due in part to the fact that women's cycle lengths vary from one cycle to the next, so that even if they correctly characterize length on average, it may be inaccurate for the study cycle. Thus, even prospectively observing a woman's cycle length prior to any clinic visits in order to appropriately schedule clinic visits during a subsequent cycle does not necessarily improve visit timing due to within-woman variability.

Using a fertility monitor is less burdensome and expensive than collecting, storing, and analyzing daily urine or blood samples. However, the fertility monitor is not suitable for use in populations with extremely short (<20 days) or long (>42 days) cycles. In the BioCycle Study, the women were fairly homogeneous and highly educated, and they were selected to have "normal" menstrual cycles. Their

adherence to the fertility monitor protocol was good despite the fact they were not trying to become pregnant, one of the traditional motivating factors in reproductive studies. Nevertheless, achieving this level of protocol adherence in other study populations might be difficult. Poor compliance would undermine the utility of the monitor, but the monitor does not need to be used in isolation. In the BioCycle Study, women who did not reach peak fertility, whether because of logistic issues related to testing or because they were anovulatory, still had clinic visits that were timed on the basis of their individualized calendars. The visits for these women may not have been timed correctly, but presumably, the timing was no worse than if we had not used a monitor at all.

Although the fertility monitor does not provide as much information as collecting daily biospecimens, it does seem to be useful for timing clinic visits, particularly compared with standardized methods of scheduling based on cycle length. In fact, the monitor gives instant results, whereas daily biospecimens are generally analyzed at the end of a study, so the monitor is advantageous if other timed measurements (e.g., test of muscle strength) are desired. In addition to identifying the LH surge, the fertility monitor could be used in studies where the initial estrogen surge in the follicular phase (corresponding to the first day of high fertility on the monitor) or a specific day of the luteal phase (corresponding to a set number of days after the monitor peak) is of interest. The monitor would be most useful for researchers who have limited funds but need to time clinic visits to particular phases of the menstrual cycle, especially if the study population includes women with flexible schedules and "regular" menstrual cycles.

ACKNOWLEDGMENTS

Author affiliations: Division of Epidemiology, Statistics, and Prevention Research, the Eunice Kennedy Shriver National Institute of Child Health and Human Development, Bethesda, Maryland (Penelope P. Howards, Enrique F.

Schisterman); Department of Social and Preventive Medicine, School of Public Health and Health Professions, University at Buffalo, Buffalo, New York (Jean Wactawski-Wende, Jennifer E. Reschke, Andrea A. Frazer, Kathleen M. Hovey); and Department of Gynecology-Obstetrics, School of Medicine and Biomedical Sciences, University at Buffalo, Buffalo, New York (Jean Wactawski-Wende).

This work was supported by the Intramural Research Program of the Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health.

The authors would like to thank Scott Littlejohn of Inverness Medical Innovations, Inc., for his assistance in improving their understanding of all aspects of the fertility monitor and Dr. James Kesner of the National Institute for Occupational Safety and Health for generously creating an idealized menstrual cycle figure.

Conflict of interest: none declared.

REFERENCES

1. Harlow SD, Ephross SA. Epidemiology of menstruation and its relevance to women's health. *Epidemiol Rev.* 1995;17(2): 265–286.
2. Rosenberg CR, Pasternack BS, Shore RE, et al. Premenopausal estradiol levels and the risk of breast cancer: a new method of controlling for day of the menstrual cycle. *Am J Epidemiol.* 1994;140(6):518–525.
3. Yu H, Shu XO, Shi R, et al. Plasma sex steroid hormones and breast cancer risk in Chinese women. *Int J Cancer.* 2003; 105(1):92–97.
4. Potischman N, Hoover RN, Brinton LA, et al. Case-control study of endogenous steroid hormones and endometrial cancer. *J Natl Cancer Inst.* 1996;88(16):1127–1135.
5. Ahmad N, Pollard TM, Unwin N. The optimal timing of blood collection during the menstrual cycle for the assessment of endogenous sex hormones: can interindividual differences in levels over the whole cycle be assessed on a single day? *Cancer Epidemiol Biomarkers Prev.* 2002;11(1): 147–151.
6. Tworoger SS, Sluss P, Hankinson SE. Association between plasma prolactin concentrations and risk of breast cancer among predominately premenopausal women. *Cancer Res.* 2006;66(4):2476–2482.
7. Sarwar R, Niclos BB, Rutherford OM. Changes in muscle strength, relaxation rate and fatiguability during the human menstrual cycle. *J Physiol.* 1996;493(pt 1):267–272.
8. Fehring RJ, Schneider M, Raviele K. Variability in the phases of the menstrual cycle. *J Obstet Gynecol Neonatal Nurs.* 2006; 35(3):376–384.
9. Small CM, Manatunga AK, Marcus M. Validity of self-reported menstrual cycle length. *Ann Epidemiol.* 2007;17(3): 163–170.
10. Windham GC, Elkin E, Fenster L, et al. Ovarian hormones in premenopausal women: variation by demographic, reproductive and menstrual cycle characteristics. *Epidemiology.* 2002; 13(6):675–684.
11. Jukic AM, Weinberg CR, Wilcox AJ, et al. Accuracy of reporting of menstrual cycle length. *Am J Epidemiol.* 2008;167(1):25–33.
12. Wactawski-Wende J, Schisterman EF, Hovey KM, et al. Bio-Cycle Study: design of the longitudinal study of the oxidative stress and hormone variation during the menstrual cycle. *Paediatr Perinat Epidemiol.* In press.
13. Behre HM, Kuhlage J, Gassner C, et al. Prediction of ovulation by urinary hormone measurements with the home use Clear-Plan Fertility Monitor: comparison with transvaginal ultrasound scans and serum hormone measurements. *Hum Reprod.* 2000;15(12):2478–2482.
14. Robinson JE, Wakelin M, Ellis JE. Increased pregnancy rate with use of the Clearblue Easy Fertility Monitor. *Fertil Steril.* 2007;87(2):329–334.