Resolving Speculations of Methodological Inadequacies in the Standardized Protocol for
Characterizing Women's Fertility: Comment on Lobmaier and Bachofner (2018)

Khandis R Blake^a

^a Evolution & Ecology Research Centre, The University of New South Wales, Sydney
Australia.

.

Author Note

Correspondence to: Khandis R. Blake, Evolution & Ecology Research Centre, The University
of New South Wales, Kensington Campus, Sydney Australia, 2052.

Abstract

Lobmaier and Bachofner (2018) suggest a series of methodological practices to increase the accuracy and reliability of determining a woman's fertile window, claiming the standardized protocol for characterizing women's fertility by Blake, Dixon, O'Dean, and Denson (2016) is inadequate. These practices include observing participants for purportedly fertile sessions a considerable time before the LH surge, and using salivary ferning and cervical mucus evaluation as real-time measures of current fertility. Here I explain that Lobmaier and Bachofner's (2018) recommendations decrease rather than increase the likelihood of observing women during peak fertility. I also summarize the pertinent literature on salivary ferning and cervical mucus evaluations, showing that neither method has sufficient sensitivity and specificity to characterize peak fertility. Using meta-analytic data of 10K menstrual cycles, I then show that the protocol provided by Blake et al. (2016) recruits women when conception probability is at its peak and is statistically higher than the window recommended by Lobmaier and Bachofner (2018).

I appreciate Lobmaier and Bachofner (2018) drawing attention to the importance of precise and careful methodology when recruiting women in the fertile phase of the menstrual cycle. Inaccuracies in their critical comments and suggestions, however, suggest a closer reading of this literature is necessary before warranting alterations to the standardized protocol for characterizing women's fertility (Blake, Dixson, O'Dean, & Denson, 2016).

Lobmaier and Bachofner (2018) state that Blake et al. (2016) does not explicitly recommend when women should attend a testing session after reporting a positive LH test, but this claim is incorrect. In the Recommendations for Researchers in Appendix A, Blake et al. stated that researchers should observe dependent variables within a maximum of 36–48 hours after an LH surge to ensure current fertility. We further noted that observations recorded on the same day as the LH surge should maximize effects, as conception probability is highest on this day. These two recommendations were the outcome of a meta-analysis on variation in the timing from LH surge to ovulation across 26 studies, which showed that most women ovulate within 24<48 hours of a detectable surge in LH (details in Blake et al., 2016). The standardized protocol by Blake et al. (2016) thus provides a two-day window in which researchers can recruit fertile participants. Depending on the time of day the LH surge is detected, some participants can also be observed early on the subsequent day (i.e., 'Day 0' the day of ovulation) and remain within the 48-hour window.

Lobmaier and Bachofner (2018) claim that conception probability (CP) drops rapidly after the LH surge, and as such, researchers wishing to recruit fertile women should observe fertile session variables a "considerable time before the LH surge", not after it. As explained below, their claim about CP rapidly dropping after the LH surge is incorrect. In addition, although "a considerable time before the LH surge" is never defined and here I assume they refer to those days at the start of the fertile window (Days -5 and -4, and perhaps the day immediately preceding the LH surge: Day -3), observations on these days do not maximize

observing women when CP is high. A meta-analysis of >10K menstrual cycles across six studies documented the probability of conception on days relative to ovulation (see Figure 1; Blake et al., 2016). Although CP drops rapidly after ovulation (Days $1 \leq$), CP on the LH surge day (Day -2, $P = .21$) is equivalent with CP on the subsequent two days (Day -1 and Day 0 $P = .19$; $t(5) < 1.00$, $p = .419$; just Day 0: $P = .16$, $t(5) = 1.98$, $p = .105$) and statistically higher than CP on the preceding three days (Day -5, -4, and -3, $P = .11$; $t(5) = 7.85$, $p = .001$). CP on the three days' prior to the LH surge are also statistically lower than CP on Days -1 and Day 0, $t(5) = -6.07$, $p = .002$ (and just Day 0: $t(5) = -3.47$, $p = .018$). Thus, the days Blake et al. (2016) recommends observing fertile variables are those with the highest CP: If researchers follow Lobmaier and Bachofner (2018), they will recruit women when they are statistically less likely to conceive.

Figure 1. Meta-analysis of >10K menstrual cycles documenting the probability of conception on days relative to ovulation.

Note. Window A = Testing days recommended Lobmaier and Bachofner (2018) to capture fertility. Window B = Testing days recommended by Blake et al. (2016) to capture fertility. Testing Window A observes women on days when CP is statistically lower than both peak fertility and Testing Window B.

It is reasonable to expect variation within and between women regarding the length of the fertile window as Lobmaier and Bachofner (2018) note, though evidence of problematic variation is not particularly compelling. Lobmaier and Bachofner (2018) cite Keulers, Hamilton, Franx, Evers, and Bots (2007), which presents variation in the length of the fertile window for 212 subfertile couples (total 212 menstrual cycles), operationalizing the start of the fertile window via sperm-mucus interaction tests. These tests have been heavily criticized, are seldom used by reproductive specialists, and provide a measure of the couple's fertility and not the individual woman's (Oei, Helmerhorst, & Keirse, 1995). Two latter attempts to replicate Kuelers et al. using populations with normal fertility and hormone assays to verify fertility (553 and 1181 menstrual cycles) both find that a higher percentage of fertile windows last 5+ days (Fehring & Schneider, 2008; Robinson, Wakelin, & Ellis, 2007). The standard length of the fertile window advocated by the American Society of Reproductive Medicine (ASRM) and the Society for Reproductive Endocrinology and Infertility is also six days (ASRM, 2017). Thus, the six-day window maxim generally holds despite variation in fertile window length. Observing fertile sessions on the LH surge day or the subsequent day further circumvents this issue, as these days have the greatest conception probability and are the least likely to be non-fertile.

Whether salivary ferning or cervical mucus evaluation in concert with LH testing improves detection of peak fertility is questionable. Salivary ferning is not especially sensitive nor specific: Superior ferning patterns indicative of peak fertility have been detected in men and in postmenopausal women, many women show ferning throughout the cycle with no discernable beginning or end to fertility, and ferning is often poorly temporally correlated with ovulation (Berardono, Melani, Ranaldi, Giachetti, & Vanni, 1993; Braat et al., 1998; Brezina, Haberl, & Wallach, 2011; Fehring & Gaska, 1998; Guida, Barbato, Bruno, Lauro, & Lampariello, 1993; Owen, 2013; Pattanasuttinont, Sereepapong, & Suwajanakorn, 2007).

One study cited in Lobmaier and Bachofner (2018) did find that ferning accurately tracked ovulation (Günther et al., 2015), though only when questionable ferning patterns were assumed to be positive. Unlike ferning, cervical mucus evaluations are highly sensitive to fertility, but like ferning they are also highly unspecific, routinely overestimating the length of the fertile window by an average of four days (Adlercreutz et al., 1982; Ecochard, Duterque, Leiva, Bouchard, & Vigil, 2015; Fehring & Schneider, 2008). LH surge detection combined with cervical mucus evaluation can increase the specificity of identifying the day of ovulation (Leiva, Bouchard, Abdullah, & Ecochard, 2017), though this day is the last in the fertile window and identifying it is often not a priority for fertility effects researchers. Determining whether a day is in the fertile window is usually more relevant, and this outcome can be identified using LH tests alone.

The disadvantage of identifying fertility using LH tests is that estradiol concentrations peak approximately one day prior to the LH surge (Porerfield & White, 2007). The method provided by Blake et al. (2016) does not observe women on the day when estradiol peaks, though it does observe women on days when estradiol is high relative to progesterone and to estradiol throughout the other phases of the menstrual cycle. Given the balancing effects progesterone has on estradiol (Porerfield & White, 2007), observing women during the window specified by Blake et al. should suit most fertility effects research purposes. However, if researchers require observations of participants on the exact day of peak estradiol, alternative methods must be employed. Designs which sample estradiol regularly throughout the menstrual cycle (e.g., Roney & Simmons, 2013) are one such alternative.

The standardized protocol for characterizing fertility provided by Blake et al. (2016) detects LH surges in >95% of all ovulatory cycles. It accommodates the observation of dependent variables during a two-day window where the probability of conception is at its peak, comprised of the LH surge day and the subsequent day (and depending on the time of

day of the LH surge, potentially also for a portion of the day after that). In practice, I have found that stipulating that participants must come into the laboratory within one day of LH surge detection has not negatively affected recruitment or data collection. To evaluate the sensitivity of any variables observed on Day +2 after the LH surge, researchers may wish to report a sensitivity analysis of dependent variables before and after excluding these women, thus amplifying confidence that detected effects are robust and reliable (see Blake, Bastian, O'Dean, & Denson, 2017). By following the protocol from Blake et al. (2016), researchers can avoid the criticisms of indirect, imprecise, and flexible methodologies that continue to muddy the important work conducted in the fertility effects field.

References

- Adlercreutz, H., Brown, J., Collins, W., Goebelsman, U., Kellie, A., Campbell, H., . . . Braissand, G. (1982). The measurement of urinary steroid glucuronides as indices of the fertile period in women. *Journal of Steroid Biochemistry*, 17(6), 695–702.
- American Society for Reproductive Medicine. (2017). Optimizing natural fertility: A committee opinion. *Fertility and Sterility*, 107(1), 52–58.
<https://doi.org/10.1016/j.fertnstert.2016.09.029>
- Berardono, B., Melani, D., Ranaldi, F., Giachetti, E., & Vanni, P. (1993). Is the salivary "ferning" a reliable index of the fertile period? *Acta Europaea Fertilitatis*, 24(2), 61–65.
- 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., Dixon, 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>
- Braat, D. D. M., Smeenk, J. M. J., Manger, A. P., Thomas, C. M. G., Veersema, S., & Merkus, J. M. (1998). Saliva test as ovulation predictor. *The Lancet*, 352(9136), 1283–1284.
- Brezina, P. R., Haberl, E., & Wallach, E. (2011). At home testing: optimizing management for the infertility physician. *Fertility and Sterility*, 95(6), 1867–1878.
<https://doi.org/10.1016/j.fertnstert.2011.01.001>
- Ecochard, Duterque, O., Leiva, R., Bouchard, T., & Vigil, P. (2015). Self-identification of the clinical fertile window and the ovulation period. *Fertility and Sterility*, 103(5), 1319–1325.
<https://doi.org/10.1016/j.fertnstert.2015.01.031>

- Fehring, R. J., & Gaska, N. (1998). Evaluation of the Lady Free Biotester® in determining the fertile period. *Contraception*, 57(5), 325–328. [https://doi.org/10.1016/S0010-7824\(98\)00039-0](https://doi.org/10.1016/S0010-7824(98)00039-0)
- Fehring, R. J., & Schneider, M. (2008). Variability in the hormonally estimated fertile phase of the menstrual cycle. *Fertility and Sterility*, 90(4), 1232–1235. <https://doi.org/10.1016/j.fertnstert.2007.10.050>
- Guida, M., Barbato, M., Bruno, P., Lauro, G., & Lampariello, C. (1993). Salivary ferning and the menstrual cycle in women. *Clinical and Experimental Obstetrics and Gynecology*, 20(1), 48–54.
- Günther, V., Bauer, I., Hedderich, J., Mettler, L., Schubert, M., van Mackelenbergh, M. T., . . . Alkatout, I. (2015). Changes of salivary estrogen levels for detecting the fertile period. *European Journal of Obstetrics, Gynecology, and Reproductive Biology*, 194, 38–42. <https://doi.org/10.1016/j.ejogrb.2015.08.007>
- Keulers, M. J., Hamilton, C. J. C. M., Franx, A., Evers, J. L. H., & Bots, R. S. G. M. (2007). The length of the fertile window is associated with the chance of spontaneously conceiving an ongoing pregnancy in subfertile couples. *Human Reproduction*, 22(6), 1652–1656. <https://doi.org/10.1093/humrep/dem051>
- Leiva, 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. <https://doi.org/10.3389/fpubh.2017.00320>
- Oei, S. G., Helmerhorst, F. M., & Keirse, M. J. (1995). When is the post-coital test normal? A critical appraisal. *Human Reproduction*, 10(7), 1711–1714.
- Owen, M. (2013). Physiological signs of ovulation and fertility readily observable by women. *The Linacre Quarterly*, 80(1), 17–23. <https://doi.org/10.1179/0024363912Z.00000000005>

- Pattanasuttinont, S., Sereepapong, W., & Suwajanakorn, S. (2007). The salivary ferning test and ovulation in clomiphene citrate-stimulated cycles. *Journal of the Medical Association of Thailand*, 90(5), 876–883.
- Porerfield, S. P., & White, B. A. (2007). *Endocrine Physiology* (3rd edition). Mosby *Physiology Monograph Series*. Mosby Elsevier: Philadelphia, PA.
- Robinson, J. E., Wakelin, M., & Ellis, J. E. (2007). Increased pregnancy rate with use of the Clearblue Easy Fertility Monitor. *Fertility and Sterility*, 87(2), 329–334.
<https://doi.org/10.1016/j.fertnstert.2006.05.054>
- Roney, J. R., & Simmons, Z. L. (2013). Hormonal predictors of sexual motivation in natural menstrual cycles. *Hormones and Behavior*, 63(4), 636–645.
<https://doi.org/10.1016/j.yhbeh.2013.02.013>