Bayesian Analysis (2012) **7**, Number 4, pp. 805–808

# Comment on Article by Kim et al.

Joseph B. Stanford *

I congratulate Kim et al. on a valuable contribution. It's gratifying to see ongoing progress in the statistical modeling of human fecundity, as well as initial results from an important data set, the Oxford Conception Study. I have a few comments regarding the comparability of these results to those from other studies in relation to day-specific probabilities of conception, model fit and inference, the impact of using different biomarkers of ovulation on the duration of the "fertile window" (including the two approaches used in this paper), and the future applications of these models.

## 1 Characteristics of the Fertile Window

Overall, these results of day-specific probabilities of conception are similar to those of previous studies, including the higher probabilities found in parous women as compared to nulliparous women (Mikolajczyk and Stanford 2006). However, the relative decreases in conception probability on LH day -2 in relation to day -3 and -1, as well as on day 0 in relation to days -1 and +1, do not admit of any plausible physiologic explanation. In a previous study using the Schwartz model, a smooth (monotonically increasing then decreasing) fertile window was found for the overall study with 434 conceptions out of 3175 cycles around basal body temperature, but many local "dips" similar to those found by Kim et al. were found within most of the analyses of various subgroups of roughly one third to one sixth of the overall sample, including subgroupings by geographic center, age, and cycle characteristics (Colombo and Masarotto 2000). Perhaps future methods work including simulations can elucidate whether such patterns are in fact artifacts related to sample size, or whether they reflect a stark heterogeneity of the pattern of the fertile window in the underlying populations (which seems less likely). Another option to address this issue when sample size limits statistical precision is to constrain the model to a unimodal shape (Dunson and Colombo 2001). The full data of the Oxford Conception Study, when available, should have roughly four times the sample size of this analysis and will likely to be able to address the impact of sample size on the shape of the trajectory of the day specific probabilities (Pyper et al. 2006).

## 2 Model Fit and Inference

The greater flexibility of the generalized t-link model in relation to exponential link functions does seem to result in a better model fit, as assessed by link functions via the Deviance Information Criterion (DIC) and the pseudo marginal likelihood (LPML). It would also be helpful to see comparison simulation results for the exponential link models.

*Department of Family and Preventive Medicine, University of Utah, Salt Lake City, UT, joseph.stanford@utah.edu

 DOI:10.1214/12-BA726B

The increased model fit comes with a trade-off for inference. In the flexible t-link model, coefficients for model covariates are essentially not interpretable of themselves. One must convert to specific day-specific probabilities of conception for various covariate values to be able to get a clear picture of their impact in the model. In the exponentially linked models, coefficients may be directly interpreted. In our application of the Dunson and Stanford (2005) model to a European multicenter data set (Colombo and Masarotto 2000), we obtained a simple logistic regression interpretation for the covariates (Bigelow et al. 2004).

For biological grounding, it would also be helpful in the future to compare model results to crude probabilities of conception in cycles in which there is only one recorded act of coitus during the fertile window. The previously mentioned study by Colombo and Masarotto (2000) found very good concordance of these probabilities to those obtained from the full (exponentially linked) Schwartz model.

## 3 Biomarkers of Ovulation and Duration of the Fertile Window

Kim et al. (by model specification) found a functional window of 7 days for two biomarkers of ovulation: urine LH and cycle length (the so-called "Ogino-Knaus" method). As discussed by the authors, the biomarker of cycle length is less precise as an indicator of the day of ovulation than LH. The original work by Ogino (1930) suggested that the luteal phase varied between 12-16 days (rather than being fixed at 14 days), and other studies have confirmed a yet wider distribution for the postovulatory (luteal) phase of the menstrual cycle (Hilgers et al. 1981; Colombo and Masarotto 2000; Colombo et al. 2006). This is not to suggest that cycle length has no value as a biomarker of ovulation, but to argue that it makes sense to acknowledge the inherent variability of this biomarker by specifying a larger potentially fertile window when using cycle length as a biomarker for the fertile window (as has been done in the contraceptive literature, e.g., by Trussell et al. (1996)), and to explicitly reference the distribution of the fertile to cycle length (i.e., days prior to next menses), rather than to a somewhat arbitrary day 0. Thus, indexing to the length of the cycle (or to the length of the previous cycle for conception cycles) yields a fertile window duration of 11-14 days (Mikolajczyk and Stanford 2005; Mikolajczyk and Stanford 2006). With a more precise marker of ovulation, such as urine LH, cervical fluid, basal body temperature, or even serial transvaginal ultrasound, referencing a day of ovulation as day 0 has more intuitive interpretability and will yield a narrow fertile window, but we must still keep in mind the variability of the ovulation biomarker (Dunson et al. 2001).

Even with a more precise biomarker of ovulation, it is likely that the functional duration of the fertile window may vary between different populations of women. In particular, it may be significantly shorter for subfertile women (Stanford et al. 2003; Keulers et al. 2007).

## 4 Future Applications of Models

Models of day-specific probabilities of conception have been used for assessing the correlation of day-specific biomarkers of reproductive function, such as women's observations of cervical fluid (Bigelow et al. 2004), as well as for potentially time-varying exposures, such as caffeine (Dunson and Colombo 2001) or emergency contraceptive drugs (Trussell et al. 2003; Stanford and Mikolajczyk 2005). Kim et al. have illustrated the application for a non-time-varying (within the fertile window) characteristic, parity. This characteristic, like time trying to conceive, is an observable point of data that reveals something about the underlying distribution of fecundity to which the woman or couple belongs (Stanford et al. 2010). They have indicated they will be working on extensions to accommodate non-time varying covariates (within the fertile window), such as body burden of chemical toxicants. In the spirit of continued flexibility and diversity of application, I would like to advocate for the continued development and testing of models that allow for both non-time varying as well as time-varying covariates.

#### Acknowledgments

David Dunson provided helpful comments on drafts of this response.

## References

Bigelow, J. L., Dunson, D. B., Stanford, J. B., Ecochard, R., Gnoth, C., and Colombo, B. (2004). "Mucus observations in the fertile window: a better predictor of conception than timing of intercourse." *Human Reproduction*, 19(4): 889–892. 806, 807

Colombo, B. and Masarotto, G. (2000). "Daily fecundability: first results from a new data base." *Demographic Research*, 3(5). 805, 806

Colombo, B., Mion, A., Passarin, K., and Scarpa, B. (2006). "Cervical mucus symptom and daily fecundability: first results from a new database." *Statistical Methods Medical Research*, 15(2): 161–180. 806

Dunson, D. B. and Colombo, B. (2001). "Bayesian modeling of the level and duration of fertility in the menstrual cycle." *Biometrics*, 57(4): 1067–1073. 805, 807

Dunson, D. B. and Stanford, J. B. (2005). "Bayesian inferences on predictors of conception probabilities." *Biometrics*, 61: 126–133. 806

Dunson, D. B., Weinberg, C. R., Baird, D., Kesner, J. S., and J.Wilcox, A. (2001). "Assessing human fertility using several markers of ovulation." *Statistics in Medicine*, 20(6): 965–968. 806

Hilgers, T. W., Bailey, A. J., and Prebil, A. M. (1981). "Natural family planning IV. The identification of postovulatory infertility." *Obstetrics and Gynecology*, 58(3): 345–350. 806

Keulers, M. J., Hamilton, C. J. C. M., Franx, A., Evers, J. L. H., and 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. 806

Mikolajczyk, R. T. and Stanford, J. B. (2005). "A new method for estimating the effectiveness of emergency contraception that accounts for variation in timing of ovulation and previous cycle length." *Fertility and Sterility*, 83(6): 1764–1770. 806

— (2006). "Measuring fecundity with standardized estimates of expected pregnancies." *Paediatric and Perinatal Epidemiology*, 20 Suppl 1: 33–40. 805, 806

Ogino, K. (1930). "Ovulationstermin and Konceptionstermin." *Zentralblatt fur Gynakologie*, 54: 464. 806

Pyper, C., Bromhall, L., Dummett, S., Altman, D. G., Brownbill, P., and Murphy, M. (2006). "The Oxford Conception Study design and recruitment experience." *Paediatric and Perinatal Epidemiology*, 20 Suppl 1: 51–59. 805

Stanford, J. B. and Mikolajczyk, R. T. (2005). "Methodological review of the effectiveness of emergency contraception." *Current Women's Health Reviews*, 1(2): 119–129. 807

Stanford, J. B., Mikolajczyk, R. T., Lynch, C. D., and Simonsen, S. E. (2010). "Cumulative pregnancy probabilities among couples with subfertility: effects of varying treatments." *Fertility and Sterility*, 93(7): 2175–2181. 807

Stanford, J. B., Smith, K. R., and Dunson, D. B. (2003). "Vulvar mucus observations and the probability of pregnancy." *Obstetrics and Gynecology*, 101(6): 1285–1293. 806

Trussell, J., Ellertson, C., and Stewart, F. (1996). "The effectiveness of the Yuzpe regimen of emergency contraception." *Family Planning Perspectives*, 28(2): 58–64. 806

Trussell, J., Ellertson, C., von Hertzen, H., Bigrigg, A., Webb, A., and Evans, M. (2003). "Estimating the effectiveness of emergency contraceptive pills." *Contraception*, 67(4): 259–265. 807