Johns Hopkins University, Dept. of Biostatistics Working Papers

10-1-2014

A BAYESIAN APPROACH TO JOINT MODELING OF MENSTRUAL CYCLE LENGTH AND FECUNDITY

Kirsten J. Lum

Johns Hopkins University, Bloomberg School of Public Health, Department of Biostatistics and Division of Intramural Population Health Research, Eunice Kennedy Shriver National Institute of Child Health and Human Development, NIH, DHHS, Rochville, Maryland, KLum@jhu.edu

Rajeshwari Sundaram

Division of Intramural Population Health Research, Eunice Kennedy Shriver National Institute of Child Health and Human Developmemnt, NIH, DHHS, Rockville, Maryland

Germaine M. Buck-Louis

Division of Intramural Population Health Research, Eunice Kennedy Shriver National Institute of Child Health and Human Development, NIH, DHHS, Rockville, Maryland

Thomas A. Louis

Johns Hopkins University, Bloomberg School of Public Health, Department of Biostatistics, U.S. Census Bureau, Suitland, Maryland

A Bayesian Approach to Joint Modeling of Menstrual Cycle Length and Fecundity

Kirsten J. Lum[1,2], Rajeshwari Sundaram[2], Germaine M. Buck Louis[2], and Thomas A. Louis[1,3]

[1]Department of Biostatistics, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD 21205
[2]Division of Intramural Population Health Research, *Eunice Kennedy Shriver* National Institute of Child Health and Human Development, NIH, DHHS, Rockville, Maryland 20852
[3]U.S. Census Bureau, Suitland, Maryland 20746, U.S.A.

Kirsten J. Lum is PhD Candidate, Department of Biostatistics, Johns Hopkins University (JHU), Baltimore, MD 21205 (e-mail: *klum@jhu.edu*) and Predoctoral Fellow, Division of Intramural Population Health Research (DIPHR), *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD), National Institutes of Health (NIH), U.S. Department of Health and Human Services (DHHS), Rockville, Maryland 20852. Rajeshwari Sundaram is Investigator, DIPHR, NICHD, NIH, DHHS, Rockville, Maryland 20852 (e-mail: *sundaramr2@mail.nih.gov*). Germaine M. Buck Louis is Director and Senior Investigator, DIPHR, NICHD, NIH, DHHS, Rockville, Maryland 20852 (e-mail: *louisg@mail.nih.gov*). Thomas A. Louis is Professor, Department of Biostatistics, JHU, Baltimore, MD 21205 (e-mail: *tlouis1@jhu.edu*) and Associate Director for Research and Methodology and Chief Scientist, U.S. Census Bureau, Suitland, MD 20746. This work was supported by the Intramural Research Program of the U.S., NIH, NICHD [Contracts N01-HD-3-3355, N01-HD-3-3356, N01-HD-3-3358]. The authors are grateful for the generosity of Principal Investigator Germaine M. Buck Louis and LIFE Study investigators for providing access to the LIFE Study data. The authors acknowledge that this study utilized the high-performance computational capabilities of the Biowulf Linux cluster at NIH, Bethesda, Maryland (`http://biowulf.nih.gov`).

Abstract

Female menstrual cycle length is thought to play an important role in couple fecundity, or the biologic capacity for reproduction irrespective of pregnancy intentions. A complete assessment of the association between menstrual cycle length and fecundity requires a model that accounts for multiple risk factors (both male and female) and the couple's intercourse pattern relative to ovulation. We employ a Bayesian joint model consisting of a mixed effects accelerated failure time model for longitudinal menstrual cycle lengths and a hierarchical model for the conditional probability of pregnancy in a menstrual cycle given no pregnancy in previous cycles of trying, in which we include covariates for the male and the female and a flexible spline function of intercourse timing. Using our joint modeling approach to analyze data from the Longitudinal Investigation of Fertility and the Environment Study, a couple based prospective pregnancy study, we found a significant quadratic relation between menstrual cycle length and the probability of pregnancy even with adjustment for other risk factors, including male semen quality, age, and smoking status.

Keywords:

Bayesian hierarchical model; Intercourse timing; Length-bias; Mixed effects accelerated failure time model; Pregnancy probability.

1. INTRODUCTION

Fecundity, or the biologic capacity of male and female for reproduction irrespective of pregnancy intentions (Buck Louis 2011), can be measured in terms of multiple endpoints such as pubertal onset and progression, menstrual cycle regularity and length, ovulation, semen quality, and pregnancy, among others. This list emphasizes the role of both the male and female partner and the complexity of factors that contribute to a couple becoming pregnant, yet few statistical models have been developed that model more than one of these endpoints. Until recently available study data on fecundity consisted mostly of retrospective pregnancy studies or prospective studies with fewer than six months follow-up, and mainly of the female partner. However, investigators of a recent prospective pregnancy study called the Longitudinal Investigation of Fertility and the Environment (LIFE) Study (Buck Louis et al. 2011) approached fecundity from a couple-based perspective, collecting measurements on exposures for both partners with the goal of assessing the effects of exposures on multiple fecundity related outcomes. We develop a joint model to investigate the relation between female menstrual cycle length and the couple's probability of pregnancy. We develop a framework that allows for the assessment of time-constant exposures on the probability of pregnancy as well as those mediated through the female menstrual cycle length. Furthermore, we incorporate male factors which may be associated with the probability of pregnancy.

Statistical models of repeated measures of menstrual cycle length data have mainly focused on extensions of mixed effects models to account for covariates, large and heterogeneous within-female variability, and extreme cycle lengths. Harlow and Zeger (1991) categorized menstrual cycle lengths as standard and non-standard and developed separate random effects models for the mean of the standard lengths and the risk of having a non-standard length. Including only standard lengths, Lin et al. (1997) ex-

tended the linear mixed model to allow for heterogeneous within-female variability (see Laird and Ware 1982; Diggle et al. 2002; Verbeke and Molenberghs 2009; Carlin and Louis 2009, for details). Allowing for nonstandard cycle lengths, Guo et al. (2006) proposed a marginal model with covariates for the population mean and variance, assuming a mixture of Gaussian and shifted Weibull error distributions.

Focusing on prediction, Bortot et al. (2010) used a state space approach to develop a predictive model of menstrual cycle length accounting for trend, autocorrelation and extremely long or short outlying lengths. McLain et al. (2012) proposed a parametric model assuming a mixture of Gaussian and Gumbel error distributions and incorporating random effects and covariates on the mean and variance parameters of the Gaussian component. Assuming log-normal error distributions to address heteroscedasticity, Huang et al. (2014) developed a hierarchical framework consisting of change point models for both the mean and variance of cycle length to study changes at late reproductive ages. Building on these works, in the menstrual cycle sub-model of our joint model we propose a Bayesian hierarchical, accelerated failure time model with a mixture error distribution to allow for skewness in cycle length. In addition, we account for length-bias in modeling the length of the cycle in which the couple is enrolled and censoring of the length of the cycle in which the couple becomes pregnant.

Statistical modeling of pregnancy attempts has focused on assessing biological risk factors in the context of the couple's day-specific intercourse behavior (see Ecochard 2006, for a review). However, many models assume there exists a fixed, narrow (fertile) window of days around ovulation outside of which there is no risk for pregnancy. This restriction may not be reasonable, as evidenced by studies which assessed the fertile window, including ovulation using the gold standard, i.e., serial vaginal ultrasound (Keulers et al. 2007). In contrast, Dominik et al. (2001) developed a method that incorporates all intercourse acts in a menstrual cycle, modeling the day-specific

probability of pregnancy as a quadratic function of distance of the intercourse act from ovulation day (assumed to be day 14).

We generalize this approach to a Bayesian hierarchical model for the day-specific probabilities of pregnancy using natural, cubic splines to model the probability of pregnancy, incorporating cycle-specific information on the couple's ovulation day as measured using a fertility monitor, and using the joint model in Section 2 investigate the association of female menstrual cycle length with the probability of pregnancy in the context of male and female risk factors.

In addition to modeling the relation between menstrual cycle length and the probability of pregnancy, we include adjustment for the male contribution. For the LIFE Study, Buck Louis et al. (2014) found significant associations between fecundity and several semen quality parameters, when each parameter was entered into the model individually. To account for the male contribution, we incorporate male age, smoking status and multiple semen quality parameters in addition to female covariates in the model for the probability of pregnancy. Here, we focus on four World Health Organization (WHO) semen quality parameters (morphology (strict criteria), semen volume, sperm concentration, and total sperm count) which are common across many studies of semen quality (see e.g. Cooper et al. 2010).

2. JOINT MODEL FOR MENSTRUAL CYCLE LENGTH AND PREGNANCY

For the i^{th} couple ($i = 1, \ldots, n$), let j ($j = 1, \ldots, n_i$) index the female menstrual cycles with lengths Y_{ij} and ovulation dates O_{ij}. For the $k^{th}(k = 1, \ldots, Y_{ij})$ day of the j^{th} cycle, let $x_{ijk} \in \{0, 1\}$ be the observed day-specific intercourse indicator. For the j^{th} cycle, collect the intercourse indicators in the vector $\mathbf{x_{ij}} = (x_{ij1}, \ldots, x_{ijY_{ij}})$. Further, A_{ij} and δ_i are the cycle-specific and δ_i couple-specific pregnancy indicators.

2.1 Sub-model for Longitudinal Menstrual Cycle Length

In the sub-model for menstrual cycle length, we model the relation between baseline covariates (e.g., age, smoking status) and the expected menstrual cycle length, while accounting for within-female correlation in longitudinal cycle lengths and extremely short and long cycle lengths. To meet these challenges, we use a hierarchical accelerated failure time model (see e.g., Kalbfleisch and Prentice 2011), which was adapted for modeling menstrual cycle length by Lum et al. (2014). For the j^{th} $(j = 2, \ldots, n_i)$ menstrual cycle length from the i^{th} $(i = 1, \ldots, n)$ female, we assume

$$
\begin{aligned}
&[\epsilon_{ij} \mid q, \mu_1, \sigma_1, \mu_2, \sigma_2] \\
&\sim q\text{Gaussian}(\mu_1, \sigma_1^2) + (1-q)\text{Gumbel}(\mu_2, \sigma_2^2),\ \mu_1 > 0, \sigma_1 > 0, \mu_2 > 0, \sigma_2 > 0; \\
&[W_i \mid \sigma_W] \sim \text{Gamma}(1/\sigma_W^2, 1/\sigma_W^2),\ \sigma_W > 0; \\
&Y_{ij} \mid W_i, \exp(\mathbf{v_i}^\top \boldsymbol{\eta}), \epsilon_{ij} = W_i \exp(\mathbf{v_i}^\top \boldsymbol{\eta})\epsilon_{ij}. \qquad (1)
\end{aligned}
$$

Here, $\mathbf{v_i}$ is an r-dimensional vector of observed covariates with corresponding unknown parameter vector η and W_i is a couple-specific random effect with support of the distribution function $\in (0, \infty)$. We assume a Gaussian/Gumbel mixture distribution for the error variables to accommodate two groups of menstrual cycle lengths. We will use $T_i = 1$ to denote normal cycle lengths from a Gaussian distribution and $T_i = 2$ to denote abnormal cycle lengths from a Gumbel distribution. As the mean of the error distribution is non-zero, we do not include an intercept in the fixed effects and we parameterize the random effect distribution such that $E[W] = 1$ and $V[W] = \sigma_W^2$. We assume the distribution of the covariates does not depend on $(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2)$ and that the cycle lengths within a female are conditionally independent given $\mathbf{v_i}$ and $W_i = w_i$. Let $F_\epsilon(\epsilon; \mu_1, \sigma_1^2, \mu_2, \sigma_2^2)$ denote the Gaussian-Gumbel cumulative distribution

function (CDF). By a transformation, the CDF of $Y_{ij}, i = 1, \ldots, n, j = 2, \ldots, n_i$ is $F(y_{ij} \mid w_i, \mathbf{v_i}; \mu_1, \sigma_1^2, \mu_2, \sigma_2^2, \eta) = F_\epsilon(y_{ij} w_i^{-1} \exp(-\mathbf{v_i}^\top \boldsymbol{\eta}); \mu_1, \sigma_1^2, \mu_2, \sigma_2^2, \eta)$ and the probability density function is

$f(y_{ij} \mid w_i, \mathbf{v_i}; \mu_1, \sigma_1^2, \mu_2, \sigma_2^2, \eta) = w_i^{-1} \exp(-\mathbf{v_i}^\top \boldsymbol{\eta}) f_\epsilon(y_{ij} w_i^{-1} \exp(-\mathbf{v_i}^\top \boldsymbol{\eta}); \mu_1, \sigma_1^2, \mu_2, \sigma_2^2, \eta).$

2.1.1 Modeling the Length of the Enrollment Cycle

In Lum et al. (2014), we proposed a method for estimating the population menstrual cycle length distribution using the enrollment cycle (i.e., the menstrual cycle during which the women enrolled), that accounts for features of the sampling plan of the LIFE Study, specifically length-bias and a selection process that is potentially a function of the time since the last menstrual period (LMP). For the LIFE study, we found that the estimated probability of enrollment was approximately constant with respect to time from LMP. Based on this finding, we account for the sampling plan by assuming a length-biased sampling distribution for Y_{i1}

$$f_{Y_1}(y_{i1} \mid w_i, \mathbf{v_i}; \mu_1, \sigma_1^2, \mu_2, \sigma_2^2, \eta) = \frac{y_{i1} f(y_{i1} \mid w_i, \mathbf{v_i}; \mu_1, \sigma_1^2, \mu_2, \sigma_2^2, \eta)}{E(Y_i \mid W_i, \exp(\mathbf{v_i}^\top \boldsymbol{\eta}), q, \mu_1, \mu_2)} \tag{2}$$

where $E(Y_i \mid W_i, \exp(\mathbf{v_i}^\top \boldsymbol{\eta}), q, \mu_1, \mu_2) = W_i \exp(\mathbf{v_i}^\top \boldsymbol{\eta})\{q\mu_1 + (1-q)\mu_2\}$.

2.1.2 Modeling the Length of the Pregnancy Cycle

Menstrual cycle lengths were prospectively measured until the pregnancy cycle, during which increased levels of progesterone preclude menstrual bleeding in preparation for implantation of the blastocyst. Therefore, the length of the pregnancy cycle (i.e., the menstrual cycle during which the couple becomes pregnant) is right censored. Let τ_{in_i} be the time (in days) from the first day of the n_i^{th} cycle to the censoring day. We assume the censoring distribution is non-informative and conditionally independent of Y given

$W = w$ and $\mathbf{v}$. For enrollment pregnancy cycles, the contribution to the likelihood is proportional to

$\{1 - F(\tau_{in_i} \mid w_i, \mathbf{v_i}; \mu_1, \sigma_1^2, \mu_2, \sigma_2^2, \eta)\}/E(Y_i \mid W_i, \exp(\mathbf{v_i}^\top \boldsymbol{\eta}), q, \mu_1, \mu_2)$; while for non-enrollment pregnancy cycles, the contribution is proportional to $1-F(\tau_{in_i} \mid w_i, \mathbf{v_i}; \mu_1, \sigma_1^2, \mu_2, \sigma_2^2, \eta)$.

2.1.3 Choice of Priors for Menstrual Cycle Length Submodel

Let $\boldsymbol{\phi_Y} = (\boldsymbol{\eta}, \sigma_W, \mu_1, \sigma_1, \mu_2, \sigma_2)$ denote the unknown population parameters of the menstrual cycle length model. We assume the components of $\boldsymbol{\phi}_Y$ are independent *a priori*, such that $[\boldsymbol{\phi_Y}] = [\sigma_W][\mu_1][\sigma_1][\mu_2][\sigma_2] \prod_{r=1}^{R} [\eta_r]$. We complete the hierarchical model with the following specification of uniform ($\mathcal{U}$) priors:

$$\begin{aligned}
[\eta_r \mid a_{\eta_r}, b_{\eta_r}] &\sim \mathcal{U}(a_{\eta_r}, b_{\eta_r}), r = 1, \ldots, R, \\
[\sigma_W] &\sim \mathcal{U}(0, b_{\sigma_W}), \\
[\mu_1] &\sim \mathcal{U}(a_{\mu_1}, b_{\mu_1}), \\
[\sigma_1] &\sim \mathcal{U}(0, b_{\sigma_1}), \\
[\mu_2] &\sim \mathcal{U}(a_{\mu_2}, b_{\mu_2}), \\
[\sigma_2] &\sim \mathcal{U}(0, b_{\sigma_2}),
\end{aligned} \tag{3}$$

where we choose values for the hyperparameters that are scaled to determine vague priors. In (3), we choose a uniform prior on the scale of the standard deviation of the female-specific parameter so as not to bias the prior away from 0 (Gelman 2006). If the posterior distribution of σ_W includes zero, for example, this would indicate no between-female variability in the mean beyond that explained by the fixed effects. We discuss alternative choices for the priors in Section 4.2.

2.2 Submodel for the Probability of Pregnancy

The observed prospective pregnancy data include at the cycle level, pregnancy indicated by Clearblue®Easy pregnancy tests which detect levels of human chorionic gonadotropin over 25 mIU/mL; and within each cycle, the ovulation day and the pattern of intercourse acts. We aim to estimate the association between the female-specific distribution of menstrual cycle length and the probability of pregnancy in a cycle, accounting for intercourse pattern in both the current and previous cycles, the preceding cycles of attempts ending without pregnancy, and male and female covariates.

For the j^{th} cycle, let k' ($k' = 1, \ldots, K'_{ij}$) index the intercourse acts in the cycle, $\nu_{ijk'}$ denote the (possibly unobserved) indicator of fertilization at the $(k')^{th}$ intercourse act, $d_{ijk'}$ denote the time difference (in days) of the intercourse day from ovulation; and let $\boldsymbol{\nu_{ij}} = (\nu_{ij1}, \ldots, \nu_{ijK'_{ij}})$ and $\mathbf{d_i} = (d_{i11}, \ldots, d_{ijK'_{ij}})$. We develop a model for the conditional probability of pregnancy in the j^{th} cycle given no pregnancy at each of the intercourse acts in previous cycles and timing of all intercourse, denoted
$Pr(A_{ij} = 1 \mid A_{i1} = \ldots = A_{i(j-1)} = 0, \boldsymbol{\nu_{i1}}, \ldots, \boldsymbol{\nu_{i(j-1)}}, K'_{ij}, \mathbf{d_{ij}}), j > 1.$

We assume intercourse cannot be successful if the ovum has already been fertilized; thus, $\nu_{ijk'}$ is dependent on $\{\nu_{i11}, \ldots, \nu_{ij(k'-1)}\}$. Let $\rho(d_{ijk'})$ be the conditional probability of fertilization at the $(k')^{th}$ intercourse given no prior fertilization. For the $(k')^{th}$ intercourse, $k' > 1$, we have

$$\begin{aligned} &\text{pr}(\nu_{ijk'} = 1 \mid \nu_{i11} = \ldots = \nu_{ij(k'-1)} = 0) = \rho(d_{ijk'}), \\ &\text{pr}(\nu_{ijk'} = 1 \mid \max(\nu_{i11}, \ldots, \nu_{ij(k'-1)}) = 1) = 0; \end{aligned} \tag{4}$$

where (4) follows from the fact that prior fertilization prevents subsequent. For the cycles preceding the j^{th} cycle, we observe $\boldsymbol{\nu_{i1}} = \ldots = \boldsymbol{\nu_{i(j-1)}} = \mathbf{0}$. However, if pregnancy occurs in the j^{th} cycle, $\boldsymbol{\nu_{ij}}$ is unobserved for $K'_{ij} > 1$. Therefore, we express

the conditional probability of pregnancy in the j^{th} cycle as the sum of the probabilities of all possible permutations of $(v_{ij1}, \ldots, v_{ijK'_{ij}})$ under the condition $\sum_{k'} v_{ijk'} = 1$. In the special case of one act of intercourse in the j^{th} cycle,

$$Pr(A_{ij} = 1 \mid A_{i1} = \ldots = A_{i(j-1)} = 0, \boldsymbol{\nu_{i1}}, \ldots, \boldsymbol{\nu_{i(j-1)}}, K'_{ij} = 1, \mathbf{d_i}) = \rho(d_{ij1}).$$

For the probability of pregnancy in the j^{th} cycle with two intercourse acts,

$$Pr(A_{ij} = 1 \mid A_{i1} = \ldots = A_{i(j-1)} = 0, \boldsymbol{\nu_{i1}}, \ldots, \boldsymbol{\nu_{i(j-1)}}, K'_{ij} = 2, \mathbf{d_i}) = \rho(d_{ij1}) + \{1 - \rho(d_{ij1})\}\rho(d_{ij2});$$

and in general, for the probability of pregnancy in the j^{th} cycle with K'_{ij} intercourse acts,

$$Pr(A_{ij} = 1 \mid A_{i1} = \ldots = A_{i(j-1)} = 0, \boldsymbol{\nu_{i1}}, \ldots, \boldsymbol{\nu_{i(j-1)}}, K'_{ij}, \mathbf{d_i}) = \rho(d_{ij1}) + \Big\{ \sum_{k'=2}^{K'_{ij}} \Big(\big[\prod_{l=1}^{k'-1} \{1 - \rho(d_{ijl})\} \big] \rho(d_{ijk'}) \Big) \Big\}^{K'_{ij} > 1}. \quad (5)$$

The conditional probability of no pregnancy in the j^{th} cycle with K'_{ij} intercourse acts is

$$Pr(A_{ij} = 0 \mid A_{i1} = \ldots = A_{i(j-1)} = 0; \boldsymbol{\nu_{i1}}, \ldots, \boldsymbol{\nu_{i(j-1)}}, K'_{ij}, \mathbf{d_i}) = \prod_{k'=1}^{K'_{ij}} \{1 - \rho(d_{ijk'})\}.$$

The expression in (5) is equivalent to

$$Pr(A_{ij} = 1 \mid A_{i1} = \ldots = A_{i(j-1)} = 0, \boldsymbol{\nu_{i1}}, \ldots, \boldsymbol{\nu_{i(j-1)}}, K'_{ij}, \mathbf{d_i}) = 1 - \prod_{k'=1}^{K'_{ij}} \{1 - \rho(d_{ijk'})\}.$$

Let k $(k = 1, \ldots, \ddot{Y}_{ij})$ index day of the menstrual cycle, where $\ddot{Y} = [Y]$ and $[\ldots]$ denotes the greatest integer function. Using x_{ijk} to indicate intercourse and $\mathbf{x_i}$ to denote the

full history of intercourse acts, we equivalently model

$$Pr(A_{ij} = 1 \mid A_{i1} = \ldots = A_{i(j-1)} = 0, \mathbf{x_i}, \mathbf{d_i}) = 1 - \prod_{k=1}^{\ddot{Y}_{ij}} \{1 - \rho(d_{ijk})\}^{x_{ijk}}.$$

One decision that must be made in fitting this model is which days of the menstrual cycle to include. Various approaches have been suggested such as assuming the intercourse act closest to ovulation is the one that fertilizes the ovum (Royston and Ferreira 1999) or assuming an intercourse act has a nonzero probability of success only if it falls within a predetermined fixed window around the ovulation day (see e.g., Weinberg et al. 1994; Dunson and Stanford 2005). We employ this second approach, but since a wide variety of fixed windows have been reported (Lynch et al. 2006), we use a very broad window, specifically excluding only intercourse acts more than 16 days before ovulation or 18 days after ovulation. Approximately 5% of all observed intercourse acts occurred outside this range. In a model of barrier contraceptive efficacy, Dominik and Chen (2006) consider a broad window consisting of 13 days prior to and 16 days after an assumed ovulation on day 14 of the cycle.

In the second level of the hierarchical model, we model the probability of pregnancy by intercourse on the k^{th} day of the cycle as a function of the latent female-specific menstrual cycle length, denoted Y_{pi}, with adjustment for male and female risk factors and difference of day k from ovulation:

$$\text{logit}[\rho\{\mathbf{m_i}, \boldsymbol{\beta}, \mathbf{z_i}, \boldsymbol{\gamma}, \alpha_0, g(d_{ijk})\}] = \mathbf{m_i}^\top \boldsymbol{\beta} + \mathbf{z_i}^\top \boldsymbol{\gamma} + \alpha_0 + g(d_{ijk}). \tag{6}$$

In (6), $\mathbf{m_i}$ is a vector composed of linear and potentially quadratic terms of Y_{pi}, which we incorporate by mixing over the posterior predictive distribution conditional on the woman's observed menstrual cycle lengths, priors and hyperpriors. The corresponding

unknown regression coefficient vector $\boldsymbol{\beta}$, links the menstrual cycle length and pregnancy sub-models.

In separate models, we also consider the relation between the probability of pregnancy and the latent female-specific menstrual cycle length conditional on normal cycles denoted $Y_{pi} \mid T_i = 1$, and the corresponding conditional means, $E(Y_{pi} \mid \mathbf{Y_i}, \phi_Y, W_i, \mathbf{v_i})$ and $E(Y_{pi} \mid \mathbf{Y_i}, \phi_Y, W_i, \mathbf{v_i}, T_i = 1)$. In Section 4.3, we present results from models with the following choices for $\mathbf{m_i}^\top \boldsymbol{\beta}$:

$$\mathbf{m_i}^\top \boldsymbol{\beta} = \beta_1 Y_{pi} \qquad \text{(Lin)}$$

$$\mathbf{m_i}^\top \boldsymbol{\beta} = \beta_1 Y_{pi} + \beta_2 (Y_{pi})^2 \qquad \text{(LinQuad)}$$

$$\mathbf{m_i}^\top \boldsymbol{\beta} = \beta_1 (Y_{pi} \mid T_i = 1) \qquad \text{(LinT=1)}$$

$$\mathbf{m_i}^\top \boldsymbol{\beta} = \beta_1 (Y_{pi} \mid T_i = 1) + \beta_2 (Y_{pi} \mid T_i = 1)^2 \qquad \text{(LinQuadT=1)}$$

$$\mathbf{m_i}^\top \boldsymbol{\beta} = \beta_1 E(Y_{pi} \mid \mathbf{Y_i}, \phi_Y, W_i, \mathbf{v_i}) \qquad \text{(LinCmean)}$$

$$\mathbf{m_i}^\top \boldsymbol{\beta} = \beta_1 E(Y_{pi} \mid \mathbf{Y_i}, \phi_Y, W_i, \mathbf{v_i}) + \beta_2 \{E(Y_{pi} \mid \mathbf{Y_i}, \phi_Y, W_i, \mathbf{v_i})\}^2 \qquad \text{(LinQuadCmean)}$$

$$\mathbf{m_i}^\top \boldsymbol{\beta} = \beta_1 E(Y_{pi} \mid \mathbf{Y_i}, \phi_Y, W_i, \mathbf{v_i}, T_i = 1) \qquad \text{(LinCmeanT=1)}$$

$$\mathbf{m_i}^\top \boldsymbol{\beta} = \beta_1 E(Y_{pi} \mid \mathbf{Y_i}, \phi_Y, W_i, \mathbf{v_i}, T_i = 1) + \beta_2 \{E(Y_{pi} \mid \mathbf{Y_i}, \phi_Y, W_i, \mathbf{v_i}, T_i = 1)\}^2 \qquad \text{(LinQuadCmeanT=1)}$$

In (6) we also adjust for male and female risk factors denoted by $\mathbf{z_i}$ with corresponding parameter vector γ. Specifically, we consider four of the WHO male semen quality parameters, male smoking status and female smoking status. As male and female age are highly correlated, we incorporate the average and difference in male and female ages. Additionally, we account for the time (in days) from intercourse to ovulation, $d_{ijk} = k - O_{ij}$. We model $g(d_{ijk})$ as a smooth function estimated by $\hat{g}(\cdot) = \sum_{l=1}^{L} \alpha_l B_l(\cdot)$, where $\{B_1(\cdot), \ldots, B_L(\cdot)\}$ are the B-spline basis functions for a natural cubic spline. The intercept is represented by the parameter α_0. Previous specifications of $g(d_{ijk})$ found in

the literature include a piecewise linear spline (Dominik et al. 2001), a quadratic function centered around ovulation (Dominik and Chen 2006), or a day-specific parameter with restriction of k to a small window of days around ovulation (see e.g., Dunson and Stanford 2005).

We denote the parameters of the pregnancy model by $\boldsymbol{\phi_A} = (\beta, \boldsymbol{\gamma}, \boldsymbol{\alpha})$. Assuming the components of $\boldsymbol{\phi_A}$ are independent and are also independent of $\boldsymbol{\phi_Y}$; we complete the model structure by choosing noninformative uniform priors for each parameter in $\boldsymbol{\phi_A}$.

2.2.1 Modeling the Probability of Pregnancy in the Enrollment Cycle

As we describe in Section 4.1, the ovulation day is estimated based on day level data collected using the Clearblue®Easy fertility monitor. Since couples were allowed to enroll on any day of the female menstrual cycle, the majority of the couples do not have sufficient monitor data throughout the enrollment cycle to estimate the ovulation day. Further, for the most fecund couples who become pregnant before the first observed bleeding event (approximately 10%), the enrollment cycle is their only cycle under observation. In order to include these couples in the analysis, we model the probability of becoming pregnant in the enrollment cycle as a function of the couple's baseline covariates, average number of intercourse acts (IntcFreq) and the female menstrual cycle length:

$$\text{logit}\{P(A_{i1} = 1; \mathbf{m_i}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\lambda}, \mathbf{z_i})\} = \mathbf{m_i}^\top \boldsymbol{\beta} + \mathbf{z_i}^\top \boldsymbol{\gamma} + \lambda_0 + \lambda_1 \text{IntcFreq}_i + \lambda_2 (\text{IntcFreq}_i)^2.$$

3. ESTIMATION

We jointly model the longitudinal menstrual cycle lengths and the conditional probability of pregnancy per cycle, assuming these processes are independent conditional

on the shared random effect W_i. Further, we assume cycle lengths, $(Y_{i1}, \ldots, Y_{in_i})$, are independent given $(W_i, \mathbf{v_i})$. Let $\mathbf{Y_i^o}$ denote the observed menstrual cycle lengths and Y_i^u the unobserved length of the pregnancy cycle. Under these assumptions, the joint distribution of $(\mathbf{Y_i}, \mathbf{A_i}, Y_{pi}, W_i)$ can be factored as

$$[\mathbf{Y_i}, \mathbf{A_i}, Y_{pi}, W_i \mid \boldsymbol{\phi_Y}, \boldsymbol{\phi_A}] = [\mathbf{A_i} \mid Y_{pi}, \mathbf{z_i}, \mathbf{x_i}, \mathbf{d_i}, \boldsymbol{\phi_A}][Y_{pi} \mid \mathbf{Y_i}, W_i, \mathbf{v_i}, \boldsymbol{\phi_Y}][\mathbf{Y_i} \mid W_i, \mathbf{v_i}, \boldsymbol{\phi_Y}][W_i \mid \boldsymbol{\phi_Y}],$$

where

$$\begin{aligned}
[\mathbf{Y_i} \mid W_i, \mathbf{v_i}, \boldsymbol{\phi_Y}] &= [\mathbf{Y_i^o}, Y_i^u \mid W_i, \mathbf{v_i}, \boldsymbol{\phi_Y}] \\
&= [Y_i^u \mid \mathbf{Y_i^o}, W_i, \mathbf{v_i}, \boldsymbol{\phi_Y}]^{\delta_i}[\mathbf{Y_i^o} \mid W_i, \mathbf{v_i}, \boldsymbol{\phi_Y}] \\
&= [Y_i^u \mid \mathbf{Y_i^o}, W_i, \mathbf{v_i}, \boldsymbol{\phi_Y}]^{\delta_i}[Y_{in_i}^o \mid W_i, \mathbf{v_i}, \boldsymbol{\phi_Y}]^{(1-\delta_i)} \prod_{j=1}^{n_i-1} [Y_{ij}^o \mid W_i, \mathbf{v_i}, \boldsymbol{\phi_Y}].
\end{aligned}$$

The unknown model parameters for couple i are: $Y_i^u, Y_{pi}, W_i, \boldsymbol{\phi_Y}$, and $\boldsymbol{\phi_A}$; and the observed data are $\mathbf{D} = (\mathbf{Y_i^o}, \mathbf{v_i}, \mathbf{A_i}, \mathbf{O_i}, \mathbf{z_i}, \mathbf{x_i}, \mathbf{d_i})$, where $\mathbf{x_i}$ is a vector composed of the stacked vectors of day-specific intercourse indicators. For a couple with $A_{in_i} = 0$ (i.e. no pregnancy), $\mathbf{Y_i^o} = \mathbf{Y_i}$ (i.e. the lengths of all cycles $j = 1, \ldots, n_i$ are observed) and the contribution to the joint posterior distribution is given by

$$\begin{aligned}
[Y_{pi}, W_i, \boldsymbol{\phi_Y}, \boldsymbol{\phi_A} | \mathbf{D}; A_{in_i} = 0] \propto \\
&\Big(\prod_{j=1}^{n_i} [Y_{ij}^o \mid W_i, \mathbf{v_i}, \boldsymbol{\phi_Y}][A_{ij} = 0 \mid A_{i1} = \ldots = A_{i(j-1)} = 0; Y_{pi}, \mathbf{x_i}, \mathbf{d_i}, \mathbf{z_i}, \boldsymbol{\phi_A}] \Big) \\
&\times [Y_{pi} \mid \mathbf{Y_i}, W_i, \mathbf{v_i}, \boldsymbol{\phi_Y}][W_i \mid \boldsymbol{\phi_Y}][\boldsymbol{\phi_Y}][\boldsymbol{\phi_A}].
\end{aligned}$$

For a couple with $A_{in_i} = 1$ (i.e. pregnant), the contribution to the joint posterior

distribution is given by

$$
\begin{aligned}
[Y_i^u, Y_{p_i}, W_i, \boldsymbol{\phi_Y}, \boldsymbol{\phi_A} | \mathbf{D}; A_{in_i} = 1] \propto \\
&\Big(\prod_{j=1}^{n_i - 1} [Y_{ij}^o \mid W_i, \mathbf{v_i}, \boldsymbol{\phi_Y}][A_{ij} = 0 \mid A_{i1} = \ldots = A_{i(j-1)} = 0; Y_{pi}, \mathbf{x_i}, \mathbf{d_i}, \mathbf{z_i}, \boldsymbol{\phi_A}] \Big)^{\mathbb{1}\{n_i > 1\}} \\
&\times [Y_i^u \mid \mathbf{Y_i^o}, W_i, \mathbf{v_i}, \boldsymbol{\phi_Y}][A_{in_i} = 1 \mid A_{i1} = \ldots = A_{i(n_i-1)} = 0; Y_{pi}, \mathbf{x_i}, \mathbf{d_i}, \mathbf{z_i}, \boldsymbol{\phi_A}] \\
&\times [Y_{pi} \mid \mathbf{Y_i}, W_i, \mathbf{v_i}, \boldsymbol{\phi_Y}][W_i \mid \boldsymbol{\phi_Y}][\boldsymbol{\phi_Y}][\boldsymbol{\phi_A}].
\end{aligned}
$$

We estimate the marginal posterior distributions of the parameters using Markov chain Monte Carlo (MCMC) integration. We describe the details of implementing our approach in Section 4.2.

4. APPLICATION TO THE LIFE STUDY, A PROSPECTIVE PREGNANCY STUDY

The aim of our analysis is to investigate the relation between female menstrual cycle length and couple fecundity, while accounting for both female and male risk factors in keeping with the couple dependent nature of human reproduction. Specifically, we use the joint model detailed in Section 2 to model the female-specific menstrual cycle length distribution jointly with the couple's fecundity, as measured by the probability of pregnancy in a menstrual cycle. We assess the association between menstrual cycle length and the probability of pregnancy adjusting for male and female smoking status, average of male and female age, difference between male and female age, and four WHO semen quality parameters; specifically semen volume, sperm concentration, total sperm count, and morphology (strict criteria) as discussed in Cooper et al. (2010). While the focus of Cooper et al. (2010) is defining cutpoints for the semen quality parameters, we choose to use the continuous measurements. Furthermore, since total sperm count

is the product of semen volume and sperm concentration, we fit two separate models adjusting for either total sperm count or both semen volume and sperm concentration.

4.1 The LIFE Study

We apply our joint modeling approach to the LIFE Study, a couple-based, prospective pregnancy study of 501 couples. Couples were accepted if they fulfilled the following inclusion criteria: married or in a committed relationship, females aged 18-40 and males over age 18 years, English or Spanish speaking, self-reported menstrual cycle lengths within 21-42 days, and no hormonal birth control injections in the past 12 months. Couples were followed until pregnancy, exit from the study or one year attempting pregnancy. Pregnancy was recognized using Clearblue®Easy pregnancy tests administered at home on the day menstruation is expected. For the length of the pregnancy attempt, both the male and female kept independent journals of daily intercourse frequency. If the intercourse question was not answered on a particular day in the female daily journal, we used the number of acts recorded in the male daily journal. If the intercourse question was left blank by both members of the couple (approximately 2.9% of days), we assume the couple did not have intercourse on that day. Women recorded daily bleeding observed on a scale of 0 (none) to 4 (heavy). The beginning of the menstrual cycle is defined by menstruation, designated as the first day of bleeding followed within one day by at least two additional days bleeding. The length of the menstrual cycle is then defined as time (in days) from the first day of menstrual bleeding to the day preceding menstrual bleeding of the next cycle. Measurements of semen quality parameters were made at the National Institute for Occupational Safety and Health's andrology laboratory and a contract laboratory (Fertility Solutions) on samples collected at home via masturbation without the use of any lubricant following 2 days of abstinence. Morphology was assessed using the strict and WHO normal

criteria (World Health Organization 1992; Rothmann et al. 2013). As both criteria measure morphology, we choose to use the strict criteria. Additional details on the collection of semen quality measurements can be found in Buck Louis et al. (2014).

Beginning on day six of the menstrual cycle, women used the Clearblue®Easy fertility monitor to measure daily levels of oestrone-$_3$-glucuronide, a metabolite of oestradiol, and luteinizing hormone (LH). The ovulation day was identified as the peak day detected using the Clearblue®Easy fertility monitor. In the case of two consecutive peak days, the latter of the two was designated as the ovulation day. If a peak was not detected and the female tested on at least 90% of the 20 testing days, we assumed the cycle was anovulatory (151 cycles) and therefore had zero risk for pregnancy; otherwise, we imputed the mean of the couple's observed ovulation days (250 cycles). If ovulation was not observed in any cycles, we imputed the ovulation day as the day with the highest LH peak (33 cycles). We excluded 20 cycles for which the ovulation day could not be detected or imputed due to non-compliance in testing.

For this analysis, we consider a subset of the LIFE Study restricted to 436 couples with complete data on the four WHO semen quality parameters so that we may investigate both male and female risk factors for pregnancy. Couples who exited the study are censored on the last day of the cycle preceding their exit. We further excluded 10 couples who exited in the enrollment cycle and thus are missing data on intercourse days, menstrual cycle length, and pregnancy success, bringing the size of the subset to 426 couples (1934 menstrual cycles). Lastly, we excluded 23 cycles shorter than 9 or longer than 89 days. As the upper limit is more than twice the study's inclusion criteria of 42 days, lengths outside this range are likely an artifact of a missing cycle stop/start.

4.2 Implementation Details

The posterior distributions of the population and couple-specific parameters of the joint model described in Section 3 are estimated using MCMC sampling methods implemented using the OpenBUGS software. Convergence of parameters is checked by visual inspection of trace plots. Based on these trace plots, we use a burn-in of 10000 iterations and a sample of 30000 iterations. While the regression coefficients of both sub-models exhibit rapid mixing, the spline parameters are slower to converge.

To assess the sensitivity to the choice of uniform priors for $(\sigma_1, \sigma_2, \sigma_W)$, we fit the model assuming uniform priors on the scale of the log standard deviations; and we found that the corresponding posterior distributions were approximately the same. As a separate sensitivity analysis, we also fit the model assuming a log-normal distribution for $W_i, i = 1, \ldots, n$. Again, there was very little difference in the posterior distributions.

For the shape of the profile of pregnancy probabilities in the fecundity model, we assume $g(\cdot)$ is a smooth function and estimate g by $\hat{g}(\cdot) = \sum_{l=1}^{L} \alpha_l B_l(\cdot)$, where $\{B_1(\cdot), \ldots, B_L(\cdot)\}$ are the B-spline basis functions for a natural cubic spline. Brown et al. (2005) suggest choosing the number of knots based on the model with the smallest deviance information criterion (DIC). We separately fit models with 4-10 knots at locations based on percentiles of the intercourse day data. We found that the DIC was smaller for the models with 8, 9, or 10 knots; therefore, we chose to use 8 knots so that the profile of pregnancy probabilities is informed by the intercourse data without the addition of too many parameters. We exclude intercourse acts occurring more than 16 days prior to ovulation or 18 days post ovulation, or on the days following a positive pregnancy test as these occur after conception; and we also exclude intercourse acts on days occurring more than 45 days after the start of a cycle.

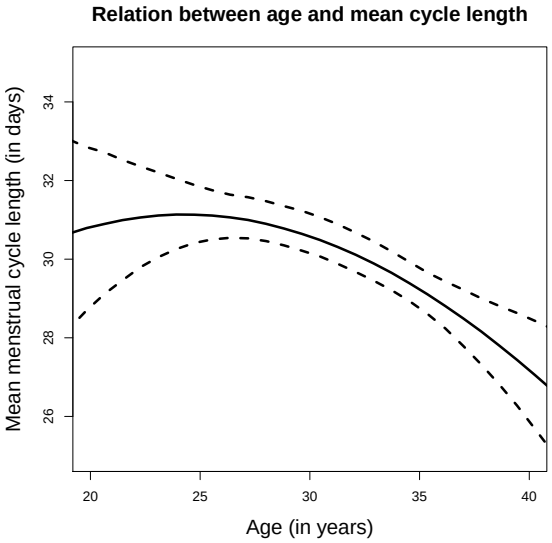

Figure 1: Median (solid) and 95% equal tail credible interval (dashed) for the posterior distribution of mean menstrual cycle length (in days) versus female age (in years).

4.3 Analysis of the LIFE Study

We first describe the relation between the female-specific mean menstrual cycle length and baseline age and smoking status. Table 1 shows the posterior medians and 95% equal tail credible intervals (CI) of the parameters of the regression model for mean menstrual cycle length (1). For the age range of the LIFE Study (18 to 40 years), we found that the female-specific mean cycle length decreased with increasing age (see Figure 1), which is in agreement with the findings by Harlow et al. (2000) in a study of menstrual cycle characteristics by age. We also included a quadratic age term, though this was not significant. For smoking status, with exposed defined as cotinine level at least 10ng/mL, the posterior median of the coefficient was negative; however, the percentage of exposed females was small and the posterior distribution included zero. In the model for the length of the enrollment cycle (2) we accounted for

Table 1: Posterior median and 95% equal-tail credible intervals for menstrual cycle length parameters of joint model displayed as: $_{\text{Lo}}\text{Median}_{\text{Up}}$ (Louis and Zeger 2009). Age is standardized and smoking status (exposed) is defined as baseline cotinine level $\geq$ 10ng/mL. Restricted to 426 couples with data on semen quality who did not exit in the enrollment cycle.

Menstrual cycle length sub-model	
Parameter	$_{\text{Lo}}\text{Median}_{\text{Up}}$
Age	$_{-0.04}-0.03_{-0.01}$
Age (squared)	$_{-0.02}-0.01_{0.00}$
Smoke (exposed/unexposed)	$_{-0.07}-0.02_{0.03}$
Probability of normal cycle	$_{0.75}0.79_{0.82}$
Mean, normal cycle	$_{28.90}29.33_{29.85}$
Standard deviation, normal cycle	$_{1.79}1.94_{2.11}$
Mean, abnormal cycle	$_{33.84}35.18_{37.02}$
Standard deviation, abnormal cycle	$_{11.06}12.34_{13.60}$
Standard deviation, random effect	$_{0.08}0.09_{0.10}$

length-bias, which can be visualized by the right shift in the histogram of the female-specific mean of the enrollment cycle lengths compared to that of the post-enrollment cycles lengths (see Figure 2). Lastly, the parameters of the Gaussian-Gumbel mixture distribution indicate that the probability of a normal cycle is approximately 79% (CI: (75%, 82%)), and that the mean and standard deviation of the Gaussian distribution are much smaller than that of the Gumbel distribution. In the supplementary materials, we assess the fit of the menstrual cycle length sub-model.

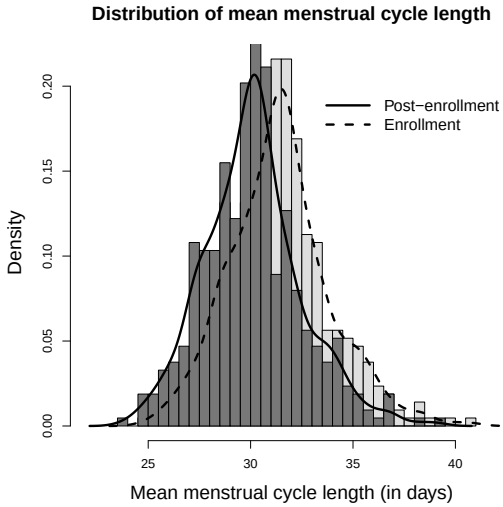

Figure 2: Histograms and density estimates of the estimated woman-specific mean menstrual cycle length (in days) for enrollment cycles (light gray bars, dashed line) compared to that of post-enrollment cycles (dark gray bars, solid line). Estimates shown are medians of the posterior distribution of woman-specific mean cycle length.

For the remainder of this section, we describe the results of the fecundity portion of the joint model. For each couple, we estimated the probability of pregnancy in a cycle conditional on no pregnancy in previous cycles of attempts, intercourse history, and couple covariates. Figure 3 shows boxplots of the conditional probability of pregnancy by cycle number. The conditional probability of pregnancy diminishes as the number of cycles without pregnancy increases.

Table 2 shows the posterior medians and 95% CIs of the parameters of the regression model for the day-specific probability of pregnancy on menstrual cycle length, without adjustment for other risk factors. We first fit a linear term for menstrual cycle length (model Lin) and then added a quadratic term (model LinQuad) which we found to be significant. For intercourse on a particular day of the menstrual cycle, we

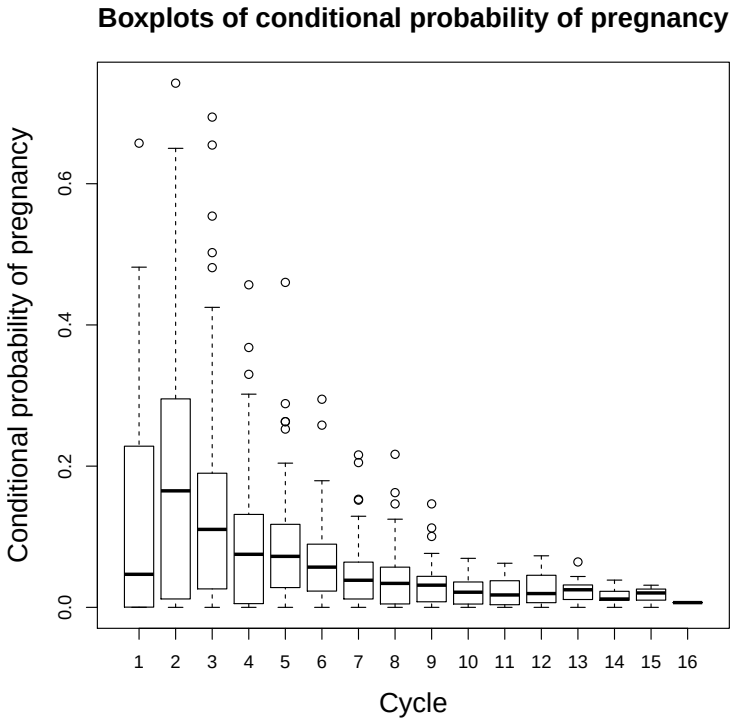

Figure 3: Box plots of conditional probability of pregnancy in a cycle given no pregnancy in previous cycles, intercourse history, menstrual cycle length, total sperm count, sperm morphology (strict criterion), mean of male and female age, difference between male and female age, male and female smoking status (exposed/unexposed).

Table 2: Posterior median and 95% equal-tail credible intervals (CI) for association between (mean) menstrual cycle length and probability of pregnancy (see models Lin-LinQuadCmeanT=1) in Section 2.2. Posterior summaries displayed as: $_{\text{Lo}}\text{Median}_{\text{Up}}$ (Louis and Zeger 2009). Restricted to 426 couples with data on semen quality who did not exit in the enrollment cycle.

Model	Parameter	$_{\text{Lo}}\text{Median}_{\text{Up}}$
Lin	Menstrual cycle length	$_{0.17}0.34_{0.55}$
LinQuad	Menstrual cycle length	$_{0.14}0.46_{0.77}$
	Menstrual cycle length (squared)	$_{-0.62}-0.35_{-0.17}$
LinT=1	Menstrual cycle length, Normal cycle	$_{0.09}0.25_{0.43}$
LinQuadT=1	Menstrual cycle length, Normal cycle	$_{0.05}0.40_{0.72}$
	Menstrual cycle length, Normal cycle (squared)	$_{-0.70}-0.45_{-0.16}$
LinCmean	Conditional mean cycle length	$_{0.05}0.21_{0.38}$
LinQuadCmean	Conditional mean cycle length	$_{0.04}0.25_{0.45}$
	Conditional mean cycle length (squared)	$_{-0.28}-0.10_{0.05}$
LinCmeanT=1	Conditional mean cycle length, Normal cycle	$_{0.06}0.21_{0.38}$
LinQuadCmeanT=1	Conditional mean cycle length, Normal cycle	$_{0.02}0.23_{0.44}$
	Conditional mean cycle length, Normal cycle (squared)	$_{-0.27}-0.10_{0.03}$

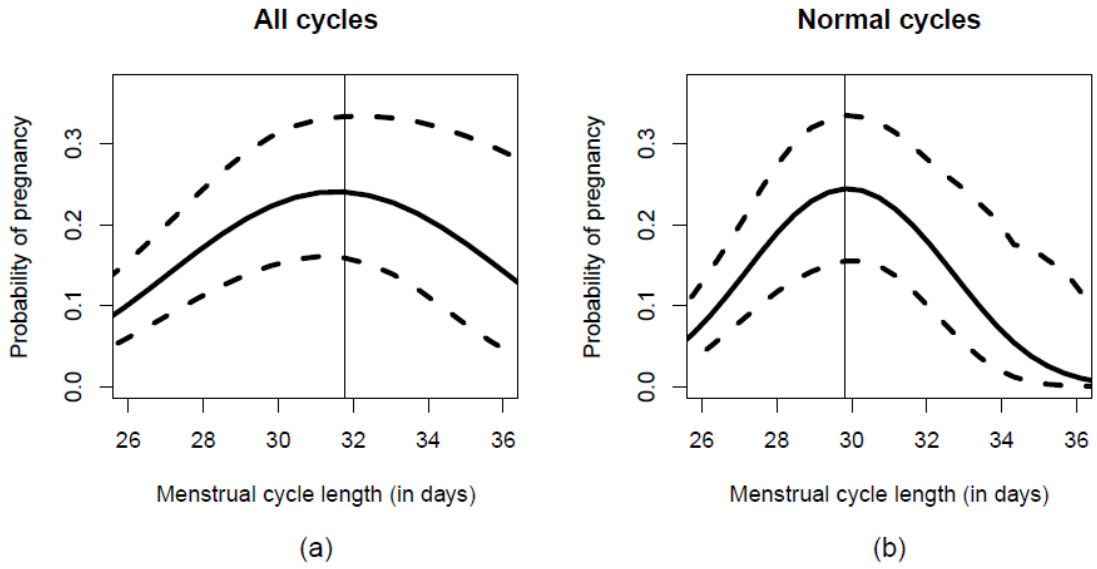

Figure 4: Median (solid) and 95% equal tail credible interval (dashed) for the unconditional probability of pregnancy due to intercourse on the day before ovulation versus menstrual cycle length (in days) for (a) all cycles (b) conditional on normal cycles.

estimate the unconditional probability of pregnancy as a function of menstrual cycle length and find the optimal cycle length. As an example, for a single intercourse act on the day before ovulation, we see that the optimal cycle length is approximately 31.8 days (see Figure 4a). In a separate model, we considered the relation between fecundity and menstrual cycle length conditional on cycles from the normal group (models LinT=1 and LinQuadT=1). For a single intercourse act on the day before ovulation, we found a similar quadratic relation with an optimal cycle length of 29.8 days (see Figure 4b). Lastly, we fit separate models to assess the relation between conditional mean menstrual cycle length and day-specific probability of pregnancy (models LinCmean-LinQuadCmeanT=1). For these models, the coefficient on the quadratic term was much smaller and was not significant (see bottom half of Table 2).

To further visualize the relation between menstrual cycle length and fecundity, we

plot the day specific probability of pregnancy for the 25th, 50th, and 75th percentiles of menstrual cycle length for all cycles and conditional on normal cycles (see Figure 5a-b). The day specific probability of pregnancy is the unconditional probability of pregnancy due to intercourse on a single day, and the values shown in Figure 5 are the medians of the posterior distributions. The x-axis shows the difference in time between the intercourse day and the ovulation day. In both plots, the profile for probability of pregnancy is non-zero starting approximately seven days before ovulation, peaks the day prior to ovulation and then drops sharply to zero one day post ovulation. The probability of pregnancy is lower for the women in the 25th percentile of menstrual cycle length.

Our finding of peak probability of pregnancy corresponding to a menstrual cycle length of approximately 31.8 days agrees with the finding of Small et al. (2006) using data from the Mount Sinai Study of Women Office Workers. In this study of both pregnancy planners and nonplanners, menstrual cycle length was observed using daily diaries and then categorized into 5 categories. To compare our results, we fit a separate joint model in which menstrual cycle length is incorporated in the fecundity model as a categorical variable using cutpoints as in Small et al. (2006). As shown in Figure 5c, the group with menstrual cycle length of 30-31 days has the highest day specific probability of pregnancy. This plot also illustrates the difficulty in choosing a priori cutpoints of cycle length and then using these cutpoints to transform menstrual cycle length into a categorical covariate as is frequently done in studies of menstrual cycle length and pregnancy.

To assess the association between menstrual cycle length and the probability of pregnancy in the context of both male and female risk factors, we incorporated covariates in the fecundity model for sperm morphology (strict criteria), total sperm

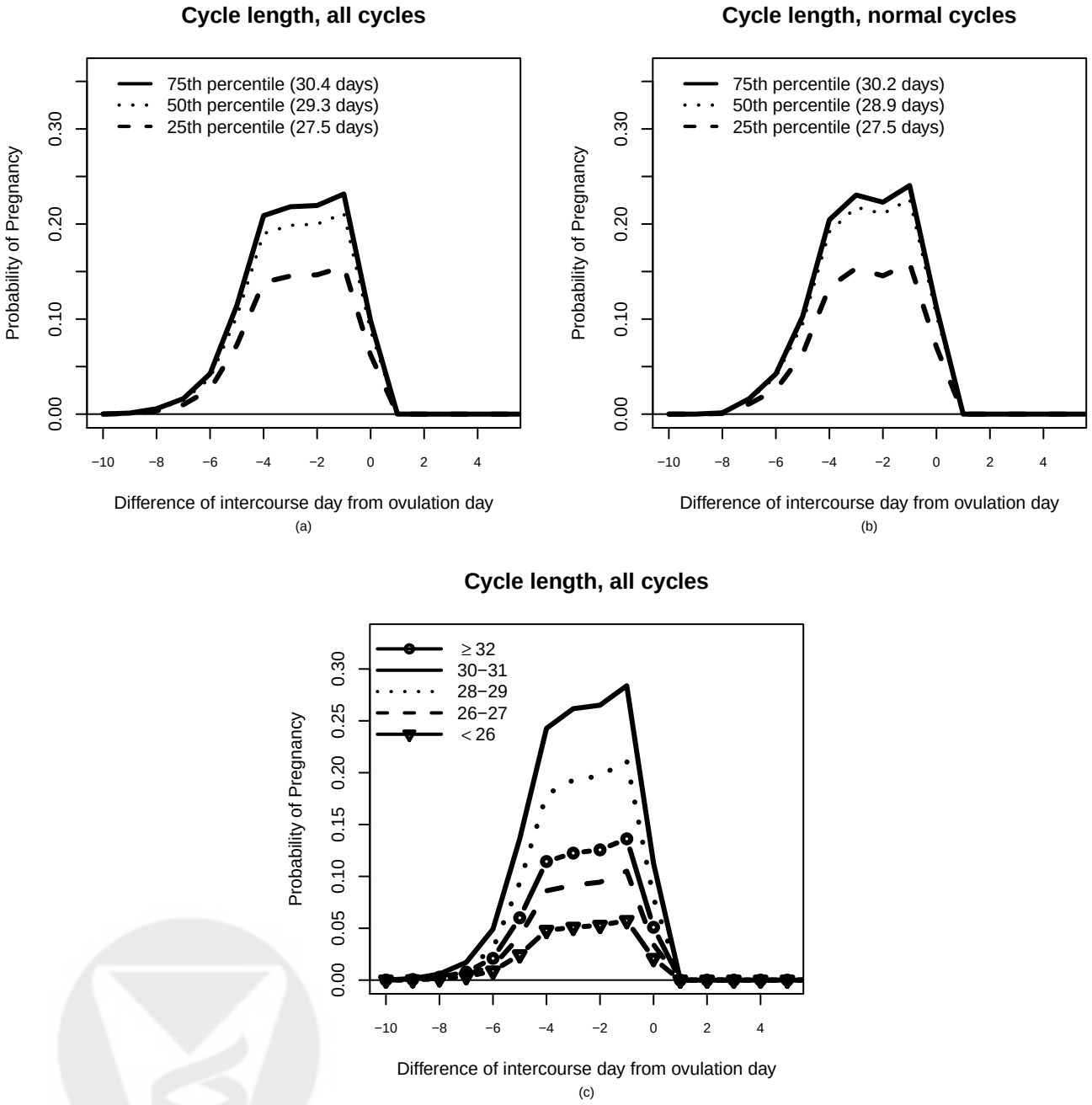

Figure 5: Percentiles of menstrual cycle length and unconditional probability of pregnancy due to a single act of intercourse by difference of intercourse day from ovulation day for (a) all cycles (b) conditional on normal cycles (c) all cycles with cycle length (in days) categorized as in Small et al. (2006)

count, semen volume, sperm concentration, mean of male and female age, difference between male and female age, male smoking status (exposed/unexposed), and female smoking status (exposed/unexposed). We first fit separate models for each covariate to determine each one's unadjusted estimated association with the probability of pregnancy. As shown in the 'Unadjusted' column in Table 3, each of the semen quality parameters had a positive association with pregnancy, while the mean couple age and smoking covariates had a negative association. Next, we fit the joint model adjusting for multiple covariates simultaneously in the fecundity model. Since total sperm count is the product of semen volume and sperm concentration, we fit two separate models (A and B), where model A adjusts for total sperm count and model B adjusts for semen volume and sperm concentration. In both adjusted models, the quadratic relation between female-specific menstrual cycle length and day-specific probability of pregnancy remained significant (see Table 3). For each of the semen quality parameters, we found a positive association with the probability of pregnancy though the coefficient was only significant for sperm morphology (strict criteria). Figure 6 illustrates the positive relation between sperm morphology and the day specific probability of pregnancy using the 25th, 50th, and 75th percentiles of morphology. We also found that the probability of pregnancy decreased with male and female smoking status and with mean age and difference in ages; however, none of these were significant in the adjusted models. Lastly, for the enrollment cycles, we found a quadratic relation between probability of pregnancy and average intercourse frequency (see Figure 7) with a peak at 6 days of intercourse, adjusted for cycle length, total sperm count, mean of male and female age, difference between male and female age, female smoking status and male smoking status.

Table 3: Posterior median and 95% equal-tail CI for pregnancy parameters of joint model displayed as: $_{\text{Lo}}\text{Median}_{\text{Up}}$ (Louis and Zeger 2009). Adjusted models A and B correspond to two combinations of uncorrelated semen quality parameters. Smoking status (exposed) is defined as baseline cotinine level $\geq$ 10ng/mL; and all other risk factors are standardized. Restricted to 426 couples with data on semen quality who did not exit in the enrollment cycle.

Parameters	Unadjusted	Model A	Model B
Menstrual cycle length	$_{0.14}0.46_{0.77}$	$_{-0.06}0.26_{0.59}$	$_{-0.05}0.27_{0.58}$
Menstrual cycle length (squared)	$_{-0.62}-0.35_{-0.17}$	$_{-0.52}-0.27_{-0.11}$	$_{-0.52}-0.27_{-0.12}$
Morphology (strict criteria)	$_{0.16}0.28_{0.40}$	$_{0.06}0.23_{0.41}$	$_{0.05}0.23_{0.40}$
Total sperm count (log)	$_{0.00}0.12_{0.25}$	$_{-0.09}0.08_{0.25}$	
Semen volume (log)	$_{-0.07}0.05_{0.17}$		$_{-0.13}0.03_{0.20}$
Sperm concentration (log)	$_{-0.02}0.11_{0.23}$		$_{-0.09}0.09_{0.27}$
Male smoke (exposed/unexposed)	$_{-0.76}-0.38_{-0.02}$	$_{-0.86}-0.37_{0.13}$	$_{-0.87}-0.37_{0.13}$
Female smoke (exposed/unexposed)	$_{-1.10}-0.37_{0.24}$	$_{-1.01}-0.13_{0.74}$	$_{-1.05}-0.15_{0.73}$
Male and female age (mean)	$_{-0.36}-0.24_{-0.11}$	$_{-0.34}-0.16_{0.02}$	$_{-0.35}-0.17_{0.01}$
Male and female age (difference)	$_{-0.12}0.01_{0.14}$	$_{-0.19}-0.02_{0.14}$	$_{-0.19}-0.03_{0.14}$
Intercourse frequency	$_{8.00}9.36_{9.94}$	$_{8.08}9.46_{9.96}$	$_{7.95}9.35_{9.96}$
Intercourse frequency (squared)	$_{-0.87}-0.79_{-0.66}$	$_{-0.87}-0.80_{-0.67}$	$_{-0.87}-0.79_{-0.66}$

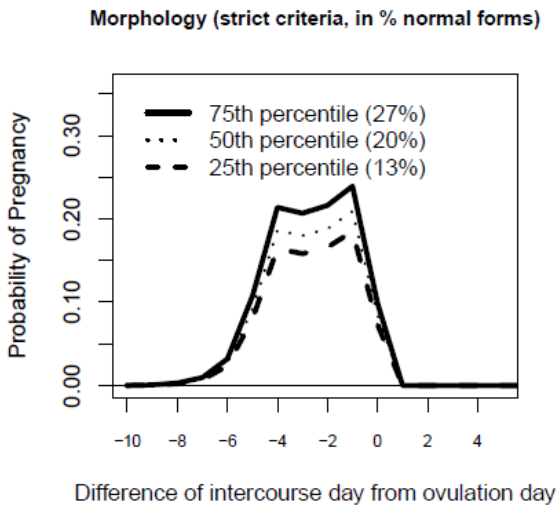

Figure 6: Percentiles of morphology (strict criteria, in % normal forms) and unconditional probability of pregnancy due to a single act of intercourse by difference of intercourse day from ovulation day. Estimates are adjusted for cycle length, total sperm count, mean of male and female age, difference between male and female age, female smoking status (exposed/unexposed), and male smoking status (exposed/unexposed).

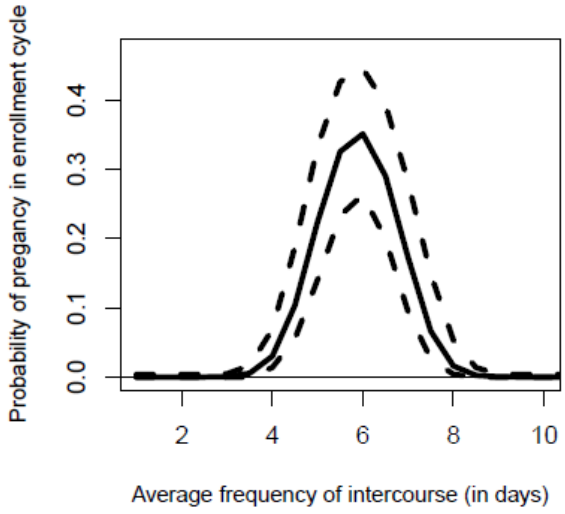

Figure 7: Median (solid) and 95% equal tail credible interval (dashed) for the probability of pregnancy in the enrollment cycle versus average number of days of intercourse. Estimates are adjusted for menstrual cycle length, total sperm count, sperm morphology (strict criteria), mean of male and female age, difference between male and female age, female smoking status (exposed/unexposed), and male smoking status (exposed/unexposed).

5. DISCUSSION

We employed a Bayesian hierarchical modeling approach for assessing the relation between female menstrual cycle length and couple fecundity, while accounting for both male and female risk factors for pregnancy. In the menstrual cycle length sub-model, we addressed several challenges including length-biased sampling of enrollment cycles, unobserved length of pregnancy cycles, and skewness of the underlying distribution. In the fecundity sub-model, we proposed a model for the probability of pregnancy conditional on no pregnancy in previous cycles of attempts. We adjusted for the couples intercourse behavior in a cycle using a flexible spline that is a function of the difference of day of intercourse from the ovulation day and gives weight to a broad window of days

of intercourse in a cycle. We also incorporated a model for the probability of pregnancy in the enrollment cycle as a function of the average frequency of intercourse.

A key aspect of our approach is the inclusion of regression models in both the mean cycle length and day-specific probability of pregnancy models. For the LIFE Study population, we found that mean cycle length decreases with increasing age and that there is a quadratic relation between menstrual cycle length and the couple's probability of pregnancy, with an optimal cycle length of about 31.8 days. These findings are consistent with others in the literature; therefore, we have developed the framework for assessing the relation between exposures and both menstrual cycle length and fecundity. The model can be extended by incorporating baseline exposure variables to investigate potential associations between environmental chemicals, menstrual cycle length, and fecundity.

It would also be of interest to develop a joint model of the three phases of the female menstrual cycle (bleeding, proliferative, and secretory) and fecundity. The length of the proliferative phase of the menstrual cycle is the time from the first day without bleeding to the ovulation day. The range of days on which the female uses the fertility monitor to test for ovulation is limited to at most 20 (e.g., day 6 through 26 of the cycle); therefore, there is the additional challenge of interval censoring. Using the predictive distribution, one could predict the unknown ovulation day as an alternative to imputation of the average ovulation day. The length of the secretory phase is the time from ovulation day to the next bleeding event. One could develop a model to study the variability of both the proliferative and secretory phases and the relation between these variabilities and fecundity.

Finally an extension of this work that focuses on the fecundity model is to incorporate an interaction between the spline function and menstrual cycle length. The trend in modeling fecundity has been to allow for flexibility in the length of the fertile

window, i.e. the window of days around ovulation for which the probability of pregnancy is nonzero. In the proposed model, we estimated the day-specific probability of pregnancy for a broad window of days in the female menstrual cycle and found that the probability of pregnancy consistently declines to zero just one day post ovulation, while pre-ovulation, we found more variation in the number of days with a non-zero probability of pregnancyacross female menstrual cycle length. Currently, we have incorporated menstrual cycle length in the fecundity model as a link-additive term, which allows the profile of day-specific probabilities to shift up or down. By incorporating an interaction with the spline function, we would allow for changes in the shape of the profile of probabilities.

References

Bortot, P., Masarotto, G., and Scarpa, B. "Sequential Predictions of Menstrual Cycle Lengths." *Biostatistics*, 11(4):741–755 (2010).

Brown, E. R., Ibrahim, J. G., and DeGruttola, V. "A Flexible B-Spline Model for Multiple Longitudinal Biomarkers and Survival." *Biometrics*, 61(1):64–73 (2005).

Buck Louis, G. M. "Fecundity and Fertility." In Buck Louis, G. M. and Platt, R. W. (eds.), *Reproductive and Perinatal Epidemiology*. New York: Oxford University Press. (2011).

Buck Louis, G. M., Schisterman, E. F., Sweeney, A. M., Wilcosky, T. C., Gore-Langton, R. E., Lynch, C. D., Boyd Barr, D., Schrader, S. M., Kim, S., Chen, Z., Sundaram, R., and on behalf of the LIFE Study. "Designing Prospective Cohort Studies for Assessing Reproductive and Developmental Toxicity During Sensitive Windows of Human Reproduction and Development - the LIFE Study." *Paediatric and Perinatal Epidemiology*, 25:413–424 (2011).

Buck Louis, G. M., Sundaram, R., Schisterman, E. F., Sweeney, A., Lynch, C. D., Kim, S., Maisog, J. M., Gore-Langton, R., Eisenberg, M. L., and Chen, Z. "Semen Quality and Time to Pregnancy: the Longitudinal Investigation of Fertility and the Environment Study." *Fertility and Sterility*, 101(2):453–462 (2014).

Carlin, B. P. and Louis, T. A. *Bayesian Methods for Data Analysis, 3rd edition*. Boca Raton, FL: Chapman and Hall/CRC Press, 3rd edition (2009).

Cooper, T. G., Noonan, E., von Eckardstein, S., Auger, J., Gordon Baker, H. W., Behre, H. M., Haugen, T. B., Kruger, T., Wang, C., Mbizvo, M. T., and Vogelsong,

K. M. "World Health Organization Reference Values for Human Semen Characteristics." *Human Reproduction Update*, 16(3):231–245 (2010).

Diggle, P., Heagerty, P., Liang, K.-Y., and Zeger, S. *Analysis of Longitudinal Data*. New York: Oxford University Press (2002).

Dominik, R. and Chen, P.-L. "Day-specific Pregnancy Probability Estimation in Barrier Contraceptive Effectiveness Trials." *Paediatric and Perinatal Epidemiology*, 20(s1):38–42 (2006).

Dominik, R., Zhou, H., and Cai, J. "A Statistical Model for the Evaluation of Barrier Contraceptive Efficacy." *Statistics in Medicine*, 20(21):3279–3294 (2001).

Dunson, D. B. and Stanford, J. B. "Bayesian Inferences on Predictors of Conception Probabilities." *Biometrics*, 61(1):126–133 (2005).

Ecochard, R. "Heterogeneity in Fecundability Studies: Issues and Modelling." *Statistical Methods in Medical Research*, 15(2):141–160 (2006).

Gelman, A. "Prior Distributions for Variance Parameters in Hierarchical Models (comment on article by Browne and Draper)." *Bayesian Analysis*, 1(3):515–534 (2006).

Guo, Y., Manatunga, A. K., Chen, S., and Marcus, M. "Modeling Menstrual Cycle Length Using a Mixture Distribution." *Biostatistics*, 7:100–114 (2006).

Harlow, S. D., Lin, X., and Ho, M. J. "Analysis of Menstrual Diary Data Across the Reproductive Life Span Applicability of the Bipartite Model Approach and the Importance of Within-woman Variance." *Journal of Clinical Epidemiology*, 53:722–733 (2000).

Harlow, S. D. and Zeger, S. L. "An Application of Longitudinal Methods to the Analysis of Menstrual Diary Data." *Journal of Clinical Epidemiology*, 44(10):1015 – 1025 (1991).

Huang, X., Elliott, M. R., and Harlow, S. D. "Modelling menstrual cycle length and variability at the approach of menopause by using hierarchical change point models." *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 63(3):445–466 (2014).

Kalbfleisch, J. D. and Prentice, R. L. *The Statistical Analysis of Failure Time Data*. New Jersey: John Wiley & Sons, 2nd edition (2011).

Keulers, M., Hamilton, C., Franx, A., Evers, J., and Bots, R. "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 (2007).

Laird, N. M. and Ware, J. H. "Random-effects Models for Longitudinal Data." *Biometrics*, 963–974 (1982).

Lin, X., Raz, J., and Harlow, S. D. "Linear Mixed Models with Heterogeneous Within-Cluster Variances." *Biometrics*, 910–923 (1997).

Louis, T. A. and Zeger, S. L. "Effective Communication of Standard Errors and Confidence Intervals." *Biostatistics*, 10:1–2 (2009).

Lum, K. J., Sundaram, R., and Louis, T. A. "Accounting for Length-bias and Selection Effects in Estimating the Distribution of Menstrual Cycle Length." *Biostatistics* (2014).

Lynch, C. D., Jackson, L. W., Louis, B., and Germaine, M. "Estimation of the Day-specific Probabilities of Conception: Current State of the Knowledge and the

Relevance for Epidemiological Research." *Paediatric and Perinatal Epidemiology*, 20(s1):3–12 (2006).

McLain, A. C., Lum, K. J., and Sundaram, R. "A Joint Mixed Effects Dispersion Model for Menstrual Cycle Length and Time-to-pregnancy." *Biometrics*, 68:648–656 (2012).

Rothmann, S. A., Bort, A.-M., Quigley, J., and Pillow, R. "Sperm Morphology Classification: a Rational Method for Schemes Adopted by the World Health Organization." In Carrell, D. T. and Aston, K. L. (eds.), *Spermatogenesis: Methods and Protocols*, 27–38. New York: Humana Press (2013).

Royston, P. and Ferreira, A. "A New Approach to Modeling Daily Probabilities of Conception." *Biometrics*, 55(4):1005–1013 (1999).

Small, C. M., Manatunga, A. K., Klein, M., Feigelson, H. S., Dominguez, C. E., McChesney, R., and Marcus, M. "Menstrual Cycle Characteristics: Associations With Fertility and Spontaneous Abortion." *Epidemiology*, 17(1):52–60 (2006).

Verbeke, G. and Molenberghs, G. *Linear Mixed Models for Longitudinal Data*. New York: Springer (2009).

Weinberg, C. R., Gladen, B. C., and Wilcox, A. J. "Models Relating the Timing of Intercourse to the Probability of Conception and the Sex of the Baby." *Biometrics*, 358–367 (1994).

World Health Organization. *WHO Laboratory Manual for the Examination of Human Semen and Sperm-cervical Mucus Interaction*. Cambridge, United Kingdom: Cambridge University Press, 3rd edition (1992).