Published in final edited form as:

Biometrics. 2016 March ; 72(1): 193–203. doi:10.1111/biom.12379.

A Bayesian Joint Model of Menstrual Cycle Length and Fecundity

Kirsten J. Lum¹, Rajeshwari Sundaram², Germaine M. Buck Louis², and Thomas A. Louis^{3,*}

¹Department of Biostatistics & Epidemiology, University of Pennsylvania, 423 Guardian Drive, Philadelphia, PA 19104, U.S.A

²Division of Intramural Population Health Research, Eunice Kennedy Shriver National Institute of Child Health and Human Development, NIH, DHHS, 6100 Executive Boulevard, Rockville, Maryland 20852, U.S.A

³Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, 615 North Wolfe Street, Baltimore, Maryland 21205, U.S.A.

Summary

Menstrual cycle length (MCL) has been shown to play an important role in couple fecundity, which is the biologic capacity for reproduction irrespective of pregnancy intentions. However, a comprehensive assessment of its role requires a fecundity model that accounts for male and female attributes and the couple's intercourse pattern relative to the ovulation day. To this end, we employ a Bayesian joint model for MCL and pregnancy. MCLs follow a scale multiplied (accelerated) mixture model with Gaussian and Gumbel components; the pregnancy model includes MCL as a covariate and computes the cycle-specific probability of pregnancy in a menstrual cycle conditional on the pattern of intercourse and no previous fertilization. Day-specific fertilization probability is modeled using natural, cubic splines. We analyze data from the Longitudinal Investigation of Fertility and the Environment Study (the LIFE Study), a couple based prospective pregnancy study, and find a statistically significant quadratic relation between fecundity and menstrual cycle length, after adjustment for intercourse pattern and other attributes, including male semen quality, both partner's age, and active smoking status (determined by baseline cotinine level 100ng/mL). We compare results to those produced by a more basic model and show the advantages of a more comprehensive approach.

Keywords

Bayesian modeling; Fecundity modeling; Joint model; Length-bias; Scaled mixture model

*tlouis@jhu.edu.

Supplementary Materials

Web Appendices referenced in Sections 3.3 and 3.4 are at the Biometrics website on the Wiley Online Library.

1. Introduction

Fecundity, the male and female biologic capacity for reproduction irrespective of pregnancy intentions (Buck Louis, 2011) can be quantified in terms of multiple measurements such as pubertal onset and progression, menstrual cycle regularity and length, ovulation, semen quality, and pregnancy. This list emphasizes roles of the male and female partners and the complex factors that contribute to becoming pregnant. However, until recently available information on fecundity was provided primarily by retrospective pregnancy studies or prospective studies with fewer than six months follow-up. Many focused only on the female partner. More recently, The Longitudinal Investigation of Fertility and the Environment (LIFE) Study (Buck Louis et al., 2011) took a couple-based perspective, and collected measurements on both partners with the goal of assessing their association with fecundity-related outcomes. Section 3 provides some details.

Using LIFE Study data, we develop a joint model to investigate the association of the distribution of menstrual cycle length (MCL), time-constant attributes of each partner, and cycle day-specific intercourse indicators with the couple's probability of pregnancy. The beginning of the menstrual cycle is defined as the first day of bleeding followed within one day by at least two additional days bleeding. MCL is then determined as time (in days) from the first day of bleeding to the day preceding menstrual bleeding that initiates the next cycle.

1.1 Modeling MCL

Statistical modeling of longitudinal MCL data has mainly focused on extensions of mixed effects models to account for covariates, within-woman variability, and skewness due to extreme cycle lengths. For example, Harlow and Zeger (1991) categorized cycle lengths as 'standard' and 'non-standard,' developed random effects models for each component, and for the probability of having a non-standard length. Murphy et al. (1995) developed a longitudinal model that accommodates length-bias and censoring, Lin et al. (1997) focused on standard lengths and extended the linear mixed model to allow for within-woman variability (see Diggle et al. (2002); Carlin and Louis (2009); Verbeke and Molenberghs (2009) for details on mixed models). Guo et al. (2006) used a mixture of Gaussian and shifted Weibull distributions to accommodate non-standard cycle lengths and developed a marginal model with covariate effects in the mean and variance. Focusing on prediction of MCL, Bortot et al. (2010) used a state space approach to account for trend, autocorrelation and extremely long or short lengths. To allow for skewness in cycle length McLain et al. (2012) used a mixture of Gaussian and Gumbel error distributions with mixed effects on the mean and variance parameters. Recently, Huang et al. (2014) modeled the association of cycle length with age using a Bayesian change point model and a log-normal error distribution.

Building on this background, we employ a Bayesian scale-multiplied mixture model with Gaussian and Gumbel components. The approach is similar to McLain et al. (2012), but as in an accelerated failure time (AFT) model, our couple-specific random effect multiplies the overall length to account for residual heterogeneity. In addition, we account for length-bias in the enrollment cycle, and censoring of the cycle in which the couple becomes pregnant at the ovulation day.

1.2 Modeling Pregnancy

Statistical modeling of pregnancy attempts has focused on assessing biological 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 that assessed the fertile window around the ovulation day using serial vaginal ultrasound (the gold standard, see Keulers et al., 2007). In contrast, Dominik et al. (2001) incorporated all intercourse acts in a menstrual cycle, modeling the day-specific probability of pregnancy using a piecewise linear spline function of distance of the intercourse act from ovulation day (assumed to be day 14). As detailed in Section 2, we generalize this approach and use natural cubic splines for the day-specific probabilities, with cycle-specific ovulation days determined by a fertility monitor which is described in Section 3. We estimate the association of the distribution of MCL with the probability of pregnancy along with other male or female attributes, including age, smoking status and semen quality. Section 3 reports on analysis of the LIFE Study, and compares results to those produced by a more basic model. Section 4 summarizes; implementation details are in the Appendix (Section 5).

2. Joint Model for MCL and Pregnancy

Couples were followed in the LIFE study from enrollment until pregnancy, exit from the study or one year of trying. Couples who exited the study before one year were censored on the last day of the cycle preceding their exit. Pregnancy was determined by human chorionic gonadotropin >25 mIU/mL measured using Clearblue®Easy pregnancy tests at home on day of expected menses. This structure contains two nested time scales: menstrual cycles and days within each menstrual cycle. Available information includes MCL, cycle-specific ovulation day determined using the Clearblue®Easy fertility monitor (see Section 3), pregnancy day indicator using the Clearblue®Easy pregnancy test, and day-specific intercourse indicators.

2.1 Sub-model for Longitudinal MCL

In the sub-model for MCL, we parameterize the relation between baseline covariates (e.g., age, smoking status) and the distribution of MCL, and build-in within-woman longitudinal correlation. We use a two-component, Gaussian/Gumbel mixture distribution to allow for both extremely short and long cycle lengths and flexible shape. While there is some scientific support for a two-population or multi-population structure, our principal motivation is to use a sufficiently flexible model. Therefore, the Gaussian and Gumbel components should not be interpreted as biology based. The mixture is embedded in a hierarchical, accelerated failure time (AFT) model (Kalbfleisch and Prentice, 2011), adapted for modeling MCL by Lum et al. (2015).

Specifically, for the i^{th} couple ($i = 1, \dots, n$), let j ($j = 1, \dots, n_i$) index menstrual cycles with lengths Y_{ij} (observed until censored at the ovulation day in a pregnancy cycle or at one year) and $\mathbf{Y}_i = (Y_{i1}, \dots, Y_{in_i})$. With $\mu_1 > 0$, $\sigma_1 > 0$, $\theta > 0$, $\zeta > 0$, $\sigma_W > 0$,

$$\begin{aligned} [Y_{ij} | W_i, \mathbf{v}_i^\top \boldsymbol{\eta}, \epsilon_{ij}] &= W_i e^{\mathbf{v}_i^\top \boldsymbol{\eta}} \times \epsilon_{ij} \\ [W_i | \sigma_W] &\sim \text{Gamma}(\text{shape}=1/\sigma_W^2, \text{rate}=1/\sigma_W^2) \end{aligned} \quad (1)$$

$$\begin{aligned} [\epsilon_{ij} | q, \mu_1, \sigma_1, \mu_2, \sigma_2] &\text{ iid } q f_1(\epsilon_{ij} | \mu_1, \sigma_1) + (1-q) f_2(\epsilon_{ij} | \theta, \varsigma) \\ f_1(\epsilon | \mu_1, \sigma_1) &= N(\mu_1, \sigma_1) \\ f_2(\epsilon | \theta, \varsigma) &= \text{Gumbel}(\theta, \varsigma) = \frac{1}{\varsigma} \exp\left\{-\frac{\epsilon-\theta}{\varsigma} - e^{-\left(\frac{\epsilon-\theta}{\varsigma}\right)}\right\}, y > 0. \end{aligned} \quad (2)$$

In (1), $\mathbf{v}_i$ is an r -dimensional vector of observed covariates (the intercept is excluded to avoid over-determination) with parameter vector $\boldsymbol{\eta}$. W_i is a couple-specific, latent random effect with mean 1.0 and standard deviation σ_W . We parameterize the Gumbel (equation 2) by location (θ) and scale (ς) which generate $E(\epsilon) = \theta + e \cdot \varsigma$, $SD(\epsilon) = \pi\varsigma / \sqrt{6}$, with “ e ” Euler’s constant. We denote the mean and standard deviation of the Gumbel by μ_2 and σ_2 , respectively, and report estimates of these in Table 1. We assume the distribution of the covariates does not depend on $(q, \mu_1, \sigma_1, \mu_2, \sigma_2, \sigma_W)$ and that cycle lengths are conditionally independent given $\mathbf{v}_i$ and $W_i = w_i$. Note that because one of the mixture components is Gaussian, there is a non-zero probability that a cycle length is negative. However, using the parameter estimates just for the Gaussian component (see Table 1), the probability is virtually zero.

Let $F_\epsilon(\epsilon; q, \mu_1, \sigma_1^2, \mu_2, \sigma_2^2)$, denote the cumulative distribution for the Gaussian/Gumbel mixture. Then, the CDF of Y_{ij} is $F_\epsilon(y_{ij} w_i^{-1} e^{-\mathbf{v}_i^\top \boldsymbol{\eta}} | q, \mu_1, \sigma_1^2, \mu_2, \sigma_2^2, \boldsymbol{\eta})$ with density,

$$f(y_{ij} | w_i, \mathbf{v}_i; q, \mu_1, \sigma_1^2, \mu_2, \sigma_2^2, \boldsymbol{\eta}) = w_i^{-1} e^{-\mathbf{v}_i^\top \boldsymbol{\eta}} f_\epsilon(y_{ij} w_i^{-1} e^{-\mathbf{v}_i^\top \boldsymbol{\eta}} | q, \mu_1, \sigma_1^2, \mu_2, \sigma_2^2, \boldsymbol{\eta}).$$

2.1.1 Priors for the MCL sub-model—Let $\boldsymbol{\phi}_Y = (\boldsymbol{\eta}, \sigma_W, q, \mu_1, \sigma_1, \mu_2, \sigma_2)$ denote the unknown population parameters of the MCL model. We assume the components of $\boldsymbol{\phi}_Y$ are independent *a priori*, with $[\boldsymbol{\phi}_Y] = [\sigma_W][q][\mu_1][\sigma_1][\mu_2][\sigma_2] \prod_{r=1}^r [\eta_r]$, and use the following uniform ($\mathcal{U}$) priors:

$$\begin{aligned} [\eta_r | a_{\eta_r}, b_{\eta_r}] &\sim \mathcal{U}(a_{\eta_r}, b_{\eta_r}), \quad [\sigma_W] \sim \mathcal{U}(0, b_{\sigma_W}), \quad [q] \sim \mathcal{U}(0, 1), \\ [\mu_1] &\sim \mathcal{U}(a_{\mu_1}, b_{\mu_1}), \quad [\sigma_1] \sim \mathcal{U}(0, b_{\sigma_1}), \quad [\mu_2] \sim \mathcal{U}(a_{\mu_2}, b_{\mu_2}), \quad [\sigma_2] \sim \mathcal{U}(0, b_{\sigma_2}), \end{aligned} \quad (3)$$

with the hyperparameters set for a broad range (vague hyper-priors). For example, in (3) we use a uniform prior for the standard deviation of σ_W^2 rather than an inverse Gamma, to better control prior mass (Gelman, 2006). We discuss alternative choices for the priors in Section 3.3.

2.2 Modeling the length of an enrollment cycle

Couples were enrolled in the LIFE Study on an arbitrary day in the female’s menstrual cycle. Therefore, the distribution of the length of this enrollment cycle is influenced by the sampling plan. In Lum et al. (2015), we focused on the enrollment cycle and developed a

method for estimating the population MCL distribution that accounts for length-bias and the possible selection effect induced by a couple's decision to or not to enroll in the study depending on 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 only for length-bias in modeling the enrollment cycles, Y_{i1} . Specifically,

$$f_{Y_1}(y_{i1} | w_i, \mathbf{v}_i; q, \mu_1, \sigma_1^2, \mu_2, \sigma_2^2, \eta) = \frac{y_{i1} f(y_{i1} | w_i, \mathbf{v}_i; q, \mu_1, \sigma_1^2, \mu_2, \sigma_2^2, \eta)}{E(Y_i | W_i, e^{\mathbf{v}_i^T \eta}, q, \mu_1, \mu_2)} \quad (4)$$

where

$$E(Y_i | W_i, e^{\mathbf{v}_i^T \eta}, q, \mu_1, \mu_2) = W_i e^{\mathbf{v}_i^T \eta} \{q\mu_1 + (1 - q)\mu_2\}.$$

2.3 Modeling the length of a pregnancy cycle

During the pregnancy cycle (the cycle in which pregnancy was detected using the Clearblue®Easy pregnancy test), in preparation for implantation of the blastocyst, levels of progesterone increase and preclude menstrual bleeding. Therefore, the cycle length is right-censored in that the (counterfactual) pregnancy-free cycle is not observed. We let τ_{in_i} be the time (in days) from the first day of the n_i^{th} cycle to censoring at the ovulation day (O_{in_i} see Section 2.4) as determined by the fertility monitor. With this censoring time $\tau_{in_i} = O_{in_i}$ but we retain the general notation. Pregnancy may have occurred a day or two later, but we follow current practice and note that τ_{in_i} is to a degree behavior and technology dependent. We assume that conditional on W_i and $\mathbf{v}_i$ the censoring distribution is non-informative and independent of the potential time Y_{in_i} . Suppressing parameters, the contribution to the

likelihood is $\frac{1 - F(\tau_{in_i})}{E(Y_i)}$ for the enrollment cycle (see equation 4) and $\{1 - F(\tau_{in_i})\}$ for subsequent cycles.

2.4 Sub-model for the Probability of Pregnancy

The LIFE Study records cycle-specific pregnancy indicators and within-cycle, day-specific intercourse indicators. Using these and other covariates, we estimate the association between the woman-specific distribution of MCL and the probability of pregnancy. To this end, let O_{ij} indicate the ovulation day, determined using the Clearblue®Easy fertility monitor as described in Section 3, and $d_{ijk} = (k - O_{ij})$, the difference between day k and the ovulation day. Let $x_{ijk} \in \{0, 1\}$ be the day-specific intercourse indicator, and $v_{ijk} \in \{0, 1\}$ denote the (usually unobserved) indicator of fertilization on day k of the j^{th} cycle, all for $(k=1, \dots, \ddot{Y}_{ij})$ with $\ddot{Y}_{ij}$ the greatest integer function. Subsequent to fertilization, intercourse is moot and so $\Pr(v_{ijk} = 1 | \text{some } v_{ij\ell} = 1, \ell < k)$. Of course, $v_{ijk} = 0$ if $x_{ijk} = 0$ (no intercourse). Furthermore, let A_{ij} be the cycle-specific pregnancy indicator with $A_{ij} \equiv 0, j < n_i$ and $A_{in_i} \in \{0, 1\}$ the couple-specific pregnancy indicator. We use,

$$\begin{aligned}
\mathbf{d}_{ij} &= (d_{ij1}, \dots, d_{ijY_{ij}}); \quad \mathbf{d}_i = (\mathbf{d}_{i1}, \dots, \mathbf{d}_{in_i}) \\
\mathbf{x}_{ij} &= (x_{ij1}, \dots, x_{ijY_{ij}}); \quad \mathbf{x}_i = (\mathbf{x}_{i1}, \dots, \mathbf{x}_{in_i}) \\
\nu_{ij} &= (\nu_{ij1}, \dots, \nu_{ijY_{ij}}); \quad \nu_i = (\nu_{i1}, \dots, \nu_{in_i}) \\
\mathcal{A}_{ij} &:= A_{i1} = \dots = A_{i(j-1)} = 0 \\
\mathcal{C}_{ijk} &:= \nu_{i1} = \nu_{i2} = \dots = \nu_{i(j-1)} = 0 = (\nu_{ij1}, \dots, \nu_{ij(k-1)})
\end{aligned} \tag{5}$$

The probability of pregnancy by intercourse on the k^{th} day of the cycle is a function of male and female attributes denoted by $\mathbf{z}_i$, d_{ijk} , and the “typical” (but unobserved) woman-specific MCL from equation 1. Personalization is via the latent variable

$Y_i^* \sim [W_i e^{\mathbf{v}_i^T \eta} \times \epsilon_{ij} \mid \mathbf{Y}_i, \phi_Y, W_i, \mathbf{v}_i]$, and model development is based on its posterior predictive distribution, $h_f(\cdot) = h_f(\cdot \mid \mathbf{Y}_i, \mathbf{v}_i, \{\text{all other data}\})$.

$$\begin{aligned}
\rho_{ijk}(Y_i^*, \mathbf{z}_i, d_{ijk}, \mathcal{C}_{ijk}) &= \Pr(\nu_{ijk}=1 \mid Y_i^*, \mathbf{z}_i, d_{ijk}, x_{ijk}=1, \mathcal{C}_{ijk}), \\
\Pr(A_{ij}=1 \mid \mathcal{A}_{ij}, Y_i^*) &= 1 - \prod_{\ell=1}^{Y_{ij}} \{1 - \rho_{ij\ell}(Y_i^*, \mathbf{z}_i, d_{ij\ell}, \mathcal{C}_{ij\ell})\}^{x_{ij\ell}} \tag{6}
\end{aligned}$$

$$\begin{aligned}
\text{logit}\{\rho_{ijk}(Y_i^*, \mathbf{z}_i, d_{ijk}, \mathcal{C}_{ijk})\} &= \alpha_0 + \mathbf{m}_i(Y_i^*)^T \beta + \mathbf{z}_i^T \gamma + g(d_{ijk}), \\
\Pr(A_{ij}=1 \mid \mathcal{A}_{ij}, h_i) &= 1 - \int \prod_{\ell=1}^{Y_{ij}} \{1 - \rho_{ij\ell}(s, \mathbf{z}_i, d_{ij\ell}, \mathcal{C}_{ij\ell})\}^{x_{ij\ell}} h_i(s) ds \tag{7}
\end{aligned}$$

Formula 6 is equivalent to $\Pr(\nu_{ij1} = 1) + \Pr(\nu_{ij2} = 1 \mid \nu_{ij1} = 0) + \dots$

In the logit model $\mathbf{m}_i(Y_i^*)$ is a vector with linear and quadratic functions of Y_i^* and

$g(\cdot) = \sum_{l=1}^L \alpha_l B_l(\cdot)$, where $\{B_1(\cdot), \dots, B_L(\cdot)\}$ are the B-spline basis functions for a natural cubic spline (α_0 serves as the intercept). Equation 7 mixes using $h_f(\cdot)$, thereby linking the MCL and pregnancy sub-models, with the random effect W_i enabling sharing of information in both directions. The B-spline generalizes previous specifications that include a piecewise linear spline (Dominik et al., 2001), a quadratic function centered around ovulation day (Dominik and Chen, 2006), and a day-specific parameter for a narrow window of days around ovulation (Dunson and Stanford, 2005). Scarpa and Dunson (2007) develop a flexible approach to determining the window.

2.4.1 Model forms for $\mathbf{m}(Y_i^*)$ —In Section 3.4, we present results from the four following specifications for including the woman-specific MCL in the model for $\rho_{ijk}(\cdot)$. For each approach we consider a linear relation and a (linear + quadratic) relation,

$$\begin{aligned}
m_{Lin}(Y_i^*) &= \beta_1 Y_i^* \\
m_{LinQuad}(Y_i^*) &= \beta_1 Y_i^* + \beta_2 (Y_i^*)^2 \tag{8}
\end{aligned}$$

In the following, we use the latent variable $T_{ij} = 1$ or 2 to indicate that a cycle length is from the Gaussian distribution (group 1) or the Gumbel distribution (group 2) in the Gaussian/Gumbel mixture (see equation 1). And, “mixing” is over the posterior distribution of W_i and model parameters.

M1 (Woman-specific MCL): Mix equations (8) over $[Y_i^* | \mathbf{v}_i, \{\text{all data}\}]$.

M2 (Woman-specific MCL conditional on $T_{ij} \equiv 1$, so all cycles come from the Gaussian distribution): Mix equations (8) over $[Y_i^* | \mathbf{v}_i, T_{ij} \equiv 1, \{\text{all data}\}]$.

M3 (Woman-specific expected MCL): Substitute $E(Y_i^* | \mathbf{v}_i, \{\text{all data}\})$ for Y_i^* in equations (8).

M4 (Woman-specific mean MCL conditional on $T_{ij} \equiv 1$, so all cycles come from the Gaussian distribution): Substitute $E(Y_i^* | \mathbf{v}_i, T_{ij} \equiv 1, \{\text{all data}\})$ for Y_i^* in equations (8).

2.4.2 Modeling the probability of pregnancy in the enrollment cycle—The ovulation day is estimated based on day-level data, and since couples were allowed to enroll on any day of the menstrual cycle, the majority of the couples do not have sufficient monitor data for the enrollment cycle to estimate the ovulation day. Furthermore, couples who become pregnant before the first observed bleeding event (approximately 10%), the enrollment cycle is the only one observed and that only partially. 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 (u_i) and MCL,

$$\text{logit} \{P_{\text{enroll}}(A_{i1}=1; \mathbf{m}_i(Y_i^*), \beta, \gamma, \boldsymbol{\lambda}, \mathbf{z}_i)\} = \lambda_0 + \mathbf{m}_i(Y_i^*)^\top \beta + \mathbf{z}_i^\top \gamma + \lambda_1 u_i + \lambda_2 u_i^2. \quad (9)$$

We use λ_0 rather than α_0 to adjust for the special nature of the enrollment cycle.

3. Application to the LIFE, Prospective Pregnancy Study

The LIFE Study enrolled 501 couples with eligibility criteria that included married or in a committed relationship, females aged 18-40 and males over age 18, English or Spanish speaking, self-reported MCLs in the interval 21-42 days, and no hormonal birth control injections in the past 12 months. Females recorded daily bleeding on a scale of 0 (none) to 4 (heavy). Pregnancy was identified using Clearblue® Easy pregnancy tests administered at home on the day menstruation is expected. The male and female partners independently kept journals of daily intercourse frequency. If the intercourse question was not answered on a particular day in the female daily journal, the male journal entry was used. If the intercourse question was not answered by either member (approximately 2.9% of days), “no intercourse” was recorded.

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). 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 6 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 monitor. If there were two consecutive peak days, the latter was used as the ovulation day. If a peak was not detected and the woman tested on at least 90% of the 20 testing days, the cycle was categorized as anovulatory (151 cycles) and so had zero risk for pregnancy. If the 90% threshold was not reached, the mean of the woman's observed ovulation days was used (required for 250 cycles). If ovulation was not observed in any cycles, the day with the highest LH peak was used as the ovulation day (required for 33 cycles). The 20 cycles for which the ovulation day could not be detected or imputed due to non-compliance in testing were excluded.

3.1 Analysis dataset and modeling

Of the 501 enrolled couples contributing 2,249 cycles, we excluded 16 couples (3%) because they withdrew in the enrollment cycle and so did not provide intercourse indicators, MCL or pregnancy success. Furthermore, 25 (1%) of cycles were excluded because they were shorter than 9 or longer than 89 days. An additional 59 (12%) couples were excluded because the first semen sample did not provide complete data on the four WHO semen quality parameters, leaving a total of 426 (85%) couples and 1911 (85%) cycles for our analysis.

We use the joint model detailed in Section 2 to connect the woman-specific MCL distribution with the couple's fecundity, as measured by the probability of pregnancy in a menstrual cycle. In the fecundity model, we adjust for male and female active smoking status determined by baseline cotinine level 100ng/mL (see Wall et al., 1988), average of male and female age, difference between male and female age, and four WHO semen quality parameters; semen volume (in *ml*), sperm concentration ($10^6/\text{ml}$), total sperm count (concentration $\times$ volume), and morphology (strict criteria) as discussed in Cooper et al. (2010). These authors define cut-points to categorize semen quality, we use the continuous measurements. Since total sperm count is the product (semen volume) $\times$ (sperm concentration), we fit separate models, one including total sperm count, the other both semen volume and sperm concentration. To deal with the high correlation between the couple-specific ages, we include the average and difference.

Choices of which days of the menstrual cycle to include in the analysis, really a determination on which days is there a non-zero probability of fertilization, range from using only the intercourse act closest to ovulation (Royston and Ferreira, 1999), to using days that fall within predetermined window around the ovulation day (Dominik and Chen, 2006; Weinberg et al., 1994; Dunson and Stanford, 2005). Days were excluded in the following order: (1) days after a positive pregnancy test, (2) days beyond day 45 of the cycle, (3) $d_{ijk} \notin [-16, 18]$, (4) if semen quality data were missing. The window in step (3) were selected so that approximately 5% for candidate days were excluded in going from step 2 to step 3.

3.2 Sensitivity analyses

To assess the sensitivity to the choice of uniform priors for $(\sigma_1, \sigma_2, \sigma_W)$, we also fit the model assuming uniform priors on the scale of the log standard deviations, and found that the corresponding posterior distributions were essentially unchanged. Also, we assumed log-normal distribution for $W_i, i = 1, \dots, n$ rather than a Gamma. Again, the posterior distributions were essentially unchanged.

For the spline model for $g(\cdot)$ 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, and found that DIC favored models with 8, 9, or 10 knots. Based on this, we use 8 knots, as a reasonable bias/variance trade-off.

3.3 Implementation Details

Posterior distributions described in Section 5 are estimated using MCMC sampling using the OpenBUGS software. Convergence of key parameters was checked by visual inspection of trace plots. Based on these trace plots (see Appendix A of the Web Supplement), we use a burn-in of 10000 iterations and a sample of 30000 iterations. While the regression coefficients of the pregnancy sub-model rapidly mix, the spline coefficients and the menstrual cycle sub-model parameters are slower to converge. Model code is presented in Appendix B of the Web Supplement. It takes approximately 52 hours to fit adjusted model A from Table 3 using OpenBUGS version 3.2.2 rev 1063 with a burn-in of 10,000 and then 30,000 iterations. In the two stage approach, the MCL stage takes 1.5 hrs and the Pregnancy stage 12 hours. Hardware specifications are at http://biowulf.nih.gov/user_guide.html#hardware.

3.4 Results

Appendix C of the Web Supplement reports on fit of the MCL sub-model. Table 1 displays the posterior median and 95% equal tail credible interval (CI) for parameters of the AFT model of MCL (equation 1). A negative slope on age or active smoking status indicates reduction in mean cycle length. The statistically significant, negative relation between age and cycle length agrees with previous studies (Harlow et al., 2000). Posterior distributions for parameters of the Gaussian/Gumbel mixture distribution indicate that the prevalence of group 1 cycles (q) is approximately 0.750.790.82, and that the mean and standard deviation of group 1 cycles are smaller than those for group 2. The estimated standard deviation of the random effect distribution (that of W_i) is $\sigma_W \approx 0.09$. The prior mean is 1.0 and so there is considerable relative variation that allows woman-specific updating of the distribution of W_i . If σ_W were near 0 ($W_i \approx 1.0$) there would be little, *a priori* between-woman variability in mean menstrual cycle mean beyond that explained by the fixed effects.

Regarding fecundity, for each couple we estimated the probability of pregnancy in a cycle conditional on no pregnancy in previous cycles, intercourse history, and MCL using the model in equation 6. Figure 1 compares empirical estimates to boxplots of model-estimated conditional probability of pregnancy in a cycle given no pregnancy in previous cycles

(having made it that far). The conditional probability of pregnancy decreases over cycles because each failed cycle induces a successively stochastically smaller distribution W_i .

We estimate day-specific probabilities using the pregnancy models in Section 2.4.1 that include the (latent) female-specific MCL, pregnancy indicators, but no other covariates are presented in Table 2. For example, in model M1 (see below equation 8), if there is a single intercourse act on the day before ovulation, the peak pregnancy probability corresponds to a cycle length of approximately 31.8 days (Figure 2a). Results from model M2 that conditions on group 1 cycles ($T = 1$) identify the peak at a cycle length of 29.8 days (Figure 2b). Models M3 and M4 assess the relation between conditional mean cycle length and pregnancy. The coefficients on $(\text{MCL})^2$ are much smaller in magnitude than those from models M1 and M2 and are not statistically significant (see Table 2).

Figures 3 (a, b) plot the relation between the probability of pregnancy (ρ_{ijk}) and the difference in days between intercourse and ovulation (d_{ijk}) for the 25th, 50th, and 75th percentiles of latent woman-specific MCL (Y_i^*) for all cycles (panel a) and conditional on group 1 cycles (panel 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 3 are medians of the posterior distribution of the day-specific probability of pregnancy. In both plots, the pregnancy probability curves are positive within approximately seven days before ovulation, peak the day prior to ovulation and drop sharply to zero one day post ovulation. The curves are lower for the 25th percentile of MCL.

Our finding of peak probability of pregnancy for a MCL of approximately 31.8 days agrees with Small et al. (2006) using data from the Mount Sinai Study of Women Office Workers. To make the comparison, we fit a model with MCL using their categorization (panel c) and the 30–31 day group has the highest day-specific probabilities. However, we recommend using percentiles of the posterior distribution to form the categories.

We expanded the model by including couple-specific, mean age and age difference, male and female active smoking status (yes/no, according as serum cotinine or $< 100\text{ng/mL}$, Wall et al., 1988), and the four WHO semen quality parameters. We first fit separate models for each covariate (the ‘Unadjusted’ column in Table 3). The quadratic term for latent cycle length is important, morphology and total sperm count are positively associated with pregnancy probability; mean age and each of the active smoking covariates are negatively associated. Next, we fit model M1, adjusting for multiple covariates. Model A adjusts for total sperm count and model B for both semen volume and sperm concentration. In both adjusted models, the quadratic term from latent cycle length remains important with a peak probability of pregnancy corresponding to a cycle length of approximately 31 days (Figure 2c), which is slightly lower than that of the unadjusted model. Of the couple-specific baseline covariates, only sperm morphology (strict criteria) remained statistically supported. Panel d of Figure 3 displays the association with the 25th, 50th, and 75th percentiles of sperm morphology.

Equations 4 and 9 are specific to the enrollment cycle. The need to account for length-bias is illustrated by the right shift in the histogram of enrollment cycle lengths compared to those

for subsequent lengths (see panel a in Figure 4). In the pregnancy model there is a quadratic association with intercourse frequency with a peak at 6 days of intercourse (panel b).

3.5 Comparison to a more basic model

Our model provides information and capabilities not available in a more basic model. For example, a model that doesn't include the couple-specific random effect (W_j) cannot produce Figure 1, and a model that doesn't take advantage of the day-specific intercourse indicators (x_{ijk}) cannot produce the fertilization probability curves in Figure 3. These capabilities are prima-facie advantages using a sufficiently complex model. Moreover, they are key for predicting a new couple's probability of pregnancy in the next cycle based on the couple's available data on MCL, intercourse pattern, baseline risk factors, and the posterior distribution of the observed parameters from the fitted model.

Other benefits of our modeling approach including avoiding the bias in the estimated role of MCL induced by failing to account for day-specific intercourse. Specifically, consider two couples, each with one intercourse act for each of 4 cycles. The first couple always has intercourse on the optimal day ($d = -1$) and the other where the fertilization curve is low ($d = -5$). Further, assume that the MCL for the couple 1 woman is at the 25th and for the couple 2 woman is at the 75th percentile of the population MCL distribution. As can be seen in the upper left panel of Figure 3, the per-cycle probability of fertilization will be basically identical for the two couples, masking the role of MCL.

Lastly, inclusion of a latent, woman-specific MCL in the pregnancy model has other advantages. Commonly, MCL is included as a single covariate such as length of the first cycle post enrollment (Kolstad et al., 1999), or length of the cycle preceding pregnancy (Small et al., 2006). The restriction to a single, fixed MCL value fails to capture the well-known complexities such as large within and between woman heterogeneity, length-bias of the enrollment cycle and censoring at the ovulation day due to pregnancy. Joint models can incorporate these features and other features, generally producing a more relevant depiction of MCL's role in the pregnancy model.

An alternative approach is to fit a two-stage version of the proposed model. In the first stage, the woman-specific MCL model is fit using only MCL data and the posterior distribution is produced. Then the posterior median MCL is used as a covariate in the pregnancy model. Table 4 compares parameter estimates and credible intervals for the joint and two-stage approaches. The estimates and the credible intervals for the coefficients corresponding to the linear and quadratic terms for cycle length differ between the two approaches. The effect of this difference is illustrated in Figure 2 which plots the probability of pregnancy for intercourse on the day $d = -1$ versus MCL. The joint model result (panel c), reports that a woman with a predicted MCL of approximately 31 days has a slightly higher probability of pregnancy than a woman with smaller or larger MCL. In contrast, the two-stage model results (panel d), report a smaller probability of pregnancy and a peak that is slightly lower at approximately 30 days. While a two-stage approach can account for some of the previously mentioned complexities in the first stage, it still relies on a single fixed value in the pregnancy model and therefore does not capture the variation in the woman-specific MCL distribution.

4. Summary

We employed a Bayesian hierarchical, joint model to assess the relation between MCL and couple-specific fecundity, accounting for both male and female attributes. The model accommodates length-bias of the enrollment cycle MCL, censoring of the MCL by pregnancy, accounts for day-specific intercourse, and “delivers” a couple-specific, latent MCL to the pregnancy model. Furthermore, the model supports computing posterior predicted, couple-specific pregnancy probabilities based on current information and future intercourse activity. These and other features, while requiring somewhat complicated modeling, provide valuable insights that are not obtainable from a basic model on the association of fecundity with MCL and other attributes. Section 3.5 documents some aspects of the value added by such complexity, including avoiding the bias induced by not accounting for intercourse pattern and the better resolution of the role of MCL.

A key aspect of our approach is inclusion of regression models for both the cycle length and day-specific probability of pregnancy. For the LIFE Study population, we found that mean cycle length decreases with increasing age and that there is a quadratic relation between MCL and the couple’s probability of pregnancy, with an optimal cycle length of about 31 days.

The model can be extended by incorporating baseline exposure variables to investigate potential associations between environmental chemicals, MCL, and fecundity. Also, additional insight would be provided by a joint model of the three phases of the menstrual cycle (bleeding, proliferative, and secretory) and fecundity. Finally extending the fecundity model to incorporate an interaction between the day-specific spline function and latent MCL would allow the shape of the function to depend on it and have to potential to yield additional insights.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

The authors thank the LIFE Study for data access. This work was supported by the Intramural Research Program of the U.S., National Institutes of Health, *Eunice Kennedy Shriver* National Institute of Child Health and Human Development [Contracts N01-HD-3-3355, N01-HD-3-3356, N01-HD-3-3358]. The authors acknowledge use of the high-performance computational capabilities of the Biowulf Linux cluster at the National Institutes of Health, Bethesda, Maryland (<http://biowulf.nih.gov>).

5. Appendix

Let $\phi_A = (a_0, \beta, \gamma)$ and assume components are independent and independent of ϕ_Y , and use non-informative, uniform priors for each parameter. Let Y_i^o denote the observed MCLs and Y_i^u the unobserved length of the pregnancy cycle. The joint distribution of (Y_i, A_i, Y_i^*, W_i) can be factored as

$$[Y_i, A_i, Y_i^*, W_i | \phi_Y, \phi_A] \\ = [A_i | Y_i^*, z_i, x_i, d_i, \phi_A] [Y_i^* | Y_i, W_i, v_i, \phi_Y] [Y_i | W_i, v_i, \phi_Y] [W_i | \phi_Y],$$

where

$$\begin{aligned} [Y_i | W_i, v_i, \phi_Y] &= [Y_i^o, Y_i^u | W_i, v_i, \phi_Y] \\ &= [Y_i^u | Y_i^o, W_i, v_i, \phi_Y]^{A_{in_i}} [Y_i^o | W_i, v_i, \phi_Y] \\ &= [Y_i^u | Y_i^o, W_i, v_i, \phi_Y]^{A_{in_i}} [Y_{in_i}^o | W_i, v_i, \phi_Y]^{1-A_{in_i}} \prod_{j=1}^{n_i-1} [Y_{ij}^o | W_i, v_i, \phi_Y]. \end{aligned}$$

The unobserved model parameters for couple i are: $(Y_i^u, Y_i^*, W_i, \phi_Y, \phi_A)$; the observed data are $D = (Y_i^o, v_i, A_i, O_i, z_i, x_i, d_i, u_i)$. If $A_{in_i} = 0$ (i.e. no pregnancy), $Y_i^o = Y_i$ (all cycles are observed), and the joint posterior distribution is,

$$\begin{aligned} &[Y_i^*, W_i, \phi_Y, \phi_A | D; A_{in_i}=0] \\ &\propto \left(\prod_{j=1}^{n_i} [Y_{ij}^o | W_i, v_i, \phi_Y] [A_{ij}=0 | \mathcal{A}_{ij}; Y_i^*, x_i, d_i, z_i, \phi_A] \right) \times [Y_i^* | Y_i, W_i, v_i, \phi_Y] [W_i | \phi_Y] [\phi_Y] [\phi_A]. \end{aligned}$$

If $A_{in_i} = 1$ (i.e. pregnant), the contribution to the joint posterior distribution is,

$$\begin{aligned} &[Y_i^u, Y_i^*, W_i, \phi_Y, \phi_A | D; A_{in_i}=1] \\ &\propto \left(\prod_{j=1}^{n_i-1} [Y_{ij}^o | W_i, v_i, \phi_Y] [A_{ij}=0 | \mathcal{A}_{ij}; Y_i^*, x_i, d_i, z_i, \phi_A] \right)^1 \{n_i > 1\} \times [Y_i^u | Y_i^o, W_i, v_i, \phi_Y] [A_{in_i}=1 | \mathcal{A}_{ij}; Y_i^*, x_i, d_i, z_i, \phi_A] > \end{aligned}$$

We estimate the posterior distributions using Markov chain Monte Carlo (MCMC), see Section 3.3 for details.

References

- Bortot P, Masarotto G, Scarpa B. Sequential predictions of menstrual cycle lengths. *Biostatistics*. 2010; 11:741–755. [PubMed: 20400622]
- Brown ER, Ibrahim JG, DeGruttola V. A flexible B-spline model for multiple longitudinal biomarkers and survival. *Biometrics*. 2005; 61:64–73. [PubMed: 15737079]
- Buck Louis, GM. Fecundity and fertility. In: Buck Louis, GM.; Platt, RW., editors. *Reproductive and Perinatal Epidemiology*. Oxford University Press; New York: 2011.
- Buck Louis GM, Schisterman EF, Sweeney AM, Wilcosky TC, Gore-Langton RE, Lynch CD, Barr DB, Schrader SM, Kim S, Chen Z, Sundaram R, 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*. 2011; 25:413–424. [PubMed: 21819423]
- Buck Louis GM, Sundaram R, Schisterman EF, Sweeney A, Lynch CD, Kim S, Maisog JM, Gore-Langton R, Eisenberg ML, Chen Z. Semen quality and time to pregnancy: the Longitudinal Investigation of Fertility and the Environment Study. *Fertility and Sterility*. 2014; 101:453–462. [PubMed: 24239161]
- Carlin, BP.; Louis, TA. *Bayesian Methods for Data Analysis*, 3rd edition. 3rd edition. Chapman and Hall/CRC Press; Boca Raton, FL: 2009.

- Cooper TG, Noonan E, von Eckardstein S, Auger J, Gordon Baker HW, Behre HM, Haugen TB, Kruger T, Wang C, Mbizvo MT, Vogelsong KM. World Health Organization reference values for human semen characteristics. *Human Reproduction Update*. 2010; 16:231–245. [PubMed: 19934213]
- Diggle, P.; Heagerty, P.; Liang, K-Y.; Zeger, S. *Analysis of Longitudinal Data*. Oxford University Press; New York: 2002.
- Dominik R, Chen P-L. Day-specific pregnancy probability estimation in barrier contraceptive effectiveness trials. *Paediatric and perinatal epidemiology*. 2006; 20:38–42. [PubMed: 17061972]
- Dominik R, Zhou H, Cai J. A statistical model for the evaluation of barrier contraceptive efficacy. *Statistics in medicine*. 2001; 20:3279–3294. [PubMed: 11746318]
- Dunson DB, Stanford JB. Bayesian inferences on predictors of conception probabilities. *Biometrics*. 2005; 61:126–133. [PubMed: 15737085]
- Ecohard R. Heterogeneity in fecundability studies: Issues and modelling. *Statistical Methods in Medical Research*. 2006; 15:141–160. [PubMed: 16615654]
- Gelman A. Prior distributions for variance parameters in hierarchical models (comment on article by browne and draper). *Bayesian Analysis*. 2006; 1:515–534.
- Guo Y, Manatunga AK, Chen S, Marcus M. Modeling menstrual cycle length using a mixture distribution. *Biostatistics*. 2006; 7:100–114. [PubMed: 16020617]
- Harlow SD, Lin X, Ho MJ. 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*. 2000; 53:722–733. [PubMed: 10941950]
- Harlow SD, Zeger SL. An application of longitudinal methods to the analysis of menstrual diary data. *Journal of Clinical Epidemiology*. 1991; 44:1015–1025. [PubMed: 1940994]
- Huang X, Elliott MR, Harlow SD. 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)*. 2014; 63:445–466.
- Kalbfleisch, JD.; Prentice, RL. *The Statistical Analysis of Failure Time Data*. 2nd edition. John Wiley & Sons; New Jersey: 2011.
- Keulers M, Hamilton C, Franx A, Evers J, Bots R. The length of the fertile window is associated with the chance of spontaneously conceiving an ongoing pregnancy in subfertile couples. *Human Reproduction*. 2007; 22:1652–1656. [PubMed: 17449509]
- Kolstad HA, Bonde JP, Hjellund NH, Jensen TK, Henriksen TB, Ernst E, Giwercman A, Skakkebaek NE, Olsen J. Menstrual cycle pattern and fertility: a prospective follow-up study of pregnancy and early embryonal loss in 295 couples who were planning their first pregnancy. *Fertility and Sterility*. 1999; 71:490–496. [PubMed: 10065787]
- Lin X, Raz J, Harlow SD. Linear mixed models with heterogeneous within-cluster variances. *Biometrics*. 1997; 53:910–923. [PubMed: 9290222]
- Louis TA, Zeger SL. Effective communication of standard errors and confidence intervals. *Biostatistics*. 2009; 10:1–2. [PubMed: 18550565]
- Lum KJ, Sundaram R, Louis TA. Accounting for length-bias and selection effects in estimating the distribution of menstrual cycle length. *Biostatistics*. 2015; 16:113–128. [PubMed: 25027273]
- McLain AC, Lum KJ, Sundaram R. A joint mixed effects dispersion model for menstrual cycle length and time-to-pregnancy. *Biometrics*. 2012; 68:648–656. [PubMed: 22321128]
- Murphy SA, Bentley GR, O'Hanesian MA. An analysis for menstrual data with time-varying covariates. *Statistics in Medicine*. 1995; 14:1843–1857. [PubMed: 8532979]
- Royston P, Ferreira A. A new approach to modeling daily probabilities of conception. *Biometrics*. 1999; 55:1005–1013. [PubMed: 11315041]
- Scarpa B, Dunson DB. Bayesian methods for searching for optimal rules for timing intercourse to achieve pregnancy. *Statistics in Medicine*. 2007; 26:1920–1936. [PubMed: 17328097]
- Small CM, Manatunga AK, Klein M, Feigelson HS, Dominguez CE, McChesney R, Marcus M. Menstrual Cycle Characteristics: Associations With Fertility and Spontaneous Abortion. *Epidemiology*. 2006; 17:52–60. [PubMed: 16357595]
- Verbeke, G.; Molenberghs, G. *Linear Mixed Models for Longitudinal Data*. Springer; New York: 2009.

- Wall MA, Johnson J, Jacob P, Benowitz NL. Cotinine in the serum, saliva, and urine of nonsmokers, passive smokers, and active smokers. *American Journal of Public Health*. 1988; 78:699–701. [PubMed: 3369604]
- Weinberg CR, Gladen BC, Wilcox AJ. Models relating the timing of intercourse to the probability of conception and the sex of the baby. *Biometrics*. 1994; 50:358–367. [PubMed: 8068836]
- World Health Organization. WHO Laboratory Manual for the Examination of Human Semen and Sperm-cervical Mucus Interaction. 3rd edition. Cambridge University Press; Cambridge, United Kingdom: 1992.

Figure 1.

Empirical estimates (triangles) of the probability of pregnancy by cycle compared with box plots of the model estimated probabilities from a model that is conditional on no pregnancy in previous cycles, intercourse history, and MCL.

Figure 2.

The probability of pregnancy due to intercourse on the day before ovulation; posterior median (—), 95% credible interval (---) versus MCL estimated by the (a) unadjusted joint model (b) unadjusted joint model conditional on group 1 cycles, (c) adjusted joint model, (d) adjusted two-stage model. Adjusted for total sperm count, sperm morphology (strict criteria), mean of male and female age, difference between male and female age, male and female active smoking (yes/no) and intercourse frequency.

Figure 3.

Unconditional probability of pregnancy due to a single act of intercourse by difference of intercourse day from ovulation day for (a) percentiles of MCL (b) percentiles of MCL conditional on group 1 cycles (c) MCL (in days) categorized as in Small et al. (2006) (d) percentiles of sperm morphology measured using strict criteria (in % normal forms).

Figure 4.

(a) Histograms and density estimates of the woman-specific MCL (in days) for enrollment cycles (light gray bars, dashed line) compared to the woman-specific means for subsequent cycles (dark gray bars, solid line). (b) Median (solid) and 95% equal tail credible interval (dashed) for the probability of pregnancy in the enrollment cycle versus mean frequency of intercourse.

Table 1

Posterior median and 95% equal-tail credible intervals for parameters of MCL sub-model displayed as: L_0 Median U_p (Louis and Zeger (2009)). Age is standardized to (mean = 0, sd = 1) and active smoking is determined by baseline cotinine level ≥ 100 ng/mL. Restricted to 426 couples with data on semen quality who did not exit before the end of the enrollment cycle.

MCL Sub-model		
Parameter	Interpretation	L_0 Median U_p
η_{Age1}	slope on female age	-0.04-0.03 -0.01
η_{Age2}	slope on (female age) ²	-0.02-0.01 0.00
η_{Smoke}	female active smoking (yes/no)	-0.07-0.02 0.03
q	$Pr(T = 1)$	0.750.79 0.82
μ_1	mean for group 1	28.9029.33 29.85
σ_1	SD for group 1	1.791.94 2.11
μ_2	mean for group 2	33.8435.18 37.02
σ_2	SD for group 2	11.0612.34 13.60
σ_W	SD for the random effect	0.080.09 0.10

Table 2

Posterior median and 95% equal-tail credible intervals for association between (mean) MCL and probability of pregnancy (see models M1–M4) in Section 2.4. Posterior summaries displayed as: L_o Median U_p (Louis and Zeger (2009)). Restricted to the 426 couples with data on semen quality who did not exit before the end of the enrollment cycle.

Pregnancy Sub-model		
Model	Parameter	L_o Median U_p
M1 _{Lin}	Cycle length	0.170.34 0.55
M1 _{LinQuad}	Cycle length	0.140.46 0.77
	Cycle length (squared)	−0.62−0.35 −0.17
M2 _{Lin}	Cycle length, group 1	0.090.25 0.43
M2 _{LinQuad}	Cycle length, group 1	0.050.40 0.72
	Cycle length, group 1 (squared)	−0.70−0.45 −0.16
M3 _{Lin}	Conditional mean cycle length	0.050.21 0.38
M3 _{LinQuad}	Conditional mean cycle length	0.040.25 0.45
	Conditional mean cycle length (squared)	−0.28−0.10 0.05
M4 _{Lin}	Conditional mean cycle length, group 1	0.060.21 0.38
M4 _{LinQuad}	Conditional mean cycle length, group 1	0.020.23 0.44
	Conditional mean cycle length, group 1 (squared)	−0.27−0.10 0.03

Table 3

Posterior median and 95% equal-tail credible intervals for parameters of the pregnancy sub-model M1 displayed as: $_{Lo}Median_{Up}$ ((Louis and Zeger, 2009)). Adjusted models A and B correspond to two combinations of uncorrelated semen quality parameters. Active smoking is defined as baseline cotinine level 100ng/mL. All other covariates factors are standardized. Restricted to 426 couples with data on semen quality who did not exit before the end of the enrollment cycle.

Pregnancy Sub-model			
	Unadjusted	Model A	Model B
Cycle length	0.140.46 0.77	-0.060.26 0.59	-0.050.27 0.58
Cycle length (squared)	-0.62-0.35 -0.17	-0.52-0.27 -0.11	-0.52-0.27 -0.12
Morphology (% normal form, strict)	0.160.28 0.40	0.060.23 0.41	0.050.23 0.40
Total sperm count (log)	0.000.12 0.25	-0.090.08 0.25	
Semen volume (log)	-0.070.05 0.17		-0.130.03 0.20
Sperm concentration (log)	-0.020.11 0.23		-0.090.09 0.27
Male active smoking (yes/no)	-0.76-0.38 -0.02	-0.86-0.37 0.13	-0.87-0.37 0.13
Female active smoking (yes/no)	-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)	-0.120.01 0.14	-0.19-0.02 0.14	-0.19-0.03 0.14
Intercourse frequency	8.009.36 9.94	8.089.46 9.96	7.959.35 9.96
Intercourse frequency (squared)	-0.87-0.79 -0.66	-0.87-0.80 -0.67	-0.87-0.79 -0.66

Table 4

Posterior median and 95% equal-tail credible intervals for parameters of the pregnancy sub-model M1 fit using a joint vs two-stage approach. Results displayed as: $_{Lo}Median_{Up}$ ((Louis and Zeger, 2009)). Active smoking is defined as baseline cotinine level $\geq 100\text{ng/mL}$. All other covariates factors are standardized. Restricted to 426 couples with data on semen quality who did not exit before the end of the enrollment cycle.

Pregnancy Sub-model		
	Joint	Two-stage
Cycle length	-0.06 ^{0.26} 0.59	-0.17 ^{0.12} 0.26
Cycle length (squared)	-0.52 ^{-0.27} -0.11	-0.31 ^{-0.20} -0.10
Morphology (strict criteria)	0.06 ^{0.23} 0.41	0.09 ^{0.23} 0.36
Total sperm count (log)	-0.09 ^{0.08} 0.25	-0.09 ^{0.05} 0.19
Male active smoking (yes/no)	-0.86 ^{-0.37} 0.13	-0.72 ^{-0.31} 0.07
Female active smoking (yes/no)	-1.01 ^{-0.13} 0.74	-0.82 ^{-0.06} 0.61
Male and female age (mean)	-0.34 ^{-0.16} 0.02	-0.34 ^{-0.21} -0.07
Male and female age (difference)	-0.19 ^{-0.02} 0.14	-0.09 ^{0.05} 0.19
Intercourse frequency	8.08 ^{9.46} 9.96	8.30 ^{9.44} 9.96
Intercourse frequency (squared)	-0.87 ^{-0.80} -0.67	-0.87 ^{-0.80} -0.69