Bayesian methods for searching for optimal rules for timing intercourse to achieve pregnancy

Bruno Scarpa[1] and David B. Dunson[2]

November 30, 2006

[1] Department of Statistical Sciences

University of Padova, Italy

[2] Biostatistics Branch,

National Institute of Environmental Health Sciences

U.S. National Institutes of Health

dunson1@niehs.nih.gov

With societal trends towards increasing age at starting a pregnancy attempt, many women are concerned about achieving conception before the onset of infertility, which precedes menopause. Couples failing to conceive a pregnancy within 12 months are classified as clinically infertile, and may be recommended for assisted reproductive therapy (ART). Because many ART procedures are expensive and may convey an increased risk of adverse outcomes for the offspring, it is advantageous to decrease time to pregnancy by natural methods. One possibility is to intentionally time intercourse during the days of the menstrual cycle having the highest conception probabilities. This article proposes a Bayesian decision theoretic approach for searching for optimal rules for timing intercourse based on cycle day, secretions and other information. Good rules result in high conception probabilities while requiring minimal targeted intercourse. A biologically-based statistical model is used to relate cycle day and biomarkers to the conception probability. A stochastic search procedure is then developed to search for rules with high expected utility, and the methods are applied to data from a recent Italian study.

Key Words: Bayesian model selection; Conception probability; Decision theory; Gibbs sampler; Loss function; Stochastic search; Time to pregnancy

1 Introduction

1.1 Motivation and Background

There has been a worldwide trend in industrialized countries towards delaying childbearing until later in a woman's reproductive years. This trend has been accompanied by growing anxiety among women that they may have waited too long to start a pregnancy attempt. Such anxiety has been heightened by reports of declines in fecundability starting in the

late 20s for women and in the late 30s for men [1,2]. In this climate, couples attempting pregnancy often become increasingly concerned as the months pass without a positive pregnancy test, sometimes pressing their clinicians for help even after 3 to 6 months of attempting. Couples are diagnosed as clinically infertile after attempting a year with no success. However, even in the absence of known causes of infertility, couples may be recommended for assisted reproductive therapy (ART) before the attempt time exceeds a year.

ART procedures can be extremely costly and may convey an increased risk of adverse pregnancy and developmental outcomes. The data on long term adverse outcomes tend to be limited, because technology in this area changes more rapidly than children develop into adulthood. For couples who could otherwise not conceive naturally, ART provides a valuable option. However, data suggest that most couples in their 20s and 30s who do not conceive within the first year of attempting could conceive naturally if attempting for longer [2,3]. The statistical explanation is that there is a high degree of heterogeneity among couples in their fecundability, the per menstrual cycle probability of conception. This heterogeneity leads to a highly skewed time to pregnancy distribution. As the attempt time increases, the distribution of fecundability among couples still at risk will increasingly concentrate on low values. However, since the proportion of sterile couples is very low (e.g., 1-3%) [4], most couples not conceiving by a year are fecund.

Unfortunately, anxiety about infertility drives many couples to seek infertility treatment even after a relatively modest attempt time. Methods for intentionally timing intercourse during the most fertile days of the menstrual cycle provide a valuable alternative to couples concerned about a long time to pregnancy. Numerous rules have been proposed based on self-monitoring of the menstrual cycle and symptoms of the fertile days [5]. Most rules are based on the identification of the ovulation day and the fertile window around it.

Traditional and widely used means of identifying the day of ovulation and the fertile window include basal body temperature [6,7,8] and calendar calculations [9]. Newer means include serial ovarian ultrasound, monitoring of hormones in urine [10,11,12], monitoring of salivary electrolytes [13], and fertility charting of vaginal discharge [14,15,16].

It is unclear which one of the available rules is the best available option. In addition, given that there has been no systematic search for optimal rules among the huge number of possibilities, there may be other rules yet to be defined that perform better than those yet proposed. A good rule for intercourse behavior is one that maximizes the probability of conception in a menstrual cycle, while minimizing the required days of intercourse. Although there are not many individuals who would characterize a high intercourse frequency as a loss, most would agree that it is appealing to limit the number of days on which intercourse is required. Requiring intercourse on specific days may be stressful for many couples.

1.2 Italian Study of Daily Fecundability

In order to assess the performance of rules for timing intercourse, it is first necessary to accurately model day-specific fecundability across the menstrual cycle, allowing for heterogeneity with measured and unmeasured predictors. Day-specific fecundability is defined here as the probability of conception in a menstrual cycle with a single act of intercourse on a specific day relative to ovulation. As noted by Barrett and Marshall [17] and by Wilcox et al. [18], because most menstrual cycles have multiple intercourse acts during the potentially fertile interval, it is necessary to assume a statistical model to estimate day-specific probabilities. Most results reported in the literature have been based on estimating probabilities in a narrow window around an estimate of the ovulation day

using the model of Schwartz et al. [19] and extensions [e.g. 20].

Unfortunately, as our goal is to assess the performance of rules for prospectively identifying fertile days on the basis of home monitoring of biomarkers, such an approach is not appropriate. What is needed is a statistical model for relating cycle day and biomarkers, such as basal body temperature and characteristics of vulvar secretions, to the day-specific probabilities of conception. Then, given intercourse on particular days of the cycle having particular biomarker levels, one can obtain a cycle-specific probability of conception, under the simplifying assumption that intercourse does not affect the biomarkers. Depending on the rule being used, the intercourse days will be altered, resulting in an altered conception probability.

In order to develop and fit such a model, it is necessary to have very complete data on biomarkers and intercourse days for a large number of menstrual cycles at risk of conception. Unfortunately, such data are difficult to obtain, as previous studies either did not collect home biomarkers [18] or collected information only for a mid-cycle interval [21]. For this reason, we focus on data from a new Italian study of users of the Billing's Ovulation method of natural family planning [22,23].

The study enrolled 193 women recruited from four Italian centers providing services on fertility awareness and natural family planning. Women enrolled were between 18 and 40 years of age, were married or in stable relationship, had at least one menses after cessation of breastfeeding or after delivery, were not taking hormonal medications or drugs affecting fertility and were experienced in the use of the Billings Ovulation Method [24] of natural family planning. The participants were followed prospectively during one or more menstrual cycles, as they collected detailed daily records of vulvar observations of the cervical mucus symptom, and recorded the days during which intercourse and menstrual bleeding occurred. The women had received training at the study centers on how to identify

different types of sensation and mucus.

Teachers classified each day of the cycle according to a five-point scale according to the type of mucus symptom described by women. As discussed in Colombo et al. [22], the two most fertile types of mucus symptom are very similar clinically. Therefore, we collapsed these into one category and used the four point scale indicating: (1) dry, (2) a humid or damp feeling, (3) thick, creamy, elastic, whitish moist mucus symptom, and (4) slippery, stretchy, watery, clear mucus. Higher scores indicate higher levels of estrogenic-type mucus and hence conditions more conducive to sperm survival and transport. Therefore, the conception probability is expected to increase monotonically with mucus score [23,25]

Focusing on menstrual cycles having complete records of mucus and excluding cycles in which mucus was not recorded on a day with intercourse, 2536 menstrual cycles from 191 women have been collected, with 161 of these cycles (from 132 women) ending in a clinical pregnancy. The average age in years of the women and men was 29.9 (standard deviation, 4.15) and 32.6 (4.8), and women contributed an average of 13.3 (12.66) cycles to the data set.

In Section 2, we propose a statistical model for the day-specific probabilities of conception across the menstrual cycle, utilizing information on timing of intercourse and biomarkers. We develop a Markov chain Monte Carlo (MCMC) algorithm for posterior computation in this non-linear hierarchical model. Using this algorithm, we can obtain draws from the posterior predictive distribution for new couples following different rules for timing intercourse. In Section 3, we develop a Bayesian decision theoretic approach and Monte Carlo algorithm for searching for optimal rules. Section 4 applies the MCMC algorithm of Section 2 to the Italian study and summaries results. Section 5 summarizes results from the rule search, and Section 6 discusses the results.

2 Modeling Daily Fecundability

2.1 Hierarchical nonlinear model

An intercourse act can result in conception only if it occurs in a mid-cycle window ending on the day of ovulation [18,26]. Thus, time-varying predictors impact the probability of conception only when one or more acts of non-contracepting intercourse occur during the fertile interval. With this biological motivation, we divide each cycle into three windows: (1) an early infertile window $I_1 = [1, \tau_1]$ during which biomarkers are considered uninformative; (2) a mid-cycle potentially fecund window $I_2 = [\tau_1 + 1, \tau_2]$ during which daily predictors can have a varying effect on the probability of conception; and (3) a late infertile window $I_{3ij} = [\tau_2 + 1, D_{ij}]$, with D_{ij} denoting the length of cycle j from couple i. The term *infertile* is used to refer to a low, but possibly non-zero, conception probability. The changepoints, τ_1 and τ_2, are treated as unknown, but constant across couples and cycles (for simplicity and identifiability). Because conception cycles have no end, we truncate the third interval at $D_{ij} = D = 40$ days for conception cycles.

Suppose that data are collected from n couples, with couple i contributing n_i cycles ($i = 1, \ldots, n$). For cycle j from couple i, let $\mathbf{v}_{ij} = (v_{ij1}, \ldots, v_{ijD_{ij}})'$ be a vector of intercourse indicators, with $v_{ijd} = 1$ denoting intercourse on day d and $v_{ijd} = 0$ otherwise, and let $y_{ij} = 1$ indicate conception and $y_{ij} = 0$ otherwise.

First, suppose that only daily intercourse records over the cycle are available to predict conception. Following a suggestion of Peter Armitage, Barrett and Marshall [17] proposed a model for the conception probability in a cycle based on the assumption that batches of sperm introduced into the reproductive tract on different days commingle and then compete independently to fertilize the ovum. This model was later generalized by a number

of authors to allow covariate effects and heterogeneity among couples. For example, Zhou and Weinberg [27] proposed an estimating equations-based approach. They also demonstrated in their paper that intercourse on the previous day does not significantly reduce the conception probability, so that it is appropriate to ignore abstinence time.

Motivated by the identifiability and computational issues presented in [28], we rely on a modification of the Dunson and Stanford [28] model. Their model focused on day-specific probabilities in a narrow fecund interval ending on the day of ovulation, so incorporated distinct probabilities for each day. To reduce dimensionality in modeling of conception probabilities across the entire menstrual cycle, we assign a different baseline parameter for each of the three intervals I_1, I_2, I_{3ij} resulting in the baseline model:

$$P\{y_{ij} = 1 \,|\, \xi_i, \mathbf{v}_{ij}\} = 1 - \prod_{d=1}^{D_{ij}} (1 - p_{id})^{v_{ijd}} = 1 - \exp\left\{-\xi_i \sum_{t=1}^{3} \sum_{d \in I_{tij}} v_{ijd} \lambda_t\right\}, \qquad (1)$$

where $I_{1ij} = I_1$, $I_{2ij} = I_2$, $p_{id} = 1 - \exp\left\{-\xi_i \sum_{t=1}^{3} \lambda_t 1_{(d \in I_{tij})}\right\}$ is the day-specific probability of conception given intercourse on only day d, λ_t $(t = 1, 2, 3)$ is the window-specific baseline parameter and ξ_i is a couple-specific random-effect measuring the ith couple's biologic fecundity, with $\xi_i < 1$ representing low fecundity, $\xi_i = 1$ for typical fecundity, and $\xi_i > 1$ for above average fecundity. To allow fecundity to vary continuously in the population, we let $\xi_i \sim \mathcal{G}\,(\nu^{-1}, \nu^{-1})$, where $\mathcal{G}(a, b)$ denotes the gamma density with mean a/b and variance a/b^2, so that $\nu = \text{var}(\xi_i)$.

Consider now the case in which $\mathbf{M}_{ij} = [\mathbf{m}'_{ij1}, \ldots, \mathbf{m}'_{ijD_{ij}}]'$, a $D_{ij} \times q$ time-varying covariate matrix for cycle j from couple i, is available. A parsimonious extension of (1) would allow the day-specific probabilities to vary by a multiplicative factor depending on the level of the predictors on the day of intercourse:

$$P\{y_{ij} = 1 \,|\, \xi_i, \mathbf{v}_{ij}, \mathbf{M}_{ij}\} = 1 - \exp\left\{-\xi_i \sum_{t=1}^{3} \sum_{d \in I_{tij}} v_{ijd} \lambda_t \exp\left\{(\mathbf{m}'_{ijd}\boldsymbol{\beta}) 1_{(d \in I_2)}\right\}\right\}, \qquad (2)$$

where $\boldsymbol{\beta}$ is a vector of regression coefficients.

It is standard practice in epidemiologic studies to categorize exposures, behavioral factors and biomarkers in order to simplify the analysis and presentation of the results. Therefore, we focus on a single M-level day specific categorical predictor $w_{ijd} \in 1, 2, \ldots, M$ (the generalization to consider multiple categorical predictors is straightforward), corresponding to the covariate matrix having rows $\mathbf{m}_{ijd} = [1_{(w_{ijd}=2)}, 1_{(w_{ijd}=3)}, \ldots, 1_{(w_{ijd}=M)}]'$, for $d = 1, 2, \ldots, D_{ij}$. We assign each day d from cycle i, j to one of $K = M + 2$ categories: $C_{ijd} = 1$ if $d \in I_1$, $C_{ijd} = w_{ijd} + 1$ if $d \in I_2$, and $C_{ijd} = K$ if $d \in I_{3ij}$.

Let $x_{ijk} = \sum_{d=1}^{D_{ij}} 1_{(C_{ijd}=k)} v_{ijd}$ denote the number of days in the jth cycle of couple i that have reported intercourse and that are in the kth category ($k = 1, \ldots, K$), with $\mathbf{x}_{ij} = (x_{ij1}, \ldots, x_{ijK})'$. The probability of conception is expressed as

$$P\{y_{ij} = 1 \,|\xi_i, \mathbf{x}_{ij}, \mathbf{M}_{ij}\} = 1 - \exp\left\{-\xi_i \sum_{k=1}^{K} x_{ijk}\lambda_k\right\}, \tag{3}$$

where $\lambda_1, \lambda_2, \lambda_K$ are the baseline parameters characterizing the probabilities of conception in the 1st, 2nd, and 3rd interval, respectively, and $\lambda_k = \exp(\beta_{k-2})$ for $k = 3, \ldots, K-1$ allow changes from λ_2 across categories of the time-varying predictor.

Note that model (3) makes a monotonicity assumption in which additional intercourse acts in a menstrual cycle can only increase the probability of conception. Although sperm concentration has been shown to decrease for 1-3 days following ejaculation (see for instance [29]), the magnitude of this decrease is not sufficient to invalidate the monotonicity assumption. In fact, Zhou and Weinberg [27] showed that the occurrence of intercourse on adjacent days did not significantly reduce the day-specific conception probability.

Model (3) modifies Dunson and Stanford [28] to allow random changepoints in the cycle to accommodate data lacking a reliable marker of ovulation day. Previously, Dominik and Chen [31] proposed to model a per cycle day pregnancy curve, avoiding the estima-

tion of the ovulation day. Dominik and Chen do not account explicitly for heterogeneity among couples in their pregnancy probabilities, instead using a robust variance estimator. Hence, because less fecund couples may contribute more cycles to a data set, the resulting pregnancy probabilities will be biased downwards. In addition, it is necessary to model heterogeneity if one wants to predict conception probabilities for new couples following different rules.

For tractability, we assume that the biomarker data and the random effects will not be altered by changing the intercourse data through targeted intercourse. This allows us to predict couple-specific changes in the conception probability with modifications to intercourse timing and frequency without imposing a model on the biomarker trajectories. In order to predict probabilities of conception for a new couple, we obtain the marginal probability of conception integrating out the couple-specific random-effect ξ_i. For model (3) we have the simple closed form $\Pr(y_{ij} = 1 \mid \mathbf{x}_{ij}, \mathbf{M}_{ij}) = 1 - (1 + \nu \sum_{k=1}^{K} x_{ijk}\lambda_k)^{1/\nu}$.

2.2 MCMC algorithm for posterior computation

Following a Bayesian approach, we modify the efficient auxiliary variables Markov chain Monte Carlo algorithm proposed by Dunson and Stanford [28]. In fact, expression (3) has an equivalent representation as an underlying Poisson variable model, with $y_{ij} = 1_{\left(\sum_{k=1}^{K} Z_{ijk} > 0\right)}$ and Z_{ijk} conditionally-independent Poisson latent variables with mean $E(Z_{ijk}) = \xi_i \, x_{ijk} \, \lambda_k$. Full conditional posterior distributions are then easily obtained for each of the parameters and latent variables.

Conditionally conjugate priors are chosen for each of the parameters, with discrete uniform priors for τ_1 and τ_2, gamma priors for $\lambda_1, \lambda_2, \lambda_K, \nu^{-1}$ and a mixture of a point mass at one (with probability π) and a gamma density, possibly truncated below or above by one,

for the $\gamma_m = \lambda_{m+1}/\lambda_m$ $(m = 1, \ldots, M-1)$ parameters quantifying the effect of increasing, in the second interval, the categorical predictor by one unit from m to $m+1$. This parameterization allows for selecting predictors of the day-specific conception probability and may improve efficiency by incorporating constraints on the values of the multiplicative increments $\{\gamma_k\}$. For example, focusing on ordered categorical predictors, if the covariate has a potentially beneficial impact on the probability of conception and an adverse effect can be ruled out *a priori*, then the constraint $\gamma_k \geq 1$ would be appropriate, and included in the prior by truncating below by one the gamma distribution in the mixture. The joint prior density for all the parameters is summarized as follows

$$\begin{array}{rcl} \pi(\theta) & = & \pi(\tau_1, \tau_2, \nu, \lambda_1, \lambda_2, \lambda_K, \gamma_1, \ldots, \gamma_{M-1}) \\ & = & \mathcal{U}(\tau_1; a_{\tau 1}, b_{\tau 1}) \cdot \mathcal{U}(\tau_2; a_{\tau 2}, b_{\tau 2}) \cdot \\ & & \mathcal{G}(\nu; c_1, c_2) \left\{ \prod_{k \in \{1,2,K\}} \mathcal{G}(\lambda_k; a_{0k}, b_{0k}) \right\} \cdot \\ & & \left\{ \prod_{h=1}^{M-1} I_1 - \mathcal{G}_{\mathcal{A}_h}(\gamma_h; \pi_{0h}, a_h, b_h) \right\} \end{array} \tag{4}$$

where $\mathcal{U}(\cdot; a, b)$ denotes the Discrete Uniform probability mass function between a and b, $\mathcal{G}(\cdot; a, b)$ denotes the Gamma density, and $I_1 - \mathcal{G}_{\mathcal{A}}(\cdot; \pi, a, b)$ denotes the density consisting of the mixture of a point mass at one (with probability π) and a Gamma density truncated to the region $\mathcal{A}$ which is typically chosen to be $\mathbb{R}^+$, $[1, +\infty)$ or $(0, 1]$, to correspond to no constraint, positive effect and negative effect, respectively. Posterior summaries of the parameters and of the latent variable are obtained using the MCMC algorithm outlined in Appendix I.

3 Searching for Rules for Timing Intercourse

3.1 Formalizing the problem

Our goal is to choose good rules for timing intercourse on the basis of cycle day and the time-varying biomarker for couples attempting conception. In particular, from a conceptual

viewpoint, a good rule is one which is simple to apply and shortens the time to pregnancy while limiting the number of days on which intercourse is prescribed. Our formal approach to the problem of identifying good rules is to first choose a list of simple rules as potential candidates. We then specify a utility function, which rewards a high conception probability while penalizing number of prescribed intercourse acts. Because this utility function necessarily involves many unknown parameters, we follow the approach of calculating the expected posterior utility averaged across the posterior distribution of the unknown parameters [32,33]. The Bayes optimal rule among those considered is then the one with the highest expected utility.

Let $\mathbf{M}_k = (M_1, \ldots, M_k)' \in \mathcal{M}_k = [1, 2, \ldots, M]^k$ denote a vector of $1 - M$ ordinal markers for cycle days 1 to k. Then, by definition, a rule R is a collection of functions $R = \{R_k, k = 1, \ldots, D\}$, with $R_k : \mathcal{M}_k \to [0, 1]$. In particular, letting $x_k = R_k(\mathbf{M}_k)$, given the markers $\mathbf{M}_k$ leading up to day k of the cycle, rule R either recommends that the couple has intercourse on that day ($x_k = 1$) or else leaves the decision up to the couple's desire ($x_k = 0$). For example, a simple calendar-based rule that prescribes intercourse only within the 10-17 day interval, would have $R_k(\mathbf{M}_k) = 1_{(k \in [10,17])}$, so that the marker data would be ignored.

In practice, there may be a large number of rules, even when one focuses on simple rules based on calendar and the history of a single biomarker. We let $\mathcal{R}$ denote the set of rules under consideration. The utility function for a rule R is defined as follows:

$$u_\delta(\theta, R, \mathbf{M}) = \Pr(y = 1|\theta, \mathbf{M}, R) - \delta B(\mathbf{M}, R), \tag{5}$$

where $\Pr(y = 1|\theta, \mathbf{M}, R)$ is the probability of conception given parameters θ, marker data $\mathbf{M}$ and rule R, $B(\mathbf{M}, R)$ is the number of days of required intercourse recommended by rule R given marker data $\mathbf{M}$, and δ is a known penalty. Note that we have not conditioned

on intercourse in the component for the conception probability, because the rule implies a particular vector of intercourse indicators given the woman's marker data, assuming for simplicity that intercourse only occurs on recommended days (extensions are straightforward). Here, δ quantifies the decrease in pregnancy probability one is willing to face in exchange for one less day of required intercourse. Hence, this utility function rewards a high conception probability for a rule, while penalizing prescribed intercourse days.

For a new couple $i = n+1$ wanting to limit their time to conception without knowledge of their marker data, the Bayes optimal rule to select is formally:

$$\begin{aligned} R^* &= \arg\max_{R\in\mathcal{R}} U_\delta(R) \qquad\qquad \text{with} \\ U_\delta(R) &= \int u_\delta(\theta, R, \mathbf{M})\, \pi(\theta \,|\, \text{data})\, \pi(\mathbf{M} \,|\, \text{data})\, d\theta\, d\mathbf{M}, \end{aligned}$$

where $\pi(\theta \,|\, \text{data})$ is the posterior distribution of the parameters in model (3) given the data, and $\pi(\mathbf{M} \,|\, \text{data})$ is the posterior predictive distribution of the marker data $\mathbf{M}$ for a new subject. The posterior distribution $\pi(\theta \,|\, \text{data})$ depends on the marker levels for the actual study subjects, whereas conceptually the marker data $\mathbf{M}$, over which the utility is averaged, is for a hypothetical new subject. Note that samples from the posterior of the parameters can be obtained using the MCMC algorithm described in Section 2.2 and the appendix. However, as we have not modeled the marker data due to the difficulty of formulating a realistic model for the complex ordered categorical time series, samples from $\pi(\mathbf{M} \,|\, \text{data})$ are not available.

We recommend plug-in approximations to the following high-dimensional integrals:

$$\begin{aligned} \Pr(y = 1 \,|\, \theta, R) &= \int \Pr(y = 1 \,|\, \theta, \mathbf{M}, R)\, \pi(\mathbf{M} \,|\, \text{data})\, d\mathbf{M}, \\ B(R) &= \int B(M, R) \pi(\mathbf{M} \,|\, \text{data})\, d\mathbf{M}. \end{aligned} \tag{6}$$

In particular, we use the empirical distribution of the marker data in the sample in place

of the predictive distribution, and rely on

$$\widehat{\Pr}(y=1|\theta,R,\mathbf{data}) \;=\; \frac{1}{n}\sum_{i=1}^{n}\left[\frac{1}{n_i}\sum_{j=1}^{n_i}\left\{1-\left(1+\nu\sum_{k=1}^{K}x_{ijk}^{(R)}\lambda_k\right)^{1/\nu}\right\}\right], \qquad (7)$$

where $x_{ijk}^{(R)}$ denotes the potential number of intercourse days in the jth cycle of couple i falling in category k given use of rule R conditional on that couple's biomarker data. To clarify, $x_{ijk}^{(R)}$ is the value of x_{ijk} that would have been observed had couple i followed rule R in cycle j. The term in $(\cdot)$ including the exponent is the marginal conception probability integrating out the couple specific random effect, but conditioning on the intercourse and marker data. Note that we can similarly obtain a plug-in estimate for B_R by averaging the intercourse days required by rule R across the data for subjects in the sample. By using this plug-in approximation we eliminate the dependence on $\mathbf{M}$ of the utility function and of the expected utility, since in this last function, the posterior distribution of θ depends on the marker levels of a hypothetical new subject and not on the marker data observed for the actual study subjects.

To calculate the expected posterior utility for a given rule R, all that remains is to integrate out θ across the posterior distribution. This can be accomplished by simply using the MCMC draws and averaging. Note that due to the approach we have used it is not necessarily to rerun the MCMC algorithm for each new rule. Instead we can reuse the samples from a single run each time we calculate the expected posterior utility for a rule. The use of Monte Carlo approximations is a common technique in many simulation-based optimal design approaches. For example, Bielza, Müller and Rios Insua [34] proposed a generic Monte Carlo method to calculate maximum expected utility in a decision analysis. Müller and Parmigiani [35] used simulation to evaluate expected utility integrals exploiting the continuity of the utility function to reduce computational effort. Carlin, Kadane and Gelfand [36] used a similar strategy for sequential decision problems. Refer also to Müller,

Sansò and Iorio [37].

When the list of rules is moderate, one can simply cycle through each rule, estimating the expected posterior utility in each case, choosing the one with the highest value. However, when the decision space $\mathcal{R}$ is very large, one cannot perform calculates for all possible rules. As an alternative to simulated annealing [38], we recommend a simple iterated hill-climber algorithm with a stochastic perturbation to avoid local maxima. The idea is to choose the steepest direction and to follow it until, in this direction, a maximum is reached. Once there, the algorithm looks for another direction with steepest increase of the expected utility function and starts again from there. In order to avoid to stopping at a local maximum, the algorithm, every time it needs to change direction, evaluates the objective function in one or more randomly chosen points over the entire space $\mathcal{R}$. By comparing these values with the one proposed by the new direction, the algorithm identifies the higher posterior integrated utility from where to start again looking for the steepest direction. Details for the algorithm are outlined in the Appendix II.

4 Application to Italian Study

In this section we apply the approach described to data presented in Section 1.2. The time-varying marker is chosen as the 1-4 mucus score, which is recorded on each day of each cycle in the study. We obtain $K = 6$ categories for the model (3). Although we expect the conception probabilities to be low in the first and third interval and in the second interval on dry days, we choose a diffuse prior for the baseline parameters $\lambda_1, \lambda_2, \lambda_6$ by letting $a_{0k} = b_{0k} = 0.01, k = 1, 2, 6$ as parameters of the gamma distribution.

It is reasonable to assume that the probabilities are nondecreasing with increases in the mucus score for days of the second interval. We incorporate this constraint by letting

Parameter	Mode	Mean	Median	SD	95% Credible Interval
τ_1	5	5.96525	5.0000	1.1653	[5, 8]
τ_2	21	20.92396	21.0000	1.0284	[19, 23]
λ_1		0.00178	0.0000	0.0055	[0.00, 0.02]
λ_2		0.01052	0.0093	0.0065	[0.001, 0.027]
γ_1		6.66306	3.8479	11.1830	[1.20, 37.01]
γ_2		2.12287	1.6683	1.4273	[1.00, 6.11]
γ_3		14.27453	13.3324	5.8750	[5.71, 28.88]
λ_6		0.00042	0.0000	0.0014	[0.000, 0.005]
ν		1.82630	1.8016	0.3492	[1.20,2.58]

Table 1: Posterior summaries of the parameters of the model fitted on the Italian study data

$\mathcal{A}_h = [1, +\infty)$ within the prior for γ's introduced in (4). We let $\pi_{0h} = 0.5^{1/3}$ for $h = 1, 2, 3$ to assign 0.5 prior probability to the global null hypothesis of no association between the mucus score and the conception probabilities, and we let $a_h = b_h = 0.01, h = 1, \ldots, 3$ to allow a high degree of uncertainty in the values of γ_h under the alternative hypothesis. Considering the magnitude of heterogeneity estimated in Dunson and Zhou [20] using a different data set and model, we choose $c_1 = 1$ and $c_2 = 2$ to specify a weakly informative prior for the frailty variance ν. Finally, we let τ_1 vary uniformly within $[5, 12]$, while letting τ_2 vary uniformly within $[17, 25]$. These choices were suggested by previous research on calendar methods.

We ran the MCMC algorithm for 12800 iterations, discarding the first 500 iterations as a burn-in. Convergence was rapid and mixing was excellent, and these burn-in and collection intervals were deemed sufficient. Posterior summaries of the parameters are presented in Table 1. While the posterior distribution τ_2 is symmetric around the day 21, the distribution of τ_1 is asymmetric, with mean of about 6, 95% credible interval of $[5, 8]$ and median and mode in 5, which was also the lower day allowed by the prior distribution. This result is not surprising, given that the prior was chosen based on day-specific probabilities

Category				Probability of conception		
k	Time interval		Mucus Type	Mean	SD	95% Credible Interval
1	$\leq \tau_1$			0.0017	0.0053	0.0000 - 0.0191
2	$(\tau_1, \tau_2]$	1	dry	0.0103	0.0063	0.0014 - 0.0258
3		2	humid or damp feeling	0.0381	0.0170	0.0115 - 0.0764
4		3	thick, creamy, elastic, whitish moist mucus symptom	0.0643	0.0216	0.0316 - 0.1189
5		4	slippery, stretchy, watery, clear mucus	0.4077	0.0520	0.3059 - 0.5094
6	$> \tau_2$			0.0004	0.0014	0.0000 - 0.0048

Table 2: Posterior summaries of probabilities of conception according to the phase of the cycle and mucus score

estimated for calendar methods that do not rely on mucus. By incorporating mucus information, we naturally estimate an earlier transition to the mid-cycle potentially fecund interval. This early transition allows for the subset of cycles with early ovulation in which the fecund interval starts soon after menses. Cycles with later ovulation will typically have low mucus scores early in the cycle, leading to low day-specific conception probabilities for these days.

Although we used a prior that assigned a moderately high probability to $\gamma_h = 1.0$ $(h = 1, 2, 3)$ in order to adjust for a possibly inflated type I error rate due to multiple testing, consistently with earlier results [25] there was clear evidence in the data in favor of $\gamma_h > 1.0$. We also obtain posterior distributions for the probabilities of conception given a single act of intercourse in the cycle, occurring in one of the three phases. For intercourse acts in the second phase, we stratify by type of mucus on the intercourse day. The resulting estimates are shown in Table 2.

In the mid-cycle interval, the probability is quite low for days with no secretions (0.01) and with a mucus score of 2 (0.038) or 3 (0.064), but then increases dramatically to 0.41 on

days with most fertile-type mucus. These differences are all statistically significant, having posterior probabilities of no difference < 0.05. On days in the first and third phase, the probability of conception is essentially zero (0.002 and 0.0004).

5 Results of the Rule Search

We compared a wide class of rules based on calendar and mucus. We focus on rules prescribing intercourse on days within a mid-cycle window, allowing the last day of the first window ϕ_1 and of the second window, ϕ_2, to vary between 5 and 12 and between 17 and 25, respectively. We included the following suggestions for intercourse: (1) every day, (2) on days with mucus score > 1, (3) mucus score > 2, (4) mucus score > 3, (5) mucus score > 1 on that day or day before, (6) mucus score > 2 on that day or day before, and (7) mucus score > 3 on that day or day before. We index these groups of rules, respectively, $r = 1, \ldots, 7$, resulting in 504 different rules.

For each rule, we considered different patterns of intercourse acts. For the first and the third interval we suppose that couples (a) never have intercourse and (b) have intercourse on one seventh of days, randomly chosen. In the mid-cycle window we suppose that couples follow the rule, having intercourse every day required.

For each scenario we consider the utility function (5) with a range of values for the penalty coefficient δ. We choose B_R in (5) to be the average number of intercourse days that each rule prescribes, while other intercourse acts during the cycle decided by the couple are not considered as a loss. We used the algorithm outlined in Section 3.1 to identify the optimal rule by finding the maximum of the approximated utility function (5) using the Markov Chain obtained in fitting the model. To reduce computational time, we thinned the chain using only 3137 iterations, one every four elements of the chain.

Table 3: Optimal rules and utility function for couples that strictly follow the proposed rule. Intercourse every day required by the rule in the mid-cycle interval and never intercourse in the others intervals.

	Rule parameters			Utility function		
δ	Interval Start $\phi_1 + 1$	Interval End ϕ_2	Mucus type	Probability of conception	Number of prescribed intercourse days	$\hat{U}_\delta(R)$
0	6	25	no	0.687	20.00	0.687
0.003	8	21	no	0.681	14.00	0.639
0.01	10	18	no	0.647	9.00	0.557
0.03	11	17	no	0.615	7.00	0.405
0.05	13	17	no	0.537	5.00	0.287
0.07	13	17	3, 4	0.469	3.92	0.195
0.1	13	17	4	0.347	2.42	0.105

Optimal decisions and the utility functions for couples strictly following the rule, for different values of δ, are presented in Table 3 considering couples with no intercourse in the cycle except when prescribed by the rule, and in Table 4, considering couples having intercourse about once a week when not differently prescribed by the rule. The first line of both Tables gives, simply, the probability of conception without an explicit penalty for intercourse days. By maximizing this probability we obtain for both the scenarios a value of 0.69 corresponding to the rule requiring intercourse every day between days 6 and 25. A more realistic frequency of intercourse is obtained, when couples choose different values of δ. In this case, if the penalty is small, optimal rules are characterized by the calendar component of the rule, while, for a high penalty the optimal rule of both scenarios prescribes intercourse considering both calendar and mucus observations.

Table 5 captures the uncertainty about Tables 3 and 4 by reporting some percentiles of the distribution of conception probabilities obtained using each optimal rule. For both scenarios, variability and skewness increase with δ, indicating that among couples applying a medium-high penalty to intercourse frequency, a relevant percentage of not conceiving

Table 4: Optimal rules and utility function for couples that strictly follow the proposed rule. Intercourse every day required by the rule in the mid-cycle interval and 1/7th of days in the other intervals

	Rule parameters			Utility function		
δ	Interval Start $\phi_1 + 1$	Interval End ϕ_2	Mucus type	Probability of conception	Number of prescribed intercourse days	$\hat{U}_\delta(R)$
0	6	25	no	0.688	20.00	0.688
0.003	8	21	no	0.683	14.00	0.641
0.01	10	18	no	0.654	9.00	0.564
0.03	12	17	no	0.605	6.00	0.425
0.05	13	17	2, 3, 4	0.546	4.45	0.323
0.055	13	17	3, 4	0.525	4.05	0.302
0.1	13	17	4	0.452	2.79	0.173

couples is still present. The last line of Table 5, indicated with $\delta = \infty$, gives the percentiles of the probability of conception when no rule is used and a frequency of intercourse of one seventh of days is supposed.

6 Discussion

This article has described a Bayesian decision theoretic approach for searching for optimal rules for timing intercourse in a cycle. The best rule maximizes the probability of conception while minimizing the number of intercourse days required. Probabilities of conception are estimated by relating biomarkers to the conception probability using a biologically-based statistical model assuming no reliable marker of ovulation is available. Previously cervical mucus has been used [e.g. 5,21,22] as a good marker of ovulation, but measurement error in identifying ovulation day using mucus data can lead to difficulties in interpreting effects of mucus as a predictor of conception probabilities. In agreement with recent studies [22,23,25] showing that mucus is an important predictor of the day-specific

Table 5: Percentiles from the distribution of the probabilities of conception for the optimal rules for couples strictly following the rule

		Percentile						
Scenario	δ	0.05	0.10	0.25	0.50	0.75	0.90	0.95
	0	0.511	0.589	0.660	0.721	0.761	0.784	0.800
mid-cycle interval	0.003	0.502	0.583	0.655	0.717	0.748	0.778	0.791
intercourse every day	0.01	0.445	0.509	0.632	0.685	0.718	0.759	0.773
	0.03	0.264	0.461	0.581	0.644	0.712	0.737	0.756
first and third interval	0.05	0.134	0.201	0.467	0.573	0.679	0.710	0.710
never intercourse	0.07	0.023	0.069	0.417	0.560	0.633	0.679	0.710
	0.1	0.000	0.000	0.060	0.422	0.560	0.632	0.676
	0	0.511	0.589	0.660	0.721	0.761	0.784	0.800
mid-cycle interval	0.003	0.502	0.587	0.655	0.717	0.748	0.779	0.791
intercourse every day	0.01	0.448	0.555	0.639	0.688	0.737	0.759	0.774
	0.03	0.240	0.451	0.574	0.644	0.711	0.737	0.742
first and third interval	0.05	0.127	0.228	0.469	0.586	0.679	0.712	0.722
intercourse 1/7th of days	0.055	0.079	0.169	0.449	0.575	0.676	0.710	0.716
	0.1	0.031	0.070	0.417	0.554	0.630	0.679	0.692
	∞	0.000	0.011	0.060	0.229	0.433	0.560	0.572

conception probabilities, we estimated that on one day within a mid-cycle window between days 7 and 21, if most-fertile type mucus is observed, the conception probability is 40 times higher than on a day with no noticeable secretions, while outside the window conception probabilities are effectively 0.

A stochastic search procedure has then been developed to search for rules with high expected utility. Using a Bayesian decision approach, we found that simple rules based on increasing the frequency of intercourse on days within a mid-cycle interval having mucus score above a threshold, have high theoretical effectiveness. As in any complex MCMC simulation, the main limitation of the proposed method is the computationally intensive implementation, even if the algorithm of Section 3.1 has been introduced. Another limitation is related to the fact that we assume that there is one single utility function for every couple, with known trade-offs between possibly competing goals, however, in our case this assumption is reasonable since it involves choices of couples, which are, usually, defined a

priori.

We considered a wide set of simple rules based only on calendar and mucus, but the same procedure can be applied to compare more complex rules, such as the ones described in Stanford, White and Hatasaka [5] or proposed by Natural Family Planning centers. Most of these rules and the ones we considered, focus on using a decision rule at a single time point. From the standpoint of minimizing a couple's time to conception, it may be better to consider a sequential rule that uses data collected in past menstrual cycles. Our procedure can be applied to search for good sequential rules, but such rules may be difficult to implement for couples, involving substantial training and record keeping.

One of the major advantages of optimizing the timing of intercourse may be the ability to condense the time required before making the diagnosis of infertility, such that appropriate further evaluation can be pursued in a timely manner when necessary. Following the optimal rule selected using the proposed combination of calendar and mucus, results in a 50% reduction in the time to pregnancy on average. There is an even greater improvement for couples of below average fecundity at risk of being diagnosed as infertile. Our results confirm previous findings [5,39] and support the suggestion by Hilgers [40] that with timed intercourse, a diagnosis of infertility can be established in 6 months.

Finally we note that our procedure could also be used by couples attempting to avoid conception using fertility awareness methods. A loss function can be defined in order to identify rules having a low pregnancy rate with a penalty for large numbers of abstinence days. The proposed procedure could also be used in comparing rules that utilize additional information, such as age of the woman or using hormone data obtained from fertility monitors. Potentially, there may be certain rules that work well for some couples but not for others. For example, the optimal mid-cycle interval may vary depending on whether a woman has long or short cycles and how regular they are. Also mucus types can have

different effects on the probability of conception in different women and in different cycles of the same woman. Incorporating such woman-specific information into the rule selection process should improve the performance of the rule, and potentially software could be developed that outputs the optimal rule when the user inputs their cycle history, age, and other characteristics, such as desired intercourse frequency.

References

1. Dunson DB, Colombo B, Baird DD. Changes with age in the level and duration of fertility in the menstrual cycle. *Human Reproduction* 2002; **17**: 1399–1403.

2. Dunson DB, Baird DD, Colombo B. Increased infertility with age in men and women. *Obstetrics & Gynecology* 2004; **103**: 51–56.

3. Bongaarts JM. A Method for the Estimation of Fecundability. *Demography* 1975;**12**: 645.

4. Trussell J, Wilson C. Sterility in a Population with Natural Fertility. *Population Studies* 1985; **39**: 269–285.

5. Stanford JB, White GL, Hatasaka H. Timing Intercourse to Achieve Pregnancy: Current Evidence. *Obstetrics and Gynecology* 2003; **100**: 1333–1341.

6. Doering GK. Ein Beitrag zur Frage der periodischen Fruchtbarkeit der Frau auf Grund der Erfahrungen bei der Zyklusanalyse mit Hilfe der Temperaturmessung. *Geburtshilfe und Frauenheilkunde* 1950; **10**: 515–521.

7. Doering GK. Bestimmung der unfruchtbahren Tage im Zyklus der Frau. *Fortschritte der Medizin - Sondedruck*; 1986; **104**(46): 941–942.

8. Marshall J. A field trial of the basal body-temperature method of regulating births *The Lancet* 1968; **2**: 8–10.

9. Ogino K. Konzeptionstermin des Weibes und seine Anwendung in der Praxis. *Zentralblatt für Gynäkologie* 1932; **56**: 721–732.

10. Thornton S, Pepperell R, Brown J. Home monitoring of gonadotropin ovulation induction using the Ovarian Monitor. *Fertility and Sterility* 1990; **54**: 1076–1082.

11. Nielsen MS, Barton SD, Hatasaka HH, Stanford JB. Comparison of several one-step home urinary luteinizing hormone detection test kits to OvoQuick. *Fertility and Sterility* 2001; **76**(2): 384–387.

12. Behre HM, Kuhlage J, Gassner C, Sonntag B, Schem C, Schneider HP, Nieschlag E. Prediction of ovulation by urinary hormone measurements with the home use ClearPlan Fertility Monitor: Comparison with transvaginal ultrasound scans and serum hormone measurements. *Human Reproduction* 2000; **15**: 2478–2482.

13. Fehring RJ. A comparison of the ovulation method with the CUE ovulation predictor in determining the fertile period. *Journal of American Academy of Nurse Practitioners* 1996; **8**: 461–466.

14. Hilgers TW, Prebil AM. The ovulation method-vulvar observations as an index of fertility *Obstetrics and Gynecology* 1979; **53**: 12–22.

15. Hume K. Fertility awareness in the 1990s - The Billings Ovulation Method of natural family planning, its scientific basis, practical application and effectiveness. *Advanced in Contraception* 1991; **7**: 301–311.

16. Stanford JB, Smith KR. Characteristics of women associated with continuing instruction in the Creighton Model Fertility Care System. *Contraception* 2000; **61**: 121–129.

17. Barrett JC, Marshall J. The risk of Fecundability on different days of the menstrual cycle. *Population Studies* 1969; **23**: 455–461.

18. Wilcox AJ, Weinberg CR, Baird DD. Timing of sexual intercourse in relation to ovulation: Effects on the probability of conception, survival of the pregnancy and sex of the baby. *New England Journal of Medicine* 1995, **333**, 1517–1521.

19. Schwartz D, MacDonald PDM, Heuchel V. Fecundability, coital frequency and the viability of ova *Population Studies* 1980; **34**: 397–400.

20. Dunson DB, Zhou H. Bayesian Modeling of Fecundability and Sterility. *Journal of the American Statistical Association* 2000; **95**: 1054–1062.

21. Colombo B, Masarotto G. Daily fecundability: first results from a new data base. *Demographic Research* 2000; **3**:5.

22. Colombo B, Mion A, Passarin K, Scarpa B. Cervical mucus symptom and daily fecundability: first results from a new data base *Statistical Methods in Medical Research* 2006; **15**: 161–180.

23. Scarpa B, Dunson DB, Colombo B. Cervical mucus secretions on the day of intercourse: An accurate marker of highly fertile days. *European Journal of Obstetrics and Gynecology and Reproductive Biology* 2006; **125**(1): 72–78.

24. Billings EL, Billings JJ, Brown JB, Burger HG. Symptoms and hormonal changes accompanying ovulation. *Lancet* 1972; **1**(7745): 282–284.

25. Bigelow J, Dunson DB, Stanford JB, Ecochard R, Gnoth C, Colombo B. Mucus observations in the fertile window: a better predictor of conception than timing of intercourse. *Human Reproduction* 2004;**19**: 889–892.

26. Dunson DB, Baird DD, Wilcox AJ, Weinberg CR. Day-specific probabilities of clinical pregnancy based on two studies with imperfect measures of ovulation. *Human Reproduction* 1999; **14**: 1835–1839.

27. Zhou H, Weinberg CR. Modeling conception as an aggregated Bernoulli outcome with latent variables via the EM algorithm. *Biometrics* 1996; *52*: 945-954.

28. Dunson DB, Stanford JB. Bayesian inferences on predictors of conception probabilities. *Biometrics* 2005; **61**: 126–133.

29. Lunenfeld LE, Weiss N, Friger M, Jar-Vardi I, Koifman A, Potashnik G. Relationship between the duration of sexual abstinence and semen quality: analysis of 9,489 semen samples. *Fertility and Sterility* 2005; **83**:1680–1686.

30. Dunson DB. Empirical Bayes density regression. *Statistica Sinica* 2006; to appear.

31. Dominik R, Chen PL. Day specific pregnancy probability estimation in barrier conceptive effectiveness trials. *Paediatric and Perinatal Epidemiology* 2006; **10**: 38–42.

32. DeGroot MH. *Optimal Statistical Decisions*, McGraw-Hill: New York, 1970.

33. Berger JO. *Statistical decision theory and Bayesian analysis.* Springer: New York, 1985.

34. Bielza C, Mueller P, Rios Insua D. Monte Carlo Methods for Decision Analysis with Applications to Influence Diagrams. *Management Science* 1999; **45**(7): 995–1007.

35. Müller P, Parmigiani G. Optimal Design via Curve Fitting of Monte Carlo Experiments. *Journal of the American Statistical Association* 1996; **90**: 1322–1330.

36. Carlin B, Kadane J, Gelfand A. Approches for Optimal Sequential Decision Analysis in Clinical Trial. *Biometrics* 1998; **54**: 964–975.

37. Müller P, Sansò B, De Iorio M. Optimal Bayesian Design by Inhomogeneous Markov Chain Simulation. *Journal of the American Statistical Association* 2004; **99**: 788–798.

38. Belisle CJP. Convergence theorems for a class of simulated annealing algorithms on R^d. *Journal of Applied Probability* 1992; **29**: 885–895.

39. Hilgers TW, Daly KD, Prebil AM, Hilgers SK. Cumulative pregnancy rates in patients with apparently normal fertility and fertility-focused intercourse. *Journal of Reproductive Medicine* 1992; **37**: 864–866.

40. Hilgers TW. *The medical applications of natural family planning: A contemporary approach to women's health care*, Pope Paul VI Institute Press: Omaha, Nebraska, 1991.

Appendix I

Details for Implementing MCMC Algorithm of model in Section 2.1

Step 1. Sample from the full conditional distribution of $Z_{ij} = \sum_{k=1}^{K} X_{ijk} Z_{ijk}$ by setting $Z_{ij} = 0$ if $Y_{ij} = 0$ and otherwise sampling sequentially from

$$\pi(Z_{ij}|Y_{ij} = 1, \theta, \xi, \text{data}) \quad = \quad \text{Poisson}\left\{\xi_i \sum_{k=1}^{K} X_{ijk}\lambda_k\right\} \text{ truncated so that } Z_{ij} > 0,$$

$$\pi(Z_{ij1},\ldots,Z_{ijK}|Z_{ij},Y_{ij},\theta,\xi,\text{data}) \;=\; \text{Multinomial}\left(Z_{ij};\frac{\xi_i X_{ij1}\lambda_1}{\xi_i\sum_{k=1}^K X_{ijk}\lambda_k},\ldots,\frac{\xi_i X_{ijK}\lambda_K}{\xi_i\sum_{k=1}^K X_{ijk}\lambda_k}\right)$$

Step 2. Sample λ_1, λ_2, λ_K from their conjugate full conditional distribution:

$$\pi(\lambda_k|\mathbf{Z}_{[\mathbf{X}\geq\mathbf{1}]},\theta_{(-\lambda_k)},\xi,\text{data}) = \mathcal{G}\left(\lambda_k; a_{0k}+\sum_{ij:X_{ijk}\geq 1} Z_{ijk}, b_{0k}+\sum_{i,j:X_{ijk}\geq 1}\xi_i\prod_{h=1}^k\gamma_h\right),$$

where $\mathbf{Z}_{[\mathbf{X}\geq\mathbf{1}]}=\{Z_{ijk}: X_{ijk\geq 1}\}$ and we integrate out $\{Z_{ijk}: X_{ijk}=0\}$.

Step 3. Sample $\gamma_1,\gamma_2,\ldots,\gamma_{M-1}$ from their conjugate full conditional distributions:

$$\pi(\gamma_h|\mathbf{Z}_{[\mathbf{X}\geq\mathbf{1}]},\theta_{(-\gamma_h)},\xi,\text{data}) = I_1-\mathcal{G}_{[1,\infty)}(\gamma_h;\tilde{\pi}_h,\tilde{a}_h,\tilde{b}_h),$$

where $\tilde{a}_h = a_h+\sum_{i,j,k:X_{ijk}\geq 1} 1_{(h<k)}Z_{ijk}$, $\tilde{b}_h = b_h+\sum_{i,j,k:X_{ijk}\geq 1} 1_{(h<k)}\xi_i\lambda_k\prod_{l:l\neq h}^k\gamma_l$,

$$\tilde{\pi}_h = \frac{\pi_{0h}\exp\left\{-\sum_{i,j,k:X_{ijk}\geq 1} 1_{(h<k)}\xi_i\lambda_k\prod_{l:l\neq h}^k\gamma_l\right\}}{\pi_{0h}\exp\left\{-\sum_{i,j,k:X_{ijk}\geq 1} 1_{(h<k)}\xi_i\lambda_k\prod_{l:l\neq h}^k\gamma_l\right\}+(1-\pi_{0h})\frac{C(a_h,b_h)}{C(\tilde{a}_h,\tilde{b}_h)}\frac{1-F(1;\tilde{a}_h,\tilde{b}_h)}{1-F(1;a_h,b_h)}}$$

Step 4. Sample ξ_i, for $i\in 1,\ldots,n$, from its full conditional distribution, which is

$$\pi(\xi_i|\mathbf{Z}_{[\mathbf{X}\geq\mathbf{1}]},\theta,\text{data}) = \mathcal{G}\,(\quad \xi_i;\nu^{-1}+\sum_{j,k:X_{ijk}\geq 1} Z_{ijk},\quad \nu^{-1}+\sum_{j,k:X_{ijk}\geq 1} X_{ijk}\lambda_k\left(\prod_{h=2}^k\gamma_h\right)^{1_{(1<k<K)}}\right),$$

Step 5. Update ν using a Metropolis step.

Step 6. Sample τ_1 and τ_2 from its full conditional distribution, which is Multinomial for τ_1 with probability of each element $t\in(a_{\tau 1},b_{\tau 1})$

$$P(\tau_1=t|\theta_{(-\tau_1)},\xi,\text{data}) = \frac{L(y|\theta_{(-\tau_1)},\xi,\tau_1=t)}{\sum_{s=a_{\tau 1}}^{b_{\tau 1}} L(y|\theta,\xi,\tau_2,\tau_1=s)}$$

and for τ_2 with probability of each element $t\in(a_{\tau 2},b_{\tau 2})$

$$P(\tau_2=t|\theta_{(-\tau_2)},\xi,\text{data}) = \frac{L(y|\theta_{(-\tau_2)},\xi,\tau_2=t)}{\sum_{s=a_{\tau 2}}^{b_{\tau 2}} L(y|\theta,\xi,\tau_1,\tau_2=s)}$$

where $L(y|\theta,\xi)$ is the Likelihood function

$$L(y|\theta,\xi) = \prod_{ij} \exp\left(-\left(X_{ijk}\lambda_k\left(\prod_{h=2}^{k}\gamma_h\right)^{1_{(1<k<K)}}\right)\right)^{Y_{ij}=0} \left\{1-\exp\left(-\left(X_{ijk}\lambda_k\left(\prod_{h=2}^{k}\gamma_h\right)^{1_{(1<k<K)}}\right)\right)\right\}^{Y_{ij}=1}$$

Step 7. Repeat steps 1-6 until apparent convergence and calculate posterior summaries based on a large number of additional iterations.

Appendix II

Algorithm for selecting decisions with lowest posterior risk among a high dimensional set of candidates

1. Initialization:

   (a) t:=0

   (b) Evaluate the posterior expected utility $U(\cdot)$ for a first parameter vector R of r discrete elements

2. Cycle for $j = 1, 2, \ldots$ until convergence:

   (a) select all new points in the neighborhood of R, that is:
   cycle for each element of the parameter vector, $i = 1, 2, \ldots, r$

      (i) obtain a new parameter vector moving only one element along one direction, $R'[i] := R[i] + 1$

      (ii) evaluate the posterior expected utility $U_i'(R')$

      (iii) update the parameters vector moving only one element along the other direction, $R''[i] := R[i] - 1$

(iv) evaluate the posterior expected utility $U_i''(R'')$

(b) in order to avoid to stack in local maxima, obtain r new parameter vectors $R_1''', \ldots, R_r'''$ at random and evaluate the posterior expected utility function $U_1''', \ldots, U_r'''$

(c) select R^* from the set of new parameters vectors $(R_1', \ldots, R_r', R_1'', \ldots, R_r'', R_1''', \ldots, R_r''')$ such that $U(R^*)$ is maximum.

(d) if $U(R^*) > U(R)$ update $R := R^*$ else stop.