Notice: This Material May Be Protected by Copyright Law (Title 17 U.S. Code). 2001 56

STATISTICS IN MEDICINE
Statist. Med. 2001; 20:3279–3294 (DOI: 10.1002/sim.965)

A statistical model for the evaluation of barrier contraceptive efficacy

Rosalie Dominik[1,*,†], Haibo Zhou[2] and Jianwen Cai[2]

[1] *Family Health International, P.O. Box 12950, Research Triangle Park, North Carolina, U.S.A.*
[2] *Department of Biostatistics, School of Public Health, University of North Carolina, Chapel Hill, North Carolina, U.S.A.*

SUMMARY

This paper describes an approach for the analysis of barrier contraceptive efficacy trials that accounts for timing frequency of intercourse and compliance. We allow exposure variables to vary for each act of intercourse and we control for timing of each act through a specific parametric function of the day of the act relative to the last day of the follicular phase of the cycle. The model can be used to examine the level of protection provided by a barrier versus no contraceptive method even when no control group of non-users is studied, as long as there are acts with no barrier use during the fertile window. We present results of a simulation study which examines performance of estimators and power under a variety of scenarios, including situations where an accurate benchmark for ovulation day is not available. As compared to the survival analysis approach commonly used in this setting, simulation results show that the new approach yields considerable gains in power to detect differences between the efficacy of contraceptive methods. An application to data from the FemCap® versus diaphragm trial show results consistent with previous findings suggesting superiority of the diaphragm but also provides new evidence of the per act protection provided by both methods. Copyright © 2001 John Wiley & Sons, Ltd.

INTRODUCTION

Barrier contraceptive methods are coital-dependent devices or drugs that provide a physical or chemical shield preventing viable sperm from entering the cervical canal. If the barrier functions as intended there is no opportunity for fertilization or subsequent pregnancy to occur. Examples of barrier contraceptive methods include male and female condoms, diaphragms, cervical caps and spermicidal gels or suppositories [1].

Barrier methods are an important category of contraceptive methods for several reasons. Some barrier contraceptive methods provide the extra benefit of reducing the risk of transmission of sexually transmitted infections (STI) as well as cervical cancer [2–4]. Barrier methods

* Correspondence to: Rosalie Dominik, Family Health International, P.O. Box 13950, Research Triangle Park, North Carolina 27709, U.S.A.
† E-mail: rdominik@fhi.org

Contract/grant sponsor: National Institute of Health; contract/grant number: R01 CA 79949, R29 HL-57444

Copyright © 2001 John Wiley & Sons, Ltd.

Received August 2000
Accepted March 2001

2

have few side-effects and are immediately reversible. In addition, since several barrier methods are available over the counter and instructions for use allow intercourse right after donning or insertion, they may be the only method of contraception available in some situations. Determining how well barrier methods prevent pregnancy is therefore an important public health research objective.

Barrier contraceptive effectiveness trials typically follow participants assigned to use a particular method until they either (i) become pregnant, (ii) discontinue participation in the trial for some other reason (for example, they are no longer in need of contraception or they decide to switch to some other method), or (iii) reach the end of some predefined follow-up period (for example, 6 or 12 months). Life-table estimators [5] of the cumulative risk of pregnancy through some fixed number of months or cycles of use are the most commonly used indicators of how well barrier contraceptive methods prevent pregnancy [6–10]. Using the life-table approach, the risk of pregnancy through interval k is estimated by $1-\hat{S}_k$ where $\hat{S}_k = \prod_{i=1}^{k} \hat{q}_k$ and $\hat{q}_k$ is an estimate of the probability of surviving interval k without pregnancy conditional on having not become pregnant prior to the start of interval k. The usual estimator for q_k is $(n_j' - d_j)/n_j'$ where n_j' is the number of women entering the interval minus half of those censored during the interval and d_j is the number with an estimated date of fertilization during the interval.

Survival analysis techniques that assume a continuous hazard, however, do not closely reflect the underlying biological process for pregnancy among barrier users. Each act of intercourse is actually a trial with a dichotomous outcome of pregnancy depending on a variety of act (for example, barrier use or cycle day), cycle (for example, egg viability) and couple specific (for example, anatomical characteristics) factors. When multiple acts occur within a cycle and pregnancy results, the act leading to the pregnancy cannot be identified and the outcome can only be measured at the cycle level. The use of analytic approaches that better reflect the biological process, with its repeated but aggregated trials, are expected to lead to improved evaluation of barrier contraceptive efficacy.

BACKGROUND

Statistical methods recently developed for reproductive toxicity research or extensions of these models may provide interpretive advantages over more traditional approaches of evaluating contraceptives [11]. These methods that model pregnancy as a mixture of two processes, cycle viability (yes/no) and pregnancy (yes/no) conditional on cycle viability, evolved from work by Barrett and Marshall [12] and Schwartz *et al.* [13].

Schwartz's extension of the Barrett–Marshall model is

$$\Pr(\text{conception in cycle } j \mid \{X_{jk}\}) = \mathscr{A}\left[1 - \prod_k (1-p_k)^{X_{jk}}\right] \tag{1}$$

where X_{jk} is an indicator variable equal to 1 if intercourse occurred on day k relative to ovulation in cycle j and 0 otherwise. The cycle viability parameter $\mathscr{A}$ represents the product of the probability a fertilizable ovule is produced and the conditional probability that if fertilized the egg will implant and develop until a pregnancy can be detected. The parameter p_k is the probability of conception resulting after intercourse on a single day k relative to ovulation

Copyright © 2001 John Wiley & Sons, Ltd.

3

conditional on the occurrence of a viable cycle. Note that the model implicitly assumes independence of outcomes of successive cycles within couples.

To identify factors affecting the probability of pregnancy among women attempting conception, Zhou and Weinberg [14] proposed a general marginal probability model incorporating covariate effects on $\mathscr{A}$ as well as on the p_k. To account for heterogeneity due to multiple cycles contributed by the same couple, they used a robust sandwich variance estimator for inference. The estimates of covariate effects on both the per day and per cycle pregnancy risk obtained from this approach are interpreted as differences in population-averaged response between groups of women with different risk factors.

To best evaluate barrier contraceptive efficacy we need a different model, one that better reflects that multiple acts of intercourse on a day can increase risk and that measures effects on the per act risk of pregnancy rather than the per day risk. In two recent barrier trials, couples reported multiple acts for over 12 per cent of the days with intercourse [9, 15]. When a barrier contraceptive is used during multiple acts on the same day, there are multiple opportunities for the barrier to malfunction, allowing sperm to break through the intended barrier. Also, exposure variables of primary interest might vary across acts within a day. For example, the assigned condom or an extra lubricant might be used during only one of two acts of intercourse on a given day.

Furthermore, the per day models mentioned above were developed in a fertility study setting where an accurate indicator of ovulation was available and where the study participants, who were trying to conceive, had a high risk of pregnancy. For example, in the Early Pregnancy Study (EPS) [16], day of ovulation was identified from results of hormonal assays on daily urine samples and pregnancy (including very early pregnancy loss) was detected in about one-third of women during their first cycle of study. In the EPS there was enough information to identify a narrow fertile window within the cycle and to estimate separate parameters for each day of the fertile window. However, pregnancy is far less frequent when barrier contraceptives are used. In barrier trials, only clinically recognizable pregnancies are typically detected and this event is usually observed in 1 per cent to 4 per cent of cycles. Also, barrier trials have not collected data necessary to accurately identify ovulation day and it may not be practical to gather such data in barrier trials. In the setting of barrier contraceptive evaluation, a robust modelling strategy with a small number of parameters is needed.

In the next section, we extend the Zhou and Weinberg [14] method to the setting of barrier contraceptive evaluation by replacing the daily probability component with a per act component and allowing a wider fertile window through parametric modelling of the day of the act relative to the end of the follicular phase (the day prior to ovulation, previously estimated to be the peak day of fertility). We subsequently present results of a simulation study which examines performance of estimators for treatment effects and power under a variety of scenarios, including situations where an accurate benchmark for day of ovulation is not available. We also provide an application of the modelling approach to data from a FemCap® versus diaphragm trial [15].

A MODEL FOR BARRIER CONTRACEPTIVE TRIALS

Suppose there are N couples in a barrier contraceptive study. For each couple i, where $i=1,2,\ldots,N$, there are binary outcomes Y_{ij} for $j=1,2,\ldots,n_i$ cycles; Y_{ij} is 1 if pregnancy

Copyright © 2001 John Wiley & Sons, Ltd.

4

occurred in cycle j and is 0 otherwise. The outcome of each cycle depends on $\mathscr{A}$ (the cycle viability parameter) and the characteristics of each act of intercourse within the cycle. Let X_{ijk} be a vector of act-level exposure variables of interest. For example, for a trial where couples are randomly assigned to use either contraceptive method A or B and they occasionally have unprotected intercourse, we are primarily interested in two exposure variables, $X_{\mathrm{A}ijk}$ equal 1 if method A is used and 0 otherwise and $X_{\mathrm{B}ijk}$ equal 1 if method B is used and 0 otherwise.

Also available for each act is information on the timing of intercourse within the cycle. The days prior to the day of ovulation comprise the follicular phase of a cycle and the days after the day of ovulation make up the luteal phase. Let Z_{ijk} be the cycle day relative to the last day of the follicular phase. Z_{ijk} is defined to be 0 for an act on the last day of the follicular phase, $-f$ for an act on the fth day prior to the last day and $+f$ for an act on the fth day after the last day of the follicular phase. We write the probability of pregnancy resulting from a single act of intercourse in a viable cycle, p_k, as $H\{\beta; X_{ijk}, g(Z_{ijk})\}$ where H is the inverse logit function yielding $p_k = \exp\{X_{ijk}\beta + g(Z_{ijk})\}/[1 + \exp\{X_{ijk}\beta + g(Z_{ijk})\}]$.

We assume a specific piecewise linear spline function for $g(\cdot) = \alpha_{\mathrm{F}} Z_{ijk}^{-} + \alpha_{\mathrm{L}} Z_{ijk}^{+}$ where $Z^{-} = -Z$ if $Z < 0$ and $Z = 0$ otherwise and $Z^{+} = Z$ if $Z > 0$ and $Z = 0$ otherwise. Note that $g(\cdot)$ equals 0 for an act of intercourse on the last day of the follicular phase. Previously obtained estimates of the probability of pregnancy per cycle day with unprotected intercourse during a viable cycle are close to zero until five days before the end of the follicular phase, then increase steadily to a probability near one for the last day of the follicular phase and then decline even more sharply back to zero [17]. We expect that this form for $g(\cdot)$ together with the inverse logit link for H will efficiently capture the relationship between timing of intercourse and pregnancy risk. For example, if X_{ijk} includes only an intercept, $\beta_0 = 3$ and $\{\alpha_{\mathrm{F}}, \alpha_{\mathrm{L}}\} = \{-1.75, -3\}$ then $H\{\beta; X_{ijk}, g(Z_{ijk})\}$ gives values which closely fit the available estimates of the risk of pregnancy by cycle day relative to ovulation for unprotected intercourse during the fertile window and near 0 probabilities for days outside this window. When only a rough benchmark of ovulation day is available and the risk per act rather than per day is modelled, the fertile window is expected to be wider (attenuated slopes for the cycle day covariates) with a slightly lower peak (smaller intercept).

Under the assumption that the probabilities of conception conditional on a viable cycle resulting from separate acts of intercourse are independent, our proposed model for conception in barrier contraceptive studies is

$$\Pr(Y_{ij} = 1 | \{X_{ijk}, Z_{ijk}^{-}, Z_{ijk}^{+}\}) = \mathscr{A}\left[1 - \prod_{k}^{n_{ij}} \{1 - \mathrm{logit}^{-1}(X_{ijk}\beta + \alpha_{\mathrm{F}} Z_{ijk}^{-} + \alpha_{\mathrm{L}} Z_{ijk}^{+})\}\right] \tag{2}$$

When the covariate vector X_{ijk} contains an intercept and the two contraceptive method use variables $X_{\mathrm{A}ijk}$ and $X_{\mathrm{B}ijk}$ defined above, then $\beta' = \{\beta_0 \ \beta_{\mathrm{A}} \ \beta_{\mathrm{B}}\}$. In barrier contraceptive research, we are interested in estimating and testing hypotheses about treatment effects included in β. The model allows us to measure and compare the reduction of risk for two barrier methods taking into account timing and frequency of intercourse through the Z_{ijk} and n_{ij}, respectively, as well as compliance through the X_{ijk}. For example, here $e^{\beta_{\mathrm{A}}}$ is interpreted as the ratio of the odds of conception resulting from a single act during a viable cycle for use of barrier A versus use of no barrier method, controlling for timing of intercourse, and $e^{\beta_{\mathrm{A}} - \beta_{\mathrm{B}}}$ is the odds ratio for use of method A versus use of method B. Note we do not include covariate

Copyright © 2001 John Wiley & Sons, Ltd.

5

effects on the cycle viability parameter $\mathscr{A}$ since barrier contraceptives are not designed or expected to affect cycle viability as it is defined above.

Asymptotically unbiased estimators for $\theta = \{\mathscr{A}, \beta_{p\times 1}, \alpha_{2\times 1}\}$ can be found by maximizing the pseudo-likelihood for the observed data

$$\prod_{i=1}^{N}\prod_{j=1}^{n_i}\left[\mathscr{A}\left\{1-\prod_{k}^{n_{ij}}\left(\frac{1}{1+\exp(X_{ijk}\beta+\alpha_{\mathrm{F}}Z_{ijk}^{-}+\alpha_{\mathrm{L}}Z_{ijk}^{+})}\right)\right\}\right]^{Y_{ij}}$$
$$\times\left[1-\mathscr{A}+\mathscr{A}\prod_{k}^{n_{ij}}\left(\frac{1}{1+\exp(X_{ijk}\beta+\alpha_{\mathrm{F}}Z_{ijk}^{-}+\alpha_{\mathrm{L}}Z_{ijk}^{+})}\right)\right]^{1-Y_{ij}} \tag{3}$$

We obtain an estimate $\hat{\theta}$ using a dual quasi-Newton–Raphson non-linear optimization subroutine (NLPQN in SAS PROC IML version 6.12). The associated gradient matrix is shown in Appendix A. To adjust for the correlation between different menstrual cycles for a given couple we used a robust variance estimator of the covariance matrix of $\hat{\theta} = M_0^{-1}M_1M_0^{-1}$ where M_0^{-1} is the usual estimator of the variance matrix when independence among all cycles is assumed, $M_1 = \sum_{i=1}^{N}\psi_i(\hat{\theta})\psi_i^{\mathrm{T}}(\hat{\theta})$, $\psi_i(\hat{\theta}) = \sum_{j=1}^{n_i}\frac{\mathrm{d}l_{ij}}{\mathrm{d}\hat{\theta}}$, and l_{ij} is the component of Y_{ij} in the pseudo-likelihood function (3). Detailed expression of l_{ij} is given in Appendix A (equation (A1)).

SIMULATION STUDY

We carried out a simulation study for a randomized trial comparing two barrier contraceptive methods through six cycles of use. We first generated cycle length, coital activity and product compliance data based on distributions described in the literature (Appendix B). The cycles generated had average follicular, luteal and total cycle lengths of 13.7, 13.9 and 28.6 days, respectively, and, on average, 9.5 acts of intercourse. The cycle viability parameter $\mathscr{A}$ was either fixed at 0.35 or varied across couples according to the distribution Beta($a = 1.46, b = 2.71$). Given the cycle viability parameter for a couple and the coital activity data relative to true day of ovulation, we assumed model (2) determines the probability of pregnancy. Specifically, we assumed $H(\beta, \alpha_{\mathrm{F}}, \alpha_{\mathrm{L}}, X_{ijk}, Z_{ijk}^{-}, Z_{ijk}^{+})$ is $\mathrm{logit}^{-1}(3+\beta_{\mathrm{A}}X_{\mathrm{A}ijk}+\beta_{\mathrm{B}}X_{\mathrm{B}ijk}-1.25Z_{ijk}^{-}-3.0Z_{ijk}^{+})$. For example, when no method is used this yields probabilities of 0.622, 0.953 and 0.500 that pregnancy results from a single unprotected act on days $-2, 0$ and $+1$ relative to the last day of the follicular phase, respectively, of a viable cycle. If barrier method A is used during the act and if $\beta_{\mathrm{A}} = -5$, these probabilities are reduced to 0.0110, 0.1192 and 0.003, respectively. If $\beta_{\mathrm{A}} = -6$ the probabilities are 0.0041, 0.0474 and 0.0001, respectively, and, if $\beta_{\mathrm{A}} = -7$, the probabilities are 0.0015, 0.0180 and <0.0001, respectively.

We examined the impact of (i) the sample size, (ii) the accuracy of the benchmark for day of ovulation, (iii) heterogeneity in cycle viability, (iv) the level of protection provided by the barrier and (v) the level of compliance on estimation and testing of efficacy parameters (β_{A} and β_{B}) in model (2). In addition, we contrast the proposed method with the usual survival analysis approach for comparing barrier contraceptive efficacy.

Copyright © 2001 John Wiley & Sons, Ltd.

6

When we assumed that the true day of ovulation was not available during analysis, we imputed ovulation day based on menstrual data alone. For all cycles that did not result in pregnancy the imputed last day in the follicular phase (that is, the day prior to ovulation) was calculated as cycle length minus 15. If the cycle resulted in pregnancy and if it was the second or a later cycle for the woman, then the imputed last day of the follicular phase was the length of the previous cycle minus 15. Otherwise (it the cycle resulted in a pregnancy and it is the first cycle for the woman), the imputed last day of the follicular phase was day 14 of the cycle.

Simulation results

Simulation results for scenarios of primary focus are provided in Table I. Model 1 includes cycle day covariates relative to the true last day of the follicular phase while model 2 includes cycle day covariates relative to the imputed day. We first present results for scenarios assuming the barrier is used in 95 per cent of acts of intercourse. Fitting the two models in scenarios assuming a fixed cycle viability parameter (Table I) and fitting the two models in scenarios where the cycle viability followed the beta distribution across couples (results not shown), on average, yielded nearly identical estimates of β and σ_β.

When the true day of ovulation was available, estimates of β_A and β_B were, on average, slightly biased away from the null, especially when a higher level of barrier effectiveness was assumed. However, increasing sample size reduced this bias considerably. Also, $\hat{\sigma}_{\beta_A}$ and $\hat{\sigma}_{\beta_B}$ were close to the standard deviations of $\hat{\beta}_A$ and $\hat{\beta}_B$, respectively, and coverage of all intervals was close to 95 per cent.

The unavailability of the true day of ovulation had a substantial impact on estimation of individual treatment effects; $\hat{\beta}_A$ and $\hat{\beta}_B$ were considerably biased towards 0, leading to poor coverage of intervals for β_A and β_B. However, since $\hat{\beta}_A$ and $\hat{\beta}_B$ were similarly attenuated, using the imputed day of ovulation had little effect on estimation of $\beta_A - \beta_B$ or the coverage of confidence intervals for this important contrast.

The assumed level of compliance, level of effectiveness and sample size together determined the number of events (that is, pregnancies) available for analysis. When compliance was assumed to be 95 per cent and sample size was 250 women per group there were on average 52, 38 and 32 pregnancies in treatment groups with assumed treatment betas of -5, -6 and -7, respectively (not shown). When compliance was 98 per cent the mean number of pregnancies dropped to 38, 23 and 16, respectively (not shown). Thus, when 98 per cent compliance was assumed, the standard errors for $\hat{\beta}_A$ and $\hat{\beta}_B$ were on average larger than when 95 per cent compliance was assumed, otherwise results were very similar to those shown in Table I.

For all scenarios in Table I, the chance of rejecting the null hypothesis of no treatment effect (for example, $H_0{:}\hat{\beta}_A = 0$) was 92 per cent or better (not shown). When neither treatment was assumed to be at all effective ($\beta_A = \beta_B = 0$) and sample size was 250 per group, the estimate of the type I error rate for testing the null of no individual treatment effect was slightly above the nominal value (about 7 per cent, not shown).

The estimated type I error rates for the test of $H_0{:}\beta_A - \beta_B = 0$ was 4.4 per cent or less for all scenarios in Table I where the two products were assumed to be equally effective. For scenarios in Table I where non-zero differences were assumed and the true day of

Copyright © 2001 John Wiley & Sons, Ltd.

1

Table I. Simulation results when cycle viability $\mathscr{A}$ is 0.35 for all couples*.

True values		Mean and SD of parameter estimates									Confidence interval % coverage		
		$\hat{\beta}_A$		$\hat{\sigma}_{\beta_A}$	$\hat{\beta}_B$		$\hat{\sigma}_{\beta_B}$	$\hat{\beta}_A - \hat{\beta}_B$		$\hat{\sigma}_{\beta_A - \beta_B}$			
β_A	β_B	Mean	SD	Mean	Mean	SD	Mean	Mean	SD	Mean	β_A	β_B	$\beta_A - \beta_B$
Model 1 results: true ovulation date (N=250 per treatment group)													
−5.0	−5.0	−5.27	(0.840)	0.833	−5.28	(0.847)	0.832	0.002	(0.320)	0.327	96.1	95.9	95.6
−5.0	−6.0	−5.26	(0.966)	0.909	−6.35	(1.067)	0.973	1.084	(0.447)	0.430	94.4	93.8	95.7
−6.0	−7.0	−6.52	(1.382)	1.227	−7.61	(1.478)	1.328	1.090	(0.797)	0.695	95.4	96.0	98.5
−7.0	−7.0	−7.64	(1.596)	1.549	−7.64	(1.618)	1.535	0.003	(0.894)	0.855	96.0	96.0	99.5
Model 1 results: true ovulation date (N=500 per treatment group)													
−6.0	−7.0	−6.17	(0.838)	0.742	−7.26	(0.892)	0.817	1.084	(0.526)	0.477	94.2	95.5	95.7
−7.0	−7.0	−7.27	(0.941)	0.882	−7.28	(0.921)	0.822	0.015	(0.666)	0.576	96.0	96.3	96.3
Model 2 results: imputed ovulation date (N=250 per treatment group)													
−5.0	−5.0	−3.74	(0.734)	0.761	−3.76	(0.709)	0.752	0.016	(0.803)	0.355	54.7	54.5	96.4
−5.0	−6.0	−3.76	(0.867)	0.801	−4.88	(0.887)	0.912	1.122	(0.895)	0.501	52.6	66.1	95.2
−6.0	−7.0	−5.00	(1.057)	1.008	−6.06	(1.357)	1.158	1.058	(1.073)	0.764	67.9	71.8	94.3
−7.0	−7.0	−6.09	(1.251)	1.248	−6.13	(1.357)	1.257	0.046	(0.924)	0.891	81.1	78.8	99.0
Model 2 results: imputed ovulation date (N=500 per treatment group)													
−6.0	−7.0	−4.77	(0.647)	0.630	−5.81	(1.019)	0.761	1.034	(1.031)	0.554	41.7	54.9	94.9
−7.0	−7.0	−5.86	(0.842)	0.766	−5.85	(0.863)	0.776	−0.013	(0.938)	0.644	60.8	61.0	97.0

*The assigned product is assumed to be used in 95 per cent of acts, otherwise no method is used. A total of 500 replicates were generated for each scenario.

Copyright © 2001 John Wiley & Sons, Ltd.

ovulation was available, the estimate of the power to reject this null was 74 per cent ($N=250$ per group, $\beta_A=-5$ and $\beta_B=-6$), 29 per cent ($N=250$ per group, $\beta_A=-6$ and $\beta_B=-7$) and 64 per cent ($N=500$ per group, $\beta_A=-6$ and $\beta_B=-7$). Power was clearly a function of the difference in the expected number of pregnancies in the two contraceptive method groups, which was 14, 6 and 12 for these three scenarios, respectively. The proposed approach yielded substantial gains in power to differentiate the efficacy of two contraceptive barrier methods compared to the traditional survival analysis techniques. Using the logrank test the power to detect differences between groups was 39 per cent, 9 per cent and 21 per cent for the three scenarios noted above. To help interpret these results as well as the implications of various assumptions about the treatment parameters β_A and β_B, it is useful to note that the 6-cycle life-table estimates of the probability of pregnancy for scenarios assuming 95 per cent compliance and treatment betas of -5, -6 and -7 were on average 20.8, 15.1 and 12.7 per 100 women. When 98 per cent compliance was assumed, the life-table estimates of the probabilities were on average 15.3, 9.1 and 6.3 per 100 women, respectively.

When ovulation day was imputed, the estimate of the power to reject the null of no difference between treatment groups was lower than the power observed when true ovulation date was available. However, the power (53 per cent), was still more than twice that observed for the life-table analysis (21 per cent) for the third scenario noted above ($N=500$ per group, $\beta_A=-6$ and $\beta_B=-7$).

In addition to examining the distribution of the estimates of the treatment parameters (β_A and β_B), it is of interest to note the impact of not knowing the true day of ovulation on the estimates of α_F, α_L and $\mathscr{A}$. When model 1 was fit to data generated under the scenario in Table I where $N=250$ per group, $\beta_A=-6$ and $\beta_B=-7$, the estimates of β_0, α_F and α_L were, on average, a bit further from zero than their true values (that is, the mean values were 3.49, -1.41 and -3.63, respectively). As expected, results suggest that the fertile window relative to the imputed day of ovulation is wider than the fertile window relative to the actual day of ovulation. When model 2 was fit to data generated under this same scenario, the estimates of β_0, α_F and α_L were on average, 1.29, -0.43 and -1.04, respectively. A plot of the probabilities of pregnancy after unprotected intercourse by cycle day relative to ovulation (assuming one act per day during a viable cycle) using these mean parameter estimates for model 2 would show a flatter curve than the one corresponding to the true underlying model relative to the actual day of ovulation (where $\beta_0=3.0$, $\alpha_F=-1.25$ and $\alpha_L=-3.0$). For this same scenario, the mean estimate of logit($\mathscr{A}$) was -0.56 when fitting model 1 and -1.035 when fitting model 2. These correspond to values of 0.362 and 0.262 for the cycle viability parameter $\mathscr{A}$.

We performed additional simulations where $\mathscr{A}$ was assumed to be known when fitting the model (either 0.25, 0.35 or 0.45), the true day of ovulation was assumed to be unavailable and only the β and α parameters in the model were estimated. Regardless of the assumed value of $\mathscr{A}$, treatment parameters β_A and β_B and their contrasts were on average biased towards the null. For example, for the scenario where N was 250 per group, $\beta_A=-6$ and $\beta_B=-7$ and we assumed $\mathscr{A}$ was 0.45, the estimates of β_A and β_B were on average -3.50 and -4.07. When we assumed $\mathscr{A}$ was 0.25, the estimates of β_A and β_B were on average -4.01 and -4.68. Regardless of the assumed value of $\mathscr{A}$, the power to detect differences between β_A and β_B was about 16 per cent for this scenario.

Copyright © 2001 John Wiley & Sons, Ltd.

9

AN APPLICATION

We applied our modelling strategy to data from a randomized trial of a new female barrier device (the FemCap®) and a diaphragm. After randomization to a device group, participants were to be followed for up to six months of product use unless they became pregnant, were discontinued for some other medical reason or chose to leave the study for personal reasons. Participants maintained daily diaries where they recorded information on the contraceptive method used during each act of intercourse and identified each day with menstrual bleeding. Details of the study design and results of the planned survival analyses have been reported elsewhere [15]. Barrier trials do not typically collect hormonal data. The first challenge in implementing any cycle based analysis involves defining the rules for identifying what will be considered a 'cycle' given available data on bleeding. We used the same rules to identify eligible cycles as those used for other cycle based analyses previously completed for the trial data, except that we also excluded cycles (4 per cent) with more than 20 coital acts. For example, any sequence of two or more bleeding days was considered a bleeding episode and any single non-bleeding day between two bleeding days was treated as a bleeding day. Cycles were defined as a sequence of days beginning with the first day of a bleeding episode and ending with the last day before the next bleeding episode. Also, rules for defining the end date for cycles resulting in pregnancy were established.

There were 28 pregnancies in 1063 eligible cycles available for users of the new device and 21 pregnancies in 1324 eligible cycles available for users of the diaphragm. Coital diary data indicated occasional use of another barrier method (for example, the male condom) instead of (4–5 per cent in both groups) or in addition to (<1 per cent in both groups) the assigned device. However, for most acts the assigned method was the only method used (89 per cent in the new method group and 92 per cent in the diaphragm group). No method was used in 4.4 per cent of acts in the new method group and 3.6 per cent of acts in the diaphragm group. For our analysis, method used during act k of cycle j for participant i was captured by two variables. The first X_{Aijk}, was equal to 1 if the woman was assigned the new barrier and used any barrier during the act, and was 0 otherwise. The second variable, X_{Bijk}, was equal to 1 if the woman was assigned the diaphragm and used any barrier during the act, and was 0 otherwise. That is, both variables were coded 0 if no barrier was used.

In this trial, an accurate benchmark for ovulation day was not available, so we first applied the same rule for imputing day of ovulation as discussed above. Except for the slight difference in our definition of the contraceptive method use variables, we attempted to fit model 2 above to the data and to estimate $\theta=\{\mathscr{A},\beta_{p\times 1},\alpha_{2\times 1}\}$ where α_F and α_L are the effects for each day prior to and after the imputed day of ovulation, respectively. The full model failed to converge. Since our main interest is in the treatment parameters and results of previous analyses on pregnancy risk among non-contracepting women give us some information about the value $\mathscr{A}$, we proceeded by fixing $\mathscr{A}$ equal to 0.35. For this model, estimates of both treatment parameters were significantly different from 0 showing evidence of the efficacy provided by each of the two barrier methods (Table II). The result of the test of whether $\beta_A-\beta_B$ was 0 neared statistical significance. We also fit a model where the algorithm for imputing day of ovulation was modified slightly by moving the imputed last day of the follicular phase one day earlier. This alternative was considered because of the reported inverse relationship between age and follicular phase length [18] and the observed age distribution of the trial population. Using the second imputation approach yielded an estimate of β_0 closer to the

Copyright © 2001 John Wiley & Sons, Ltd.

10

Table II. Parameter estimates obtained for the FemCap® versus diaphragm trial data.

Parameter	Interpretation	Estimate obtained (or *fixed* value assumed)			
		Model using original ovulation day algorithm		Model using modified ovulation day algorithm	
		Parameter	SE	Parameter	SE
$\mathscr{A}$	Chance that cycle is viable	*0.35*	fixed	*0.35*	fixed
β_0	Log-odds of pregnancy for single act of intercourse on day prior to ovulation in a viable cycle	−1.852	0.7380	−1.073	0.6270
β_A	Change in log odds for FemCap versus no method	−1.276	0.5999	−2.288	0.5977
β_B	Change in log-odds for diaphragm versus no method	−1.892	0.6337	−2.955	0.6099
α_F	Change in log-odds for each day prior to last day of follicular phase	−0.190	0.0950	−0.445	0.2591
α_L	Change in log-odds for each day past last day of follicular phase	−13.600	1.670	−0.283	0.0952
$\beta_A-\beta_B$	Change in log-odds for FemCap versus diaphragm	0.615	0.3456	0.667	0.3592

Null hypothesis	Interpretation of null hypothesis	*p*-value*	*p*-value*
$\beta_A=0$	FemCap provides no protection against pregnancy during an act of intercourse	0.0167	0.0001
$\beta_B=0$	Diaphragm provides no protection against pregnancy during an act of intercourse	0.0014	<0.0001
$\beta_A-\beta_B=0$	There is no difference in protection provided by FemCap and diaphragm	0.0750	0.0632

*For one-sided tests of $\beta_A=0$ and $\beta_B=0$ and a two-sided test of $\beta_A-\beta_B=0$.

Copyright © 2001 John Wiley & Sons, Ltd.

value expected given previous data on pregnancy risk for the peak day of fertility among non-contracepting women. However, when this modified imputation approach was used the estimate of α_L suggested a much slower decline in the risk of pregnancy throughout the luteal phase than the estimate of α_L obtained in the original analysis. Such a change in the estimate of α_L is not entirely unexpected since previous studies have shown that the risk of pregnancy drops quickly after it peaks on the last day of the follicular phase. Although the estimate of α_L was quite sensitive to the imputation rule used, statistical conclusions for tests of hypotheses about β_A and β_B were unchanged. In addition, parameter estimates changed little (not shown) when we assumed other values of $\mathscr{A}$ (that is, 0.25, 0.30, 0.40).

The estimate of the odds ratio for pregnancy resulting from a single act of intercourse with the FemCap® versus the diaphragm was 1.85 using the original imputation algorithm and 1.95 using the alternative algorithm. Since the estimate of the risk of pregnancy is quite low for acts with the diaphragm regardless of the cycle day relative to ovulation, it is reasonable to interpret this odds ratio as a relative risk.

It is of interest to note that a logrank test of the difference in life-table estimates of the pregnancy curves through six cycles yielded nearly identical statistical conclusions for the test of no difference between groups ($p = 0.070$). Life-table estimates of the risk of pregnancy through six cycles using the cycles included in this analysis were 14.2 and 8.8 per 100 women for the new device and the diaphragm, respectively. In summary, the application of our modelling approach yields conclusions about the relative effectiveness of the FemCap and diaphragm that are similar to those obtained in the primary analysis of the trial [15] but also provides new evidence of the per act protection against pregnancy provided by each barrier method.

DISCUSSION

We proposed a barrier contraceptive analysis strategy that better reflects the underlying biological process of pregnancy than the traditional survival analysis strategy. Our simulation results suggest that by using our model which accounts for timing and frequency of intercourse and barrier method use, we can learn more about the absolute and relative efficacy of barrier contraceptive methods than the usual survival analysis allows, especially if an accurate indicator for ovulation day is available. The gains in power to detect a difference in the protection provided by two barrier methods for this approach relative to the usual survival approach were quite large when data were assumed to arise from a model of the same form as the fitted model or from a model that assumed each couple has their own cycle viability parameter. Since other plausible models for pregnancy among barrier contraceptive users would follow a similar form as the model we evaluated, we expect that reasonable gains in power would still result even if other assumptions of our model do not hold exactly. Compared to the traditional survival approach, this analysis approach offers the advantage of estimating the absolute efficacy of a barrier method even when there is no control group of non-users as long as there are some acts with no method use during the fertile window of the cycle.

We examined estimates obtained under realistic sample size and compliance assumptions. Treatment betas within the interval -5 to -7 were of primary focus because along with the other assumptions they lead to six-cycle pregnancy probabilities varying from about 0.06 to 0.21. This interval includes the 6-month pregnancy probability estimates observed for typical

Copyright © 2001 John Wiley & Sons, Ltd.

12

users of a number of barrier methods. On average, estimates of the difference between efficacy parameters for the two barrier methods being compared were close to their true values across most scenarios examined and coverage was excellent for the confidence interval of this contrast of primary interest. Estimates of the effects of barrier use on the per act risk of pregnancy were invariant to assumptions about heterogeneity of the cycle viability parameter which is consistent with results observed in secondary analyses of the EPS data [19].

For some data sets generated, parameter estimates did not converge and this was more likely to occur for scenarios yielding lower numbers of pregnancies compared to scenarios yielding higher numbers of pregnancies. Also, convergence problems were more likely when ovulation day was imputed. For the scenarios in Table I where $N = 250$ per group, $\beta_A = -6$ and $\beta_B = -7$, convergence was reached for about 91 per cent of the data sets generated when fitting model 1 (which included true ovulation day) and for about 67 per cent of data sets when fitting model 2 (which included imputed ovulation day). Doubling the sample size substantially improved the chance of convergence to 82 per cent when fitting model 2. When convergence problems are encountered in an application, it may be helpful to assume that the cycle viability parameter is known. When the analysis objective is to assess or compare the per act level of protection provided by two different barrier methods, the cycle viability parameter $\mathscr{A}$ is in a sense a nuisance parameter and inference about $\mathscr{A}$ is at most of secondary interest. Even though simulation results suggested that treatment effects and their contrasts are underestimated when $\mathscr{A}$ is assumed to be known, our application of this approach to the data from a trial comparing the FemCap and the diaphragm showed strong evidence of the absolute efficacy of both products and some evidence of a difference between products.

Our simulation study results demonstrate that it would be preferable to gather data that allow more accurate identification of the ovulation day in barrier contraceptive studies if the analysis will control for timing of intercourse using our proposed modelling approach. When an accurate indicator of day of ovulation is not available, further extensions of the model, rather than simply using the best guess for day of ovulation, may better reflect the biological process. For example, one might fit a mixed distribution model where follicular phase length is treated as random. However, given the number of pregnancies typically observed in samples sizes considered reasonably attainable for barrier trials, such trials may not provide sufficient information to reliably estimate the additional parameters needed in these more complex models. The results we obtained when we applied a simple imputation approach for day of ovulation suggested there may be value in using a model of this form even when the only information on timing of ovulation is in the cycle length data. Future research on multiple imputation techniques may identify an even better strategy for situations where day of ovulation is not measured directly. Also, our simulations assumed accurate coital activity data and data for identifying cycles. Inaccuracies in coital activity data and cycle length data will likely reduce potential gains in power and more research is needed to examine the impact of these other types of measurement error. Procedures for assuring the quality of coital activity data are needed.

We accounted for correlation between cycles contributed by the same couple through the use of the population average model with a robust variance estimator. Another approach for accounting for heterogeneity in cycle viability for barrier trials would have been to extend the random-effects model previously proposed for fertility studies [20]. However, the interpretation of the act-specific effects in such a model would be the same as for the population average model.

Copyright © 2001 John Wiley & Sons, Ltd.

13

Our model takes the independence assumption of the Barrett–Marshall model a bit further by assuming independence of all acts within a cycle, even if they occurred on the same day. This is actually not much different than the Barrett–Marshall assumption since two acts on adjacent days may be in fact closer in time than two acts on the same day. In the FemCap® versus diaphragm trial, multiple acts were reported on about 12 per cent of the days with coital acts. We recommend that applications of this approach consider sensitivity analyses where, for example, acts within days are collapsed or a covariate indicating whether the act was the first of the day is added. In addition, use of chemical barriers (or devices such as the FemCap or the diaphragm that are used with a spermicide) during an act of intercourse may have a carry-over effect on the risk of pregnancy for subsequent acts and additional sensitivity analyses may be useful in this setting. Occasional use of emergency contraceptive back-up in a barrier trial introduces additional complexities. However, if the only reason for EC back-up in a trial is occasional non-use of the assigned barrier, simply excluding cycles with EC back-up may still allow valid estimation of barrier efficacy parameters.

We recognize that any analysis incorporating compliance data is inconsistent with an intention-to-treat (ITT) strategy and should thus be exercised with caution. If participant characteristics related to compliance are also related to cycle viability or the per act probability of pregnancy during unprotected intercourse on a specific day relative to ovulation, then it may be necessary to identify and control for such characteristics as well. For example, including an additional act-level covariate for a variable related to compliance in the barrier trial data set and also shown in prior studies to be related to fertility (for reasons other than timing and frequency of intercourse) should help to reduce bias. If we were solely concerned with barrier contraceptive *effectiveness* (a function of the inherent protection provided by the barrier if used as well as the coital and use patterns) rather than *efficacy* (a function of the inherent protection alone), analysis strategies that incorporate timing and frequency of intercourse would not be necessary. However, our interest goes well beyond the simpler effectiveness measures, especially since typical use patterns are most likely a function of perceived level of protection provided by a method.

In conclusion, analysis strategies incorporating compliance as well as timing and frequency of intercourse should be employed, at least as supporting approaches in barrier trials. If a barrier provides reasonable protection against pregnancy, the signal from a six-cycle trial with 250 complete observations per group will almost certainly be strong enough for the model to detect that effect even if day of ovulation is imputed based on cycle length data alone. Furthermore, collecting data that would allow accurate identification of the day of ovulation could substantially improve our ability to sort out the efficacy of two barrier methods and should be considered for future barrier trials.

APPENDIX A

Define $V_{ijk}=\{X_{ijk}\ Z^{-}_{ijk}\ Z^{+}_{ijk}\}, \theta^{*}=\{\beta_{p\times 1},\alpha_{2\times 1}\}$ and $\theta=\{\mathscr{A},\beta_{p\times 1},\alpha_{2\times 1}\}$. Under model (2), the pseudo-likelihood for the observed data is

$$L=\prod_{i=1}^{N}\prod_{j=1}^{n_i}\left[\mathscr{A}\left(1-\prod_{k}^{n_{ij}}\left(\frac{1}{1+\exp(V_{ijk}\theta^{*})}\right)\right)\right]^{Y_{ij}}\left[1-\mathscr{A}+\mathscr{A}\prod_{k}^{n_{ij}}\left(\frac{1}{1+\exp(V_{ijk}\theta^{*})}\right)\right]^{1-Y_{ij}}$$

Copyright © 2001 John Wiley & Sons, Ltd.

14

and the corresponding log-likelihood is $l=\sum_{i=1}^{N}\sum_{j=1}^{n_i} l_{ij}$, where the contribution from cycle j of couple i is

$$\begin{aligned} l_{ij} = {} & Y_{ij}\ln\mathscr{A} + Y_{ij}\ln\left[1-\prod_{k}^{n_{ij}}\left(\frac{1}{1+\exp(V_{ijk}\theta^*)}\right)\right] \\ & +(1-Y_{ij})\ln\left[1-\mathscr{A}+\mathscr{A}\prod_{k}^{n_{ij}}\left(\frac{1}{1+\exp(V_{ijk}\theta^*)}\right)\right] \end{aligned} \tag{A1}$$

The gradient vector is

$$\frac{\mathrm{d}l}{\mathrm{d}\theta} = \begin{bmatrix} \frac{\mathrm{d}l}{\mathrm{d}\mathscr{A}} \\ \frac{\mathrm{d}l}{\mathrm{d}\theta^*_{(1)}} \\ \frac{\mathrm{d}l}{\mathrm{d}\theta^*_{(2)}} \\ \cdot \\ \cdot \\ \frac{\mathrm{d}l}{\mathrm{d}\theta^*_{(p+2)}} \end{bmatrix}$$

where $\frac{\mathrm{d}l}{\mathrm{d}\mathscr{A}}=\sum_{i=1}^{N}\sum_{j=1}^{n_i}\frac{\mathrm{d}l_{ij}}{\mathrm{d}\mathscr{A}}$ and $\frac{\mathrm{d}l}{\mathrm{d}\theta^*_{(h)}}=\sum_{i=1}^{N}\sum_{j=1}^{n_i}\frac{\mathrm{d}l_{ij}}{\mathrm{d}\theta^*_{(h)}}$ $h=1,2,\ldots,p+2$.

In the above

$$\frac{\mathrm{d}l_{ij}}{\mathrm{d}\mathscr{A}} = \frac{Y_{ij}}{\mathscr{A}} + (1-Y_{ij})[1-\mathscr{A}+\mathscr{A}Q_{ij}]^{-1}[Q_{ij}-1]$$

and

$$\begin{aligned} \frac{\mathrm{d}l_{ij}}{\mathrm{d}\theta^*_{(a)}} = {} & Y_{ij}[1-Q_{ij}]^{-1}Q_{ij}\sum_{k=1}^{n_{ij}}\frac{V_{ijk(a)}\exp(V_{ijk}\theta^*)}{1+\exp(V_{ijk}\theta^*)} \\ & -(1-Y_{ij})[1-\mathscr{A}+\mathscr{A}Q_{ij}]^{-1}\mathscr{A}Q_{ij}\sum_{k=1}^{n_{ij}}\frac{V_{ijk(a)}\exp(V_{ijk}\theta^*)}{1+\exp(V_{ijk}\theta^*)} \end{aligned}$$

with $Q_{ij}=\prod_{k}^{n_{ij}}\left(\frac{1}{1+\exp(V_{ijk}\theta^*)}\right)$.

Note that in all analyses, the cycle viability parameter $\mathscr{A}$ in likelihood was replaced by $\mathrm{logit}^{-1}(\mathbf{a})$ to force $\mathscr{A}$ to be between 0 and 1. Derivatives of the likelihood were modified accordingly as $\frac{\mathrm{d}l_{ij}}{\mathrm{d}\mathbf{a}}=\frac{\mathrm{d}l_{ij}}{\mathrm{d}\mathscr{A}}\frac{\mathrm{d}\mathscr{A}}{\mathrm{d}\mathbf{a}}$.

APPENDIX B

We generated data assuming a randomized trial comparing the effectiveness of two barrier contraceptive methods through six cycles of use and an underlying model of the form of (2) determines whether or not a given cycle results in pregnancy. We used cycle length, coital

Copyright © 2001 John Wiley & Sons, Ltd.

15

activity and product compliance assumptions based on distributions described in the literature [6–10, 15, 18, 21]. The steps below are carried out for each replicate of a scenario.

1. For each subject, age was generated from a Uniform [20, 35] distribution.
2. Also, for scenarios assuming a subject specific cycle viability parameter, $\mathscr{A}_i$ was randomly generated for each subject from the distribution $\text{Beta}(\mu = 0.433, \sigma^2 = 0.044) = \text{Beta}(a = 1.46, b = 2.71)$ [20].
3. Six pairs of luteal and follicular phase lengths were generated for each subject. First, a log mean follicular phase length (LOGF_i) was generated for each subject from a normal distribution conditional on age, $\text{Normal}(\mu = (2.653 - (\text{Age} - 20) \times 0.01), \sigma = 0.24)$. Next, six log follicular phase lengths were generated for each participant from a $\text{Normal}(\mu = \text{LOGF}_i, \sigma = 0.15)$. Also, each subject was randomly assigned a subject specific probability (PL_i) that she had a cycle with a luteal phase coming from distribution 1 which was $\text{Normal}(\mu = 9.21, \sigma = 1.41)$ instead of the more common distribution 2 which was $\text{Normal}(\mu = 14.13, \sigma = 1.41)$. If age was less than 25 years old then PL_i was generated from $\text{Beta}(\mu = 0.09, \sigma^2 = 0.005)$. If age was 25 or older then PL_i was generated from $\text{Beta}(\mu = 0.04, \sigma^2 = 0.005)$. Given PL_i, six luteal phase lengths were generated for each subject by randomly determining whether each specific luteal phase length comes from distribution 1 or 2 and then generating the actual luteal phase length from the applicable distribution. Six cycle lengths were created for each woman by summing the corresponding luteal and follicular phase lengths plus one for day of ovulation.
4. For each day of each cycle, whether or not the subject had 0, 1 or 2 acts of intercourse was randomly determined assuming an average of 10 acts per 30 day interval and that 10 per cent of the days with intercourse had two acts of intercourse.
5. For each act, whether the assigned contraceptive method was used or whether no method was used was randomly determined given the assumed value of the per cent of acts with compliance for the scenario.
6. Given the coital data for each act k of cycle j of woman i, a pregnancy probability p_{ij} was determined for each cycle using model (2) and the assumed value of β for the scenario. ($\mathscr{A}_i$ replaced $\mathscr{A}$ in model (2) for scenarios that assumed heterogeneity in cycle viability). Given this pregnancy probability, whether or not a pregnancy occurred in the cycle (Y_{ij}) was randomly generated. All cycles generated for a woman through her first pregnant cycle were used for data analysis.

ACKNOWLEDGEMENTS

The authors thank Elizabeth Raymond and James Trussell for their helpful comments and the Contraceptive Research and Development (CONRAD) Program for use of the FemCap® versus diaphragm trial data. Acknowledgment goes to the US Agency for International Development for their support of several barrier contraceptive trials that motivated this work. This research was supported in part by the National Institute of Health grants R01 CA 79949 (H.Z.) and R29 HL-57444 (J.C).

REFERENCES

1. Hatcher RA, Trussell J, Stewart F, Cates Jr. W, Stewart GK, Guest F, Kawal D. *Contraceptive Technology*. 17th revised edn. Ardent Media, Inc.: New York, 1998.

Copyright © 2001 John Wiley & Sons, Ltd.

16

2. Slattery ML, Overall JC Jr, Abbott TM, French TK, Robison LM, Gardner J. Sexual activity, contraception, genital infections, and cervical cancer: support for a sexually transmitted disease hypothesis. *American Journal of Epidemiology* 1989; **130**(2):248–258.
3. Cates W Jr, Stone KM. Family planning, sexually transmitted diseases and contraceptive choice: a literature update-Part I. *Family Planning Perspectives* 1992; **24**(2):75–84.
4. Feldblum PJ, Morrison CS, Roddy RE, Cates W Jr. The effectiveness of barrier methods of contraception in preventing the spread of HIV. *AIDS* 1995; **9**(Suppl A):S85–S93.
5. Cutler SJ, Ederer F. Maximum utilization of the life table in analyzing survival. *Journal of Chronic Diseases* 1958; **8**(6):699–713.
6. Farr G, Gabelnick H, Sturgen K, Dorflinger L. Contraceptive efficacy and acceptability of the female condom. *American Journal of Public Health* 1994; **84**(12):1960–1964.
7. Bernstein GS, Clark V, Coulson AH, Frezieres RG, Kilzer L, Moyer D, Nakamura RM, Walsh T. Use effectiveness of cervical caps. Final Report. National Institute of Child Health and Human Development: Washington, D.C., 1986.
8. Mauck C, Glover LH, Miller E, Allen S, Archer DF, Blumenthal P, Rosenzweig BA, Dominik R, Sturgen K, Cooper J, Fingerhut F, Peacock L, Gabelnick HL. Lea's Shield: a study of the safety and efficacy of a new vaginal barrier contraceptive used with and without spermicide. *Contraception* 1996; **53**(6):329–335.
9. Raymond E, Dominik R and the Spermicide Trial Group. Contraceptive effectiveness of two spermicides: a randomized trial. *Obstetrics and Gynecology* 1999; **93**(6):896–903.
10. Nelson A, Bernstein GS, Frezieres R, Walsh T, Clark V, Coulson A. Study of the efficacy, acceptability and safety of a non-latex (polyurethane) male condom. Final Report. National Institute of Child Health and Human Development: Bethesda, MD, 1997.
11. Weinberg CR, Zhou H. Model-based approaches to studying fertility and contraceptive efficacy. *Advances in Contraception* 1997; **13**(2–3):97–103.
12. Barrett JC, Marshall J. The risk of conception on different days of the menstrual cycle. *Population Studies* 1969; **23**:455–461.
13. Schwartz D, MacDonald PDM, Heuchel V. Fecundability, coital frequency and the viability of ova. *Population Studies* 1980; **34**:397–400.
14. Zhou H, Weinberg CR. Modeling conception as an aggregated Bernoulli outcome with latent variables via the EM algorithm. *Biometrics* 1996; **52**(3):945–954.
15. Mauck C, Callahan M, Weiner DH, Dominik R and the FemCap® Investigators' Group. A comparative study of the safety and efficacy of FemCap®, a new vaginal barrier contraceptive, and the Ortho All-Flex® diaphragm. *Contraception* 1999; **60**(2):71–80.
16. Wilcox AJ, Weinberg CR, O'Connor JF, Baird DB, Schlatterer JP, Canfield RE, Armstrong EG, Nisula BC. Incidence of early loss of pregnancy. *New England Journal of Medicine* 1988; **319**(4):189–194.
17. Rodríguez G, Trussell J. Fitting the Schwartz-Barrett–Marshall model. Working Paper No. 97-3 ed. Princeton University Office of Population Research: Princeton, NJ, 1997.
18. Lenton EA, Landgren B-M, Sexton L, Harper R. Normal variation in the length of the follicular phase of the menstrual cycle: effect of chronological age. *British Journal of Obstetrics and Gynecology* 1984; **91**(7): 681–684.
19. Zhou H, Weinberg CR. Potential for bias in estimating human fecundability parameters: a comparison of statistical models. *Statistics in Medicine* 1999; **18**(4):411–422.
20. Zhou H, Weinberg CR, Wilcox AJ, Baird DD. A random-effects model for cycle viability in fertility studies. *Journal of the American Statistical Association* 1996; **91**(436):1413–1422.
21. Lenton EA, Landgren B-M, Sexton L. Normal variation in the length of the luteal phase of the menstrual cycle: identification of the short luteal phase. *British Journal of Obstetrics and Gynecology* 1984; **91**(7):685–689.

Copyright © 2001 John Wiley & Sons, Ltd.

17