

Methodologic and Statistical Approaches to Studying Human Fertility and Environmental Exposure

Candace Tingen,¹ Joseph B. Stanford,² and David B. Dunson¹

¹Biostatistics Branch, National Institute of Environmental Health Sciences, National Institutes of Health, Department of Health and Human Services, Research Triangle Park, North Carolina, USA; ²Health Research Center, Department of Family and Preventive Medicine, University of Utah, Salt Lake City, Utah, USA

Although there has been growing concern about the effects of environmental exposures on human fertility, standard epidemiologic study designs may not collect sufficient data to identify subtle effects while properly adjusting for confounding. In particular, results from conventional time to pregnancy studies can be driven by the many sources of bias inherent in these studies. By prospectively collecting detailed records of menstrual bleeding, occurrences of intercourse, and a marker of ovulation day in each menstrual cycle, precise information on exposure effects can be obtained, adjusting for many of the primary sources of bias. This article provides an overview of the different types of study designs, focusing on the data required, the practical advantages and disadvantages of each design, and the statistical methods required to take full advantage of the available data. We conclude that detailed prospective studies allowing inferences on day-specific probabilities of conception should be considered as the gold standard for studying the effects of environmental exposures on fertility. *Key words:* conception, fecundability, menstrual cycle, ovulation, reproductive epidemiology, statistical methods, study design, time to pregnancy. *Environ Health Perspect* 112:87–93 (2004). doi:10.1289/ehp.6263 available via <http://dx.doi.org/> [Online 24 September 2003]

There is increasing concern about the effects of environmental exposures on human fertility (Baird and Strassmann 2000). At least 10% of couples in the United States have had difficulty achieving pregnancy (Chandra and Stephen 1998). Investigators are worried that fertility may be declining, and there is corresponding concern in the general public (Carlsen et al. 1992; Pearce et al. 1999; Swan et al. 2000; United Nations 1997). The increased public focus on fertility problems has resulted partly from the increasing numbers of women who delay attempting pregnancy until their middle to late 30s, ages at which a substantial proportion of couples will fail to conceive within a year and hence be categorized as clinically infertile (Dunson et al. In press). Many of these couples will resort to assisted reproduction techniques, which pose potential concerns about safety and impact on perinatal and child health (Mitchell 2002). Despite broad interest in the scientific community and in the general public, surprisingly little is known about key factors related to human fertility and fecundity, such as age, environmental exposures, sexual behavior, and lifestyle (Joffe 2003; Olsen and Rachootin 2003). In this article we first review broadly the factors known to affect fertility. We then discuss methodologic and statistical issues involved in studying fecundity, with an emphasis on the advantages, necessary design elements, and statistical methods for detailed prospective preconception cohort studies. We also comment on the need to integrate the study of human fecundity with the study of other aspects of human reproduction and development.

Throughout the article, we use the term “fecundity” to refer to a couple’s probability of pregnancy with regular intercourse without the use of contraception. In other words fecundity is the inherent capacity to conceive. Depending on the context, fecundity can be assessed for women, for men, or for couples. The related term from demography, “fecundability,” is the specific probability of conception within a single menstrual cycle with noncontracepted intercourse. We use the term “fertility” to refer to the ability of a couple to achieve a pregnancy that survives to birth.

Factors Affecting Fertility

Age and environmental exposures. It is generally accepted that female fecundity declines with age (Sauer 1998). However, limited data are available on the rate of decline (Schwartz and Mayaux 1982; Stovall et al. 1991; van Noord-Zaadstra et al. 1991) and on factors contributing to the decline (Abdalla et al. 1997; Rosenwaks et al. 1995). Even less is known about aging effects on male fecundity, with the available data pertaining mostly to declines in the elderly years (Kidd et al. 2001). A recent study reported that female fecundity starts to decline in the late 20s and male fecundity in the late 30s, controlling for timing of intercourse (Dunson et al. 2002), but more data are needed to validate this result and investigate causes. In particular, little is known about the impact of environmental exposures on the variability in fecundity among young couples and in the rate of decline with age. Some studies have reported lower fecundity associated with environmental factors, such as parental consumption of contaminated fish (Buck et al.

2000) and exposure to lead (Apostoli et al. 2000; Sallmen et al. 1995), pesticides (Curtis et al. 1999; Larsen et al. 1998; Thonneau et al. 1999), organic and chemical solvents (Sallmen et al. 1998; Wennborg et al. 2001), and cigarette smoking (Weinberg et al. 1989). However, in studies to date, exposure has been assessed only retrospectively, and these results were based mostly on small sample sizes.

Sexual behavior. One of the main difficulties in studying human fertility is the large behavioral component. There is a tremendous interplay between behavior and biology, both of which need to be considered when assessing etiologic end points. The ages at which couples attempt conception vary substantially between different socioeconomic and ethnic groups (Morabia and Costanza 1998; O’Connell and Rogers 1982; Pearce et al. 1999; Taffel 1977). Over the last several decades there has been a steady increase in the age of the mother at first birth (Morabia and Costanza 1998; Pearce et al. 1999; Ventura et al. 2000, 2001), largely due to women delaying childbirth while focusing on careers. Such trends may be more prevalent among couples in certain demographic groups, making it important to carefully adjust for age and behavior in analyses of environmental effects. In particular, including only age as a covariate in a time to pregnancy (TTP) model may not adequately adjust for differences between groups in the timing and frequency of intercourse. Fertility data analysis is also biased by the “survival” effect, where more fertile couples conceive early in their reproductive years, resulting in an age-dependent increase in the proportion of subfertile couples among the couples attempting pregnancy.

Because there are no realistic animal models (Amann 1982; Working 1988) or universally accurate biomarkers (Barnhart and Osheroff

This article is part of the mini-monograph “Understanding the Determinants of Children’s Health.”

Address correspondence to D. Dunson, Biostatistics Branch, NIEHS, MD A3-03, NIH, DHHS, PO Box 12233, Research Triangle Park, NC 27709 USA. Telephone: (919) 541-3033. Fax: (919) 541-4311. E-mail: dunson1@niehs.nih.gov

We acknowledge the helpful comments of members of the Fertility and Early Pregnancy Working Group of the National Children’s Study.

The authors declare they have no competing financial interests.

Received 6 February 2003; accepted 2 September 2003.

1998; 1999; Berardono et al. 1993; Scott and Hofmann 1995) of human fecundity, it is necessary to study humans attempting pregnancy. The number of menstrual cycles of noncontracepting intercourse required to achieve conception, or the TTP, is a useful, commonly employed measure of a couples' fecundity. However, there are a number of important statistical and methodological issues to consider. In particular, complete assessment of the effects of sexual behavior requires the collection of prospective daily information about the occurrence of vaginal–penile sexual intercourse and the timing of ovulation, and the use of this information to estimate day-specific probabilities of conception relative to ovulation. These issues are summarized below, in the context of methodologic design options to study human fecundity (Table 1).

Time to Pregnancy

Time to pregnancy is generally defined as the number of menstrual cycles it takes a couple with regularly occurring, noncontracepted intercourse to achieve pregnancy. Since the 1980s TTP has been used in epidemiologic studies as a measure of fecundability, the probability of conception in a menstrual cycle for a couple at risk of conception (Baird et al. 1986). TTP can be obtained retrospectively, by asking pregnant women how long it took to become pregnant, or prospectively, either by enrolling couples at the time they stop contraception to attempt conception or by following couples at risk for pregnancy, ideally regardless of pregnancy intentions at enrollment. Most of the studies of environmental exposures have been based on retrospective studies. However, significant biases can occur in retrospective studies.

Retrospective studies of time to pregnancy. In retrospective interviews women are asked to recall the number of menstrual cycles or the number of calendar months it took them to conceive from the cessation of contraception.

More precisely, women are asked to recount their contraceptive and sexual history, from which the number of noncontracepted cycles to conception can be derived. Other data on environmental exposures, smoking and alcohol use, medical history, family income, education level, and pregnancy history may also be collected (Baird 1988). Interviews can take place during a pregnancy, near the time of birth, or several years after a birth.

Bias in recruitment, recall, and behavior or exposure trends are all possible in retrospective studies of TTP. Recruitment for retrospective studies is often done when women present to obstetric clinics for prenatal care. This method introduces selection bias into the study if differences in prenatal care are linked to the investigated environmental exposure. For example, if women who were heavily exposed to an environmental factor were more worried about their pregnancies, this group would be more apt to present early for prenatal services (and use them more frequently over a longer period of time) and could be overrepresented in the study. Conversely, if a decrease in fecundability (or an increased probability of early spontaneous abortion) were linked to an increase in exposure, heavily exposed women would be underrepresented among those using prenatal care services, making the effect harder to detect in the study. Juul et al. (2000) warn against selection bias in choosing only pregnant samples because of error it causes when studying age-dependent effects on fecundity. Their study found that using only completed pregnancies, a common practice in retrospective studies, could lead to the incorrect assumption that TTP decreases with age. In addition, because typically only pregnant women are recruited, no allowance is made in studies for a sterile subpopulation. Therefore, associations between environmental exposures and sterility cannot be studied using such a design. Further, early pregnancy outcomes such as spontaneous

abortion or ectopic pregnancy, which may be related to some environmental exposures, cannot be accurately assessed, and this also introduces confounding with regard to TTP.

In addition to obtaining information on current pregnancies, investigators in retrospective studies may also interview women about previous pregnancies. A longer time until recall may lead to information bias, although a high level of accuracy in recall has been reported (Joffe et al. 1993). Digit preference, bias in which women are inclined to choose a rounded digit such as 3 or 6 when retrospective studies ask them to remember how many menstrual cycles occurred before they conceived, has been noted in some fertility studies (Jain 1969; Linn et al. 1982). Additionally, in retrospect, couples may change how they feel about a pregnancy and say it was planned even if the pregnancy resulted from a birth control failure, leading to the inclusion of data from a pregnancy that occurred during the use of contraception. Currently, there is no method to adjust for the effect of the use of contraception on fecundity, and therefore pregnancies that occur during the use of contraception must be excluded.

Women who experience longer TTP and suspect themselves to be subfertile may change their behavior (quit smoking, decrease caffeine or sugar intake) in a way they believe is more conducive to conception. Bias is thus introduced if exposure is analyzed using day of conception as the index day. In the same manner, a time trend bias is introduced if a woman's exposure to an environmental factor increases over time. A woman with a shorter TTP will report less exposure, whereas one with a longer TTP will report a greater exposure, even if the exposure had no direct effect on TTP.

Despite obvious bias, retrospective studies are often used because of the ease and low cost of collecting data. They may be particularly suitable for exploratory studies or for ongoing population surveillance (Joffe 2003; Olsen and Rachootin 2003). However, because retrospective methods are subject to these biases and do not account for sexual behavior, they are inadequate to definitively assess the effects of environmental exposures on human development.

Conventional prospective studies of time to pregnancy. In conventional prospective studies investigators follow women from the time of discontinuation of contraception until conception or until a set time if conception does not occur. Study participants in conventional prospective studies are often asked to give data on intercourse frequency, menstrual bleeding, contraception history, and exposure(s) of interest. This approach enables investigators to study fecundity, impaired fecundity (e.g., pregnancy loss, ectopic pregnancy), and infertility (i.e., absence of pregnancy).

The prospective design corrects many problems inherent in retrospective studies.

Table 1. Research designs to study human fecundity.

Type of study	Advantages	Disadvantages
Retrospective study of TTP	Inexpensive; suitable for exploratory studies	Multiple potential biases in recruitment, recall, behavior, and exposure trends. Outcomes generally limited to completed or advanced pregnancy.
Conventional prospective study of TTP	Fewer biases than retrospective studies; can accurately assess outcomes of sterility and spontaneous abortion	Higher cost and time commitment than retrospective studies. Some potential biases remain, including biases arising from planning, recognition, medical intervention, and the "unhealthy worker" phenomenon. Cannot adjust for timing of intercourse.
Detailed prospective study of fecundity, with day-specific probabilities of conception	Can assess full spectrum of reproductive outcomes including early pregnancy (embryonic) loss. Can fully adjust for sexual behavior including the timing of intercourse.	Cost may be more than conventional prospective studies. Higher burden for subject participation. Participants might be less representative of target population.

Problems with recall such as digit preference are no longer factors. Because a prospective study is based on conception attempts, not successes, a sterile subpopulation may be present and later accounted for in the analysis (Weinberg and Gladen 1986). Information on exposures would be collected for the duration of the study, allowing investigators to account for any change in prevalence.

Prospective studies can also accurately ascertain a much broader array of pregnancy outcomes, such as ectopic pregnancy, spontaneous abortion, and stillbirth. This allows for a more complete assessment of potential outcomes from environmental exposure, as well as a more accurate portrayal of TTP.

Some of the potential biases inherent in retrospective studies such as pregnancy planning bias, pregnancy recognition bias, medical intervention bias, and unhealthy worker bias may still be present in conventional prospective TTP studies (Baird 1988; Baird et al. 1986; Weinberg et al. 1993, 1994a). The practice of only using planned pregnancies in TTP studies (Baird 1988) could introduce pregnancy planning bias if exposed couples are more or less likely to attempt conception than unexposed couples. If exposed couples are less fecund, they will be less likely to experience unexpected pregnancy during the use of contraception and therefore may ultimately be more likely to seek to plan a pregnancy. Bias in pregnancy recognition can occur if an exposed group is more likely to have irregular menstrual cycles or less likely to buy home pregnancy kits. If recognition of pregnancy is delayed by an exposure, the TTP may seem longer, even though the exposure has no direct link to TTP. Additionally, if early spontaneous abortions go unnoticed, participants may have two or more pregnancies before a pregnancy is detected, leading the group to appear less fecund. Assisted reproductive techniques may increase the odds of conception for some couples, so any medical intervention may lead to a higher fertility rate among couples who seek assistance. Some couples may enhance their probabilities of conception by using aids for selecting timing of intercourse, such as urine luteinizing hormone (LH) testing or urinary estrogen metabolite testing (both available over the counter) or monitoring signs of fertility, such as vaginal mucus discharge or basal body temperature (Stanford et al. 2002). It is therefore essential that any interventions used by the couple to enhance conception be identified and be accounted for in analysis. (Such intervention cycles could either be excluded from analysis or included with the interventions noted as covariates.) Finally, the unhealthy worker bias is particularly problematic in studies of occupational exposures. Women who are successful in conceiving quickly may leave the workforce earlier than those with

longer TTP, leading more fecund women to have fewer occupational exposures.

Conventional prospective TTP studies do not allow for confounding effects of timing of intercourse and the increased chance of conception on days close to ovulation (Wilcox et al. 1995), the timing of which varies substantially from cycle to cycle (Wilcox et al. 2000). This problem is addressed in more detailed prospective studies, described below. Overall, prospective designs allow for more accurate, time-specific data on exposure, contraceptive method, intercourse frequency, and menstrual pattern than does the retrospective design (Bonde et al. 1998). Measurement of environmental exposures, including exposures to both parents, can be studied prospectively, adding important insight in areas where knowledge is currently very limited. When more definitive assessments are sought for the effects of environmental exposures, the advantages of the prospective design outweigh the logistical drawbacks, particularly the higher cost and larger time commitment than retrospective studies.

Statistical models for conventional time to pregnancy studies. Retrospective and prospective TTP studies generally obtain the same type of data on contraception and cycles until conception. If pregnancies resulting from contraceptive failure are excluded, both retrospective and conventional prospective data may be analyzed in the same manner (Weinberg and Gladen 1986). Because each menstrual cycle provides one conception opportunity, it is considered the natural time unit for TTP analysis. If duration is more easily remembered in months in retrospective studies, the length of the average menstrual cycle can be used to estimate the interval in menstrual cycles.

Because models that allow for heterogeneity in fecundity among the population are more realistic than those that assume homogeneity,

Sheps (1964) proposed a model for TTP that assumed a beta distribution for the probability of pregnancy per menstrual cycle.

Weinberg and Gladen (1986) extended this beta-binomial model to include couple-specific covariates and allow for a sterile subpopulation (those with a zero probability of conception). Ridout and Morgan (1991) proposed an extension that allowed for digit preference among women. Boldsen and Schaumburg (1990) suggested an alternative model that treats TTP as a continuous variable.

None of the previously mentioned models allow for time-dependent covariates such as age or accruing environmental exposure. In response to this problem, discrete time survival models have been proposed (Clayton and Ecochard 1997; Dunson and Neelon 2003; Scheike and Jensen 1997).

Detailed Prospective Studies

Table 2 summarizes data necessary for a detailed prospective study of human fecundity. These include data collected at the level of the couple, the cycle, and each day in the study. The couple-level data include well-established factors that can impact human fecundity and have been described adequately in previous reviews (Baird and Strassman 2000; Baird et al. 1986). The cycle-level and day-level data are described in more detail below.

It is highly desirable to obtain more precise data on intercourse frequency and timing to adjust for the differences in conception probabilities by day in the cycle. From the time of Ogino (1930) and Knaus (1929), who estimated that ovulation occurred approximately 14 days before the start of the next menstrual cycle, it has been known that most of the variation in cycle length (both between women and within the same woman) occurs in the preovulatory (follicular) phase of the menstrual cycle,

Table 2. Data elements to include in detailed prospective studies of human fecundity.

Couple-level data
Age and other demographic data
Reproductive, gynecologic, andrologic, and other medical history
Past use of contraception and hormonal therapies
Occupational exposure
History of alcohol, tobacco, and other drug use
Semen analysis (with repeated time-referenced measures)
Cycle-level data
Occurrence and estimated day of ovulation (derived from day-level data) ^a
Occurrence of conception (derived from day-level data) ^a
Reproductive intentions
Use of assisted reproduction techniques or home techniques to enhance conception
Day-level data
Menstrual or other vaginal bleeding
Sexual intercourse or genital contact ^a
Use of barriers, withdrawal, or spermicide ^a
Marker of ovulation, such as urinary hormones, vaginal mucus observation, or basal body temperature ^a
Marker of conception, such as serum or urine human chorionic gonadotropin ^a
Estrogen or metabolites (serum, urine, or saliva)
Progesterone or metabolites (serum, urine, or saliva)
Biological exposure assessments (serum, urine, or saliva) with reference to day of collection
Use of alcohol, tobacco, drugs, or herbs

^aIndicate essential core elements for analysis of day-specific probabilities of conception.

whereas the postovulatory (luteal) phase is relatively constant at approximately 14 days. The later finding that, while in the reproductive tract, the human ovum can only be fertilized for a window of approximately 12–48 hr (Siegler 1944) led to the hypothesis that an increased chance of conception would occur in the days surrounding ovulation. A prospective study in the 1960s (Barrett and Marshall 1969) of 221 married British couples was among the first to test this hypothesis. Chances of conception were low in the early part of the menstrual cycle; conception probabilities increased to a peak 2 days before the estimated day of ovulation. After the day of ovulation the conception probabilities decreased to near zero. A later study by Wilcox et al. (1995), which collected first morning urine data on each day of the menstrual cycle for hormonal analysis, found that intercourse was unlikely to result in a conception unless it occurred in the 6-day interval ending on the day of ovulation. Detailed prospective TTP studies such as the European Study of Daily Fecundability (Colombo and Masarotto 2000) have confirmed the relatively narrow interval of days immediately preceding ovulation when intercourse may result in pregnancy. These studies allow one to adjust for the confounding effects of the timing and frequency of intercourse in studying biological effects of covariates such as age (Dunson et al. 2002).

Such studies require daily data collection to determine the days of ovulation and intercourse (Table 2). Methods of estimating day of ovulation include direct ultrasonographic monitoring to determine time of follicular rupture, the use of surrogates such as the LH surge in urine or serum, the last day of hypothermia prior to the postovulatory rise in basal body temperature (BBT), the cervical mucus peak day (the last day of slippery or stretchy vaginal discharge), and the rapid decline in the ratio of estrogen to progesterone metabolites in the urine (day of luteal transition) (Baird et al. 1991). These methods differ in their accuracy, cost, and time commitment. Ultrasound monitoring is the gold standard, but cost is prohibitive in larger studies. Detection of the LH surge and the day of luteal transition require assaying and collection of daily first morning urine samples. Some studies have attempted to use calendar calculations based on previous or expected cycle length to estimate the daily probabilities of conception, but estimates from such calculations are very imprecise, even among women with a history of regular menstrual cycles (Wilcox et al. 2000). BBT and mucus-based methods are somewhat less accurate than hormonal measures but much more accurate than calendar calculations, making them cost effective for large studies (Guida et al. 1999). Vaginal observation of mucus discharge for purposes of predicting the fertile days of the cycle and the cervical

mucus peak can be learned easily by women from a variety of socioeconomic backgrounds (World Health Organization 1981). In addition to having low cost, a clear advantage of mucus-based methods is that the presence and quality of mucus discharge provides additional information about the probability of sperm survival and conception, independent of the timing of ovulation (Dunson et al. 2001a; Hilgers and Prebil 1979; Stanford et al. 2003)

Studies that attempt to identify day of ovulation must consider important differences between methods. For example some of the methods of determining the timing of ovulation will alert couples to the days they are more likely to be fertile, whereas other methods will not. If couples are alerted to the days of fertility, this may alter their sexual behavior and hence their TTP (Hilgers et al. 1992; Stanford et al. 2002; World Health Organization 1983). A number of studies have used daily urine collections analyzed in a central laboratory for the occurrence and timing of ovulation without feedback to couples during the study, thus eliminating this bias (Waller et al. 1996; Wang et al. 2003; Wilcox et al. 1985). If accurate records of the days with intercourse relative to the identified ovulation day are collected and inferences are based on day-specific probabilities of pregnancy (further described below), methods that prospectively inform the couples about their fertile days can then be used without biasing the results. Methods that provide information to the couples about their fertile days should be preferable for couples attempting pregnancy and those who wish to avoid pregnancy without using hormonal or barrier methods of contraception.

Daily analysis of urine, serum, or saliva for ovarian hormones, pituitary reproductive hormones, or their metabolites has been used in prospective studies to assess ovarian function beyond the simple occurrence of ovulation. Such hormonal profiles are predictive of both maternal outcomes (Waller et al. 1996) and reproductive outcomes (Baird et al. 1999).

Although some researchers suspect that exposure to alcohol, tobacco, and caffeine decreases human fertility (Curtis et al. 1997; Dunson 2001; Hakim et al. 1998; Weinberg et al. 1989; Wilcox et al. 1988a), the exact nature of these associations is hard to characterize in retrospective studies and has been the source of some controversy. In addition, despite widespread use of herbal products in the United States (Eisenberg et al. 1993), almost nothing is known about the effects of these products on human fertility and human development. A detailed prospective study could collect such exposure information on a daily basis and allow a more precise examination of these effects.

An additional advantage of detailed prospective studies is the opportunity for a

rigorous assessment of the outcome of conception itself. Many conceptions result in spontaneous pregnancy loss, which may not be recognized as a spontaneous abortion. A number of studies have assessed these outcomes using sensitive assays for human chorionic gonadotropin, an early marker of implantation (Wang et al. 2003; Wilcox et al. 1988b). Early pregnancy factor, now identified as a protein that is a close homolog to chaperonin-10, has been used to identify conception prior to implantation, but it is not yet sufficiently specific for use in large epidemiologic studies (Cavanagh 1996; Morton et al. 1992). Because of the relatively high proportion of conceptions that end prior to clinical recognition of pregnancy, a complete assessment of reproductive outcomes should include the measurement of a biochemical marker of conception in the postovulatory phase of the menstrual cycle. At present, the best candidate is human chorionic gonadotropin.

A final advantage of the detailed prospective TTP design is the ability to examine interactions between exposure effects and the age of the gametes. Such interactions are plausible, as aged gametes may be more susceptible to exposures. Wilcox et al. (1998) noted an increase in early pregnancy loss among conceptions that occurred when the ovum had the opportunity to age prior to conception. Potentially, gametes damaged by exposure to a toxicant may degrade more quickly with age. It is also plausible that sperm damaged by an exposure may have a higher probability of surviving and transporting themselves to the ovum if introduced on days with high levels of estrogenic mucus, whereas only the most progressively motile sperm have a chance of fertilizing the egg on days with sub-optimal mucus. This hypothesis is plausible, because sperm survival and transport are regulated by cervical mucus secretions which vary during the menstrual cycle (Katz 1991). However, daily records of mucus and intercourse are necessary to investigate this.

Methods of analysis. Taken together, the data necessary for a prospective study of human fertility necessitate statistical methods developed specifically to accommodate the complex and multilevel nature of the data structure. The concept of day-specific probabilities of pregnancy allows the integration of these data into a meaningful measure of human fecundity that can be used to assess the effects of various exposures, demographic factors, behavioral factors, and their interactions. Day-specific probabilities have the advantage of not depending on intercourse behavior, unlike the per-menstrual-cycle probabilities of conception and the TTP. Thus, the day-specific probabilities provide a more direct measure of biologic fecundity. Knowledge of the biology of the menstrual cycle can be used in developing statistical models for the daily

probabilities. In particular, ovulation is the key event in the menstrual cycle that determines the timing of the fecund interval during which intercourse can result in a pregnancy with non-negligible probability. If intercourse occurred only once in each menstrual cycle under study, it would be straightforward to estimate day-specific probabilities and to relate these probabilities to covariates using logistic regression, ideally with a couple-specific random effect included to account for within-couple dependency. However, in a menstrual cycle with multiple acts of intercourse occurring within the fertile window, it is not possible to attribute conception to a single act.

To account for this problem, Barrett and Marshall (1969) applied a model suggested by Peter Armitage, which assumes that batches of sperm introduced into the reproductive tract on different days commingle and then compete independently in attempting to fertilize the egg. Their model has the following mathematical form:

$$P_{ij} = 1 - \prod_k (1 - p_k)^{X_{ik}}, \quad [1]$$

where P_{ij} is the probability of conception for couple i in cycle j , k is the day in the cycle relative to ovulation ($k = 0$ on day of ovulation), X_{ik} is an indicator variable that equals 1 if intercourse occurred on day k and 0 otherwise, and p_k is the probability of conception in a cycle with intercourse only on day k .

Schwartz et al. (1980) modified this model by including a susceptibility multiplier A to allow menstrual cycle characteristics other than intercourse to have an effect on the probability of conception:

$$P_{ij} = A \left\{ 1 - \prod_k (1 - p_k)^{X_{ik}} \right\}, \quad [2]$$

where A is typically referred to as the cycle viability probability and p_k is the probability of conception in a viable cycle with intercourse only on day k . The term cycle viability probability is somewhat misleading because it implies that A includes only woman-specific factors such as uterine receptivity and oviduct function. However, like p_k , cycle viability (A) also includes male factors (e.g., the presence of motile sperm) and interaction effects (ability of sperm to fertilize ovum, survival of embryo to detection), making it difficult to distinguish which biological factors relate directly to A and which relate to p_k (Dunson 2001).

Variations of the model of Schwartz et al. have been proposed to allow covariate effects on A (Weinberg et al. 1994b), covariate effects on p_k (Zhou and Weinberg 1996), heterogeneity among couples in A (Dunson and Zhou 2000; Zhou et al. 1996), missing data on intercourse (Dunson and Weinberg 2000a), and measurement error in identifying the true

day of ovulation (Dunson et al. 2001b; Dunson and Weinberg 2000b). In addition, Royston (1982) and Weinberg and Wilcox (1995) developed parametric versions of the Schwartz model by assuming distributions for the survival times of the sperm and egg. Royston and Ferreira (1999) later proposed an approximation of Equation 2 that assumes that sperm introduced into the reproductive tract on any given day have no chance of fertilizing the ovum and thus no effect on the conception outcome if intercourse occurs on a more fertile day. Potentially, these models could be generalized to accommodate time-varying exposure effects by using a time-varying coefficient model (Hastie and Tibshirani 1993; Verweij and van Houwelingen 1995).

As previously stated, the incorporation of male and female factors into both p_k and A makes it difficult to determine which biological factors relate directly to each. In addition, it tends to be difficult to separately estimate A and the maximum p_k because of colinearity in these two parameters. In the special case where there is a single intercourse act in each cycle, the Schwartz model is not estimable, and one of the parameters must be fixed to fit the model. As the highest p_k is close to one for each of the available data sets, a reasonable modification of the Schwartz model that solves the estimability and colinearity problems is to set the highest p_k equal to one. Dunson (2001) proposed such an approach within the framework of a Bayesian hierarchical model that also incorporates the constraint that the p_k s increase to a peak and then decrease. The Dunson (2001) approach accommodates variability among couples in the day-specific conception probabilities and covariate effects on both the maximum day-specific probability and the duration of the fertile interval. A further advantage of this model is that it enhances statistical power to study the effects of covariates such as follicular phase length or age on fecundity. Hence, it is a useful approach in applications (Dunson et al. 2002; Stanford et al. 2002).

An important issue in designing and analyzing studies of day-specific pregnancy probabilities is sample size. The two classic studies in this area, the Barrett and Marshall (1969) study and the study by Wilcox et al. (1995), had data for slightly more than 200 couples, a sample size that has proven sufficient to produce many important results. However, the number of couples is not the only important issue, as estimation of day-specific probabilities relies on the availability of conception and nonconception cycles having a variety of intercourse days within the fecund window. Menstrual cycles with no acts of intercourse within the fecund window do not contribute to the analysis. In addition, cycles with multiple acts of intercourse contribute less information to the estimation of day-specific

probabilities than cycles with a single intercourse act. In the latter case, the intercourse act responsible for the conception is known, so there is less uncertainty.

In the study by Wilcox et al. (1995), the sample size was sufficient to obtain precise estimates of covariate effects on the cycle viability probability (Dunson and Zhou 2000), but there was low power to detect interactions between timing in the fecund window and the effect of a covariate. Even in analyses that did not adjust for covariates, 95% confidence limits for the day-specific conception probabilities ranged almost from zero to one. The high degree of uncertainty is partly due to the large proportion of cycles with multiple intercourse acts, as the couples in the Wilcox et al. study (1995) were attempting conception. However, another important factor is the type of statistical methods used to analyze the data. Prior to the approach of Dunson (2001), analyses did not incorporate constraints on the p_k s and hence were subject to the problems discussed above. As illustrated by Dunson (2001), by incorporating biologically reasonable parameter constraints on the p_k s, one can greatly reduce uncertainty in the estimates and increase power to assess covariate effects. Applying this approach to the Wilcox et al. (1995) data revealed evidence of an interaction between the effect of caffeine exposure on reducing fecundability and the timing of intercourse.

Standard formulas used for sample-size calculations do not apply here, and it is difficult to formulate general guidelines because of complex interactions between the sample size needed to obtain a given power, the couples' intercourse behavior, the numbers of cycles of follow-up, the distribution of fecundability in the population, and the prevalence of the exposure(s). As a rule of thumb, small studies involving fewer than 100 couples are not recommended unless one is willing to use Bayesian methods with informative priors chosen based on historical studies in the analysis. For couples attempting conception (assuming that one does not want to incorporate historical data from previous studies that might be informative), an excess of 100 couples followed until conception or at least 6–12 months if not conceiving is needed to investigate common exposures possibly associated with overall fecundability. The use of day-specific probabilities for the analysis will adjust for the effects of sexual behavior (i.e., timing of intercourse) while allowing for an overall assessment of the effect of exposure on fecundability. To investigate more detailed interactions between the exposure effect and timing of intercourse by day relative to ovulation, sample sizes need to be much larger. To obtain estimates of power under a given scenario, one can conduct a simulation study. Although it did not assess environmental exposures per se, the European

Study of Daily Fecundability (Colombo and Masarotto 2000), with 881 couples and 7,017 menstrual cycles, had sufficient sample size to examine the effect of some demographic and reproductive factors (age, parity, prior use of oral contraceptives, and follicular phase length) on day-specific probabilities.

The role of pregnancy planning. Most prospective studies of TTP exclude couples who are not planning pregnancy, an exclusion that may lead to the pregnancy planning bias described earlier. Demographic research indicates that about half of all pregnancies in the United States are considered unintended (Henshaw 1998). In addition, a significant proportion of pregnancies that occur during the use of contraception are nevertheless considered by the woman to be intended (Trussell et al. 1999). If unintended pregnancies are excluded from prospective study, it might lead to bias in considering the effects of various exposures on human development. The actual effects of this potential bias are currently unknown. An innovative approach to address this problem would be to prospectively follow couples at risk for pregnancy, regardless of their current pregnancy intention status (Miller 1994). A recently published study prospectively followed a cohort of 1,357 couples who kept daily menstrual and fertility diaries, identifying the point at which they started seeking to become pregnant by means of a question that the couples answered at the beginning of each menstrual cycle (Gnoth et al. 2003).

Integrating the study of environmental effects on human fertility and human developmental outcomes. There is increasing concern that preconception and periconception exposures may profoundly impact not only reproductive health, but also perinatal and child development outcomes and even some adult diseases (Chapin et al. 2004; Eaton 2002). In mice, a variety of agents have significant effects only when present at the critical window of implantation (Rutledge et al. 1992). Timing of exposure has hardly been studied in humans to date because most studies have been retrospective with regard to conception and implantation. A full understanding of the effect of environmental exposures on human development is possible only if detailed information is available on a complete range of reproductive and developmental outcomes and on the timing and level of exposures. For example, delays in TTP are reported to increase the risk of adverse perinatal outcomes such as low birth weight or preterm delivery (Henriksen et al. 1997; Joffe and Li 1994; Williams et al. 1991). Importantly, an agent that causes adverse perinatal or child health outcomes at one dose may cause infertility at a higher dose. Thus, couples with the highest exposure levels may be underrepresented and severe bias introduced into developmental studies if reproductive outcomes

such as sterility or spontaneous abortion are not included among the outcomes assessed. For example, it is possible that an environmental exposure could be misclassified as having a weak effect or no effect on the continuum of human reproduction when in fact it has a strong effect. In addition, if an exposure tends to differentially affect those embryos with a higher overall susceptibility to adverse outcomes, as seems likely, then there can even be an apparent beneficial effect of an adverse exposure on later developmental outcomes (Dunson and Perreault 2001). Therefore, to accurately assess the lifetime effects of environmental exposures on human development, studies must follow couples prospectively, starting prior to conception. Otherwise, effects can be missed entirely or attributed to an incorrect pathway. Day-specific, period-specific, or cycle-specific effects of exposures could be modeled not only for outcomes of conception but also for later reproductive and developmental outcomes. This joint approach could potentially be implemented within a model that allows the parameter measuring a couple's biologic fertility to impact the probabilities of adverse later outcomes, which in turn are linked through shared parameters (Dunson and Perreault 2001).

Discussion

Human fertility is of vital importance to human health and the survival of the species. Only recently have the effects of environmental factors on human fertility begun to be studied systematically. Retrospective TTP studies have been widely used for studying environmental factors that may affect any of the stages of reproduction without leaving a couple sterile (Baird et al. 1986), and they may be well suited for exploratory studies or population surveillance (Joffe 2003; Olsen and Rachootin 2003). However, they are subject to serious limitations and biases, reviewed in this article and by previous authors, and cannot be used to establish the effects of environmental exposures on human fertility. In addition to various biases inherent in retrospective assessment, a major limitation is the inability to accommodate the effects of sexual behavior, namely, the association between the conception probability and the timing of intercourse in relation to ovulation. To study factors related to biological fecundity and sterility, independent of behavioral factors, well-designed prospective studies of TTP are needed. The optimal study design begins prior to conception and collects detailed data on the timing of intercourse and ovulation. The analysis of day-specific probabilities of conception relative to ovulation allows an assessment of environmental and demographic factors on fecundity that is independent of sexual behavior. As this article describes, statistical methods with a variety of methodologic enhancements have been developed to analyze

day-specific probabilities of conception. Prospective studies, particularly detailed prospective studies outlined here, will be necessary to expand our understanding of the effects of environmental exposures on human fertility. The prospective designs also have an important role in addressing the growing interest in effects of early exposure on later outcomes of human development: large cohort studies such as the National Children's Study and others (Eaton 2002) are currently being proposed. The methodological and statistical methods reviewed in this article should prove useful in these lines of inquiry.

REFERENCES

- Abdalla HI, Wren ME, Thomas A, Korea L. 1997. Age of the uterus does not affect pregnancy or implantation rates; a study of egg donation in women of different ages sharing oocytes from the same donor. *Hum Reprod* 12:827–829.
- Amann RP. 1982. Use of animal models for detecting specific alterations in reproduction. *Fundam Appl Toxicol* 2:13–26.
- Apostoli P, Bellini A, Porru S, Bisanti L. 2000. The effect of lead on male fertility: a time to pregnancy (TTP) study. *Am J Ind Med* 38:310–315.
- Baird DD. 1988. Using time-to-pregnancy data to study occupational exposures: methodology. *Reprod Toxicol* 2:205–207.
- Baird D, Strassman B. 2000. Women's fecundability and factors affecting it. In: *Women and Health* (Goldman MB, ed). New York:Academic Press, 126–137.
- Baird DD, Weinberg CR, Wilcox AJ, McConaughy DR, Musey PI. 1991. Using the ratio of urinary oestrogen and progesterone metabolites to estimate day of ovulation. *Stat Med* 10:255–266.
- Baird DD, Weinberg CR, Zhou H, Kamel F, McConaughy DR, Kesner JS, et al. 1999. Preimplantation urinary hormone profiles and the probability of conception in healthy women. *Fertil Steril* 71:40–49.
- Baird DD, Wilcox AJ, Weinberg CR. 1986. Use of time to pregnancy to study environmental exposures. *Am J Epidemiol* 124:470–480.
- Barnhart K, Osheroff J. 1998. Follicle stimulating hormone as a predictor of fertility. *Curr Opin Obstet Gynecol* 10:227–232.
- . 1999. We are overinterpreting the predictive value of serum follicle-stimulating hormone levels. *Fertil Steril* 72:8–9.
- Barrett JC, Marshall J. 1969. The risk of conception on different days of the menstrual cycle. *Popul Stud* 23:455–461.
- Berardono B, Melani D, Ranaldi F, Giachetti E, Vanni P. 1993. Is the salivary "ferning" a reliable index of the fertile period? *Acta Eur Fertil* 24:61–65.
- Boldsen JL, Schaumburg I. 1990. Time to pregnancy—a model and its application. *J Biosoc Sci* 22:255–262.
- Bonde JP, Hjollund NH, Jensen TK, Ernst E, Kolstad H, Henriksen TB, et al. 1998. A follow-up study of environmental and biologic determinants of fertility among 430 Danish first-pregnancy planners: design and methods. *Reprod Toxicol* 12:19–27.
- Buck GM, Vena JE, Schisterman EF, Dmochowski J, Mendola P, Sever LE, et al. 2000. Parental consumption of contaminated sport fish from Lake Ontario and predicted fecundability. *Epidemiology* 11:388–393.
- Carlsen E, Giwercman A, Keiding N, Skakkebaek NE. 1992. Evidence for decreasing quality of semen during past 50 years. *Br Med J* 305:609–613.
- Cavanagh AC. 1996. Identification of early pregnancy factor as chaperonin 10: implications for understanding its role. *Rev Reprod* 1:28–32.
- Chandra A, Stephen EH. 1998. Impaired fecundity in the United States: 1982–1995. *Fam Plann Perspect* 30:34–42.
- Chapin RE, Robbins WA, Schieve LA, Sweeney AM, Tabacova SA, Tomashek KM. 2004. Off to a good start: the influence of pre- and periconceptional exposures, parental fertility, and nutrition on children's health. *Environ Health Perspect Environ Health Perspect* 112:69–78.
- Clayton DG, Ecochard R. 1997. Artificial insemination by donor: discrete time survival data with crossed and nested random effects. In: *First Seattle Symposium in Biostatistics: Survival Analysis* (Lin DY, Fleming TR, eds). New York:Springer-Verlag, 99–122.

- Colombo B, Masarotto G. 2000. Daily fecundability: first results from a new data base. *Demogr Res* 3(5).
- Curtis KM, Savitz DA, Ar buckle TE. 1997. Effects of cigarette smoking, caffeine consumption, and alcohol intake on fecundability. *Am J Epidemiol* 146:32–41.
- Curtis KM, Savitz DA, Weinberg CR, Ar buckle TE. 1999. The effect of pesticide exposure on time to pregnancy. *Epidemiology* 10:112–117.
- Dunson DB. 2001. Bayesian modeling of the level and duration of fertility in the menstrual cycle. *Biometrics* 57:1067–1073.
- Dunson DB, Baird DD, Colombo B. In press. Increased infertility with age in men and women. *Obstet Gynecol*.
- Dunson DB, Colombo B, Baird DD. 2002. Changes with age in the level and duration of fertility in the menstrual cycle. *Hum Reprod* 17:1399–1403.
- Dunson DB, Neelon B. 2003. Bayesian inference on order-constrained parameters in generalized linear models. *Biometrics* 59:286–295.
- Dunson DB, Perreault SD. 2001. Factor analytic models of clustered multivariate data with informative censoring. *Biometrics* 57:302–308.
- Dunson DB, Sinai I, Colombo B. 2001a. The relationship between cervical secretions and the daily probabilities of pregnancy: effectiveness of the Two Day Algorithm. *Hum Reprod* 16:2278–2282.
- Dunson DB, Weinberg CR. 2000a. Accounting for unreported and missing intercourse in human fertility studies. *Stat Med* 19:665–679.
- . 2000b. Modeling of human fertility in presence of measurement error. *Biometrics* 56:288–292.
- Dunson DB, Weinberg CR, Baird DD, Kesner JS, Wilcox AJ. 2001b. Assessing human fertility using several markers of ovulation. *Stat Med* 20:965–978.
- Dunson DB, Zhou H. 2000. A Bayesian model for fecundability and sterility. *J Am Stat Assoc* 95:1054–1062.
- Eaton WW. 2002. The logic for a conception-to-death cohort study. *Ann Epidemiol* 12:445–451.
- Eisenberg DM, Kessler RC, Foster C, Norlock FE, Calkins DR, Delbanco TL. 1993. Unconventional medicine in the United States. Prevalence, costs, and patterns of use. *N Engl J Med* 328:246–252.
- Gnoth C, Godehardt D, Godehardt E, Frank-Herrmann P, Freundl G. 2003. Time to pregnancy: results of the German prospective study and impact on the management of infertility. *Hum Reprod* 18:1959–1966.
- Guida M, Tommaselli GA, Palomba S, Pellicano M, Moccia G, DiCarlo C, et al. 1999. Efficacy of methods for determining ovulation in a natural family planning program. *Fertil Steril* 72:900–904.
- Hakim RB, Gray RH, Zaccaro H. 1998. Alcohol and caffeine consumption and decreased fertility. *Fertil Steril* 70:632–637.
- Hastie T, Tibshirani R. 1993. Varying-coefficient models with discussion. *J R Stat Soc Ser B* 55:757–776.
- Henriksen TB, Baird DD, Olsen J, Hedegaard M, Secher NJ, Wilcox AJ. 1997. Time to pregnancy and preterm delivery. *Obstet Gynecol* 89:594–599.
- Henshaw SK. 1998. Unintended pregnancy in the United States. *Fam Plann Perspect* 30:24–29, 46.
- Hilgers TW, Daly KD, Prebil AM, Hilgers SK. 1992. Cumulative pregnancy rates in patients with apparently normal fertility and fertility-focused intercourse. *J Reprod Med* 37:864–866.
- Hilgers TW, Prebil AM. 1979. The ovulation method—vulvar observations as an index of fertility/infertility. *Obstet Gynecol* 53:12–22.
- Jain AK. 1969. Fecundability and its relation to age in a sample of Taiwanese women. *Popul Stud* 23:69–85.
- Joffe M. 2003. Invited commentary: the potential for monitoring of fecundity and the remaining challenges. *Am J Epidemiol* 157:89–93.
- Joffe M, Li Z. 1994. Association of time to pregnancy and the outcome of pregnancy. *Fertil Steril* 62:71–75.
- Joffe M, Villard L, Li Z, Plowman R, Vessey M. 1993. Long-term recall of time-to-pregnancy. *Fertil Steril* 60:99–104.
- Juul S, Keiding N, Tvede M. 2000. Retrospectively sampled time-to-pregnancy data may make age-decreasing fecundity look increasing. *European Infertility and Subfecundity Study Group. Epidemiology* 11:717–719.
- Katz DF. 1991. Human cervical mucus: research update. *Am J Obstet Gynecol* 165:1984–1986.
- Kidd SA, Eskenazi B, Wyrobek AJ. 2001. Effects of male age on semen quality and fertility: a review of the literature. *Fertil Steril* 75:237–248.
- Knaus H. 1929. Eine neue Methode zur Bestimmung des Ovulationstermines. *Zentralbl F Gynak* 53:2193.
- Larsen SB, Joffe M, Bonde JP. 1998. Time to pregnancy and exposure to pesticides in Danish farmers. *ASCLEPIOS Study Group. Occup Environ Med* 55:278–283.
- Linn S, Schoenbaum SC, Monson RR, Rosner B, Ryan KJ. 1982. Delay in conception for former ‘pill’ users. *JAMA* 247:629–632.
- Miller W. 1994. Reproductive decisions: how we make them and how they make us. *Adv Popul* 2:1–27.
- Mitchell AA. 2002. Infertility treatment—more risks and challenges. *N Engl J Med* 346:769–770.
- Morabia A, Costanza MC. 1998. International variability in ages at menarche, first livebirth, and menopause. *World Health Organization Collaborative Study of Neoplasia and Steroid Contraceptives. Am J Epidemiol* 148:1195–1205. Erratum in: *Am J Epidemiol* 1999 150:546.
- Morton H, Rolfe B, Cavanagh A. 1992. Early pregnancy factor. *Semin Reprod Endocrinol* 10:72–82.
- O’Connell M, Rogers CC. 1982. Differential fertility in the United States: 1976–1980. *Fam Plann Perspect* 14:281–284.
- Ogino K. 1930. Ovulationstermin und Konzeptionstermin. *Zentralbl F Gynak* 54:464–479.
- Olsen J, Rachootin P. 2003. Invited commentary: monitoring fecundity over time—if we do it, then let’s do it right. *Am J Epidemiol* 157:94–97.
- Pearce D, Cantisani G, Laihonen A. 1999. Changes in fertility and family sizes in Europe. *Popul Trends* 95:33–40.
- Ridout MS, Morgan BJ. 1991. Modelling digit preference in fecundability studies. *Biometrics* 47:1423–1433.
- Rosenwaks Z, Davis OK, Damaro MA. 1995. The role of maternal age in assisted reproduction. *Hum Reprod* 10(suppl 1): 165–173.
- Royston JP. 1982. Basal body temperature, ovulation and the risk of conception, with special reference to the lifetimes of sperm and egg. *Biometrics* 38:397–406.
- Royston P, Ferreira A. 1999. A new approach to modeling daily probabilities of conception. *Biometrics* 55:1005–1013.
- Rutledge JC, Generoso WM, Shourbaji A, Cain KT, Gans M, Oliva J. 1992. Developmental anomalies derived from exposure of zygotes and first-cleavage embryos to mutagens. *Mutat Res* 296:167–177.
- Sallmen M, Anttila A, Lindbohm ML, Kyyronen P, Taskinen H, Hemminki K. 1995. Time to pregnancy among women occupationally exposed to lead. *J Occup Environ Med* 37:931–934.
- Sallmen M, Lindbohm ML, Anttila A, Kyyronen P, Taskinen H, Nykyri E, et al. 1998. Time to pregnancy among the wives of men exposed to organic solvents. *J Occup Environ Med* 55:24–30.
- Sauer MV. 1998. The impact of age on reproductive potential: lessons learned from oocyte donation. *Maturitas* 30:221–225.
- Scheike TH, Jensen TK. 1997. A discrete survival model with random effects: an application to time to pregnancy. *Biometrics* 53:318–329.
- Schwartz D, MacDonald PDM, Heuchel V. 1980. Fecundability, oital frequency and the viability of ova. *Popul Stud* 34:397–400.
- Schwartz D, Mayaux MJ. 1982. Female fecundity as a function of age: results of artificial insemination in 2193 nulliparous women with azoospermic husbands. *N Engl J Med* 306:404–406.
- Scott RT Jr, Hofmann GE. 1995. Prognostic assessment of ovarian reserve. *Fertil Steril* 63:1–11.
- Sheps MC. 1964. On the time required for conception. *Popul Stud* 18:85–97.
- Siegler SL. 1944. Fertility in Women: Causes, Diagnosis and Treatment of Impaired Fertility. Philadelphia:J.B. Lippincott and Co.
- Stanford JB, Smith KR, Dunson DB. 2003. Vulvar mucus observations and the probability of pregnancy. *Obstet Gynecol* 101:1285–1293.
- Stanford JB, White GL, Hatasaka H. 2002. Timing intercourse to achieve pregnancy: current evidence. *Obstet Gynecol* 100:1333–1341.
- Stovall DW, Toma SK, Hammond MG, Talbert LM. 1991. The effect of age on female fecundity. *Obstet Gynecol* 77:33–36.
- Swan SH, Elkin EP, Fenster L. 2000. The question of declining sperm density revisited: an analysis of 101 studies published 1934–1996. *Environ Health Perspect* 108:961–966.
- Taffel S. 1977. Trends in fertility in the United States. *Vital Health Stat* 21 28i–iv, 1–41.
- Thonneau P, Abell A, Larsen SB, Bonde JP, Joffe M, Clavert A, et al. 1999. Effects of pesticide exposure on time to pregnancy: results of a multicenter study in France and Denmark. *ASCLEPIOS Study Group. Am J Epidemiol* 150:157–163.
- Trussell J, Vaughan B, Stanford J. 1999. Are all contraceptive failures unintended pregnancies? Evidence from the 1995 National Survey of Family Growth. *Fam Plann Perspect* 31:246–247.
- United Nations. 1997. *World Fertility Patterns 1997*. New York:United Nations.
- van Noord-Zaadstra BM, Looman CW, Alsbach H, Habbema JD, teVelde ER, Karbaat J. 1991. Delaying childbearing: effect of age on fecundity and outcome of pregnancy. *Br Med J* 302:1361–1365.
- Ventura SJ, Martin JA, Curtin SC, Mathews TJ, Park MM. 2000. Births: final data for 1998. *Natl Vital Stat Rep* 48(3):1–100.
- Ventura SJ, Mosher WD, Curtin SC, Abma JC, Henshaw S. 2001. Trends in pregnancy rates for the United States, 1976–97: an update. *Natl Vital Stat Rep* 49(4):1–9.
- Verweij PJM, van Houwelingen HC. 1995. Time-dependent effects of fixed covariates in Cox regression. *Biometrics* 51:1550–1556.
- Waller K, Reim J, Fenster L, Swan SH, Brumback B, Windham GC, et al. 1996. Bone mass and subtle abnormalities in ovulatory function in healthy women. *J Clin Endocrinol Metab* 81:663–668.
- Wang X, Chen C, Wang L. 2003. Conception, early pregnancy loss, and time to clinical pregnancy. *Fertil Steril* 79:577–584.
- Weinberg CR, Baird DD, Rowland AS. 1993. Pitfalls inherent in retrospective time-to-event studies: the example of time to pregnancy. *Stat Med* 12:867–879.
- Weinberg CR, Baird DD, Wilcox AJ. 1994a. Sources of bias in studies of time to pregnancy. *Stat Med* 13:671–681.
- Weinberg CR, Gladen BC. 1986. The beta-geometric distribution applied to comparative fecundability studies. *Biometrics* 42:547–560.
- Weinberg CR, Gladen BC, Wilcox AJ. 1994b. Models relating the timing of intercourse to the probability of conception and the sex of the baby. *Biometrics* 50:358–367.
- Weinberg CR, Wilcox AJ. 1995. A model for estimating the potency and survival of human gametes in vivo. *Biometrics* 51:405–412.
- Weinberg CR, Wilcox AJ, Baird DD. 1989. Reduced fecundability in women with prenatal exposure to cigarette smoking. *Am J Epidemiol* 129:1072–1078.
- Wennborg H, Bodin L, Vainio H, Axelsson G. 2001. Solvent use and time to pregnancy among female personnel in biomedical laboratories in Sweden. *Occup Environ Med* 58:225–231.
- Wilcox AJ, Dunson DB, Baird DD. 2000. The timing of the “fertile window” in the menstrual cycle: day specific estimates from a prospective study. *Br Med J* 321:1259–1262.
- Wilcox AJ, Weinberg CR, Baird DD. 1988a. Caffeinated beverages and decreased fertility. *Lancet* 2:1453–1456.
- . 1998. Post-ovulatory ageing of the human oocyte and embryo failure. *Hum Reprod* 13:394–397.
- . 1995. Timing of sexual intercourse in relation to ovulation. Effects on the probability of conception, survival of the pregnancy, and sex of the baby. *N Engl J Med* 333:1517–1521.
- Wilcox AJ, Weinberg CR, O’Connor JF, Baird DD, Schlatterer JP, Canfield RE, et al. 1988b. Incidence of early loss of pregnancy. *N Engl J Med* 319:189–194.
- Wilcox AJ, Weinberg CR, Wehmann RE, Armstrong EG, Canfield RE, Nisula BC. 1985. Measuring early pregnancy loss: laboratory and field methods. *Fertil Steril* 44:366–374.
- Williams M, Goldman M, Mittendorf R, Monson R. 1991. Subfertility and the risk of low birth weight. *Fertil Steril* 56:688–671.
- Working PK. 1988. Male reproductive toxicology: comparison of the human to animal models. *Environ Health Perspect* 77:37–44.
- World Health Organization. 1981. A prospective multicentre trial of the ovulation method of natural family planning. I. The teaching phase. *Fertil Steril* 36:152–158.
- . 1983. A prospective multicentre trial of the ovulation method of natural family planning. III. Characteristics of the menstrual cycle and of the fertile phase. *Fertil Steril* 40:773–778.
- Zhou H, Weinberg CR. 1996. Modeling conception as an aggregated Bernoulli outcome with latent variables via the EM algorithm. *Biometrics* 52:945–954.
- Zhou H, Weinberg CR, Wilcox AJ, Baird DD. 1996. Random effects model for cycle viability in fertility studies. *J Am Stat Assoc* 91:1413–1422.