Prospective Pregnancy Study Designs for Assessing Reproductive and Developmental Toxicants

Germaine M. Buck,1 Courtney D. Lynch,1 Joseph B. Stanford,2 Anne M. Sweeney,3 Laura A. Schieve,4 John C. Rockett,5 Sherry G. Selevan,6 and Steven M. Schrader 7

The determinants of successful human reproduction and development may act as early as periconceptionally, underscoring the need to capture exposures during these critical windows when assessing potential toxicants. To identify such toxicants, couples must be studied longitudinally prior to conception without regard to a couple’s ability to ascertain a clinically recognized pregnancy. We examined the utility and feasibility of prospective pregnancy study designs by conducting a systematic review of the literature to summarize relevant information regarding the planning, implementation, and success of previously published prospective pregnancy studies. Information concerning design elements and participation was abstracted from 15 eligible studies (from a total of 20 identified studies) using a standardized form. The primary author of each study was contacted to review our summary of their work and obtain missing information. Our findings confirm the ability to recruit women/couples from diverse populations using a variety of recruitment strategies. Among the studies we reviewed, 4–97% of eligible individuals were successfully contacted, with enrollment rates ranging from 42 to 100%. Length of follow-up varied from 3 to 12 months. A high percentage of women provided urine (57–98%) and blood (86–91%) specimens and most male partners (94–100%) provided semen samples. These data support the feasibility of this design. Key words: design, development, fetal, preconception, pregnancy, prospective, reproduction, toxicity. Environ Health Perspect 112:79–86 (2004). doi:10.1289/ehp.6262 available via http://dx.doi.org/ [Online 24 September 2003]

A growing body of evidence challenges traditional thinking that only in utero exposures are of concern for the health of the developing fetus. Specifically, reproductive biologists, epidemiologists, and toxicologists recognize the potential importance of parental exposures at critical periconceptional windows, in addition to exposures during organogenesis (Chapin et al. 2004; Selevan et al. 2000). A spectrum of human health endpoints can be conceptualized for study, as reflected in the evaluative guidelines set forth by various regulatory agencies or organizations (California Environmental Protection Agency 1991; European Commission 2002; International Programme on Chemical Safety 2001; Moore et al. 1995; U.S. Environmental Protection Agency 1991, 1996). Recent biomedical advances offer promise for population-based studies of this type that can potentially address the many critical data gaps that confront this field.

Strategies for weighing scientific evidence regarding reproductive and developmental toxicity highlight study design as a criterion for evaluating the strength of available evidence (Subcommittee on Reproductive and Developmental Toxicology 2001). Although experimental study designs present the strongest data, they are not an ethical option for assessing the effect(s) of potentially toxic exposures on human reproductive and developmental end points. Hence, observational designs are the sole choice for epidemiologic investigation. Among observational studies, data from properly designed and implemented prospective cohort studies usually receive more weight than data obtained via retrospective cohort or case–control studies. This is mainly because of the investigator’s ability to ensure a temporal ordering between exposure(s) and outcome(s), measure exposure more accurately, measure relevant covariates at multiple time points, and minimize potential information biases (e.g., recall bias) (Adams 2001; Andersson et al. 2000; Reichman and Hade 2001; Werler et al. 1989). A recent example of recall bias in retrospective design is in a study that found poor reliability and recall bias in women’s retrospective reports of exposure to chemicals during pregnancy (Till et al. 2002).

Several cohort studies have followed human development by studying pregnant women (Golding et al. 2001; Niswander and Gordon 1972). These studies, however, could not ascertain exposures (or collect biospecimens) at critical periconceptional windows and could not assess early reproductive outcomes (before clinically recognized pregnancy).

The most comprehensive and informative observational design is a prospective cohort study that measures exposures longitudinally (on both parents) beginning prior to pregnancy and continuing throughout pregnancy (if it occurs) and beyond. This study design, which we call a prospective pregnancy study with preconception enrollment, allows for the assessment of early exposures and a complete range of reproductive and developmental outcomes, key information for avoiding bias in evaluating effect(s) of potential toxicants (Tingen et al. 2004).

Prospective pregnancy studies are often described as difficult, intensive, and expensive to conduct, with limited overall yield. In this article, we examine the empirical evidence on the utility and feasibility of prospective pregnancy study designs for identifying reproductive and developmental toxicants. Although

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

Address correspondence to G. Buck, Epidemiology Branch, NICHD, NIH, DHHS, 6100 Executive Blvd., Rm. 7B03, Rockville, MD 20852 USA. Telephone: (301) 496-6155. Fax: (301) 402-2084. E-mail: gb156i@nih.gov

We thank the principal investigators of the prospective pregnancy studies included in this review for their pioneering work and generous time devoted to answering our many inquiries. In addition, we acknowledge members of the Fertility and Early Pregnancy Working Group, National Children’s Study, for their critical review of this work.

The views in this article reflect those of the authors and not necessarily those of their affiliated institutions. The information in this document has been subjected to review by the National Health and Environmental Effects Research Laboratory (U.S. EPA) and approved for publication. Approval does not signify that the contents reflect the views of the Agency, nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

The authors declare they have no competing financial interests.

Received 6 February 2003; accepted 2 September 2003.
most prospective pregnancy studies focus on the determinants of sensitive end points (e.g., time to pregnancy and early pregnancy loss), a review of these issues is beyond the scope of this article. Our work is based on a systematic literature review to summarize relevant information on the planning, implementation, and relative success of this design.

**Methods**

**Search Strategy**

We conducted a MEDLINE (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi) search in May 2002 to locate published prospective pregnancy studies using the following search terms: prospective studies [MeSH term] AND (fertility OR fecundity OR time to pregnancy OR urine OR pregnancy). We sought to identify all large epidemiologic prospective pregnancy studies with preconception enrollment and at least a 3-month follow-up period. We reviewed the references cited by each study investigator to ensure that all relevant published works had been identified. Our initial search yielded 18 studies, of which 13 were selected for review. Five studies were excluded for the following reasons: a) clinical study focusing on postimplantation pregnancy (Miller et al. 1980); b) small sample size ($n = 24, 13,$ and 13, respectively) (Hilgers et al. 1978; Li et al. 2002b; Sanders and Bruce 1997), and c) prospective study comprising only women with clinically recognized pregnancy (Hilgers et al. 1997). Nine studies did not enumerate a denominator because of their reliance on community volunteers responding to recruitment advertisements or other such attempts to solicit participation (Colombo and Masarotto 2000; de Mouzon et al. 1988; France et al. 1984; Sweeney et al. 1988, 1989; Vartiainen et al. 1994; Wang et al. 2003, Wilcox et al. 1988; Zinaman et al. 1996).

Participants have been recruited from a number of diverse referent populations (general or medical communities, job sites, population-based registries), and on the basis of recreational exposures (e.g., anglers). Most investigators studied women, with only four studies focusing on couples (Bonde et al. 1998; Colombo and Masarotto 2000; de Mouzon et al. 1988; France et al. 1997). Nine studies did not enumerate the size of their target population (Bonde et al. 1998; Brown et al. 1997; Buck et al. 2002; Ellish et al. 1996; Eskenazi et al. 1995; Hakim et al. 1995). When recruitment details were available (Table 2), the percentage of women/couples who were successfully contacted ranged from 2% in a population-based study of first pregnancy planners (Bonde et al. 1998) to 67% in a study of women working in the semiconductor industry (Eskenazi et al. 1995). Of particular note is the high percentage of contacts (46%) achieved by one group of investigators by mailing questionnaires to women of reproductive age who were listed in the New York State registry of licensed drivers (Ellish et al. 1996). The percentage of women/couples successfully contacted or eligible for enrollment could not be determined for every study because of the lack of available denominator information.

The percentage of eligible women/couples among those who were contacted ranged from 4% in a population-based study that targeted women of reproductive age (Ellish et al. 1996) to 95% in a volunteer community-based sample of couples desiring pregnancy (Zinaman et al. 1996) and 97% in a group of newly married textile workers in China (Wang et al. 2003). It should be noted that the number of women contacted and deemed eligible appeared to vary according to the recruitment strategy (i.e., those who publicized their

**Data Collection**

We developed a standardized data abstraction form that included author and year of primary (or methodologically oriented) publication; size of the target population; number of individuals contacted; number of eligible individuals; number of study participants; length of follow-up; type(s) of data collection, specifically, use of daily diaries and biopsychosocial assessment (namely, urine and blood); semen collection; number of people dropping out of the study; and type(s) of incentives offered for participation. Requests for specific information were sent to all primary authors in June 2002, with 100% response. The authors were asked to review and approve our summaries of their work and to provide missing information if possible. Both published and unpublished data obtained from the authors were summarized for our review. Several investigators were unsure or unable to enumerate the exact size of the target population, given the sampling strategy employed. Thus, the eligibility and participation percentages presented here should be regarded as best estimates.

**Table 1.** Target population, sampling unit, and recruitment strategy among the selected prospective pregnancy studies.

<table>
<thead>
<tr>
<th>Primary author (year)</th>
<th>Target population</th>
<th>Sampling unit</th>
<th>Recruitment strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bonde (1998)</td>
<td>Trade union members</td>
<td>Couples</td>
<td>Letters</td>
</tr>
<tr>
<td>Brown (1997)</td>
<td>HMO women of reproductive age</td>
<td>Women</td>
<td>Letters to female HMO members (also media and health providers)</td>
</tr>
<tr>
<td>de Mouzon (1988)</td>
<td>Community</td>
<td>Couples</td>
<td>Media and letters</td>
</tr>
<tr>
<td>Eskenazi (1995)</td>
<td>Semiconductors</td>
<td>Women</td>
<td>Letters (also informational meetings)</td>
</tr>
<tr>
<td>Hakim (1995)</td>
<td>Semiconductors</td>
<td>Women</td>
<td>Women who are outreach talks and posters</td>
</tr>
<tr>
<td>Sweeney (1988)</td>
<td>Community</td>
<td>Women</td>
<td>Women who are media and letters</td>
</tr>
<tr>
<td>Sweeney (1989)</td>
<td>Motor vehicle registry and telephone directory</td>
<td>Women</td>
<td>Letters</td>
</tr>
<tr>
<td>Vartiainen (1994)</td>
<td>Community</td>
<td>Women</td>
<td>Media</td>
</tr>
<tr>
<td>Wilcox (1988)</td>
<td>Community</td>
<td>Women</td>
<td>Media</td>
</tr>
<tr>
<td>Zinaman (1996)</td>
<td>Community</td>
<td>Couples</td>
<td>Media, physician, and acquaintance referral</td>
</tr>
</tbody>
</table>

*The sampling units were not required to be free of known fecundity or fertility impairments. Media include television, radio, and newspaper/poster announcements. *Men were enrolled after study was implemented; baseline questionnaire data available from approximately two-thirds of husbands. (Personal communication with authors.) Female partner of couple had to be free of fecundity impairments.
Table 2. Recruitment details for the selected prospective pregnancy studies.*

<table>
<thead>
<tr>
<th>Primary author (country)</th>
<th>Year</th>
<th>Context</th>
<th>Eligibility criteria</th>
<th>Target (n)</th>
<th>Contacts</th>
<th>Eligible</th>
<th>Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bonde (Denmark)</td>
<td>1998</td>
<td>Couples recruited from Danish trade unions; office workers, nurses, and daycare workers</td>
<td>No children Cohabiting Age 20–35 (nurses 23–39) years Planning to discontinue contraception Access to a telephone Working home freezer Partner within ± 10 years Excluded if either partner had previous reproductive experience</td>
<td>52,255</td>
<td>1,113 (2%)</td>
<td>851 (76%)</td>
<td>430 (51%)</td>
</tr>
<tr>
<td>Brown (United States)</td>
<td>1997</td>
<td>Participants recruited from Group Health, Inc., a large HMO in Minneapolis and St. Paul, Minnesota</td>
<td>Female HMO member Age 18–35 years Planning to attempt pregnancy in &lt; 3 months Nulliparas were recruited first; multiparas included later Multiparas excluded if planned attempt was within 12 months of most recent delivery Excluded if history of &gt; 1 loss at less than 20 weeks, history of &gt; 2 abortions, history of infertility, hypertension, diabetes, heart or kidney disease, or if pregnancy plans changed or pregnancy occurred before planned</td>
<td>28,000</td>
<td>2,840 (10%)</td>
<td>1,649 (58%)</td>
<td>1,152 (70%)</td>
</tr>
<tr>
<td>Buck (United States)</td>
<td>2002</td>
<td>Female members of the New York State Angler Cohort Study who indicated at enrollment in 1991 that they had not yet completed childbearing</td>
<td>Indicated that they had not yet started or completed childbearing Age 18–35 years Absence of infertility or fecundity problems (self-reported)</td>
<td>2,637</td>
<td>1,031 (39%)</td>
<td>244 (24%)</td>
<td>102 (42%)</td>
</tr>
<tr>
<td>Colombo (Europe)</td>
<td>2000</td>
<td>782 women recruited at seven European centers (Milan, Verona, Lugano, Duesseldorf, Paris, London, Brussels); most were trying to avoid pregnancy</td>
<td>Women experienced in the use of natural family planning Married or in a stable relationship Age 18–40 years In multiparas, must have had at least one menstruational cycles after delivery/ breastfeeding Excluded if women were taking hormonal meds that could affect fertility or if either partner was sterile or had an endocrine disorder</td>
<td>( \text{b} ) 782</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>de Mouzon (France)</td>
<td>1988</td>
<td>Couples in France without a history of infertility who desired pregnancy</td>
<td>Absence of contraception during the study Existence of at least one eligible menstrual cycle Interpretable cycles Knowledge of smoking status</td>
<td>( \text{b} ) 1,887</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ellish (United States)</td>
<td>1996</td>
<td>Women randomly selected from the 1987–1988 New York State Department of Motor Vehicles database of licensed drivers who were living in Albany County when they applied for or renewed their license</td>
<td>Off contraception for ≤ 12 months Planning to discontinue contraception within 6 months Residence within the prescribed area Regular menstrual cycle length (± 5 days)</td>
<td>16,800</td>
<td>7,649 (46%)</td>
<td>293 (4%)</td>
<td>227 (77%)</td>
</tr>
<tr>
<td>Eskenazi (United States)</td>
<td>1995</td>
<td>Women were recruited from seven silicon wafer fabrication sites in five U.S. companies</td>
<td>Not currently pregnant Menstruated within the past 2 months Intercourse within the past 2 months Working home freezer No plan to leave the company within the next 3 months Ability to speak in English, Spanish, Vietnamese, or Tagalog Excluded if sterilized, using oral contraceptives, IUDs, had a sterilized partner, or using hormonal steroids that might affect fertility</td>
<td>3,915</td>
<td>2,639 (67%)</td>
<td>739 (28%)</td>
<td>481 (65%)</td>
</tr>
<tr>
<td>France (New Zealand)</td>
<td>1984</td>
<td>Couples contemplating pregnancy in Auckland, New Zealand, from 1979 to 1985 who had the desire to preselect the sex of their child</td>
<td>Proven fertility</td>
<td>( \text{b} ) 239</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hakim (United States)</td>
<td>1995</td>
<td>Female employees at two semiconductor facilities in Vermont and New York from May 1989 to August 1991</td>
<td>Proven fertility Woman desiring pregnancy 42 years of age or younger Excluded if using oral contraceptives, using an IUD, or surgically sterilized</td>
<td>Over 5,000</td>
<td>( \text{b} ) 148</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(Continued on next page)
eligibility criteria during the recruitment process contacted fewer women, but more of their contacts were eligible for participation).

Participation rates seemed to be influenced by both the recruitment strategy and the study design features, with rates ranging from 42% of women originally enrolled in a larger cohort study with a less intense protocol to 100% of community volunteers meeting eligibility criteria at one urban medical center. One group of investigators examined the degree to which pregnancy intentions influenced a woman’s decision to participate (Sweeney et al. 1989). They found that only 2% of enrolled women reported actively trying to conceive during the 3-month study period, 46% reported using oral contraceptives or intrauterine devices, 24% reported using barrier methods or monitoring their cervical mucus and basal body temperature (BBT) to avoid pregnancy, 18% reported being sexually inactive, 8% reported being sexually active but not using contraception, and 2% reported being infertile.

Table 3 summarizes the follow-up and specimen collection details for each selected study. The length of follow-up varied by study purpose and intensity of the data collection. Study durations ranged from 3 to 12 months. The least intensive protocols included a minimum of baseline interviews with some prospective recording of relevant study factors with or without the collection of biologic specimens.

Daily diaries were used by 12 (80%) studies for varying periods of time ranging from 1 month to 12 at-risk menstrual cycles. The type of data collected with these diaries varied but typically included exposure(s) of interest, menstruation, fecundity signs (namely, vaginal mucus discharge and/or BBT), sexual intercourse, lifestyle behaviors (e.g., cigarette smoking, alcohol, caffeine, or vitamin/mineral consumption, illnesses, medications), and home pregnancy test results. Among those studies in which compliance rates were available, rates ranged from 80 to 98%, with the exception of one study that reported a 38% completion rate for the entire study protocol (France et al. 1984).

Four types of biospecimens have been collected in prospective pregnancy studies: urine, blood, semen, and breast milk. Biospecimen compliance rates were quite high among the studies for which information was available, ranging from 57 to 98% for urine, 86 to 93% for blood, 94 to 100% for semen, and 97% for a single postpartum breast milk sample. Our review suggests that once enrolled, women (and male partners, if applicable) will provide a variety of specimens for study purposes.

The reported study dropout rates varied widely, in part depending on how withdrawals were handled. (Some authors counted withdrawals as ineligible.) Moreover, some investigators requested that women/couples participate as long as possible, while others asked a priori for participation for a set period of time (e.g., 6 months). The lowest dropout rate (3%) was reported by one group of investigators in their 6-month prospective pregnancy study of community volunteers desiring pregnancy (Wilcox et al. 1988). France and associates reported the highest dropout rate (62%) in their study of couples desiring pregnancy who wished to preselect the sex of their child (France et al. 1984). Of the 148 women who dropped out of that study, 28% cited a change in pregnancy plans, 18% stated that the study was too demanding, 12% felt the study was too stressful, and 7% failed to become pregnant. Among other studies reporting reasons for dropout, the most common reasons were changes in pregnancy plans or health status (Bonde et al. 1998; Brown et al. 1997; Buck et al. 2002; Ellish et al. 1996; Sweeney et al. 1988).

Prospective pregnancy studies have offered varying levels of incentives for study participation. Notably, four (27%) authors reported offering no incentives for participation (Colombo and Masarotto 2000; Sweeney et al. 1988, 1989; Vartiainen et al. 1994). The largest incentive was US$500, which was given to couples upon completion of a protocol that required multiple clinic visits and sensitive procedures such as midcycle postcoital tests (Zinaman et al. 1996). Among U.S. studies reporting the use of incentives, the smallest was US$10, which was given either weekly (Wilcox et al. 1988) or every 2 months (Ellish et al. 1996) to women who

### Table 3

<table>
<thead>
<tr>
<th>Primary author (country)</th>
<th>Year</th>
<th>Context</th>
<th>Eligibility criteria</th>
<th>Target n</th>
<th>Contacts</th>
<th>Eligible</th>
<th>Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sweeney (United States)</td>
<td>1988</td>
<td>University of Pittsburgh employees and other area volunteers who were trying to become pregnant from October 1988 to October 1989</td>
<td>Age 16–44 years</td>
<td>88</td>
<td>b</td>
<td>b</td>
<td>82</td>
</tr>
<tr>
<td>Sweeney (United States)</td>
<td>1989</td>
<td>Women were recruited from a particular section of Pittsburgh, Pennsylvania, in which 89% of people were employed in service-oriented and blue-collar jobs</td>
<td>Not currently pregnant</td>
<td>a</td>
<td>b</td>
<td>b</td>
<td>104</td>
</tr>
<tr>
<td>Vartiainen (Finland)</td>
<td>1994</td>
<td>First-time pregnancy planners were recruited from the Kuopio area in Eastern Finland</td>
<td>Healthy couples without a history of infertility Planning to have a baby</td>
<td>a</td>
<td>443c</td>
<td>b</td>
<td>191</td>
</tr>
<tr>
<td>Wang (China)</td>
<td>2003</td>
<td>Newly married female textile workers in China who intended to conceive</td>
<td>Full-time employment Newly married Age 20–34 years Had permission to conceive</td>
<td>a</td>
<td>1,006c</td>
<td>971 (97%)</td>
<td>961 (99%)</td>
</tr>
<tr>
<td>Wilcox (United States)</td>
<td>1988</td>
<td>Women in the Research Triangle Park, North Carolina, area who were planning to discontinue contraception</td>
<td>Age 18 years or older Not currently pregnant Excluded if they had a history of fertility problems or chronic illness</td>
<td>a</td>
<td>b</td>
<td>b</td>
<td>221c</td>
</tr>
<tr>
<td>Zinaman (United States)</td>
<td>1996</td>
<td>Couples discontinuing contraception to become pregnant</td>
<td>Men age 21–37 years Men age 21–60 years Regular menstrual cycles (25–33 days) Men willing to provide semen samples Excluded if couples had been without contraception for &gt; 3 months or if either partner had a history of infertility or early pregnancy loss</td>
<td>a</td>
<td>210</td>
<td>200 (95%)</td>
<td>200 (100%)</td>
</tr>
</tbody>
</table>

Abbreviations: HMO, health maintenance organization; IUD, intrauterine device.

* Cycles refer to menstrual cycles, whereas months refer to calendar time. * Information not available. * Personal communication with author(s).
participated in a protocol that included daily diaries and urine collection (the former had an attrition rate of 3% and the latter 7%). A recent study conducted in China paid women US$1 per three urine samples provided (Wang et al. 2003). Only two studies reported providing feedback to participants in the form of summarized menstrual cycle information (Buck et al. 2002; Hakim et al. 1995).

**Discussion**

This review suggests that prospective pregnancy studies are a relatively new, powerful, and feasible design for examining the relation between biological, environmental, and lifestyle exposures and various reproductive and developmental outcomes. The utility of prospective pregnancy studies has greatly furthered our understanding of human reproduction and development, including notable advances such as estimates of the incidence of

---

### Table 3. Follow-up and specimen collection details for the selected prospective pregnancy studies.

<table>
<thead>
<tr>
<th>Primary author (country)</th>
<th>Year</th>
<th>Participants</th>
<th>Length of follow-up</th>
<th>Daily diaries</th>
<th>Urine samples</th>
<th>Blood sample(s)</th>
<th>Semen sample(s)</th>
<th>Dropped out of the study</th>
<th>Incentive(s) to participate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bonde (Denmark)</td>
<td>1998</td>
<td>430 (51%)</td>
<td>6 cycles or until pregnancy occurred</td>
<td>1.32/1,957 completed (80%)</td>
<td>Women: 9,671 (83%) Men: 820 (69%)</td>
<td>Women: 288/317 (91%) Men: 350/376 (93%)</td>
<td>418/430 (97%)</td>
<td>35 (8%)</td>
<td>$200 for specimens; participants entered in a lottery for $3,000 &lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Brown (United States)</td>
<td>1997</td>
<td>1,152 (70%)</td>
<td>12 cycles, second miscarriage, or delivery if pregnancy occurred &lt;sup&gt;a&lt;/sup&gt;</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>$100 for completion and small gifts (e.g., pencils) with newsletters &lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Buck (United States)</td>
<td>2002</td>
<td>102 (42%)</td>
<td>12 cycles, or until first postpartum visit or cessation of breastfeeding (if nursing)</td>
<td>7 women missing 1+ weekly cards &lt;sup&gt;b&lt;/sup&gt;</td>
<td>N/A</td>
<td>88/102 (86%)</td>
<td>N/A</td>
<td>20 (20%)</td>
<td>$50 for completion through postpartum blood and breast milk (97% of women provided a postpartum breast milk; 74% provided a second sample upon weaning) &lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Colombo (Europe)</td>
<td>2000</td>
<td>782</td>
<td>Average of 8.6 cycles per woman</td>
<td>80.6% included BBT, 85.2% included cervical mucus score &lt;sup&gt;b&lt;/sup&gt;</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>300 (25%); 84 of these women re-entered &lt;sup&gt;b&lt;/sup&gt;</td>
<td>None &lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>de Mouxon (France)</td>
<td>1988</td>
<td>1,887</td>
<td>12 cycles or until delivery if pregnancy occurred</td>
<td>1,304/1,516 (86%) completed &lt;sup&gt;a&lt;/sup&gt;</td>
<td>95%</td>
<td>N/A</td>
<td>N/A</td>
<td>138 (38%)</td>
<td>$10 for every 2 months of participation; lab results were forwarded to physicians &lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Ellish (United States)</td>
<td>1996</td>
<td>227 (77%)</td>
<td>12 cycles or until pregnancy occurred</td>
<td>403 (84%) completed at least one cycle &lt;sup&gt;a&lt;/sup&gt;</td>
<td>84% completed at least one cycle &lt;sup&gt;b&lt;/sup&gt;</td>
<td>N/A</td>
<td>N/A</td>
<td>7 (16%)</td>
<td>$35 for each month of participation; eligible for a prize drawing for a trip to Hawaii or other local resort &lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Eskenazi (United States)</td>
<td>1995</td>
<td>481 (65%)</td>
<td>6 cycles or clinical pregnancy</td>
<td>91 (38%) completed the entire protocol &lt;sup&gt;a&lt;/sup&gt;</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>148 (62%)</td>
<td>Personal instruction regarding fertility awareness and the Shettles theory of sex selection &lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>France (New Zealand)</td>
<td>1984</td>
<td>239</td>
<td>6 months or until pregnancy occurred</td>
<td>91 (38%) completed the entire protocol &lt;sup&gt;a&lt;/sup&gt;</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>148 (62%)</td>
<td>Personal instruction regarding fertility awareness and the Shettles theory of sex selection &lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Hakim (United States)</td>
<td>1995</td>
<td>148</td>
<td>At least 6 months or until pregnancy occurred (mean = 7 cycles)</td>
<td>Near 100% &lt;sup&gt;b&lt;/sup&gt;</td>
<td>90%</td>
<td>N/A</td>
<td>N/A</td>
<td>24 (16%) &lt;sup&gt;b&lt;/sup&gt;</td>
<td>$100 for completion of the study; feedback on menstrual cycles &lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Sweeney (United States)</td>
<td>1988</td>
<td>82</td>
<td>12 months or until pregnancy occurred</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>45 (55%) &lt;sup&gt;b&lt;/sup&gt;</td>
<td>None &lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Sweeney (United States)</td>
<td>1989</td>
<td>104</td>
<td>3 cycles or until the end of the first trimester if pregnancy occurred</td>
<td>81%</td>
<td>80%</td>
<td>N/A</td>
<td>N/A</td>
<td>None &lt;sup&gt;b&lt;/sup&gt;</td>
<td>None &lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Vartiainen (Finland)</td>
<td>1994</td>
<td>191</td>
<td>6 months or until pregnancy occurred</td>
<td>88% first f/u; 59% second f/u; 39% third f/u &lt;sup&gt;b&lt;/sup&gt;</td>
<td>N/A</td>
<td>N/A</td>
<td>180 (94%) &lt;sup&gt;b&lt;/sup&gt;</td>
<td>11 (6%)</td>
<td>None &lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Wang (China)</td>
<td>2003</td>
<td>961 (99%)</td>
<td>12 months or clinically confirmed pregnancy after stopping contraception</td>
<td>545 (57%) completed at least one cycle &lt;sup&gt;a&lt;/sup&gt;</td>
<td>545 (57%) completed at least one cycle &lt;sup&gt;a&lt;/sup&gt;</td>
<td>N/A</td>
<td>N/A</td>
<td>35 (4%)</td>
<td>$1 per three urine samples &lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Wilcox (United States)</td>
<td>1988</td>
<td>221 &lt;sup&gt;b&lt;/sup&gt;</td>
<td>6 months or until pregnancy occurred</td>
<td>98% &lt;sup&gt;b&lt;/sup&gt;</td>
<td>98%</td>
<td>N/A</td>
<td>N/A</td>
<td>6 (3%)</td>
<td>$10/week for urine collection &lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Zinaman (United States)</td>
<td>1996</td>
<td>200 (100%)</td>
<td>12 months or until pregnancy occurred</td>
<td>Over 90% &lt;sup&gt;b&lt;/sup&gt;</td>
<td>Over 90% &lt;sup&gt;b&lt;/sup&gt;</td>
<td>100% (participation requirement)</td>
<td>8 (4%)</td>
<td>$500 for completion of the study &lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: f/u, follow-up; N/A, information not applicable.  
<sup>a</sup>Cycles refer to menstrual cycles, whereas months refer to calendar time.  
<sup>b</sup>Personal communication with author(s).  
*Information not available.
likely to be employed, and from higher socioe-
nancies. However, the possibility of differing
women with intended versus unintended preg-
Schnare 2001). Little empirical evidence exists
because of inaccurate contact information.
Targeted sampling is also possible by weight-
radioactive waste sites and industrial sources of
environmental pollution are often located in
low-income communities (Farber and Krieg
2002; Morello-Frosch et al. 2002; Wilson et al.
2002) whose residents typically do not partici-
pate in research studies in the absence of tar-
ged recruiting. Further, lifestyle factors such as
cigarette smoking, alcohol use, and medica-
tions can influence the effects of environmental
chemical exposures in humans (Anwar 1993;
McCauley 1998). Given the potential for dif-
ferring exposure profiles among pregnant
women, coupled with the likelihood that some
behaviors will be modified during pregnancy,
the possible interactive effects of toxic agents
and divergent lifestyle profiles during the peri-
conceptional period (including those that are
paternally mediated) must be evaluated.
Prospective pregnancy study designs are the
only reliable approach for such inquiry.

Additional concerns have been raised
garding the generalizability of prospective
pregnancy studies because of research suggest-
ing that women with intended pregnancies
have fewer adverse pregnancy outcomes com-
pared with mothers with unintended pregnan-
cies (Piccinino and Mosher 1998). However,
data from the National Longitudinal Survey of
Youth suggest that differences in pregnancy
outcomes by pregnancy intentions might be
explained by the women’s socioeconomic
status rather than by planning status per se
(Joyce et al. 2000a).

The conceptualization and measurement of
intended or planned pregnancies has recently
come under intense scrutiny, with many
researchers in the field suggesting that more
accurate measures are needed (Klerman 2000;
Luker 1999; Sable 1999; Stanford et al. 2002;
Trussell et al. 1999). For example, one study
reported that 25% of women gave discordant
responses to questions designed to assess preg-
nancy intentions in two large population-based
surveys (Kaufmann et al. 1997). Discrepancies
in pregnancy intention responses were asso-
ciated with age, marital status, income, edu-
cation, parity, time since pregnancy, and
pregnancy outcome.

As with any epidemiologic investigation,
researchers must weigh the relative importance
of external validity in relation to internal valid-
ity (Grimes and Schulz 2002; Rothman and
Greenland 1998). Given the difficulty in defin-
ing the exact size of the population from which
participants in prospective pregnancy studies
are recruited, empirical evaluation of external
validity is often not possible. Although results
from prospective pregnancy studies may not be
generalizable to all women of reproductive age,
they are likely to yield important observations
that prompt additional studies.

As demonstrated in other pregnancy-
related studies (Wyatt et al. 2002), prospective
pregnancy studies with semen collection
were successful in obtaining specimens from
most male participants (Bonde et al. 1998;
Couple-based studies permit exploration of
developmental toxicants that may be medi-
ated through exposure of the embryo or fetus
to the components of seminal fluid via intra-
canicular exposure or by absorbance of semi-
1998; Sandberg et al. 1968). Semen collection
provides the opportunity to measure biological
and chemical components of the seminal fluid
(Yay et al. 2001; Younglai et al. 2002), per-
form standard sperm analyses, and even examine
spermatozoal gene expression profiles
(Ostermeier et al. 2002). The routine collec-
tion of semen specimens would further the
assessment of human reproductive function,
as these data could identify paternally medi-
ated developmental effects. Semen analyses
afford an opportunity to identify biomarkers
that could delineate causal mechanisms of
paternal toxicant exposure and/or fertility.

Our review suggests that study participants
were generally willing to participate in studies
even when they included time-consuming
and/or invasive protocols for extended periods
of time. Future studies may yield even higher
rates of participation as technologic advances
are incorporated into study protocols. Examples
of relatively inexpensive technologies that could
be implemented include specially programmed
handheld devices to record menstrual cycle
symptoms (Wyatt et al. 2002), home fertility
monitors based on daily urine dipsticks (Behre
et al. 2000) or salivary or vaginal probes
(Fehring and Schlaff 1998), one-step luteiniz-
hormone tests (Nielsen et al. 2001), finger-
prick blood spots (Worthman and Stallings
1997), home semen collection (Royster et al.
2000), and mouthwash methods for collecting
 genomic DNA (Lum and Le Marchand 1998).
These technologies will be a useful addition to
the biomarkers of fecundity and ovulation cur-
rently in use (e.g., vaginal mucus and BBT)
(Stanford et al. 2002). For example, one recent
study suggests early pregnancies can be detected
with home pregnancy test kits (Buck et al.
2002). These kits have high sensitivity and
specificity for detecting hCG concentrations
of 25 mIU/mL, the level anticipated on the day
following expected menstruation when concep-
tion has occurred (Ehrenkranz 2002). Because
the timing of ovulation can vary in healthy
women, this approach would be most accurate
if used with a marker for ovulation (Wilcox
et al. 2001).

Our assessment of the utility and feasibility
of prospective pregnancy studies has several
limitations. Only published prospective
pregnancy studies were summarized for review. Though we made every effort to learn of all large-scale prospective pregnancy studies undertaken to date, both published and unpublished, the possibility remains that we may have missed some studies. Further, although it would have been valuable to be able to include estimates of study costs and personnel, most investigators were unable to provide us with that information.

In summary, recruiting women/couples for prospective pregnancy studies prior to conception is feasible for both those planning pregnancy and those at risk of pregnancy. Among the population-based studies of women of reproductive age examined in this review (Bonde et al. 1998; Brown et al. 1997; Ellish et al. 1996), the number of participants divided by the size of the target population ranged from 0.8 to 4%. Using a conservative estimate, it therefore appears that about 120 women of reproductive age would need to be approached to identify one eligible woman/couple planning pregnancy who might be willing to participate in a study of this type. Our review suggests that once recruited, women/couples are often willing to complete very intensive protocols, even if only a modest incentive is provided. In one study, when an urban sample of women was presented with the choice of four protocols that ranged in intensity, 74% opted to participate in the most intense protocol (Sweeney et al. 1989).

As previously noted, individuals from underrepresented minority or economically disadvantaged groups should be targeted for recruitment, given their potentially higher risk of exposure to toxicants and possibly greater susceptibility (Sexton 1997). In so doing, investigators should consider factors reported to enhance participation, such as building trust with community participants (Shavers et al. 2002). Finally, couples experiencing infertility-related impairments, including those undergoing assisted reproductive technologies, might represent another group suitable for study, in that exposure(s) to toxicants may be impairing their ability to conceive or carry a pregnancy to term.

References


urinary hormone measurements with the use ClearPlan Fertility Monitor: comparison with transvaginal ultrasound scan and serum hormone measurements. Hum Reprod 15:478–482.
Luker KC. 1998. A reminder that human behavior frequently refuses to conform to models created by researchers. Fam Plann Perspect 31:248–249.
Sibley MR. 1999. Pregnancy intentions may not be a useful measure for research on maternal and child health outcomes. Fam Plann Perspect 31:249–250.


