

Epidemiologic Tools to Study the Influence of Environmental Factors on Fecundity and Pregnancy-related Outcomes

Rémy Slama*, Ferran Ballester, Maribel Casas, Sylvaine Cordier, Merete Eggesbø, Carmen Iniguez, Mark Nieuwenhuijsen, Claire Philippat, Sylvie Rey, Stéphanie Vandentorren, and Martine Vrijheid

* Correspondence to Dr. Rémy Slama, Team of Environmental Epidemiology Applied to Reproduction and Respiratory Health, Institut Albert Bonniot (U823), INSERM and Joseph Fourier University, F-38042 Grenoble Cedex 09, France (e-mail: remy.slama@ujf-grenoble.fr).

Accepted for publication September 13, 2013.

Adverse pregnancy outcomes entail a large health burden for the mother and offspring; a part of it might be avoided by better understanding the role of environmental factors in their etiology. Our aims were to review the assessment tools to characterize fecundity troubles and pregnancy-related outcomes in human populations and their sensitivity to environmental factors. For each outcome, we reviewed the possible study designs, main sources of bias, and their suggested cures. In terms of study design, for most pregnancy outcomes, cohorts with recruitment early during or even before pregnancy allow efficient characterization of pregnancy-related events, time-varying confounders, and in utero exposures that may impact birth outcomes and child health. Studies on congenital anomalies require specific designs, assessment of anomalies in medical pregnancy terminations, and, for congenital anomalies diagnosed postnatally, follow-up during several months after birth. Statistical analyses should take into account environmental exposures during the relevant time windows; survival models are an appropriate approach for fecundity, fetal loss, and gestational duration/preterm delivery. Analysis of gestational duration could distinguish pregnancies according to delivery induction (and possibly pregnancy-related conditions). In conclusion, careful design and analysis are required to better characterize environmental effects on human reproduction.

birth weight; cohort studies; congenital abnormalities; environment; fecundity; fetal membranes, premature rupture; pregnancy; preterm birth

Abbreviations: PROM, premature rupture of the fetal membranes; TTP, time to pregnancy.

INTRODUCTION

Human reproduction occurs through a complex chain of behavioral and biological events involving 3 individuals. Although apparently inessential to the health of individuals, reproduction is crucial for the survival of the species. In addition, perturbations in specific stages of reproduction can have an important health impact: A major health concern related to pregnancy is the rate of perinatal and maternal deaths. There have been important declines in these rates over the last decades in Western but not in all countries (1). Another major health concern is preterm delivery, which is associated with increased neonatal morbidity and mortality and also has long-term consequences, such as increased risk of neurodevelopmental and behavioral adverse events (cerebral palsy and cognitive and school difficulties), or altered pulmonary function in childhood and adolescence (2). Fecundity

represents another important issue, because involuntary infertility concerns a large proportion of pregnancy attempts (probably around 15%–25% of pregnancy attempts in the case of 12-month involuntary infertility) (3, 4), entails psychological suffering (5), and may be associated with medical treatments with limited efficiency and potential negative effects on maternal health. Birth weight is associated with many adult illnesses (6). This association may not be causal but rather may be explained by restricted fetal growth and adult diseases sharing common causes such as environmental exposures during development, as considered in the context of the developmental origins of health and disease (DOHaD) hypothesis (6). This hypothesis, which highlights the importance of development as a window of heightened sensitivity to stressors, as well as the burden entailed by adverse fecundity and pregnancy outcomes, warrants for a better understanding of the impact of environmental exposures during pregnancy.

Hundreds of studies (reviewed, e.g., in 7–10) have identified environmental factors that possibly impact the occurrence of fecundity or pregnancy-related outcomes in humans. The study of such potential impacts is made difficult by methodological challenges linked to the complexity of human reproductive function (11–15). Such challenges include issues related to the following: 1) the identification of the population at risk (the “right denominator”). For example, data on the number of couples who resort to in vitro fertilization are difficult to use to estimate trends in fecundity troubles without knowledge on the number of couples at risk, that is, the number of pregnancy attempts; these couples cannot easily be identified; 2) the fact that many reproductive life events are not illnesses and, consequently, do not systematically require contact with the health-care system; and 3) the interrelation between the events from the start of a pregnancy attempt until the delivery of a child, entailing complex selection phenomena and “competing risks.” For example, an exposure increasing the risk of spontaneous abortion will possibly limit the proportion of surviving fetuses sensitive to the environmental factor, which might in turn reduce the apparent impact of this exposure on gestational duration, fetal growth, or frequency of congenital anomalies.

Our aim was to review and discuss the assessment tools to evaluate fecundity and pregnancy-related outcomes in human populations and their sensitivity to environmental factors. The health outcomes covered include fecundity troubles, fetal loss and growth, congenital anomalies, premature rupture of membranes, and preterm birth (Figure 1). Male (16) or female fecundity parameters, sex ratio, pregnancy complications, and postnatal health are out of the scope of this review. Generally, exposure assessment can rely on biomarkers assessed in (preferably prospectively collected) biological samples (possibly in conjunction with toxicokinetic models, when available), environmental models (for exposures occurring via 1 main pathway, such as air or drinking water, generally in conjunction with questionnaires on behaviors

influencing exposure), personal dosimeters, job-exposure matrices, and questionnaires (17). Issues related to exposure assessment (17) are discussed only if they are specific to the field of reproduction.

FECUNDITY AND TIME TO PREGNANCY

Definitions

Fecundity corresponds to the biological ability to obtain a livebirth. We follow the demographic terminology that distinguishes fecundity from *fertility*, the actual number of livebirths of a woman (18). A quantitative estimate of fecundity is *fecundability*, the cycle-specific probability for a couple not using any birth control method and with regular sexual intercourse to obtain a pregnancy (13, 19, 20). Time to pregnancy (TTP) (20) corresponds to the number of months (or menstrual cycles) elapsed between the start and the end of the period of unprotected intercourse (because of a pregnancy, because of the end of the relationship, or because the couple no longer wishes to obtain a pregnancy). Note that some authors define TTP for only the periods of unprotected intercourse leading to a pregnancy, which we do not recommend, following the common practice of survival analysis in which time periods not ending with the event of interest should not be excluded, but rather censored. *Involuntary infertility* is used to indicate that a couple has unsuccessfully tried to obtain a pregnancy; it should be mentioned together with a duration (e.g., 12-month involuntary infertility) (18). Usually, fecundability (estimated by the proportion of cycle 1 conceptions) is in the 20%–30% range (Figure 2), while 15%–25% of couples are expected to suffer from 12-month involuntary infertility (3, 4). *Sterility* corresponds to a biological inaptitude to obtain a pregnancy. The diagnostic of sterility requires medical examinations that can be difficult to systematically perform in the context of large population-based cohorts; this outcome will not be considered here.

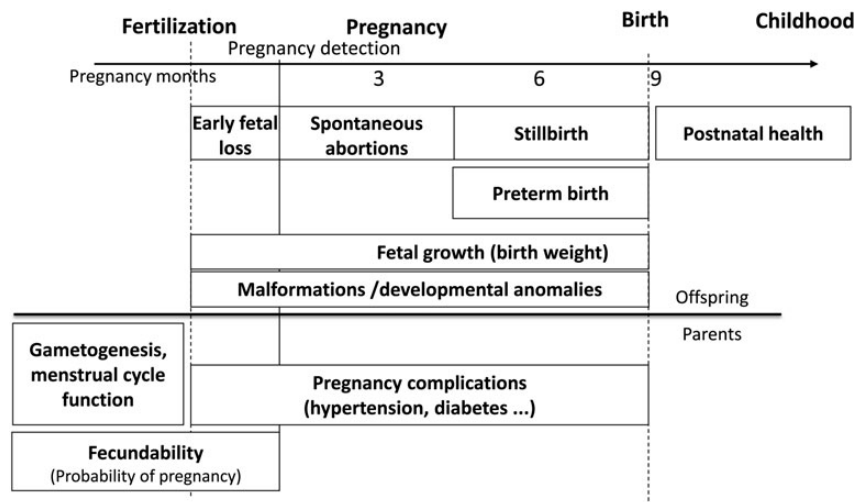


Figure 1. Overview of the timeline of pregnancy-related outcomes covered.

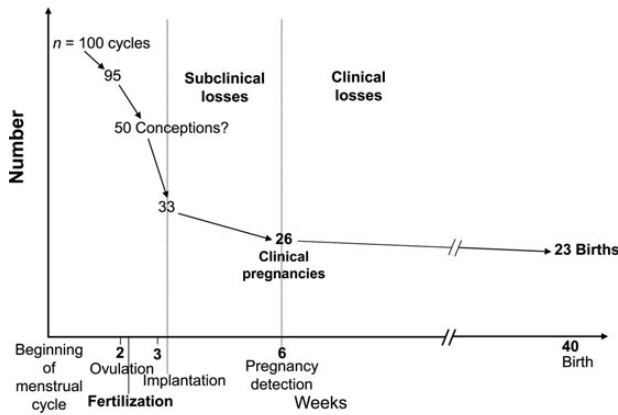


Figure 2. Evolution of the number of ongoing pregnancies, from the start of a menstrual cycle, among 100 cycles (assuming that ovulation occurs in 95 of these cycles). Approximate figures; slightly adapted from Baird et al. (20).

Options for study design and outcome assessment

A straightforward estimate of fecundability at the population level is the proportion of couples for which a pregnancy starts during the first month of unprotected intercourse. This proportion, as well as the survival curve describing the cumulative proportion of couples who obtain a pregnancy over time, can be derived by assessing TTP (20). TTP also allows one to estimate the frequency of involuntary infertility (e.g., using a Kaplan-Meier survival curve) if TTP has been assessed for all pregnancy attempts, including those not leading to a livebirth. TTP is usually assessed by a questionnaire. The TTP study designs can be distinguished according to whether couple identification takes place before, during, or after the period of unprotected intercourse (Figure 3). The latter

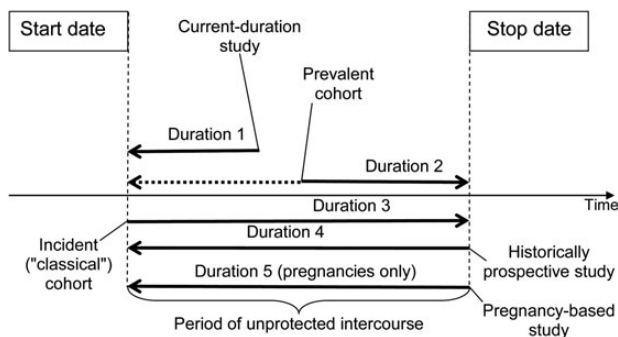


Figure 3. Main designs for time to pregnancy studies (27). The start date may be either the discontinuation of a method to avoid pregnancy or the end of a pregnancy not followed by use of methods to avoid pregnancy. The stop date may be the beginning of a pregnancy (detected later), the resumption of any method to avoid pregnancy (contraception, sexual abstinence), or the initiation of medical treatment for infertility. These last 2 situations correspond to censoring events. Duration 5 (pregnancy-based design) is assessed only if a pregnancy is detected at the end of the period of unprotected intercourse.

corresponds to the most frequent design and has among others been used in the context of cohorts of pregnant women or birth cohorts or of cross-sectional samples of women recruited at the maternity clinic or sometime after delivery. This so-called *pregnancy-based design* (21) has the advantage of being the simplest from a logistic point of view, allowing a large sample size, but the disadvantages that TTP is assessed retrospectively (inducing measurement error) (22) and that couples remaining infertile are excluded (giving rise to a potential selection bias and underestimation of involuntary infertility rates). To avoid this, information on past periods of unprotected intercourse not leading to a pregnancy can be retrospectively collected. Alternatively, recruitment can also be performed before the start of the period of unprotected intercourse (23, 24). In this incident cohort setting, an option is the detailed prospective design; it implies collection of information on timing of intercourse within each month, menstrual bleeding, and a marker of ovulation (possibly using devices monitoring urinary hormone metabolites, such as home fecundity monitors) (24), thereby allowing the making of inferences regarding the impact of environmental factors on day-specific probabilities of conception (25). Another design implies recruitment of couples during the period of unprotected intercourse. If the couples are followed up, then this corresponds to a prevalent cohort design; if not, then the design is termed *current duration* (26, 27). Whatever the design, an option relates to whether or not the study should be restricted to first pregnancy planners (see below). The case-control study may be particularly prone to selection bias when fecundity is the outcome in focus, because if cases correspond to couples resorting to infertility treatments, controls should be sampled from the source population of these couples, which is difficult to identify.

Main threats to validity

Sources of bias (21, 27–29) and possible remedies are indicated in Table 1. The exclusion of infertile couples occurring in the pregnancy-based design was formerly considered as a minor issue (20); however, this exclusion can in some situations strongly bias the estimated effect of exposure variables on time to pregnancy (30–34). Several biases are related to the complex notion of pregnancy planning (Table 1) (35). The main issue is that, in the context where a relatively large proportion of pregnancies are unplanned and hence excluded from time to pregnancy studies because the outcome cannot be assessed, bias may result if these couples with unplanned pregnancy have a different distribution in terms of fecundity (which is likely) and in terms of environmental exposures.

Recommendations

Design. The case-control design, if not nested within a larger well-defined cohort, should be avoided. The preferred design for fecundability studies corresponds to recruitment before the start of the pregnancy, allowing including couples eventually remaining childless; this is achieved in the incident cohort, prevalent cohort, and current duration designs (see above). A pregnancy-based design might entail bias and a more limited statistical power for a given sample size

Table 1. Main Sources of Bias in Studies of Environmental Impacts on Time to Pregnancy

Source of Bias	Description of Potential Bias	Possible Approaches to Limit Bias
Study design	Exclusion of infertile couples (e.g., in a pregnancy-based design) (30–34)	Consider alternative designs (incident or prevalent cohort, current duration approach). In a population-based retrospective setting, try to collect information on unsuccessful attempts at pregnancy.
	Truncation (of short/long TTPs at the beginning/end of the study period, respectively) (40, 106)	Define the study period with respect to the date of the beginning (and not the end) of the period of unprotected intercourse; include current attempts.
	Pregnancy planning bias 1: exclusion of unplanned/mistimed pregnancies	Ascertain exposures for nonplanners as well and conduct sensitivity analyses (21, 107).
Outcome assessment	Pregnancy planning bias 2: unplanned pregnancies can retrospectively be described as planned (also termed “wantedness bias”)	Reanalyze excluding declared cycle 1 conceptions. Define inclusion criteria on contraceptive use rather than pregnancy wish.
	Pregnancy planning bias 3: exclusion of couples who have unprotected intercourse without planning to become pregnant	Try to include periods of unprotected intercourse corresponding to couples not planning to become pregnant and reanalyze.
	Pregnancy recognition bias: couples in whom pregnancy is diagnosed very early may identify pregnancies ending with an early loss that would not have been identified by other couples (21)	Record when and how pregnancy was recognized. Restrict analyses to pregnancies leading to a livebirth.
	Use of oral contraception may vary with exposure and might be associated with decreased fecundability in the first cycles	Assess the last contraceptive method used; ask if couples used abstinence after discontinuation of pill; treat pill use as time-varying covariate.
	In a retrospective setting, couples may recall TTP with some error (e.g., digit preference) (22)	Focus on pregnancies leading to a livebirth, for which recall may be better; group consecutive TTP values (e.g., 5–7, 11–13 months).
Medical intervention	Medical intervention bias: 1) medical treatments may modify the probability of pregnancy; 2) if couples remaining childless are not recruited and if exposure is associated with the delay before couples resort to medical help, bias may occur	Treat infertility treatment as a censoring mechanism; try to assess if infertility treatments depend on exposure (conditionally on waiting time).
Exposure assessment	Bias due to time trends in exposure and TTP: if, for example, TTP tended to be longer at the start of the pregnancy period, and if exposure tended to decrease over time, then a spurious association (corresponding here to an apparently longer TTP in association with exposure) might be induced (28)	Simulate by using external data on the time trends in exposure (28).
	Exposure is assessed during or after pregnancy instead of at the start of the period of unprotected intercourse. Bias can occur if exposure varies quickly over time (e.g., from month to month) or increases with time (creating a spurious association with TTP)	Assess exposure at the start of the period of unprotected intercourse (38).
	Behavior modification bias (if exposure is influenced by behaviors; for example, couples may stop smoking after a delay of unsuccessful attempt at pregnancy)	Use date of discontinuation of contraception as a reference date to define exposure and confounders. Assess modifications of behavior and use a model with time-dependent covariates. Censor TTP (e.g., at 12 months).
	Toxicokinetic bias (in studies relying on exposure biomarkers): bias due to toxicokinetic changes in the body burden of the environmental factor following a previous pregnancy (108)	Stratify analyses on parity (37).
Statistical analysis	Proportional hazard hypothesis not verified	Test for different effects of exposure during months 1–3 and 4–12 (or other cutoff) of the pregnancy attempt.

Abbreviation: TTP, time to pregnancy.

(32). However, if bias is limited (which is hard to know a priori), the loss in statistical power could be compensated by a larger sample size, which is easier to achieve because of the higher eligibility rate for this design compared with the other ones; information on periods of unprotected intercourse not leading to a pregnancy should also be collected in this retrospective design, which can be done if recruitment is not done at a maternity clinic and if inclusion criteria are based on having had a pregnancy or a period of unprotected intercourse started in a given time period.

Effort should be made to collect information on couples with unplanned or mistimed pregnancies, including on their exposures and behaviors before the start of the pregnancy, to discuss possible bias related to pregnancy planning (Table 1); this is probably easier in the context of a pregnancy-based approach than for other designs.

It might be relevant to restrict time to pregnancy studies to couples with no previous pregnancy or pregnancy attempt; reasons for such a choice include 1) the fact that recruiting already fertile couples will tend to overrepresent the most fecund ones—even in societies with wide access to contraception, there is a rather weak but real association between time to pregnancy and family size (36)—and 2) that, in studies in which exposure is assessed with biomarkers, past pregnancies and breastfeeding before the period of unprotected intercourse under study will impact the measured levels of biomarkers before the period of unprotected intercourse in a way that is dependent on past reproductive history and possibly fecundity level. This issue has been highlighted in a study of perfluorinated compounds on fecundity, in which an increased risk of subfecundity in association with exposure has been observed in parous but not nulliparous couples (37). Blood levels of perfluorinated compounds are known to decrease during pregnancy and lactation as the result of transfer to the fetus (baby), as well as to a lesser extent through increased renal excretion (increased glomerular filtration rate during pregnancy). The authors' interpretation of the observed association was that, in parous couples, the longer the interpregnancy interval, the higher the potential for concentrations of perfluorinated compounds to increase again. Because, in parous couples, a long TTP will tend to increase the interpregnancy interval, a bias toward an association between long TTP and high concentrations of perfluorinated compounds (corresponding to reverse causality) could be expected (37).

The main points on which TTP studies should collect information are indicated in Web Appendix 1 available at <http://aje.oxfordjournals.org/>.

Exposure assessment. The event of interest being the occurrence of a pregnancy, the relevant exposure window is some time before the start of the pregnancy (38). This also applies to potential confounders (e.g., one should adjust for smoking or age at the start of rather than after the period of unprotected intercourse or during pregnancy). Information on variations in exposure and in confounding factors during the period of unprotected intercourse should be collected.

Statistical analysis. Statistical analysis of time to pregnancy data requires specific statistical approaches (e.g., discrete-time Cox model) and many sensitivity analyses (21). This setting theoretically allows the handling of time-

varying covariates (e.g., if exposure varies during the period of unprotected intercourse), although this has been applied very little in TTP studies. Special attention should be given to truncation bias (39, 40) and other sources of bias leading to temporal trends in TTP (28); these may occur if inclusion is based on the year of birth rather than the year of the start of the period of unprotected intercourse. If couples with previous history of pregnancy or pregnancy attempt are included, specific considerations are required; adjustment for this previous history may not be relevant (41). Analyses stratified on parity may be warranted to handle the above-mentioned specific issues related to overselection of the most fecund couples and to biomarker-based exposure assessment (37).

FETAL LOSS

Definitions

Fetal losses are defined as a spontaneous end of pregnancy, without living birth, occurring between conception and the end of labor; in practice, because pregnancies cannot currently be detected until the time of implantation, this window should be narrowed to the period between implantation and the end of labor. Induced abortions and medical termination of pregnancy are not considered as fetal losses, and neither are ectopic pregnancies (corresponding to embryos implanted outside the uterus).

Fetal losses encompass a very heterogeneous and broad continuum. The one end of this continuum corresponds to a silent event occurring around the time of implantation (an early fetal loss), while the other end of this continuum corresponds to a stillbirth, with usually major psychological consequences for the couple. The proximal causes of fetal losses are also very heterogeneous with, for example, chromosomal anomalies being much more frequent in first compared with second trimester spontaneous abortions (42).

The fetal loss continuum is subdivided into categories defined according to gestational age (Figure 4), starting from early (or subclinical) fetal losses (between implantation and the time of detection of the pregnancy, or within 6 weeks from the last menstrual bleeding), spontaneous abortions (sometimes referred to as miscarriages), and stillbirths. The gestational age cutpoint between spontaneous abortions and stillbirths has different legal definitions according to national rules for registering births (43) and varies between 20 and 28 weeks. In epidemiologic studies, the cutpoints at 20 or 22 gestational weeks are most often used.

Because conception is difficult to identify prospectively outside the setting of in vitro fertilization, and because early fetal losses are not accompanied by clinical signs other than vaginal bleeding, the frequency of early fetal loss is difficult to quantify. Overall, fetal losses (from implantation onward) may occur in about 20%–30% of all conceptions (44–46); most of these (about two thirds) remain undetected by couples (Figure 2). The frequency of spontaneous abortions typically corresponds to 10%–15% of identified pregnancies in women younger than 35 years (11). In Europe in 2010, stillbirths corresponded to about 0.3%–0.9% of detected pregnancies if medical terminations are counted as stillbirths (47).

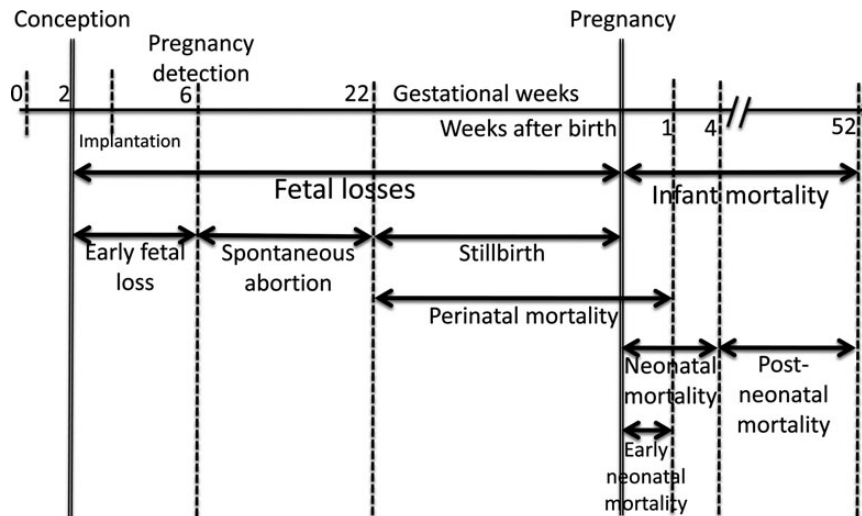


Figure 4. Definition of fetal loss and infant mortality categories expanded from Wilcox et al. (14).

Options for study design and outcome assessment

Study design will depend on the type of fetal loss in focus. Except for stillbirths, cohort studies are a relevant setting. Questionnaires are not adapted to assess early fetal loss occurrence, as they identify a very small proportion of the fetal losses occurring before 6 gestational weeks, among women who detected very early that they were pregnant, who may constitute a selected group. For this outcome, regular (e.g., daily) urine sampling of women attempting to become pregnant to assay human chorionic gonadotropin levels can be performed to identify a rise in its level (which typically occurs 7–11 days after conception (48)) and possibly a drop in the level (usually followed by vaginal bleeding), as a biological measure of early pregnancy loss (46, 49). A group of sterile women may be used to efficiently define the threshold of human chorionic gonadotropin concentration indicative of a pregnancy start (50).

Outside the context of assisted reproduction techniques, some studies relying on couples planning to start a pregnancy attempt have been conducted to study early losses (46, 49, 51, 52); such couples are difficult to recruit, apart from exceptional situations such as that of China, where couples have to register before planning a pregnancy (49).

Spontaneous abortions and stillbirths are usually detected by either the woman herself or her clinician, for example, via ultrasound, and a questionnaire-based approach can be used. Studies relying on clinical data from the maternity ward to identify fetal losses only allow identification of the subgroup of fetal losses leading to an admission to the hospital. Prospective studies require identification of women (or at least a large number of them) before the end of the first trimester, if possible between 6 and 10 weeks after the last menstrual bleeding, and their follow-up (53). For stillbirths, recruiting women before the end of the second trimester may be sufficient.

Main threats to validity

The main sources of bias are listed in Table 2. Many of them are related to the fact that there are variations in the gestational age at pregnancy recognition or at inclusion. These variations may, for example, be due to women pregnant as a result of a failure in birth control methods detecting the pregnancy later than those actively seeking to become pregnant. These variations in the gestational age at pregnancy recognition will lead to some losses not being included (corresponding to left truncation in the terminology of survival modeling) and to differences in the observed time at risk between women; this may lead to bias, in particular if these variations are related to exposure (54).

Recommendations

Study design. Cohorts with recruitment as early as possible during (or before) pregnancy appear as a relevant design for studies of environmental influences on the risks of spontaneous abortion and stillbirth. If early fetal losses are the focus, then recruitment should take place before the beginning of the pregnancy, and biological samples should be collected on a daily or almost daily basis. In this setting, trying to obtain information on couples who became pregnant without planning it, by separately recruiting a group of such couples with unplanned pregnancies, may be relevant to discuss any bias resulting from their exclusion from the main analysis done on the couples recruited before the beginning of the pregnancy.

Exposure assessment. Only exposures before or during the time period when the type of fetal losses considered can occur should be considered; because the duration of the pregnancy varies according to its outcome, the probability for exposure to occur or to be detected may differ according to the

Table 2. Main Sources of Bias and Possible Cures in Studies of Environmental Impacts on Fetal Loss^a

Source of Bias	Description of Potential Bias	Possible Approaches to Limit Bias
Study design	Gestational age at start of follow-up or at pregnancy detection varies between women and may be associated with exposure (54). Study is based on the recall of the last pregnancy only, leading to an underestimation of fetal loss rate (109). Elective or therapeutic abortions may be described as spontaneous abortions.	Assess date of pregnancy recognition and gestational age at start of follow-up and analyze data with a survival model with delayed entry (54, 56). Alternatives include collecting information on all past pregnancies or using a prospective design. Design study questionnaire with great care to distinguish pregnancy outcomes. Collect medical records.
Outcome assessment	The fetal loss may occur several weeks after the death of the fetus.	Identify the gestational age and results of all ultrasound examinations. Consider using a discrete-time (interval-censored) survival approach allowing consideration that the fetal death occurred in a time window rather than during a specific week.
Exposure assessment	Exposure window varies with outcome. Exposure varies during pregnancy (or probability of exposure depends on pregnancy duration), leading to noncausal differences in exposure between pregnancies with short (fetal losses) and longer (livebirths) durations.	Assess temporal variations in exposure and analyze data with a survival model with time-varying covariates. Consider only very early exposures (e.g., during first month of pregnancy), if there is a biological relevance.
Statistical analysis	Previous history of fetal loss is adjusted for in the model, whereas a previous fetal loss may be due to the exposure under study. Pregnancies ending as an induced abortion or medical termination may have otherwise ended as a spontaneous abortion; these pregnancies are not considered in the denominator when assessing the frequency of fetal loss.	Build a directed acyclic graph to discuss relevance of adjustment for previous reproductive history (41, 57). Use a survival analysis; treat induced abortions and medical terminations of pregnancy as competing events.

^a Early fetal losses are not considered.

pregnancy outcome, possibly leading to bias, as discussed for another pregnancy outcome (55). Similar issues will occur if exposure is not assessed as a binary but as a continuous variable, such as a pregnancy average. This time-related bias can be handled in the statistical analysis (see below) if information on variations in exposure during pregnancy has been collected.

Statistical analysis. Statistical analysis of fetal loss must use survival models to take into account left truncation (due to losses occurring prior to enrollment) and right censoring (54, 56). Right censoring (or possibly competing risk models) is required to handle induced abortions and medical pregnancy terminations. The use of such approaches requires information on the date of identification of the pregnancy.

Adjustment for reproductive history (e.g., previous history of spontaneous abortion) should not be systematically performed and may even induce a bias (41, 57), nor is a variable identifying “usual aborters” relevant in analyses (58). Although including past reproductive history in a regression model can generally increase its ability to predict future reproductive history (and hence the model fit), past reproductive history does not always correspond to a confounder (in particular if a past adverse pregnancy outcome had been caused by a persistent exposure corresponding to that under study) (59). Howards et al. (57) illustrate that a careful definition of the research question and a representation of the hypothesized causal relations among exposure, outcome, and past reproductive history (e.g., through a directed acyclic graph) are helpful to identify when adjustment for past reproductive history, and possibly past exposures, is required. Additional analyses describing any association between exposure levels and the gestational age of detection of the pregnancy can be useful to discuss potential bias.

CONGENITAL ANOMALIES

Definitions

The definition of congenital anomalies classically includes structural malformations, syndromes, functional defects, and chromosomal anomalies present in the fetus or the newborn, as defined by the *International Classification of Diseases*, 10th Revision (ICD-10), codes Q00.0–Q89.9.

Embryos affected by an anomaly may spontaneously abort before clinical recognition of pregnancy, be identified during prenatal screening (followed or not by pregnancy termination), or be detected at birth (in live- or stillbirths) or later in life. The presence of anomalies in aborted embryos is expected to be frequent because malformations are a major cause of abortion (60) and because chromosomal anomalies (a proximal cause of malformations) are also a major cause of abortion (42). Depending on the study design, only a subset of all conceptuses affected by anomalies are usually identified.

Because of this potential for missed diagnoses, the frequency of anomalies is usually measured by a prevalence rate, expressed as percent births (even if anomalies are not all detected at birth). The overall prevalence is around 2% when only major anomalies are selected and may be much higher when minor forms are included (15, p. 1079; 61). Specific types of major anomalies (neural tube defects, oral clefts, severe hypospadias . . .) touch a few births per thousand.

Options for study design and outcome assessment

Because of the low prevalence of single categories of congenital anomalies (a few per thousand births), the preferred study design has traditionally been a case-control approach. Registers of congenital anomalies have been set up in many countries and generally provide the basis for identification of cases. Registers distinguish major and minor anomalies and often register only the major ones. The definition, diagnosis, and reporting of minor anomalies vary considerably across registries (62). As an example, until recently, glanular (or coronal) hypospadias (a malformation in which the urinary opening is misplaced on the glans) was not considered a major anomaly by the European Registration of Congenital Anomalies and Twins (EUROCAT) and consequently not uniformly registered (63). Similar difficulties exist for cryptorchidism, the most frequent congenital anomaly in boys (64, 65). Both sensitivity and accuracy of the diagnosis of congenital anomalies will depend upon the strategy of prenatal screening, the classification used (66), and the sophistication of medical investigations. Screening procedures include ultrasonography and biochemical tests with a timing during pregnancy that is usually defined by the national screening policies. The final diagnosis is based on clinical examination, autopsy or surgical reports, and complementary investigations such as echocardiogram, tomodensitometry, and cytogenetics.

As usual, the recruitment of controls needs to be carefully designed. Consider a case-control study of air pollution effects on congenital anomaly. If cases and controls are matched on birth date, possible differences in exposure distribution between cases and controls will be underestimated because the date is a strong determinant of air pollution level. A similar issue may occur if cases and controls are matched on the maternity of birth, which will possibly limit the spatially driven exposure contrasts between cases and controls. Mother-child cohorts recruiting pregnant women early during pregnancy can constitute the source population for nested case-control analyses if cases of congenital anomalies have been diagnosed during the follow-up or have been identified through record linkage with congenital anomalies registries. This design offers the greatest opportunities for the study of the most frequent anomalies, such as cryptorchidism (67). For persistent chemicals, measurement of exposure through assay of chemicals in biological media likely to reflect exposure during pregnancy, such as cord blood, placenta (68), or meconium, can be applied to case-control studies (ideally, nested in cohort studies to allow collection of biological samples before occurrence of outcomes). Related challenges include possible bias due to any impact of pregnancy duration or from time since the last pregnancy on the biomarkers' levels, as exemplified for fecundity studies (37), as well as difficulties in collecting biological samples from pregnancy terminations and children with anomalies diagnosed several weeks after birth. Family triads (cases and their 2 parents) have also been used to estimate gene-environment interactions (69).

In countries where registration of congenital anomalies among births, stillbirths, and pregnancy terminations covers a large population and is accurate and reasonably exhaustive, ecological correlation (registry-based) studies at small geographical scale have been conducted (70). In addition to

issues in terms of exposure assessment, a limitation of such studies is that few individual characteristics and behaviors are recorded besides knowledge of the defect, sex of the baby, parental postal address (usually at birth), and maternal age. The time-series design generally allows one to characterize short-term effects of exposure, which, because the window of sensitivity to pollutants is more likely to be in the first months of pregnancy (see below), does not represent an appealing approach.

Main threats to validity

Potential biases are summarized in Table 3. Like for fetal loss, a potential bias lies in variations in the duration of the period of follow-up, a duration influencing the measured prevalence of anomalies.

Recommendations

Study design. In cohort studies, ideally, timing of inclusion in the cohort should occur before the median gestational age of prenatal diagnosis in the corresponding country. Diagnosis of congenital anomalies should be performed at least in medical terminations of pregnancy and preferably also in stillbirths, in addition to livebirths. In some areas covered by a birth defects registry, linkage of the cohort members with records from this registry can provide additional data to identify and characterize congenital anomalies.

Outcome assessment. To limit heterogeneity in outcome assessment, cases of mild anomalies such as cryptorchidism can be distinguished according to whether or not this condition requires surgical repair, although there may also be variability in medical practices. Congenital anomalies diagnosed after birth should be recorded with a follow-up of at least 1 year. In the special case of some anomalies not leading to fetal death or medical termination of pregnancy, and present at birth, a standardized protocol of assessment and reporting is advised.

Exposure assessment. The embryonic period of intrauterine development is highly sensitive to teratogens. Germinal cells, especially spermatozoa, may be affected before conception and result in affected fetuses. Assessing exposure in early pregnancy and if possible before conception is therefore necessary. Moreover, when studying malformations of specific organs, exposures during the time window of development of the organ should be specifically considered (e.g., gestational weeks 3–7 in the case when heart defects are of interest) (71).

GESTATIONAL DURATION AND PRETERM BIRTHS

Definition

Gestational duration corresponds to the time between the first day of the last menstrual period and birth. The median duration of pregnancy has been estimated to be 40 gestational weeks and 2 days, or 282 days (72). Dichotomous outcome measures of gestational duration exist (Table 4), the most frequently used being preterm births (below 37 completed gestational weeks). Note that, in the 1919–1961 period and

Table 3. Main Sources of Bias and Possible Cures in Studies of Environmental Impacts on Congenital Anomalies

Source of Bias	Description of Potential Bias	Possible Approaches to Limit Bias
Study design	The overall prevalence of anomalies is conditioned by the duration of observation (from conception up to a few years after birth).	Ascertain anomalies in therapeutic, spontaneous abortions and stillbirths; collect information on anomalies diagnosed during the first year of life.
Outcome assessment	<p>Definition, diagnosis, and reporting of minor anomalies vary across registries (62–65).</p> <p>Prevalence of specific anomalies such as cryptorchidism decreases with gestational age (64).</p> <p>Identification of minor anomalies may vary between clinicians and between malformation registers.</p> <p>Sensitivity and specificity of identification of a given anomaly may vary with the classification used (66).</p>	<p>Review individual medical records. Study with particular caution observations with multiple anomalies (62).</p> <p>Take into account gestational duration in analysis.</p> <p>Use a standardized protocol to assess and report the anomalies of interest.</p> <p>Use fine classifications; review individual medical records. Study with particular caution observations with multiple anomalies (62).</p>
Medical interventions	Access to prenatal diagnosis may vary with couples' characteristics and possibly exposures (110).	Assess if access to prenatal diagnosis is associated with exposure in controls. In studies in large areas, identify differences in screening policies (110).
Exposure assessment	The biological window of sensitivity to environmental factors may be early in pregnancy and narrow for many anomalies.	Identify the a priori relevant time window from biological and toxicological knowledge; finely assess exposure in this time period.

Table 4. Main Cutoffs Used in the Study of Gestational Duration

Definition	Cutoff ^a	
	No. of Weeks ^b	No. of Days
Extremely premature birth	<28	<196
Very preterm birth	<32	<224
Preterm birth	<37	<259
Term birth	37–41	259–293
Postterm birth	≥42	294

^a Since the end of the last menstrual period.

^b Completed gestational weeks, corresponding to the integer part of (date of birth – date of last menstrual period)/7.

beyond, a birth weight below 2,500 g was used as an operational definition of preterm births (14).

Preterm birth occurs in 4%–13% of all births (while very preterm births account for about 1% of all births) and is associated with strongly increased perinatal mortality and long-term morbidity (73). Preterm deliveries are heterogeneous in terms of assumed proximal causes or associated maternal conditions (ischemic placental diseases, infectious or inflammatory context, . . .) and of clinical context (premature rupture of the chorioamniotic membranes, preterm labor, induced labor, or cesarean section for medical indications, e.g., to try to protect the mother or the fetus) (74). A classification of preterm births relying on 5 components has been proposed (75). These components are as follows: 1) the maternal conditions present before presentation for delivery (e.g., infection, preeclampsia); 2) the fetal conditions present before presentation for delivery (intrauterine growth restriction, polyhydramnios, . . .); 3) placental pathologic conditions; 4) signs of the initiation of parturition (cervical shortening, premature rupture of membranes, . . .); and 5) pathway to delivery (spontaneous or induced).

Options for study design and outcome assessment

Outcome assessment. Outside the context of medically assisted reproduction, fertilization is a silent event, so that the date of conception is most often estimated retrospectively and rarely known with accuracy. Gestational age can be defined in 3 ways: 1) from the date of the last menstrual period, which is usually collected retrospectively by questionnaire during pregnancy or after birth, by the medical staff or specific questionnaires; 2) from ultrasound measures in early pregnancy, making use of the property that specific fetal measures vary linearly and rather homogeneously in the very beginning of pregnancy, and that these ultrasound measures (e.g., biparietal diameter or crown-heel length) can be used as a proxy for gestational duration, using growth standards that are now recorded in the ultrasound devices; and 3) from a third (mixed) approach, in which ultrasound measures are used to correct the questionnaire-based date of the last menstrual period, when the discrepancy between the 2 approaches exceeds a given threshold (e.g., 2 weeks). Not treating preterm delivery as a single entity is relevant; as an illustration, maternal smoking seems to decrease the risk of very preterm delivery

with a hypertensive context but to increase the risk of preterm delivery associated with other pregnancy-related conditions (76). Finer distinctions between subtypes of preterm deliveries, possibly relying on the components outlined above, may be a relevant option for future studies.

Study design. The case-control approach should be reserved to the rarest forms of preterm delivery, such as very preterm birth. Cohorts of pregnant women allow the greatest flexibility in terms of statistical analysis and assessment of exposures and underlying conditions. Time series are another option (77), depending on the exposure considered, if this exposure has short-term variations and is expected to have short-term effects. The use of case-crossover methods to study acute triggers of preterm birth has also been proposed (78, 79).

Main threats to validity

The main sources of bias in studies of gestational duration are summarized in Table 5. The last menstrual period-based approach suffers from (generally random) recall errors in the estimation of the date of the last menstrual period and from misclassification errors due to between-women or within-woman variations in the duration of the first phase of the menstrual cycle (80). If the studied exposure influences menstrual cycle duration or regularity, this error will be differential according to exposure, so that measurement error in gestational duration based on the last menstrual period may also differ according to exposure. Although the ultrasound-based definition of the date of fertilization is the most efficient one in predicting the date of birth (81), it is probably not the most relevant for etiologic studies aiming at identifying the influence of environmental factors on gestational duration (15). Indeed, this approach relies on the assumption that fetal growth (e.g., of biparietal diameter or femoral length) is similar for all fetuses during the first trimester of pregnancy. It has been shown that this assumption does not hold (82). If the environmental factor studied impacts on fetal growth during the first trimester of pregnancy, then the error in the estimated gestational age based on ultrasound measures will possibly depend on exposure (Figure 5), so that the less accurate but also less prone to bias measure based on the last menstrual period could then be preferred (15).

Recommendations

Outcome assessment. Information on gestational duration should be collected by using both the definition based on the date of the last menstrual period and on the ultrasound measures, and sensitivity analyses relying on these different approaches should be conducted. When one has to rely on the hospital records to estimate gestational duration, it is important to understand which approach is generally used by the local practitioners. Gestational duration is counted in completed weeks (i.e., rounded to the closest lower integer) in the clinical setting, but this is a considerable loss in information, and a continuous measure should be preferred (in particular if gestational duration is to be analyzed as a “continuous” outcome, or as an adjustment factor in the study of birth

Table 5. Main Sources of Bias and Possible Cures in Studies of Environmental Impacts on Gestational Duration and Preterm Delivery

Source of Bias	Description of Potential Bias	Possible Approaches to Limit Bias
Outcome assessment	Assessment of gestational duration based on last menstrual period is subject to random error. Bias can arise if exposure is associated with menstrual cycle function.	Use alternative definitions of gestational duration in sensitivity analyses.
Medical treatments	Assessment of gestational duration based on early ultrasound dating is subject to systematic error (bias) if exposure is associated with fetal growth.	Perform ultrasound dating as early as possible in pregnancy (before gestational week 12) if exposure is likely to impact fetal growth, the gestational duration estimate based on the date of the last menstrual period should be used (15).
	The handling of medically induced deliveries before 37 gestational weeks requires specific consideration, because it is not clear when the delivery would have taken place in the absence of induction. Exposures after week 37 cannot impact the risk of preterm delivery. Because the probability of occurrence of an exposure or its mean level may vary with the length of the time window considered, any comparison of preterm and term births in terms of exposure during pregnancy as a whole may be biased. In time-series analyses of preterm delivery, if exposure varies over time, then seasonal patterns in the characteristics of pregnancies may induce a bias (111).	Consider using a survival modeling in which induced deliveries (possibly with specific indications) are handled as censoring or competing events. Consider only exposures during the first 37 gestational weeks (or less). Use a survival modeling with time-varying exposure (55). Consider only exposures during the first 37 gestational weeks (or less). Use a survival modeling with time-varying exposure (55). Check for seasonal patterns in factors influencing preterm delivery.

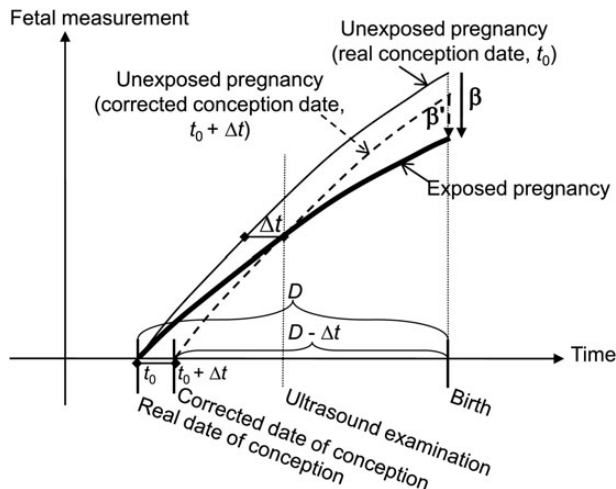


Figure 5. Illustration of the bias possibly occurring in studies of environmental effects on fetal growth, when gestational age is defined by early ultrasound measures (96). Hypothetical evolution of a fetal measurement (e.g., fetal length) during pregnancy, for a pregnancy exposed or unexposed to an environmental factor affecting fetal growth from early pregnancy. The ultrasound examination leads the obstetrician to correct the date of conception of the exposed pregnancy by Δt , so this exposed pregnancy is not compared with unexposed pregnancies with the same gestational age D (solid curve) as should be the case but with gestational age $D - \Delta t$ (dashed curve). Consequently, the estimated difference in the gestational age-specific fetal measurement at birth between exposed and unexposed pregnancies is not the correct value β but a smaller value β' .

weight). Detailed information on the onset of labor (natural or induced; reason of induction) should also be collected.

Issues related to exposure. When preterm birth is the outcome of interest, exposures after gestational week 37 in noncases should not be taken into account (as they are posterior to the possible period of occurrence of the outcome of interest).

Statistical analysis. Because what is actually studied is the occurrence of a delivery and because censoring events (e.g., specific medical intervention) occur, survival models (rather than logistic regression of preterm delivery as a binary outcome) should be considered as the default option in studies not relying on a case-control design (55). Induced deliveries with specific causes can be handled as censoring (or competing) events in the survival model. Preterm deliveries with different underlying pregnancy-related conditions could be considered separately. The continuous (linear regression) analysis is tempting when the number of cases of preterm deliveries is limited; one should, however, keep in mind that mean gestational duration (which is modeled in linear regression) will be strongly influenced by variations in gestational duration in the range of 38–41 weeks, where most pregnancies are. These variations may be driven by factors unrelated to the frequency of preterm delivery, such as local habits in terms of delivery induction and of handling of postterm births, which strongly vary between countries. Quantile regression, allowing the modeling of, for example, the 10th percentile of gestational duration, might be a better option

than linear regression. Sensitivity analyses using the various definitions of gestational durations should be performed, to see if the estimated impact of environmental factors on gestational duration or on the risk of preterm birth varies according to the definition used.

PREMATURE RUPTURE OF THE FETAL MEMBRANES

Definition

Premature rupture of the fetal membranes (PROM) is defined as the spontaneous rupture of the amniotic membranes before 37 weeks of gestation, at least an hour before the onset of contractions. PROM occurs in approximately one third of all preterm births (83). If left untreated, PROM will lead to onset of labor, possibly causing miscarriage or preterm delivery. Apart from the risk posed by preterm delivery, amniotic infections due to ascending microbes from the genital tract are a major possible consequence of PROM. Furthermore, the absence of amniotic fluid increases the risk of fetal postural deformities and pulmonary hypoplasia.

Options for study design and outcome assessment

The diagnosis is set by a vaginal inspection where fluid containing meconium or vernix is observed coming through the cervix. Also, increased pH in the upper vagina (a sign of possible presence of amniotic fluid) may be tested by nitrazine yellow. Immunoassays for amniotic-specific proteins are also available. A reduced amount of amniotic fluid (oligohydramnion) cannot be used as a diagnostic criterion because it is not specific to PROM.

Main threats to validity

Related to study design. Retrospective studies (i.e., with postnatal recruitment) suffer from classification errors, in particular if researchers have no access to medical records to identify the existence of PROM.

Related to technique used to assess outcome. Misclassification occurs because of the subjectivity of the symptoms and the low sensitivity of diagnostic techniques for confirmation of ruptured membranes. Leakage of urine, as well as vaginal discharge, is commonly experienced during pregnancy and may be mistaken for amniotic fluid. Reciprocally, leakage of amniotic fluid may occur slowly and in small amounts and then be misinterpreted as urine or cervical discharge (84). Errors tied to estimation of gestational age, previously outlined, may cause errors in classification of PROM.

Recommendations

Outcome assessment. Immunology-based techniques should be used; a physician diagnosis corresponds to the minimum requirement. Maternal reports should not be relied on because of the high likelihood of misclassification.

Statistical analysis. The heterogeneity of PROM (e.g., infectious, physical, cone-surgical related) may make it relevant to perform separate subanalysis of mothers according to the existence of these predisposing factors; physical

(polyhydramnion, twins) and cone-chirurgical-related risk factors should be taken into account. Survival analysis is an option for statistical analysis. As for studies on gestational duration, gestational age should preferably be estimated by using both last menstrual period and ultrasound information, because of the possibility of errors tied to both methods.

Issues related to exposure. In cases of slow amniotic leakage, the starting point of PROM may be difficult to assess, and this must be kept in mind when evaluating whether exposure truly occurred before (and not after) the onset of PROM.

FETAL GROWTH AND SIZE AT BIRTH

Definition

Fetal growth is a continuous phenomenon taking place throughout the pregnancy. The fetal weight gain is approximately 50 g during the first trimester, 700 g during the second trimester, and 2,500 g in the last trimester, to reach a median birth weight at term around 3,000–3,500 g (85). The proportion of low birth weights (<2,500 g) in Europe and North America is in the range of 3%–10%, while it is in the range of 10%–20% in the other continents; in Europe, low birth weight accounts for about 1% of term births. Birth weight is correlated with health status in life. Birth weight has indeed been associated (after control for gestational duration) with mortality from several chronic diseases, including circulatory diseases or causes other than cancer, as well as with occurrence of neurological disorders in childhood and later. Relations exhibited a J shape, the risk of disease decreasing until about 4,000 g and increasing for higher birth weights (14, 86). Cancer mortality, on the other hand, increases continuously with birth weight (86). As discussed by Wilcox (14), this association may not be causal, which does not diminish the validity of birth weight as a nonspecific marker of exposures and mechanisms taking place during development (6, 87–89).

Different metrics have been used to identify fetal growth determinants. All need to be corrected for the gestational duration, which can be done directly or at the step of regression modeling with weight determinants by statistical adjustment. These metrics include continuous birth weight, low (<2,500 g) or very low (<1,500 g) birth weight (possibly restricted to term births), small for gestational age (corresponding to a birth weight lower than the 10th percentile of a suitable sex- and gestational age-specific weight reference distribution), intrauterine growth restriction (generally used equivalently to small for gestational age), and fetal growth restriction. The entity of “small for gestational age” newborns fails at distinguishing “constitutionally small” infants—those small because of parental anthropometric characteristics—from those who are “growth retarded,” that is, smaller than what their parental characteristics, parity, sex, and gestational duration would predict. The definition of fetal growth restriction is meant to address this limitation. Although a consensus is lacking, a newborn is defined as fetal growth retarded if it is in the lower tail of the size distribution of fetuses with the same gestational age, sex, parity, and maternal characteristics.

Birth weight is a summary measure that does not identify the respective contributions of the different body parts. Other measures of neonatal anthropometry, such as length and head and abdominal circumferences, provide additional information.

Birth weight is at the endpoint of intrauterine growth and, as such, does not provide information on fetal growth or growth velocity in specific gestational periods (90). In order to identify possible periods of fetal growth restriction, serial measures of fetal parameters, such as biparietal diameter, head circumference, femoral length, and abdominal circumference by ultrasound, can be performed (91, 92).

Options for study design and outcome assessment

Study design. Unlike most previously discussed outcomes, the quality of information on birth weight from hospital records or questionnaire to the mother is usually good, limiting outcome misclassification in retrospective studies; these studies may still suffer from exposure and confounder misclassification and, importantly, from heterogeneity and error in assessment of gestational duration (refer to the Conclusion). Cohorts with prenatal recruitment (as early as possible, depending on the toxicologically relevant exposure window) therefore represent the preferred option to limit potential bias. Ultrasound examinations can be used to monitor fetal growth (91, 92).

Main threats to validity

Related to study design. Compared with cohorts of pregnant women or of couples recruited before a pregnancy, studies with a postnatal recruitment might entail classification errors on potential confounders that need to be assessed during pregnancy. This is, for example, the case for maternal smoking during pregnancy (better assessed during than after pregnancy) or prepregnancy weight.

Related to technique used to assess outcome and confounders. The approach used to define gestational age may impact on the estimated effect of environmental factors on birth weight; the last menstrual period-based definition has been advocated if the exposure may impact fetal growth (Figure 5).

Recommendations

Outcome assessment. Birth weight should be measured just after birth, because from 24 to 72 hours after birth babies tend to lose about 3%–10% of their weight. Newborn length and head circumference, on the other hand, should be measured from 24 to 48 hours after birth as they could be distorted by labor (recording also the exact date of the measure in order to correct for the time elapsed between birth and the measurements). Assessment of fetal growth by ultrasound measures can be useful to detect possible transitory effects of exposures on fetal growth. Information on who performed the measure should be collected and, if possible, the design should allow limiting the number of ultrasound operators and devices used. Standardization workshops among ultrasonographers prior to the study start are also encouraged. In addition to fetal measures, the placenta can be weighed at birth to study

any environmental influence on placental function (93–95), a strong determinant of fetal growth and development.

Statistical analysis. Sensitivity analyses using different definitions of gestational duration are advised (91, 96). Analysis of fetal growth using repeated ultrasound measures can be done using a longitudinal approach (97, 98), possibly in addition to repeated cross-sectional analyses (91, 99); the 2 main options are presented in Web Appendix 2. According to toxicological data and biological plausibility, analyses stratified on sex of the offspring can be considered a priori.

CONCLUSION

We have reviewed some of the tools allowing study of the impact of environmental factors on human fecundity and reproduction. These may also, of course, apply to the study of nonenvironmental factors. Our focus was on clinical outcomes, but biomarkers of effect can also be assessed in specific tissues or with medical examinations, which can allow, among others, discussion of pathways whereby environmental factors could impact fecundity and reproduction (e.g., inflammatory or epigenetic pathways). Some examples of possibly useful tissues and examinations are given in Table 6.

Because the collection and analysis of biological samples are logistically cumbersome and expensive (depending on the exposure considered and the biochemical assays available), they may not be the first option for very rare exposures.

In terms of study design, cohorts with prenatal enrollment constitute a good approach allowing characterization of potential confounders, exposures (assuming that the relevant time window is not before fertilization), and most pregnancy-related outcomes. If children are followed up, this cohort approach allows study of the longer-term consequences of environmental exposures during the developmental window. Recruitment in such cohorts is a crucial issue; it should take place as early as possible, which is a logistical challenge in many areas where women are not referred to a specific hospital in early pregnancy. Recruitment before conception would be a good option to ensure efficient characterization of periconceptional exposures (and would additionally allow study of fecundity in a prevalent cohort design), but recruitment of couples shortly before a pregnancy is also a logistic challenge. The cohort design may be at odds for rare outcomes, such as specific congenital anomalies, very preterm delivery, or perinatal mortality, for which case-control, time-series (for the 2 latter outcomes), and register-based studies (with more

Table 6. Biological Samples and Medical Examinations of Relevance for Studies on Environmental Impacts on Fecundity and Reproduction

Tissue or Examination	Related Outcomes or Measures	Examples of References
Semen	Male fecundity (sperm morphology, mobility, concentration, DNA quality)	(112)
Ultrasound examination of ovaries and ovarian reserve	Ovarian/menstrual function, ovarian structure (polycystic ovaries), and fecundity	
Urine	Exposure biomarkers Hormonal markers Metabolomics	(49)
Maternal blood	Exposure biomarkers Markers of inflammation, immunological and hormonal status, glucose tolerance tests DNA methylation, gene expression, metabolomics	(113)
Vaginal sample	Vaginal microbiota, vaginosis	
Ultrasound examination	Fetal size and growth; possibly placental morphology Congenital malformations	(114)
Maternal blood pressure, electrocardiography	Maternal cardiovascular function	(115, 116)
Doppler examination of uterine, umbilical, or fetal arteries	Blood flow, oxygen and nutrient supply	(95)
Cord blood	Exposure biomarkers Markers of inflammation, immunological and hormonal status DNA methylation, gene expression, metabolomics	(117)
Amniotic fluid (if available)	Exposure biomarkers	(118)
Placental tissue	DNA methylation	(119)
	Gene expression	(120)
	DNA content	(121)
	Vascularization/histology	(122)
	Exposure biomarkers	(123)
Meconium	Exposure biomarkers	(124)

ecological assessment of exposure) can be preferred to increase the number of observations or cases. Electronic birth registers, which exist in some areas, can indeed serve as a basis for studies on specific pollutants whose levels can be assessed by using simple information such as the home address (100), but not when biomarkers are required to assess exposure; they are particularly useful for the rare events mentioned above. It can, however, be noted that information on home address may have poor accuracy (e.g., limited to zip code or to the birth address without considering changes of address during pregnancy); that the availability of potential confounders is generally limited in such registers, which may or may not (101) be an issue depending on the association studied; and that information on clinical interventions (such as elective cesarean sections, which occur in up to one third or half of pregnancies in some areas (47)) may not be available in such designs, which is a potential source of bias as in studies on preterm delivery (see above). Two-phase approaches can be used to limit issues related to missing information in register-based studies (102). Overall, on the one hand, cohorts with prenatal enrollment have a larger potential for selection bias (because of intense follow-up and collection of biological samples), whereas register-based and retrospective studies are more prone to measurement error and confounding while allowing larger sample sizes. The choice between these designs can be seen as an illustration of the bias-variance trade-off and needs to be carefully discussed in relation to the considered exposure(s).

In terms of confounding, just like for other outcomes, all factors statistically or biologically associated with the reproductive outcomes should not be systematically adjusted for; in particular, factors that are possible consequences of the exposure or of the health outcome in focus are not confounders (103). Following this line, whether a previous history of the adverse pregnancy outcome being studied is a potential confounder should be discussed (41, 57). Directed acyclic graphs constitute a relevant tool to discuss the choice of confounders in reproductive epidemiology (57, 104).

We did not review exposure assessment, which is an issue not specific to reproductive epidemiology, but listed biological samples of relevance for assays of exposure biomarkers (Table 6) and discussed specific sources of bias related to exposure assessment. It can be noted that pregnancy and delivery are associated with strong changes in physiological parameters (e.g., in terms of immunological status, energy needs, metabolism, blood viscosity and volume, cardiovascular function (105)), body weight and composition, and behaviors (e.g., smoking, home address, work and occupational exposures, physical activity, diet), which may impact exposure levels and may drastically increase the amplitude of exposure misclassification when exposure has not been assessed in the “right” time window. Some of the factors may also impact the body burden of specific environmental contaminants, increasing exposure misclassification if, for example, there is heterogeneity in previous reproductive history or in the gestational age at which biological samples used to assess exposure biomarkers are collected.

Serious consideration of these issues is relevant in order to better characterize the impact of environmental factors on human reproductive function.

ACKNOWLEDGMENTS

Author affiliations: Team of Environmental Epidemiology Applied to Reproduction and Respiratory Health, Institut Albert Bonniot (U823), INSERM, Grenoble, France (Rémy Slama, Claire Philippat, Sylvie Rey); Institut Albert Bonniot, Grenoble University, Grenoble, France (Rémy Slama, Claire Philippat, Sylvie Rey); Centre for Public Health Research (CSISP), Valencia, Spain (Ferran Ballester, Carmen Iniguez); Spanish Consortium for Research on Epidemiology and Public Health (CIBERESP), Madrid, Spain (Ferran Ballester, Maribel Casas, Carmen Iniguez, Mark Nieuwenhuijsen, Martine Vrijheid); Faculty of Nursing, University of Valencia, Valencia, Spain (Ferran Ballester, Carmen Iniguez); Centre for Research in Environmental Epidemiology (CREAL), Barcelona, Spain (Maribel Casas, Mark Nieuwenhuijsen, Martine Vrijheid); Universitat Pompeu Fabra (UPF), Barcelona, Spain (Maribel Casas, Mark Nieuwenhuijsen, Martine Vrijheid); U1085-IRSET, National Institute of Health and Medical Research (INSERM), Rennes, France (Sylvaine Cordier); University of Rennes I, Rennes, France (Sylvaine Cordier); Division of Epidemiology, Norwegian Institute of Public Health, Oslo, Norway (Merete Eggesbø); Department of Environmental Health, French Institute of Public Health Surveillance (InVS), Saint-Maurice, France (Stéphanie Vandentorren); and Pôle Epidémiologie et Santé Publique, Observatoire du Samu Social de Paris, Paris, France (Stéphanie Vandentorren).

This work was supported by the European Union (grant agreement no. 226285, ENRIECO project FP7). The Team of Environmental Epidemiology Applied to Reproduction and Respiratory Health (INSERM U823) is supported by an AVENIR/ATIP grant from INSERM.

The contribution of all members of the Environmental Health Risks in European Birth Cohorts (ENRIECO) project working group on pregnancy outcomes is acknowledged, in particular Anne-Marie Nybo-Andersen. We thank Dr. Diana van Gent for coordination of the ENRIECO project and Dr. Thierry Debillon for useful comments.

Conflict of interest: none declared.

REFERENCES

1. United Nations Children’s Found (UNICEF). Progress for children. A report card on maternal mortality. New York, NY: UNICEF; 2008. (http://www.childinfo.org/files/progress_for_children_maternalmortality.pdf). (Accessed September 11, 2013).
2. Saigal S, Doyle LW. An overview of mortality and sequelae of preterm birth from infancy to adulthood. *Lancet*. 2008; 371(9608):261–269.
3. Guzik DS, Swan S. The decline of infertility: apparent or real? *Fertil Steril*. 2006;86(3):524–526; discussion 534.
4. Slama R, Hansen OK, Ducot B, et al. Estimation of the frequency of involuntary infertility on a nation-wide basis. *Hum Reprod*. 2012;27(5):1489–1498.
5. Mahajan NN, Turnbull DA, Davies MJ, et al. Changes in affect and state anxiety across an in vitro fertilization/ intracytoplasmic sperm injection cycle. *Fertil Steril*. 2010; 93(2):517–526.

6. Hanson MA, Gluckman PD. Developmental origins of health and disease: new insights. *Basic Clin Pharmacol Toxicol*. 2008;102(2):90–93.
7. Wigle DT, Arbuckle TE, Turner MC, et al. Epidemiologic evidence of relationships between reproductive and child health outcomes and environmental chemical contaminants. *J Toxicol Environ Health B Crit Rev*. 2008;11(5-6):373–517.
8. Slama R, Cordier S. Environmental contaminants and impacts on healthy and successful pregnancies. In: Woodruff TJ, Janssen S, Guillette LJ Jr, et al, eds. *Environmental Impacts on Reproductive Health & Fertility*. New York, NY: Cambridge University Press; 2010:125–144.
9. Windham G, Fenster L. Environmental contaminants and pregnancy outcomes. *Fertil Steril*. 2008;89(2 suppl): e111–e116; discussion e117.
10. Slama R, Cordier S. Impact of chemical and physical environmental factors on the course and outcome of pregnancy (in French). *J Gynecol Obstet Biol Reprod (Paris)*. 2013;42(5):413–444.
11. Savitz DA, Hertz-Picciotto I, Poole C, et al. Epidemiologic measures of the course and outcome of pregnancy. *Epidemiol Rev*. 2002;24(2):91–101.
12. Weinberg CR, Wilcox AJ. Reproductive epidemiology. In: Rothman KJ, Greenland S, Lash TL, eds. *Modern Epidemiology*. 3rd ed. Philadelphia, PA: Lippincott, Williams & Wilkins; 2008:585–608.
13. Leridon H. *Human Fertility: The Basic Components*. Chicago, IL: The University of Chicago Press; 1977.
14. Wilcox AJ. *Fertility and Pregnancy: An Epidemiologic Perspective*. New York, NY: Oxford University Press, Inc; 2010.
15. Olsen J, Basso O. Reproductive epidemiology. In: Ahrens W, Pigeot I, eds. *Handbook of Epidemiology*. Berlin, Germany: Springer Verlag; 2005:1043–1109.
16. Bonde JP, Giwercman A, Ernst E. Identifying environmental risk to male reproductive function by occupational sperm studies: logistics and design options. *Occup Environ Med*. 1996;53(8):511–519.
17. Nieuwenhuijsen M, Paustenbach D, Duarte-Davidson R. New developments in exposure assessment: the impact on the practice of health risk assessment and epidemiological studies. *Environ Int*. 2006;32(8):996–1009.
18. Habbema JD, Collins J, Leridon H, et al. Towards less confusing terminology in reproductive medicine: a proposal. *Fertil Steril*. 2004;82(1):36–40.
19. Gini C. Premières recherches sur la fécondabilité des femmes (in French). In: Fields JC, ed. *Proceedings of the International Mathematical Congress, 1924*. Toronto, Canada: Toronto University Press; 1928:889–892.
20. Baird DD, Wilcox AJ, Weinberg CR. Use of time to pregnancy to study environmental exposures. *Am J Epidemiol*. 1986;124(3):470–480.
21. Joffe M, Key J, Best N, et al. Studying time to pregnancy by use of a retrospective design. *Am J Epidemiol*. 2005;162(2): 115–124.
22. Baird DD, Weinberg CR, Rowland AS. Reporting errors in time-to-pregnancy data collected with a short questionnaire. Impact on power and estimation of fecundability ratios. *Am J Epidemiol*. 1991;133(12):1282–1290.
23. Bonde JP, Ernst E, Jensen TK, et al. Relation between semen quality and fertility: a population-based study of 430 first-pregnancy planners. *Lancet*. 1998;352(9135):1172–1177.
24. Buck Louis GM, Sundaram R, Schisterman EF, et al. Persistent environmental pollutants and couple fecundity: the LIFE study. *Environ Health Perspect*. 2013;121(2): 231–236.
25. Tingen C, Stanford JB, Dunson DB. Methodologic and statistical approaches to studying human fertility and environmental exposure. *Environ Health Perspect*. 2004; 112(1):87–93.
26. Keiding N, Kvist K, Hartvig H, et al. Estimating time to pregnancy from current durations in a cross-sectional sample. *Biostatistics*. 2002;3(4):565–578.
27. Slama R, Ducot B, Carstensen L, et al. Feasibility of the current-duration approach to studying human fecundity. *Epidemiology*. 2006;17(4):440–449.
28. Weinberg CR, Baird DD, Rowland AS. Pitfalls inherent in retrospective time-to-event studies: the example of time to pregnancy. *Stat Med*. 1993;12(9):867–879.
29. Weinberg CR, Baird DD, Wilcox AJ. Sources of bias in studies of time to pregnancy. *Stat Med*. 1994;13(5-7):671–681.
30. Jensen TK, Scheike T, Keiding N, et al. Selection bias in determining the age dependence of waiting time to pregnancy. *Am J Epidemiol*. 2000;152(6):565–572.
31. Juul S, Keiding N, Tvede M. Retrospectively sampled time-to-pregnancy data may make age-decreasing fecundity look increasing. European Infertility and Subfecundity Study Group. *Epidemiology*. 2000;11(6):717–719.
32. Slama R, Kold-Jensen T, Scheike T, et al. How would a decline in sperm concentration over time influence the probability of pregnancy? *Epidemiology*. 2004;15(4):458–465.
33. Sallmen M, Lindbohm ML, Nurminen M. Paternal exposure to lead and infertility. *Epidemiology*. 2000;11(2):148–152.
34. Sallmen M, Lindbohm ML, Anttila A, et al. Time to pregnancy among the wives of men occupationally exposed to lead. *Epidemiology*. 2000;11(2):141–147.
35. Trussell J, Vaughan B. Contraceptive failure, method-related discontinuation and resumption of use: results from the 1995 National Survey of Family Growth. *Fam Plann Perspect*. 1999;31(2):64–72, 93.
36. Joffe M, Key J, Best N, et al. The role of biological fertility in predicting family size. *Hum Reprod*. 2009;24(8):1999–2006.
37. Whitworth KW, Haug LS, Baird DD, et al. Perfluorinated compounds and subfecundity in pregnant women. *Epidemiology*. 2012;23(2):257–263.
38. Stolwijk AM, Straatman H, Zielhuis GA, et al. Seasonal variation in the time to pregnancy: avoiding bias by using the date of onset. *Epidemiology*. 1996;7(2):156–160.
39. Jensen TK, Keiding N, Scheike T, et al. Declining human fertility (letter)? *Fertil Steril*. 2000;73(2):421–423.
40. Scheike TH, Petersen JH, Martinussen T. Retrospective ascertainment of recurrent events: an application to time to pregnancy. *J Am Stat Assoc*. 1999;94(447):713–725.
41. Weinberg CR. Should we adjust for pregnancy history when the exposure effect is transient? *Epidemiology*. 1995;6(3): 335–337.
42. Boue A, Boue J, Gropp A. Cytogenetics of pregnancy wastage. *Adv Hum Genet*. 1985;14:1–57.
43. Gourbin G, Masuy-Stroobant G. Registration of vital data: Are live births and stillbirths comparable all over Europe? *Bull World Health Organ*. 1995;73(4):449–460.
44. Wang X, Chen C, Wang L, et al. Conception, early pregnancy loss, and time to clinical pregnancy: a population-based prospective study. *Fertil Steril*. 2003;79(3):577–584.
45. Zinaman MJ, Clegg ED, Brown CC, et al. Estimates of human fertility and pregnancy loss. *Fertil Steril*. 1996;65(3): 503–509.
46. Wilcox AJ, Weinberg CR, O'Connor JF, et al. Incidence of early loss of pregnancy. *N Engl J Med*. 1988;319(4):189–194.
47. Euro-PERISTAT Project with SCPE and EUROCAT. European Perinatal Health Report. The health and care of

- pregnant women and babies in Europe in 2010. Paris, France: INSERM; 2013. (http://www.europeristat.com/images/EuropeanPerinatalHealthReport_2010.pdf). (Accessed May 2, 2013).
48. Wilcox AJ, Baird DD, Weinberg CR. Time of implantation of the conceptus and loss of pregnancy. *N Engl J Med*. 1999; 340(23):1796–1799.
 49. Venners SA, Korrick S, Xu X, et al. Preconception serum DDT and pregnancy loss: a prospective study using a biomarker of pregnancy. *Am J Epidemiol*. 2005;162(8):709–716.
 50. Weinberg CR, Hertz-Picciotto I, Baird DD, et al. Efficiency and bias in studies of early pregnancy loss. *Epidemiology*. 1992;3(1):17–22.
 51. Venners SA, Wang X, Chen C, et al. Paternal smoking and pregnancy loss: a prospective study using a biomarker of pregnancy. *Am J Epidemiol*. 2004;159(10):993–1001.
 52. Hjollund NH, Jensen TK, Bonde JP, et al. Spontaneous abortion and physical strain around implantation: a follow-up study of first-pregnancy planners. *Epidemiology*. 2000;11(1): 18–23.
 53. Swan SH, Waller K, Hopkins B, et al. A prospective study of spontaneous abortion: relation to amount and source of drinking water consumed in early pregnancy. *Epidemiology*. 1998;9(2):126–133.
 54. Howards PP, Hertz-Picciotto I, Poole C. Conditions for bias from differential left truncation. *Am J Epidemiol*. 2007;165(4): 444–452.
 55. O'Neill MS, Hertz-Picciotto I, Pastore LM, et al. Have studies of urinary tract infection and preterm delivery used the most appropriate methods? *Paediatr Perinat Epidemiol*. 2003; 17(3):226–233.
 56. Slama R, Bouyer J, Windham G, et al. Influence of paternal age on the risk of spontaneous abortion. *Am J Epidemiol*. 2005;161(9):816–823.
 57. Howards PP, Schisterman EF, Heagerty PJ. Potential confounding by exposure history and prior outcomes: an example from perinatal epidemiology. *Epidemiology*. 2007; 18(5):544–551.
 58. Gladen BC. On the role of 'habitual aborters' in the analysis of spontaneous abortion. *Stat Med*. 1986;5(6):557–564.
 59. Howards PP, Schisterman EF, Poole C, et al. "Toward a clearer definition of confounding" revisited with directed acyclic graphs. *Am J Epidemiol*. 2012;176(6):506–511.
 60. Haxton MJ, Bell J. Fetal anatomical abnormalities and other associated factors in middle-trimester abortion and their relevance to patient counselling. *Br J Obstet Gynaecol*. 1983; 90(6):501–506.
 61. Holmes LB, Westgate MN. Inclusion and exclusion criteria for malformations in newborn infants exposed to potential teratogens. *Birth Defects Res A Clin Mol Teratol*. 2011;91(9): 807–812.
 62. EUROCAT. EUROCAT Guide 1.3 and reference documents. Instructions for the registration and surveillance of congenital anomalies. Antrim, Northern Ireland, United Kingdom: EUROCAT Central Registry; 2005. (<http://www.eurocat-network.eu/content/EUROCAT-Guide-1.3.pdf>). (Accessed September 12, 2013).
 63. Dolk H, Vrijheid M, Scott JE, et al. Toward the effective surveillance of hypospadias. *Environ Health Perspect*. 2004; 112(3):398–402.
 64. Boisen KA, Kaleva M, Main KM, et al. Difference in prevalence of congenital cryptorchidism in infants between two Nordic countries. *Lancet*. 2004;363(9417):1264–1269.
 65. Toppari J, Kaleva M, Virtanen HE. Trends in the incidence of cryptorchidism and hypospadias, and methodological limitations of registry-based data. *Hum Reprod Update*. 2001; 7(3):282–286.
 66. Strickland MJ, Riehle-Colarusso TJ, Jacobs JP, et al. The importance of nomenclature for congenital cardiac disease: implications for research and evaluation. *Cardiol Young*. 2008;18(suppl 2):92–100.
 67. Pierik FH, Klebanoff MA, Brock JW, et al. Maternal pregnancy serum level of heptachlor epoxide, hexachlorobenzene, and beta-hexachlorocyclohexane and risk of cryptorchidism in offspring. *Environ Res*. 2007;105(3): 364–369.
 68. Ren A, Qiu X, Jin L, et al. Association of selected persistent organic pollutants in the placenta with the risk of neural tube defects. *Proc Natl Acad Sci U S A*. 2011;108(31): 12770–12775.
 69. Lie RT, Wilcox AJ, Taylor J, et al. Maternal smoking and oral clefts: the role of detoxification pathway genes. *Epidemiology*. 2008;19(4):606–615.
 70. Nieuwenhuijsen MJ, Toledano MB, Bennett J, et al. Chlorination disinfection by-products and risk of congenital anomalies in England and Wales. *Environ Health Perspect*. 2008;116(2):216–222.
 71. Ritz B, Yu F, Fruin S, et al. Ambient air pollution and risk of birth defects in Southern California. *Am J Epidemiol*. 2002; 155(1):17–25.
 72. Jukic AM, Baird DD, Weinberg CR, et al. Length of human pregnancy and contributors to its natural variation. *Hum Reprod*. 2013;28(10):2848–2855.
 73. Goldenberg RL, Culhane JF, Iams JD, et al. Epidemiology and causes of preterm birth. *Lancet*. 2008;371(9606):75–84.
 74. Savitz DA, Murnane P. Behavioral influences on preterm birth: a review. *Epidemiology*. 2010;21(3):291–299.
 75. Villar J, Papageorgiou AT, Knight HE, et al. The preterm birth syndrome: a prototype phenotypic classification. *Am J Obstet Gynecol*. 2012;206(2):119–123.
 76. Burguet A, Kaminski M, Abraham-Lerat L, et al. The complex relationship between smoking in pregnancy and very preterm delivery. Results of the Epipage study. *BJOG*. 2004;111(3): 258–265.
 77. Sagiv SK, Mendola P, Loomis D, et al. A time-series analysis of air pollution and preterm birth in Pennsylvania, 1997–2001. *Environ Health Perspect*. 2005;113(5):602–606.
 78. Basu R, Malig B, Ostro B. High ambient temperature and the risk of preterm delivery. *Am J Epidemiol*. 2010;172(10): 1108–1117.
 79. Darrow LA. Invited commentary: application of case-crossover methods to investigate triggers of preterm birth. *Am J Epidemiol*. 2010;172(10):1118–1120; discussion 1121–1122.
 80. Savitz DA, Terry JW Jr, Dole N, et al. Comparison of pregnancy dating by last menstrual period, ultrasound scanning, and their combination. *Am J Obstet Gynecol*. 2002; 187(6):1660–1666.
 81. Lynch CD, Zhang J. The research implications of the selection of a gestational age estimation method. *Paediatr Perinat Epidemiol*. 2007;21(suppl 2):86–96.
 82. Bukowski R, Smith GC, Malone FD, et al. Fetal growth in early pregnancy and risk of delivering low birth weight infant: prospective cohort study. *BMJ*. 2007;334(7598):836.
 83. ACOG Practice Bulletin No. 80. Premature rupture of membranes. Clinical management guidelines for obstetrician-gynecologists. *Obstet Gynecol*. 2007;109(4):1007–1019.
 84. Management of premature prelabor rupture of the membranes *Ann N Y Acad Sci*. 2010;1205:123–129.
 85. Cunningham FG, Leveno KJ, Bloom SL, et al, eds. *Williams Obstetrics*. 23rd ed. New York, NY: McGraw-Hill; 2010.

86. Risnes KR, Vatten LJ, Baker JL, et al. Birthweight and mortality in adulthood: a systematic review and meta-analysis. *Int J Epidemiol*. 2011;40(3):647–661.
87. Gluckman PD, Hanson MA. Maternal constraint of fetal growth and its consequences. *Semin Fetal Neonatal Med*. 2004;9(5):419–425.
88. Barker DJ. The origins of the developmental origins theory. *J Intern Med*. 2007;261(5):412–417.
89. Wilcox AJ. On the importance—and the unimportance—of birthweight. *Int J Epidemiol*. 2001;30(6):1233–1241.
90. Hemachandra AH, Klebanoff MA. Use of serial ultrasound to identify periods of fetal growth restriction in relation to neonatal anthropometry. *Am J Hum Biol*. 2006;18(6):791–797.
91. Slama R, Thiebaugeorges O, Goua V, et al. Maternal personal exposure to airborne benzene and intrauterine growth. *Environ Health Perspect*. 2009;117(8):1313–1321.
92. van den Hooven EH, Pierik FH, de Kluienaar Y, et al. Air pollution exposure during pregnancy, ultrasound measures of fetal growth, and adverse birth outcomes: a prospective cohort study. *Environ Health Perspect*. 2012;120(1):150–156.
93. Rahmalia A, Giorgis-Allemand L, Lepeule J, et al. Pregnancy exposure to atmospheric pollutants and placental weight: an approach relying on a dispersion model. *Environ Int*. 2012;48C:47–55.
94. Yorifuji T, Naruse H, Kashima S, et al. Residential proximity to major roads and placenta/birth weight ratio. *Sci Total Environ*. 2012;414(1):98–102.
95. van den Hooven EH, Pierik FH, de Kluienaar Y, et al. Air pollution exposure and markers of placental growth and function: the generation R study. *Environ Health Perspect*. 2012;120(12):1753–1759.
96. Slama R, Khoshnood B, Kaminski M. How to control for gestational age in studies of effects of environmental factors on fetal growth? *Environ Health Perspect*. 2008;116(7):A284.
97. Gurrin LC, Blake KV, Evans SF, et al. Statistical measures of foetal growth using linear mixed models applied to the foetal origins hypothesis. *Stat Med*. 2001;20(22):3391–3409.
98. Royston P, Altman DG. Design and analysis of longitudinal studies of fetal size. *Ultrasound Obstet Gynecol*. 1995;6(5):307–312.
99. Aguilera I, Garcia-Esteban R, Iiguez C, et al. Prenatal exposure to traffic-related air pollution and ultrasound measures of fetal growth in the INMA Sabadell cohort. *Environ Health Perspect*. 2010;118(5):705–711.
100. Wilhelm M, Ritz B. Local variations in CO and particulate air pollution and adverse birth outcomes in Los Angeles County, California, USA. *Environ Health Perspect*. 2005;113(9):1212–1221.
101. Darrow LA, Woodruff TJ, Parker JD. Maternal smoking as a confounder in studies of air pollution and infant mortality. *Epidemiology*. 2006;17(5):592–593.
102. Hoggatt KJ, Greenland S, Ritz BR. Adjustment for response bias via two-phase analysis: an application. *Epidemiology*. 2009;20(6):872–879.
103. Rothman KJ, Greenland S, Lash TL. *Modern Epidemiology*. 3rd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2008.
104. Hernan MA, Hernandez-Diaz S, Werler MM, et al. Causal knowledge as a prerequisite for confounding evaluation: an application to birth defects epidemiology. *Am J Epidemiol*. 2002;155(2):176–184.
105. Kaaja RJ, Greer IA. Manifestations of chronic disease during pregnancy. *JAMA*. 2005;294(21):2751–2757.
106. Scheike TH, Rylander L, Carstensen L, et al. Time trends in human fecundability in Sweden. *Epidemiology*. 2008;19(2):191–196.
107. Baird DD, Weinberg CR, Schwingl P, et al. Selection bias associated with contraceptive practice in time-to-pregnancy studies. *Ann N Y Acad Sci*. 1994;709:156–164.
108. Fromme H, Mosch C, Morovitz M, et al. Pre- and postnatal exposure to perfluorinated compounds (PFCs). *Environ Sci Technol*. 2010;44(18):7123–7129.
109. Weinberg CR, Baird DD, Wilcox AJ. Bias in retrospective studies of spontaneous abortion based on the outcome of the most recent pregnancy. In: Campbell KL, Wood JW, eds. *Human Reproductive Ecology: Interactions of Environment, Fertility and Behavior*. New York, NY: The New York Academy of Sciences; 1994:280–286.
110. Boyd PA, Devigan C, Khoshnood B, et al. Survey of prenatal screening policies in Europe for structural malformations and chromosome anomalies, and their impact on detection and termination rates for neural tube defects and Down's syndrome. *BJOG*. 2008;115(6):689–696.
111. Darrow LA, Strickland MJ, Klein M, et al. Seasonality of birth and implications for temporal studies of preterm birth. *Epidemiology*. 2009;20(5):699–706.
112. Hauser R, Meeker JD, Singh NP, et al. DNA damage in human sperm is related to urinary levels of phthalate monoester and oxidative metabolites. *Hum Reprod*. 2007;22(3):688–695.
113. Lee PC, Talbott EO, Roberts JM, et al. Particulate air pollution exposure and C-reactive protein during early pregnancy. *Epidemiology*. 2011;22(4):524–531.
114. Hansen CA, Barnett AG, Pritchard G. The effect of ambient air pollution during early pregnancy on fetal ultrasonic measurements during mid-pregnancy. *Environ Health Perspect*. 2008;116(3):362–369.
115. Hampel R, Lepeule J, Schneider A, et al. Short-term impact of ambient air pollution and air temperature on blood pressure among pregnant women. *Epidemiology*. 2011;22(5):671–679.
116. van den Hooven EH, de Kluienaar Y, Pierik FH, et al. Air pollution, blood pressure, and the risk of hypertensive complications during pregnancy: the generation R study. *Hypertension*. 2011;57(3):406–412.
117. Govarts E, Nieuwenhuijsen M, Schoeters G, et al. Birth weight and prenatal exposure to polychlorinated biphenyls (PCBs) and dichlorodiphenyldichloroethylene (DDE): a meta-analysis within 12 European birth cohorts. *Environ Health Perspect*. 2012;120(2):162–170.
118. Philippat C, Wolff MS, Calafat AM, et al. Prenatal exposure to environmental phenols: concentrations in amniotic fluid and variability in phenol urinary concentrations during pregnancy. *Environ Health Perspect*. 2013;121(10):1225–1231.
119. Janssen BG, Godderis L, Pieters N, et al. Placental DNA hypomethylation in association with particulate air pollution in early life. *Part Fibre Toxicol*. 2013;10(1):22.
120. Adibi JJ, Whyatt RM, Hauser R, et al. Transcriptional biomarkers of steroidogenesis and trophoblast differentiation in the placenta in relation to prenatal phthalate exposure. *Environ Health Perspect*. 2010;118(2):291–296.
121. Janssen BG, Munters E, Pieters N, et al. Placental mitochondrial DNA content and particulate air pollution during in utero life. *Environ Health Perspect*. 2012;120(9):1346–1352.
122. Veras MM, Damasceno-Rodrigues NR, Caldini EG, et al. Particulate urban air pollution affects the functional morphology of mouse placenta. *Biol Reprod*. 2008;79(3):578–584.
123. Vilahur N, Molina-Molina JM, Bustamante M, et al. Male specific association between xenoestrogen levels in placenta and birthweight. *Environ Int*. 2013;51:174–181.
124. Zhang Y, Lin L, Cao Y, et al. Phthalate levels and low birth weight: a nested case-control study of Chinese newborns. *J Pediatr*. 2009;155(4):500–504.