Registries for Evaluating Pregnancy Outcomes

Draft White Paper for Third Edition of

"Registries for Evaluating Patient Outcomes: A User's Guide"

A pregnancy exposure registry is an observational prospective cohort of women receiving a biopharmaceutical product(s) of interest as part of their routine clinical care who are enrolled voluntarily during gestation, before outcomes can be known. Participants are followed until the end of pregnancy or longer to systematically collect information on specific pregnancy outcomes and evaluate their frequency relative to a scientifically valid reference population(s).[1-3] Specific examples of pregnancy registries can be found on the Food and Drugs Administration (FDA) website.[4]

This paper reviews the "why, how, and who" of conducting pregnancy registries. The paper first discusses why pregnancy registries are needed to assess risks and benefits of medications during pregnancy. Second, the paper describes the distinctive methodological aspects of these registries, including design, study population, enrollment and follow-up of pregnant women, ascertainment and definition of exposures and outcomes, reference groups, statistical power, and validity issues. Third, the paper takes a more pragmatic approach and presents key operational aspects such as protocol structure, recruitment and retention of participants, methods of data collection, when to release findings, role of advisory boards, and challenges of global designs. Finally, characteristics to consider when evaluating pregnancy registries are described.

1. Justification

All patients in need of treatment should have access to medications that have been adequately studied and be provided with information to assess the risks versus the benefits of using the medication. Collecting postmarketing data on the safety of medications during pregnancy is commonly done through the use of pregnancy exposure registries or pregnancy disease registries that collect treatment information. Pregnancy registries are prospective observational studies specifically designed to collect clinically relevant data and provide information for treating or counseling not only women who are pregnant but also women of childbearing potential. In 2002, the FDA published its guidance for pregnancy registries[2]

with a goal of encouraging the regular use of more formal, prospective study designs to obtain clinically relevant human data that can be used in product labeling. Similar guidelines were published by the EMA in 2005.[5] In 2007, the Food and Drug Amendments Act (FDAAA) provided the authority under Title IX[6] to require pregnancy registries as a post marketing requirement (PMR). Pregnancy registries are now required at the time of a new drug approval when there is a safety concern or when there is a need to gather data on the use of the product in pregnancy based on the following circumstances: 1) prior knowledge of the product suggests a safety concern based on the pharmacologic or chemical class, or on data from animal studies or clinical trials; 2) the product will be indicated for use during pregnancy (e.g., vaccines and medications for chronic illness); or 3) there is a high likelihood of use in females of reproductive age such that inadvertent exposure during pregnancy may be expected.

For non-pregnant individuals, safety and efficacy data that yield such information are derived from well-controlled clinical trials conducted prior to a drug's approval. When it comes to pregnant women, however, the situation is different. Clinical trials rarely include pregnant women because there is a lack of safety information on the drug's use in pregnancy.[7,8] As a consequence, most information regarding the safety/risk profile of drugs during pregnancy is collected after the drug has been approved and used by pregnant women intentionally or unintentionally – intentionally because some conditions require treatment during pregnancy, and unintentionally because approximately half of all pregnancies in the United States are unplanned,[9] so embryo/fetal exposure to medications can occur before pregnancy is detected.

Tests in animal models are a regulatory requirement for new drugs and biologics prior to approval. In some cases, these animal toxicology studies can provide a means to detect teratogenic effects. Often however, results are not easily translated into human risk because of variations in teratogenic response among species.[10] In addition, animal toxicology studies are designed so that at least one dose tested will provoke an adverse toxic response. The results at those dose levels may not predict those that might be observed at the intended therapeutic doses used in humans.[11]

In humans, passive data collection such as the FDA Adverse Event Reporting System (AERS) can suggest potential drug safety issues for further study. The system includes any patient population that may have been exposed to the drug. Reporting to the database is voluntary (although required for manufacturers) and underreporting is a significant issue with extent of reporting thought to vary substantially depending on the drug, the indication for use, and the nature of the adverse event.[12,13] There is no reference group and no information on the number of individuals taking the drug who did not have an adverse event. It is often not clear whether the adverse event reported to AERS is a medication-related

event, an event resulting from the underlying illness, or a coincidence. Since adverse pregnancy outcomes such as pregnancy losses and congenital malformations are relatively common, they will inevitably occur among exposed individuals; selected reporting of exposed cases can lead to false alarms. Consequently, although data from AERS can be useful for identifying initial signals of adverse events, it cannot be used to quantify risks for a particular product or to compare risks between drugs.[14] Similarly, case series published in scientific journals cannot distinguish chance from causation or be used to quantify and assign teratogenic risks.

Information on human teratogens must come from adequately controlled epidemiologic studies, which include case-control and cohort designs. Case control studies identify births with the outcome of interest (e.g., a specific birth defect) and compare their frequency of exposures to that in a control group without this outcome. This design offers advantages in detection and confirmation of associations between prenatal exposure to the medication and the risk for rare events.[15] However, case control studies have some limitations. They collect information on exposure retrospectively, rarely have enough sample size to evaluate infrequently used medications, and can estimate relative risks but not the absolute risks associated with the drug.

Follow-up studies of pregnant women have the advantage of identifying drug exposure before the adverse outcomes are recognized. In non-pregnant populations, health care utilization databases such as Medicaid claims files or records in large health maintenance organizations have become a standard source of information for drug safety studies.[16] These databases are a resource for large-scale observational post-marketing studies because they offer the ability to study rare consequences of drug use. Some of these databases have limitations for the study of pregnancy outcomes because they do not routinely record evidence of pregnancy (e.g., estimated date of conception) or provide child-mother linkages. Important reproductive information, such as gestational age at birth, birth weight, and maternal reproductive history, is rarely available.[17] Moreover, when exposure to the specific drug of interest involves a small fraction of the pregnant population, even these large cohorts are constrained in their statistical power. In this scenario, concentrating on women exposed to selected drugs through a pregnancy registry can increase efficiency. However, the FDA-funded Medication Exposure in Pregnancy Risk Evaluation Program (MEPREP) has established collaboration among a selected group of large administrative databases with the ability to link mothers and babies and with linkage to birth certificates, which contain additional information.[18] This resource will represent an important tool to study outpatient dispensing of medications during pregnancy and a number of validated pregnancy outcomes.

2. Pregnancy Registry Objectives

The overall purpose of pregnancy exposure registries is to provide human data on the safety of biopharmaceutical products during pregnancy.[1, 19] Pregnancy registries should have specific primary and secondary objectives defined *a priori* in a scientifically sound study protocol.[2] Many exposure registries have as their primary objective to "assess the risk of major congenital malformations" in the offspring of women exposed to a given drug just before or during pregnancy. Implicit in this objective is to determine whether that risk is higher or lower than expected. Registries can evaluate multiple maternal, obstetrical, fetal, and infant outcomes, from spontaneous abortions to developmental delays. Moreover, they may provide an opportunity to evaluate not only the safety, but also the effectiveness of drugs, as well as the risks associated with untreated diseases during pregnancy. They can also evaluate the effect of dose, gestational timing of exposure, and effect modification by maternal characteristics.[1]

Since the ultimate goal is to inform the decisions of medical care providers and patients, it is in the common interest of all the parties to initiate the registry as soon as possible after marketing authorization, use proactive enrollment strategies (i.e., if possible, broaden the source population to obtain, for example, 1,000 exposed women in one year rather than 100 per year for 10 years), and analyze the data and report findings on a regular basis.[2] As more data accumulate over time, the registry can provide narrower boundaries of uncertainty around the point estimates, which leads to increasing assurance of relative safety or more precise quantification of relative risks. These issues are discussed more extensively in the Operations and Stopping Rules sections below, respectively.

3. Design

Whether stated or not, one scientific question in the evaluation of drugs during pregnancy is often "*what would have been the outcome of this pregnancy had the woman not been exposed?*" Since the counterfactual outcome for a given pregnancy is unknown, the closest strategy to respond to this question would be to randomize a group of women periconceptionally to either the drug of interest or a reference group and follow them in a blinded manner until the outcome of interest is fully assessed. This design would ensure that the groups only differ in their drug exposure and that the same methodology and clinical judgment are used to assess the outcomes. The reference group could be randomized to placebo, if not-treating were clinically acceptable, or to an alternative therapy, and the comparative efficacy and safety of two or more therapeutic options could be evaluated. However, since for ethical reasons pregnant women are rarely included in sufficient numbers in randomized controlled trials (RCT) during the process

of drug development, the safety and effectiveness of drugs during pregnancy needs to be studied in post-marketing non-randomized studies.[3, 19]

The specific design of a pregnancy registry is determined by its purpose.[1] To rule out strong teratogenic effects (e.g., over 20% risk of malformations after prenatal exposure to thalidomide), enrollment of 100 exposed pregnancies in a simple uncontrolled cohort might suffice. The effect of major teratogens is so large as to overwhelm the potential impact of common methodological biases on relative risks. However, most known teratogens are associated with a more moderate increase in the risk of relatively rare malformations.[20] To detect moderate teratogens, registries need to enroll a larger number of gestations and be carefully designed.[19]

Although the same rigor and most principles of RCTs can be applied to any observational study, the lack of randomization calls for additional epidemiological methods.[21] This paper discusses in the next sections how pregnancy registries differ from RCTs; how deviation from RCTs standards can compromise the validity of results; and how biases can be minimized through collection of detailed data on exposure and other maternal characteristics, close follow-up, accurate assessment of outcomes, and inclusion of comparable reference groups.

4. Study Population: Who and When

Through clear inclusion and exclusion criteria, registries target a well-defined study population that, ideally, should be closer to real clinical practice than the selected populations of a RCT. However, to the extent that pregnancy registries rely on patients or their health care providers to hear about the registry, to contact the registry, and to agree to participate, there is the potential for selective enrollment. Women who do participate might differ systematically on factors related to the pregnancy outcome from the population of exposed pregnant women who are not part of the registry. Therefore, baseline risks in the registry population might differ from those in the general population of women using the drug of interest.

In assessing the relative risks associated with the drug, there is often a tradeoff between generalizability and validity. A population-based sampling strategy can be logistically complicated and would arguably estimate a similar relative risk since a teratogenic effect found in participants would probably apply to "non-volunteers" as well. On the other hand, inclusion of non-motivated individuals might increase losses to follow-up, misclassification of information, and the variability and impact of confounders.

A peculiarity of pregnancy registries is that the population can be defined based on women, pregnancies, or fetuses. A woman might have more than one pregnancy, and she might enroll in the same registry more than once. Clustered analyses are often used in this situation. In addition, multifetal gestations result in more than one fetus "enrolled" within the same pregnancy. Although there may be several ways of dealing with multiple gestations, it is prudent to collect information from all the fetuses. When reporting risks, whether using fetuses or pregnancies as the unit of analysis, both the numerator and denominator should be consistent with the choice.[22]

5. Enrollment and Follow-up

An ideal pregnancy cohort would include women at conception and follow them for months beyond delivery. However, that rarely happens for logistical reasons, and, consequently, pregnancy cohorts have some degree of unintended truncation on both sides of the ideal follow-up. Left truncation occurs because follow-up can only start after women realize they are pregnant (in patient-initiated enrollment in registries) or health care providers identify the pregnancy in a patient (in clinician-initiated enrollment in registries), and the process of enrollment itself can further delay the inclusion. Right truncation occurs because follow-up would end with unknown outcomes when there are losses to follow-up or pregnancy terminations without fetal autopsy.

As a result, time from enrollment to end of follow-up can range from 1 month to over 1 year. In any study, longer follow-up periods naturally lead to higher opportunities for diagnosis and therefore larger cumulative risk estimates and statistical power. As discussed later, inclusion of either prenatal diagnoses or outcomes identified during infancy would result in higher risks than restriction to delivery hospital discharge diagnoses. More worrisome in pregnancy registries is that selection bias can be introduced if the outcome explicitly or implicitly affects enrollment (e.g., a known outcome affects eligibility or influences self-selection) or retention (e.g., exclusion of study subjects after an abortion or neonatal death) into the cohort.

5.1. Enrollment

Registries should include women as soon as possible after conception, or even earlier at pregnancy planning stages, to allow the evaluation of early pregnancy events. For instance, pregnancies enrolled earlier in gestation would result in higher risk estimates for spontaneous abortions and terminations than those enrolled later. Still, unless periodic pregnancy tests are conducted, studies will never pick up fetal losses that occur before pregnancy is known.

Women should be enrolled before the pregnancy outcome is known to avoid a selection into the study affected by the outcome. Retrospective enrollment of women after prenatal screening (i.e., nuchal translucency, chorionic villous sampling, amniocentesis, alpha fetoprotein measurements and second-trimester ultrasound (first trimester dating ultrasounds do not assess malformations)), whether the test is normal or abnormal, can introduce bias towards a lower or higher risk of malformations.[1] Underestimation of the risk might occur if enrollment after informative screening tests are conducted selects a survivor cohort of women with uneventful pregnancies (e.g., women might be less willing to contact a "pregnancy registry" after a major malformation diagnosis, clinicians enrolling patients might miss women who had a therapeutic abortion, and some registries do not allow enrollment of women with abnormal prenatal tests or pregnancy losses when an abnormality has been identified). Overestimation of the risk might occur if participation is allowed after an abnormal test and there is a preferential enrollment of women with a diagnosis (e.g., the diagnosis prompts the exposed woman to look for information, find the registry, and enroll). These two scenarios can coexist and even occur differentially in unexposed and exposed women, leading to spurious associations. For example, a reference group of unexposed women might be more willing to volunteer if they are enjoying an uneventful gestation, while women with chronic conditions (e.g., rheumatoid arthritis) might contact a pregnancy registry after receiving an adverse pregnancy diagnosis, seeking both personal support and contact with peers with the same medical condition. Because this bias is difficult to identify and correct in the analytic phase, it needs to be prevented in the design by enrolling subjects prospectively before the outcome is known.

5.2. Follow-up

Another peculiarity of pregnancy studies is that follow-up needs to go beyond the onset of the outcome under study. For most major structural congenital anomalies, the theoretical follow-up would be from conception to the end of fetal organogenesis (i.e., on average three gestational months). However, one often learns about the fetal outcomes only after birth. Longer follow-ups will identify more congenital problems since some structural and many functional malformations might become clinically apparent only months or years later.[19, 23] In pregnancy registries, follow-up typically stops at three to 12 months after the end of pregnancy. Some registries restrict the cohort to infants with a minimal standard follow-up time, such as at least three months after birth (of note, fetal deaths and infant deaths before the minimal follow-up should also be included in the assessment to avoid excluding lethal malformations). Efforts should be made to minimize losses to follow-up and to obtain outcome information for all participants. Information needs to be collected on the number of losses and, if possible, on their reasons, in order to assess whether they are similarly distributed among exposed and unexposed.

6. Exposure Ascertainment

How registries collect information can affect the accuracy of the data. Some registries obtain information from the woman herself. Women often know more about their habits, occupations, medical and obstetrical history, and compliance with the drug than individual health care providers. Other registries rely on reporting by the clinician and have no contact with the patient. Clinicians can provide more complete and accurate information regarding diagnoses and indications.[24] However, there is a risk of exposure misclassification if women stop or incompletely comply with the prescription drug regimen during pregnancy. A woman might be more willing to tell an interviewer not related to her care than to tell her doctor that she decided to reduce her dose or quit taking a medication. Including unexposed subjects in the exposed group (i.e., false positives) can dilute any potential association.

In order to maximize the quality of data, the combination of several sources of information is an optimal strategy. With adequate help, women can recall exposure during pregnancy. As noted previously, they should be enrolled before the pregnancy outcome is known to reduce selection bias, as well as to obtain reliable prospective information on exposure and other characteristics not affected by the outcome. One interview should take place at enrollment (i.e., during the first trimester), and at least one interview should take place post-partum (e.g., 2 months after end of pregnancy). Additional interviews (e.g., mid pregnancy to update exposure information, 12 months after delivery to evaluate development) might be useful depending on the objectives of the pregnancy registry. Data from treating physicians can document the medical condition and confirm the prescription. Confirmation of exposure should be blinded to the outcome.

7. Exposure Definition

If one were concerned about structural malformations, the etiologically relevant period of exposure would be the first trimester of pregnancy. To identify this period, one needs to establish gestational timing. Obstetricians typically time pregnancies from the first day of the last menstrual period (LMP), determined by maternal recall or, preferentially, by a more accurate early pregnancy ultrasound. Depending on the pharmacological characteristics of the drug, the specific defect, and the accuracy of timing information, the window of interest might be the second and/or third month, any time during first four months after the LMP, or even weeks before LMP for drugs with long half-lives or unspecified period of effect (e.g. vaccines). Exposures later in pregnancy can adversely affect other outcomes.[25-27]

Treatment strategies change during pregnancy, and doses are commonly adjusted. Therefore, it is recommended that detailed information on start and stop dates, dose, frequency, duration, and indication be collected.[28] Some medications (e.g., anticonvulsants) are prescribed in combination with other drugs. The effect of polytherapy resulting from concomitant treatments, or from switching drugs, within first trimester should be explored in the analyses. Not only the number of drugs, but which drugs are combined, might affect the outcome of interest. Although the power is usually limited, analyses of dose response can inform recommendations, analyses of specific timing within first trimester can assess biological plausibility, and analyses of indications might help explore confounding. If possible, maternal body mass index should be considered when evaluating dose effects.

8. Covariates: What Else to Collect?

There is always a tension between simplicity and a desire to be comprehensive. Although there is no general rule, one widely accepted principle is that quality is more important than quantity (if the information is not trustworthy, do not collect it). Minimizing the effort and time from participants can increase both the willingness and the quality of participation. Necessary information on exposure, outcome, and key confounders (e.g., history, status, severity, and management of the indication) should take preference over desirable but less useful information. Most registries collect information on demographics, concomitant illnesses and medications, and reproductive history. A list of variables commonly collected in pregnancy studies is provided in Table 1.

Table 1. Variables Commonly Collected in Exposure Pregnancy Registries

Identification	Exposure	Outcome	Covariates
Study ID Date at enrollment LMP date or EDD gestational age at enrollment Date first data collection Date(s) of follow-up data collection(s) Date end of follow-up Follow-up status (e.g.	Drug (for each drug, each episode of use and each dose if changed): -Start date -Stop date -Indication and measure of disease severity (e.g. CD4 count for HIV patients, type of epilepsy and convulsions during pregnancy for	Sex Status (livebirth, elective termination, spontaneous abortion, late fetal death) Number of fetuses (singleton or multiples) For fetal loss: Date end of pregnancy Reason for termination	Demographics: Maternal age, race, occupation, education level. Pre-pregnancy weight and height. Reproductive history: number of previous completed pregnancies and miscarriages, fertility interventions for past and

Identification	Exposure	Outcome	Covariates
complete, loss to follow-up, withdrawal) Contact information -Woman - Alternative contact(s) -Obstetrician / prenatal health care provider -Specialist -Pediatrician	anticonvulsants) -Dose -Route -Frequency -Duration	For Livebirth: Date of birth Birth weight Birth length Head circumference Gestational age at birth Conditions at birth (e.g. admission to intensive care unit, drug withdrawal syndrome) Congenital anomalies: Specific defects Date of diagnosis Methods of diagnosis Date and results of any prenatal testing Obstetric outcomes: Delivery (vaginal, C-section type) Preeclampsia Premature labor Preterm delivery Gestational diabetes	current pregnancy. Family history of defects (specific defect and degree of relationship). Habits: cigarette smoking, alcohol intake, and use of illicit drugs. Chronic medical conditions: diabetes, pregravid obesity, hypertension, epilepsy, depression, other psychiatric disorders, hepatitis, thyroid disease, autoimmune disease, asthma, sexual transmitted disorders, AIDS. Concomitant medications including folic acid supplementation and potential teratogens.

NOTE: LMP: Last menstrual period. EDD: Estimated date of delivery

9. Outcome Ascertainment

The source of information for outcomes in pregnancy registries is critical. Although registry designs that incorporate interviews with pregnant women can provide an initial source of data regarding the results of prenatal diagnosis, postnatal events, and malformations that are recognized during longer term follow-up periods, validation of any maternally-reported diagnosis by the health care provider is an important

criterion for inclusion of specific outcomes such as major congenital anomalies in most registries (see Outcomes Definitions below). In registries where the outcome is routinely collected from the woman and validated with health care providers, the mother's report can correct potential false negative reports from one clinician; in comparison, registries where the outcomes are only reported by the provider rarely include maternal validation. In some circumstances, a more stringent level of validation might be required to confirm an outcome (e.g., an echocardiogram to validate the presence of specific heart defects). Requiring commitment at enrollment to provide consent and medical release of information from obstetrician and specialists might select a motivated patient population and can minimize loss to follow-up and maximize access to medical data.[2]

Whether the exposure is ascertained from the patient or from the clinician, it is important to obtain delivery data from the obstetrician or hospitalization records to ascertain the outcome accurately. It is also important to follow-up with subsequent providers, such as the infant's pediatrician and other specialists, because those treating a woman for a non-pregnancy related condition often know little about obstetric or pediatric outcomes (e.g., a woman's neurologist might not know about the patient's preeclampsia), and obstetricians often know little about the infant after delivery.[1, 19] Obtaining this information might require woman's consent and therefore contact with the patient. In addition, treating physicians might have a legal or ethical conflict of interest if they are asked to report on pregnancy outcome when they were responsible for the exposure.

It is important to recognize that maternal report alone can result in misclassification of the presence or absence or any defect and/or the presence or absence of a specific defect. Although specificity is most relevant for the evaluation of infrequent outcomes, medical validation may be required for all births and not just those in which the mother reports an abnormality to maximize sensitivity.

Other primary or supplemental methods of ascertainment of birth defect outcomes can be employed. These include linkage to claims data, public birth registries, or birth defects surveillance systems. In any case, comparable methods for ascertainment of outcomes must be used in the exposed and reference groups.

10. Outcome(s) Definition

Pregnancy registries must set *a priori* criteria for defining outcomes as part of the study design. Outcomes are typically ranked in order of importance to the registry objectives. For example, most pregnancy registries select major structural birth defects as the "primary" outcome of interest. This

outcome is frequently defined as primary because identification of an increase in major malformations, particularly specific major malformations, in association with a given gestational exposure may indicate a teratogenic effect. In addition, some pregnancy registries incorporate a measure of "minor" structural defects as an outcome representing a broader spectrum of potential structural differences that may be attributable to an exposure.

Other "secondary" endpoints frequently include measures of fetal growth deficiency, preterm delivery, spontaneous abortion or stillbirth, and elective terminations. Some pregnancy registries incorporate longer-term measures of outcomes that can include, for example, postnatal growth deficiency, cognitive and behavioral development, or measures of immune function, depending on the characteristics of the exposure under study in the registry. However, the cost and logistics of following children over time usually are prohibitive.

Using the primary outcome of major structural defects as an example, but relevant to all outcomes in any pregnancy registry, the following definitions must be determined *a priori*.

10.1. Inclusion / Exclusion Criteria for a Defect to be Defined as "Major"

Criteria for defining defects as "major" must be established for a pregnancy registry. For example, major structural defects might be defined as abnormalities in structural development that are medically or cosmetically significant, are present at birth, and persist in postnatal life unless or until repaired. Similarly, criteria must be established for defects that will be excluded. For example, those that are transient and maturational such as a patent ductus arteriosus or an inguinal hernia that might occur in a preterm infant simply as an artifact of shortened gestational age but do not represent an abnormality in embryonic or fetal development might be excluded. Another example of an excluded defect might be a small muscular ventricular septal defect that may spontaneously close with no consequences for the infant. Attention must be paid to the comparability of definitions for inclusion and exclusion between the exposed and the reference groups.

Frequently, pregnancy registries employ an existing standard coding system for inclusion and exclusion of structural defects, such as that developed by the U.S. Centers for Disease Control and Prevention's (CDC) Metropolitan Atlanta Congenital Defects program.[29] This system was created and is maintained for use in an ongoing population-based surveillance program for birth defects that are identifiable up to one year of age. For some pregnancy registries, additional definitions or modified inclusionary/exclusionary criteria may need to be employed, depending on the length of follow-up or the specific outcomes of interest. Even with standard coding criteria, the information available for some

defects that are reported may be insufficient or ambiguous for classification such that an additional level of review is required to classify the defect appropriately. A method for expert adjudication of defect classification, blinded to exposure status, is an important component of a pregnancy registry.

Minor malformations are defined as those uncommon structural differences in the infant that have no serious medical or cosmetic consequences (e.g., an extra hair whorl on the head). Although these minor structural defects are of potential interest as more subtle measures of outcome, they may not be reliably assessed and therefore are frequently excluded unless they are uniformly evaluated in all patients. Similarly, positional deformities are often excluded (e.g., abnormal head shape or plagiocephaly that spontaneously resolves shortly after birth).

The source of information regarding a major structural defect must also be defined as meeting the criteria for inclusion. For example, maternal report of a malformation with no validation by a physician or postnatal diagnostic test may be defined as insufficient for inclusion of a major defect. Another example of a situation that might be defined as exclusionary is a defect that is suggested through prenatal diagnostic tests but for which no postnatal validation is available.

10.2 Timeframe of Diagnosis

A specified period for follow-up through which standardized efforts to collect outcome information on major birth defects should be determined as part of the pregnancy registry design. As length of follow-up increases, the baseline risk of major structural defects is expected to increase because not all structural defects are reliably recognized at birth. Ideally, the longer the follow-up, the more complete an assessment of major birth defects could be. Specific outcomes of interest for the exposure under study may require longer follow-up to be appropriately assessed. However, the length of follow-up selected for the registry may be influenced by the availability of resources and ability to maintain contact with registry participants and/or health care providers over a longer term. For comparability reasons, it is essential that the timeframe for diagnosis be identical in the exposed and the reference population or group.

In addition, the case of major birth defects that occur in pregnancies that end in embryonic or fetal demise must be considered in the registry design. Major defects might be identified in spontaneous pregnancy losses by post-natal pathology, and criteria for inclusion of those defects must be established, as not all spontaneously aborted pregnancies or stillbirths will be uniformly evaluated for the presence or absence of defects. In the special case of elective terminations, criteria for ascertainment of malformations are critical because terminations for defects that are identified on prenatal diagnosis may represent an important subset of outcomes within the registry. Elective terminations for social reasons may not be

uniformly evaluated by prenatal diagnosis for major structural defects. For these reasons, many pregnancy registries treat malformations identified in pregnancies ending in live birth separately from malformations included in the overall sample of registry pregnancies including terminations, spontaneous abortions and stillbirths. Failure to include defects detected among terminations can decrease power and introduce bias, particularly for defects for which termination is often chosen after prenatal diagnosis (e.g., neural tube defects).

10.3. Analytical Approach

No known teratogen increases the risk of all major birth defects. Typically, a specific defect or pattern of defects occurs with increasing frequency following a teratogenic exposure in the critical gestational window for susceptibility. However, specific major defects are rare events in the general population – the most common occur no more frequently than ~1 in 1,000 live births. Pregnancy registries usually do not have sufficient sample size/power to evaluate increased risks for specific defects unless the relative risks are quite large. Therefore, most registries compare the overall proportion of all major defects combined in the exposed group to the overall proportion in the reference group. The rationale for this approach is that if specific defects are increased following exposure, these specific defects will incrementally inflate the overall proportion of malformations in the exposed sample and therefore reflect the excess risk associated with that exposure. For example, if the baseline risk of major malformations is 3% and the risk of neural tube defects is 0.1%, a five-fold increased risk of neural tube defects would inflate the overall risk to about 3.4%. Although the analytic approach for a registry may be based on a comparison of all malformations combined, it is important for pregnancy registries to evaluate any potential excess of specific defects in the exposed group, even if it is descriptive.

In some registry designs, it is argued that malformations of known etiology should be excluded from the overall proportion of major structural defects as they do not have the potential to have been “caused” by the exposure. Examples might be chromosomal anomalies, those defects that have a known single gene cause, or occur in families with a positive history for that defect. The rationale for these exclusions is that inclusion of defects not thought to be caused by teratogens can decrease power to detect an overall difference in risk between exposed and unexposed fetuses. Arguments against exclusion of such defects in the analysis are that the true cause of the defect may not be known and that it is possible that exposure to a medication modifies the genetic predisposition, or other risks for that defect, in a multifactorial manner. Therefore, by excluding such defects one could miss the effect.[30] Inclusion or exclusion of chromosomal defects or those of other known etiology may also be driven by the inclusion or exclusion of those defects in the reference population that is selected for comparison.

11. Reference Group(s): Internal or External, Exposed or Unexposed?

A critical element for pregnancy exposure registries is the choice of comparator groups. A valid reference group needs to have comparable 1) outcome definition (e.g., exclusion of minor anomalies); 2) outcome assessment (e.g., intensity of screening, frequency of terminations, inclusion of prenatal diagnoses, availability of diagnostic tests, start and stop of follow-up)[23]; 3) selection of subjects into the study (e.g., gestational age at enrollment); and 4) baseline risk (e.g., distribution of risk factors, including indication). Ideally, each registry is constructed to include an internal reference group. When this is not possible, inclusion of an external reference group must be selected with care. Each comparison group has its advantages and disadvantages. For example, an external population-based reference can provide stable estimates for specific malformations, while internal comparison group would have limited sample size to asses specific malformations but can provide more comparable estimates for malformations overall. More than one comparison group can be contemplated in the protocol to improve the overall validity and statistical power. In anticipation of potential conflicting results, however, primary comparison groups should be identified and justified *a priori*.

External references are rarely considered valid in medical research; for example, a RCT or observational study on the relative risk of stroke in an elderly population exposed to a drug would never use an external reference. However, pregnancy registries often compare their estimates with the background risk in "standard populations" such as the CDC's Metropolitan Atlanta Congenital Defects Program, where the frequency of malformations among pregnancies of 20 weeks or greater is 2.1% when diagnosed prenatally or within the first week of life, and 2.6% if infants are followed until the first birthday;[19, 23] or EUROCAT data in Europe where the prevalence at birth is 2.0%.[31] One advantage of the use of available data from large external reference populations is that it avoids the costs in time and money of enrolling unexposed subjects and provides stable risk estimates for common specific malformations in the general population.

External reference groups must be used with caution since the estimated risk of "major malformations" can vary widely depending on the population, on the definition, and on the ascertainment methodology. When external references are the only alternative, they must be appropriate to the population being studied. Analyses should at least take into account the characteristics of the surveillance program and use the same methodology in their exposed group. For example, characteristics to be considered are whether prenatal diagnoses and terminations were counted, were malformations identified during the delivery hospitalization or was there follow-up for a number of months, or were chromosomal malformations and minor malformations included. Many registries use external references as a necessity. Information from

these sources can be helpful, particularly when there are no other data available, and as long as findings are interpreted with caution. For example, external references can identify major teratogens (like thalidomide), generate hypotheses when unusual patterns of malformations are identified, and inform the need for additional, targeted, epidemiologic studies.[24]

Some registries enroll an internal reference group of unexposed women who undergo the same processes as the exposed and then adjust for potential confounding by matching or adjusting for key covariates. There is still some risk of differential gestational time at enrollment (exposed women may tend to enroll earlier) and lack of comparability (unexposed women rarely have the underlying condition for which the drug was indicated). To make groups more comparable, some registries use women exposed to other non-teratogenic drugs as the reference (e.g., OTIS registries)[32] or enroll pregnant women with a common condition or indication treated with various drugs or untreated (e.g., the multi-sponsor North American Antiepileptic Drug [NAAED] Pregnancy Registry,[33] and the human immunodeficiency virus [HIV] Anti-retroviral Pregnancy Registry[34]). More recently, the scientific community has moved towards the evaluation of the comparative safety and efficacy of different treatments with similar indications used in similar populations, whenever possible. The comparability of groups is enhanced by this approach, although sometimes more severe conditions are channeled to specific treatment, while milder ones can remain untreated, thus confounding by severity or type of disease is still possible. In addition to improving validity, using alternative treatments as a reference would answer the clinically relevant question of "how to treat" rather than whether to treat. Moreover, since there is often need for safety data for multiple drugs with the same indication, evaluation of a variety of drugs used to treat the same condition could be most efficient. The feasibility of multiple-drug registries depends in part on the sponsorship.

Taking advantage of the etiologically relevant periods of exposure during pregnancy, some studies compare first trimester with second or third trimester use of the drug. These comparisons are only possible for non-chronic treatments and should consider that the opportunity for exposure after first trimester can be affected by the outcome. For example, if the pregnancy is terminated because of a malformation, there would not be second/third trimesters exposures; therefore, later exposures might be artificially associated with lower risks.

12. Analysis of Registry Data

Pregnancy registries frequently include multiple outcomes as endpoints and may have more than one comparison group; in addition, as stated above, major birth defects among live births may be evaluated separately from major birth defects among all pregnancies. To address the problem of multiple comparisons in analysis and interpretation of registry data, it is essential to establish an analysis plan that identifies the primary hypothesis being tested, typically the proportion of pregnancies involving a major birth defect, and to specify which are the primary groups being compared.

Similarly, the analysis plan should attend to the design of the registry and the expected sample size. For example, registries that involve multiple centers/multiple countries should use appropriate conditional methods of analysis that account for center. For outcomes with low frequencies, e.g., major structural defects, sample size projections should indicate when exact methods must be used for analysis. To evaluate the robustness of findings in the registry, the analysis plan can incorporate a plan for *post-hoc* sensitivity analyses under various scenarios.

13. Statistical Power, Registry Size, and Duration

The projected sample size for a specific pregnancy registry is affected by the frequency with which the medication is used by women of reproductive age, the proportion of exposed pregnancies that it is estimated are possible to identify and recruit into the registry, and the scope of the registry (local, national, international). The power of the study to detect an effect at or above a certain level is affected by the sample size and the baseline risk for the outcome in the population.[19] The estimated losses to follow-up will affect the useful sample size. The duration of follow-up for outcome assessment will affect the cumulative risk estimate. Missing therapeutic abortions will affect both.

If the medication under study typically is only taken for a few days or intermittently, this fact should be considered in calculations of power and sample size. The effective sample size of pregnancies exposed only in specific gestational windows of time (any one of which may be a risk period while others are not) may be much smaller than the overall projected sample size of pregnancies exposed for the registry. These factors must be balanced against the amount of time needed to accumulate a sample size that is sufficient to produce a clinically relevant result.

For example, in a pregnancy registry that compares the overall proportion of pregnancies resulting in an infant with a major birth defect among exposed women to an external reference population with a baseline prevalence of major defects of 3%, a sample size of 200 exposed live born infants would be sufficient to detect a 2.2 fold relative risk with 80% power at an alpha of 0.05. However, the same sample size of exposed live births would only be sufficient to detect a 10.4 or greater relative risk for cleft lip with or without cleft palate.[34]

Because of the limits of power associated with typical sample sizes for pregnancy registries, especially in detecting risks for specific birth defects, an approach that has been used in some registries is the "rule of three": an alert is triggered when three or more specific defects are reported for a specific exposure. This rule is based on the <5% likelihood of a chance finding of three or more of the same specific, relatively rare, defect in a cohort of 600 or fewer subjects. Although some defects or defect groups occur frequently enough that the "rule of three" would not apply, this method can be used to flag an unusual finding for further review.[34]

14. Biases

14.1. Selection Biases

Because pregnancy registries typically enroll women only after recognition of pregnancy and in some cases much later in pregnancy, the group of enrollees is a selective group of pregnancies that have survived to that point in gestation and may have a shortened remaining period at risk of incurring the outcome of interest. There is "left truncation" of the registry cohort such that it is devoid of women who have already had a spontaneous abortion, an elective termination, or a stillbirth depending on the gestational age of enrollment. Although statistical methods can be used to address left-truncation, survivor bias threatens ability to evaluate risk for pregnancy outcomes including birth defects and calls for a registry design that encourages recruitment of participants as early in gestation as possible. In the extreme, one cannot study infertility in a cohort of pregnancies because, by definition, women conceived. Nor would one estimate the incidence of spontaneous abortions in women enrolled after 20 weeks of pregnancy. Spontaneous abortions can still be evaluated when women are enrolled during the first trimester by assessing the rate of miscarriages per gestational month, thus taking into account the decreasing trend as gestation progresses and the gestational time at enrollment.

Early enrollment is also of benefit with respect to biases potentially introduced by prenatal diagnosis. Prospective registry enrollment before any prenatal diagnostic test for major birth defect avoids bias in

the direction of an increased risk for defects so identified, or in the other direction if women enroll preferentially after prenatal diagnosis that has shown no defect. Pregnancy registry guidelines recommend that only participants enrolled prior to prenatal screening be included.[2] However, as prenatal diagnosis becomes feasible earlier in gestation, this becomes more difficult to achieve.

Many registries also collect data retrospectively, but these data should be analyzed separately. Women enrolled after an abnormal pregnancy test can be analyzed as case series in passive surveillance systems of spontaneous adverse event reports.[2] Malformations can be evaluated for biological plausibility, and specific patterns of malformations or distinct congenital abnormalities can generate hypotheses. On the other hand, retrospectively enrolled subjects without malformations would offer limited information, therefore, the benefit (and ethics) of including these women is questionable.

Registries should report the gestational age at enrollment for their exposed and reference groups. If enrollment time differs, methods that adjust for left censoring such as restriction to prospective pre-tests enrollees should be applied.

Another bias might be associated with right truncation of the registry cohort. This occurs when pregnancies with unknown outcomes are considered ineligible for analyses. By excluding terminations, spontaneous abortion and losses to follow-up, one assumes that the exposure had the same effect in these pregnancies as in those that remain under observation. Such assumption is less plausible if the exposed group has a higher frequency of these outcomes than the reference group. The frequency of spontaneous and therapeutic abortions, losses to follow-up, and withdrawals should be reported for the exposed and reference groups. Of note, a higher frequency of terminations among exposed women might reflect a higher risk of malformations, as well as more fear of malformations with consequent abortion if the drug is suspected of being teratogenic.

Bias may also occur with events that shorten the follow-up (e.g., preterm delivery that cuts the possible number of weeks of exposure to the drug of interest). The assessment of transient exposures (e.g., vaccines) during pregnancy in relation to outcomes associated with shorter gestations (e.g., preeclampsia, prematurity, pregnancy weight gain) needs to take into account the smaller opportunity for exposure in shorter pregnancies.

Biases can also be introduced in the analysis by stratification or adjustment for covariates that are themselves affected by the exposure of interest and are affected by the outcome or share common causes

with the outcome (e.g,. adjusting for gestational age at birth when studying the effect of a pharmaceutical on structural malformations).[35-37]

14.2. Information Bias

As noted above, pregnancy registries are preferentially "prospective" in design. Therefore, the outcome cannot directly affect the accuracy of exposure information, and any misclassification of drug exposure would tend to be non-differential with respect to the outcome. However, non-differential misclassification of exposure is still problematic for safety evaluation since it would tend to bias any potential effect towards the null. It is therefore crucial to maximize the quality of drug exposure information.

When information on exposure is provided by the women, the accuracy of the recall can be maximized by using structured questionnaires, detailed questions, and calendars to help establish gestational timing and enhance recall of dates. Maternal reports on drug utilization are sometimes cross validated with medical records. When the health care provider is the only source of data, information on prescriptions might not reflect the real use of the drug during pregnancy if patient compliance is incomplete. Since pregnancy registries typically focus on uncommonly used medications, the impact of false positives among exposed is much greater than the occasional inclusion of false negatives among the unexposed reference group.

At the time of enrollment in the registry, women are reporting their medical history retrospectively, knowing that they are enrolled in the registry because there is lack of information on the safety of the drug under study. This potentially could influence the accuracy of recall for baseline covariates (i.e., recall might be more accurate for exposed than for unexposed groups because there is a different motivation to recall a medical history) and affect the ability to control for confounding. For this reason, methods for validation of key covariates are of benefit.

Diagnostic bias, or outcome misclassification, could also occur in pregnancy registries. These biases could be either non-differential or differential between the exposed and the reference group and could bias the estimate of an effect towards or away from the null. For example, participation in the registry itself or concern that a drug exposure might pose a risk could lead to more access to or uptake of prenatal diagnostic measures such as ultrasound and to more careful examination of infants for defects postnatally, potentially leading to differential accuracy in detection and classification of defects among exposed and unexposed. The risk of differential outcome classification among the exposed and the reference is greatest when external control reference groups are used.

As mentioned above, reported major malformations must be validated. Although neither treating physicians nor women are blinded to the treatment, registries can do blinded validation and adjudication of outcomes. To detect malformations not reported by the patients, records for all pregnancies would have to be reviewed. For example, women might be less likely to volunteer information regarding male genital malformations in their infants, which could result in under-ascertainment of these malformations. To detect malformations detected by screening and frequently terminated (e.g., anencephaly), therapeutic abortions need to be included. To detect an increase in abnormalities incompatible with life, it is important to collect information on autopsy results at stillbirth and, if possible, on examinations of the fetus after spontaneous or induced abortion. The study of spontaneous abortions and pregnancy terminations itself presents an additional methodological challenge, especially in countries where abortions are illegal or when studies include pediatric populations, because induced fetal losses are sometimes categorized as spontaneous in medical records, and questions about terminations might be considered of sensitive nature.

14.3. Confounding

Several sources of confounding may affect pregnancy registries, as in any other observational study. Socioeconomic status, maternal age, tobacco and alcohol use, illegal drug use, maternal body mass index, vitamin use are examples of potential confounders that might be related to, or impact, the exposure under study and are also risk factors for some pregnancy outcomes.

With medications, there is the concern of confounding by indication (e.g., the association of a weight loss drug with birth defects might be confounded by the maternal obesity the drug is used to treat). Confounding by indication is difficult to address in non-randomized studies for common chronic conditions such as depression, asthma, epilepsy, HIV, and autoimmune diseases where separating the effect, if any, of the drug from the underlying disease can be challenging. A related form of confounding that may be of concern is channeling bias (i.e., women with more severe underlying disease may be most likely to be selected to receive the drug of interest).

Each of these potential confounding concerns provides a strong rationale for inclusion of actively recruited comparison groups that are matched to the exposed registry participants on the maternal underlying disease and, to the extent possible, similar in distribution of maternal disease severity.

15. External Validity or Generalizability

As mentioned above, pregnancy registries typically rely on volunteers to participate. With appropriate selection of comparison groups and control for sources of confounding, pregnancy registries can make assertions about internal validity. However, there is usually little known about the characteristics of the entire population of exposed pregnant women from whom the sample is drawn. For this reason, it is difficult to make conclusions about external validity or generalizability for any pregnancy registry. If the participants represent a select group, it may be difficult to generalize the findings from the registry. However, the registry volunteers would have to differ from non-participants on some characteristic that modifies the effect of the drug on the pregnancy outcome (e.g. a teratogenic effect might vary by race or by baseline folate levels in the population).

16. Operations

16.1. Study Protocol

Pregnancy registries are scientifically rigorous studies designed to monitor safety of product use in pregnancy. As such, they have formal written protocols based on epidemiologic principles and regulatory authority guidance documents for pregnancy registries.[2, 5] The protocol should include a brief and cogent review of the literature, registry objectives, study design, detailed data collection procedures including sources of data, inclusion and exclusion criteria for patient enrollment, operational aspects of enrollment and retention, definitions of relevant endpoints, analytic considerations, statistical plan, regulatory and ethical considerations, reporting and publication plans, governance, and criteria for registry termination. While protocols of RCTs are registered (e.g., on ClinicalTrials.gov), it is still controversial whether those from observational studies should also be registered.[38]

Pregnancy registries have procedural differences that are distinct from RCTs and from other observational study designs. Thus, they often require specific standard operating procedures that clearly document their unique processes. Each pregnancy registry should also have its own registry management plan that serves as the registry team's roadmap to conducting the registry. For more details see Chapter 2 on Developing a Project Plan in the Registry Users Guide.[39]

16.2. Human Subjects, Informed Consents, and Medical Records Release Forms

Pregnancy registries involve human subjects research and thus involve institutional review board approval, informed consent of participants, and protection of confidentiality in data collection, data

storage, and publication. Pregnant women are considered a 'vulnerable' population in legal/ethical terms but, importantly, unlike other populations considered in this section of regulations (mentally challenged patients, inmates in penal institutions and children), they are able to give consent. Therefore, registries must adhere to guidelines for human research protection in pregnant women. Pregnancy registries are typically considered minimal risk protocols. As with participation in any research of this type, there is the potential for loss of the confidentiality of the information provided. Methods to manage this risk must be in place and communicated to the potential participant. There is also the psychological risk of participation in a pregnancy registry that may be involved when a woman becomes more aware of the potential fetal effects of exposure. Resources for expert counseling on fetal risks are available in the U.S. and Canada through the Organization of Teratology Information Specialists (http:www.otispregnancy.org).

Informed consent is typically required for participation in a pregnancy registry. Some registries collect data directly from health care providers without collecting patient identifiers. These types of registries may qualify for a waiver of informed consent. For pregnancy registries in which the patient is identified and data are collected from multiple sources, informed consent is required. However, under several conditions, signed informed consent may be waived and verbal consent allowed. Additionally, in order for pregnancy registries to collect data from a woman's health care provider(s) and/or her newborn's pediatric health care provider, the woman must complete medical release forms for each clinician who will report data or provide medical records to the registry. Chapter 8 in the Registry User's Guide [39] thoroughly discusses the application of ethical principles for registries.

16.3. Recruitment and Enrollment

In some registries, pregnant women self-identify and self-enroll calling a toll-free number that can be found in key websites, printed materials in specialists' offices, or drug inserts. In other registries, women are referred by general practitioners, specialists treating the condition of interest, or obstetricians. Direct enrollment of women may allow inclusion of participants earlier in gestation since they are usually the first to know about their pregnancy. Enrolling women directly may also facilitate pregnancy and postnatal follow-up. However, for some drugs, it might be unrealistic to expect self-enrollment and, for some conditions, it might be more efficient to identify eligible women through specialists. Some countries do not allow self-enrollment. Who initiates enrollment has implications for data collection and informed consent processes.

Pregnancy registries can be conducted by regulatory or other government agencies, academic centers, contract research organizations (CRO), or drug manufacturers. Who sponsors and who conducts the

registry can affect the participation of health care centers, some of which might have barriers for collaborations with industry. Biopharmaceutical companies sometimes identify exposed pregnancies as outcomes reported to their surveillance system for drugs not recommended for pregnant women. If identified and enrolled before prenatal tests, these pregnancies can be followed and yield a pregnancy registry nested in a passive surveillance system.

To maximize enrollment and maintain efficiency, pregnancy registries typically do not use a traditional "site-based" approach. Rather they employ a single coordinating center that recruits and enrolls all eligible pregnant women as soon as possible after conception. The number of coordinating centers that would make the study feasible varies among countries and among registries. Therefore, awareness campaigns must reach out to pregnant women and their health care providers in a broad variety of settings. A carefully constructed awareness plan should be designed specifically for each pregnancy registry, accounting for the aims of the registry, the target population, and the geographic scope. In general, the plan should incorporate a variety of persistent awareness strategies to ensure broad coverage, including announcements of the registry with contact information posted in the following:

- Product label
- Registry and/or product Internet sites
- Personal mailings to health care specialists – particularly high prescribers and high-risk obstetricians
- Professional journals
- Exhibits or scientific presentations at professional meetings
- Lay magazines
- Advocacy group newsletters and/or Internet sites
- Social media

For pregnancy registries sponsored by individual biopharmaceutical companies, promotional materials are subject to 21 CFR 314.81(b)(3) or 601.12(f)(4), and should be submitted to the FDA Division of Drug Marketing, Advertising, and Communications (DDMAC). In general, registry recruitment and awareness

materials should not promote use of the product in pregnancy nor imply that the product is safe and effective in pregnancy unless sufficient scientific evidence exists to support these claims.

Enrollment might also be affected by the lag time between approval of the pharmaceutical and launching of the registry. Awareness and interest might peak right after approval, which is also the ideal time to collect safety data in the post-marketing setting. Streamlined procedures for informed consent and data privacy/HIPAA authorization have been demonstrated to increase enrollment in pregnancy registries.[40]

16.4. Retention of Participants during Follow-up

Participant retention is crucial to achieving the registry goals. Retention can be encouraged by engaging registry participants, including pregnant women and their health care providers in the reporting process and making it as easy as possible to report data to the registry. Registry staff should be trained to collect data for observational studies from both patients and health care providers who do not usually participate in research activities. They should develop a rapport with the reporters that facilitates data collection and promotes retention through relationship building.

Streamlined data collection processes and simple, concise data collection forms are essential for reducing the burden of reporting. For registries sponsored by pharmaceutical companies, duplication of work because of safety reporting should be avoided as much as possible. A small monetary stipend may also encourage retention but may not be universally accepted by registry reporters. For health care providers, especially obstetric care providers, registry data are often more valuable than a stipend. Giving interim data reports to health care providers who report data to the registry is a powerful recruitment and retention incentive.

16.5. Data Collection

Like any other epidemiological study, registries can benefit from the technologic advances in communications (e.g., ability to enroll subjects through social networks and collect data through confidential websites). Most registries enter the data directly into an electronic database from records or questionnaires, or using computer-assisted telephone interviews. To assure confidentiality, identifiers should not be included in the database containing clinical information. The goal is to make reporting data to the pregnancy registry as easy and unobtrusive as possible. Providing a variety of data-reporting mechanisms that are simple to use and fit the reporters' preferred communication practices facilitates recruitment and retention. If feasible, a pregnancy registry should be designed to allow participants to report data through telephone interviews, paper data collection forms that can be mailed, scanned, or

faxed to the registry, simple Internet-based electronic data capture systems, and mobile applications. For details on these data capture systems, see Chapter 10 in the Registry User's Guide.[39]

Pregnant women are ideally enrolled into a pregnancy registry before or soon after the exposure of interest and then followed throughout pregnancy. If a live infant is born, the infant may be followed for a period typically ranging from 3 to 12 months. Data may be collected at several different time points.

- At enrollment, information is collected on product exposure(s), maternal characteristics, prenatal testing, and other baseline data.
- During the second and/or third trimester of pregnancy, a brief update on the status of the pregnancy and exposures may be obtained.
- At end of pregnancy, information is collected on additional exposures during pregnancy, risk factors, and details regarding the pregnancy outcome, including any congenital anomalies, gestational age and birth weight, and perinatal complications.
- For live births, data may be collected at several points in time (typically 3, 6, and/or 12 months) on infant characteristics and health outcomes including congenital anomalies, functional or developmental deficits, and other outcomes pertinent for the drug of interest.
- Targeted follow-up may be needed to collect additional detailed data on specific outcomes of interest.

The source of baseline and follow-up data is an important consideration. It is best to evaluate all options for obtaining data to determine the most appropriate source for the specific data requested. The pregnant woman can provide detailed data on drug use including exposures to the drug of interest as well as other drugs, on relevant risk factors, and on the pregnancy outcome. Her health care providers, such as the prescriber, specialists and/or obstetrician should be able to provide or verify this information. The collection of data from a variety of sources throughout pregnancy and during infancy contributes to the accuracy and comprehensiveness of pregnancy registry data. Requiring information from multiple health care providers, in addition to the patient, is a distinct challenge of pregnancy registries.

Data quality is of the utmost importance in a pregnancy registry, and multiple levels of quality assurance should be employed beginning with the design of data collection instruments. Instruments should be

designed with care, thoroughly vetted, and pilot tested to ensure ease of reporting valid, reliable data. Instructions for self-reporting should be clear and succinct. Electronic data collection (EDC) systems should include validity checks. If interviews are to be used, the interviewers should be thoroughly trained to conduct neutral, unbiased interviews using detailed interview scripts. Once captured, the data should go through a rigorous cleaning and quality assurance process to reduce errors, missing data, and misclassifications. When possible, patient-reported data should be verified by health care providers or medical records.

16.6. Adjudication of Outcomes

Many pregnancy registries enlist the services of a clinical geneticist or dysmorphologist to review and classify all congenital anomalies. A standardized classification system should be used, such as the Centers for Disease Control and Prevention's Metropolitan Atlanta Congenital Defects program.[29] The assessor and method of assessment should be the same for both the exposed group and comparison group and the assessor should be blinded to the exposure of interest. The assessor may also evaluate likely causes for the particular birth defect, such as family history, genetic factors, and/or exposure to known teratogens.[28] If the exposure of interest is made available, the assessor may examine the timing of the exposure relative to the origin of the birth defect (e.g., to attempt to determine if the timing of the exposure is relevant to the formation of the birth defect).[28] Registries can engage the Scientific Advisory Board and/or a subgroup of the Board to review each case and the assessor's classification and reach consensus on the classification.

16.7. Process of Releasing Findings

To Regulatory Agencies

The FDA considers individual case reports from a pregnancy registry as derived from active solicitation of patient information (US FDA 1997), including reports from participants enrolled both prospectively, where the exposure is reported prior to knowledge of the pregnancy outcome, and retrospectively, where the pregnancy outcome is already known at the time the exposure is reported.[2] A company sponsor holding marketing authorization for an approved drug or licensed biological product must submit to the FDA, within 15 calendar days, reports of pregnancy registry adverse events that are both *serious* and *unexpected* by regulatory definition and where a reasonable possibility exists that the drug or biological product caused the adverse event. Current reporting requirements in the regulations consider any congenital anomaly a serious adverse event.

Pregnancy registries run independently of any sponsors holding marketing authorizations are not subject to postmarketing regulatory reporting requirements. However, investigators running such registries may

forward reports of any serious adverse events including congenital anomalies to the sponsor of the medical product or report directly to the FDA MedWatch office (1-800-FDA-1088 or http://www.fda.gov/medwatch).

Any company conducting a pregnancy registry required by the FDA must submit an annual status report to the agency. Companies conducting pregnancy registries not subject to annual reporting requirements are encouraged to include a status report in the periodic safety report. The status/interim report should describe the study design and summarize the status of the planned, initiated, in progress, or completed pregnancy registry conducted by or otherwise obtained by the sponsor during the reporting period. The status report should also provide a descriptive summary of progress to date, interpretation of findings and appropriate analyses with comments on the clinical significance of the findings. Copies of full reports may be appended, if appropriate. Any publications based on data from the registry should be included.

The registry status report should include the following, presented separately for prospective and retrospective reports: number of pregnant women enrolled to date, number of pregnancies with outcome known (stratified by live birth, spontaneous abortions, elective terminations, fetal deaths/stillbirths), number of pregnancies with outcome pending, and the number of pregnancies lost to follow-up (p 16, FDA guidance).[2] For pregnancies with known outcome, line listings and summaries of: demographics, obstetrical, and medical history of mothers, weeks of gestational age at exposure, dose and duration of exposure, whether multiple gestation, and weeks of gestational age at completion or termination of pregnancy. For live births and deaths/stillbirths, small for gestational age, preterm delivery, and congenital anomalies or other fetal abnormalities, and for spontaneous abortions and elective terminations, abnormalities in products of conception, if known.

Reporting to Sponsors

There are no specific rules around reporting events to the sponsor company. The sponsor company negotiates this with the academic center or CRO. Some pregnancy registries have regular (semi-annual or annual) data cut-off points and issue corresponding periodic interim reports summarizing the aggregate data. The sponsor can then use this interim report for any regulatory reporting requirements.[1] Others, such as the NAAED have pre-specified criteria for release of results for a positive association.[33]

Most pregnancy registries administered by CRO or academic institutions report all adverse pregnancy outcomes (birth defects, spontaneous fetal losses, induced abortions), maternal and non-defect fetal events, regardless of attribution or seriousness, to the sponsor within a few business days. Some sponsors only request specific outcomes, or serious and attributable events, reported to them. The sponsor reports

these events to regulatory agencies. Pregnancy is not considered an adverse event (AE). The registry might also report spontaneous AEs (not registry related outcomes) to the sponsor, although these may be received infrequently. If the sponsor requires further information on a pregnancy related AE, they can contact the registry who would contact the health care provider. The sponsor follows up on any non-registry cases as needed.

Publication Policy

Some pregnancy registries have formal publication committees.[34] In others, such as the NAAED, the independent Scientific Advisory Committee has final approval over the announcement and publication of registry data and conclusions.[33] Others simply publish the findings when they have them. Registry results should be published as soon as the number of women in the registry permits to allow dissemination of the results to the scientific and clinical communities.

16.8. Role of (Scientific) Advisory Board

The scientific conduct and analysis of the Registry are usually overseen by an independent Scientific Advisory Committee. The committee can advise and participate in establishing a registry and its design, as well as assist in the review of data, classification of any birth defects and the dissemination of information to ensure that results are interpreted and reported accurately. The role and duties of the committee should be specified in the protocol. Members of the committee could include experts in obstetrics, embryology, teratology, pharmacology, epidemiology, pediatrics, clinical genetics, any relevant therapeutic areas and consumers representing the disease state being treated and may include members from the CDC, the National Institutes of Health (NIH), academia and the private sector. The Advisory Committee might review the Registry data, develop Consensus Statements, provide recommendations on modifications or enhancements to the Registry, and/or assist in the dissemination of information and formulation of strategies to encourage enrollment in the Registry. In addition to the Scientific Advisory Committee, multi-company sponsored pregnancy registries may also include a Steering Committee composed of representatives of the sponsoring pharmaceutical companies.

16.9. Stopping Rules

The criteria to determine when to end the study may be predetermined and specified in the protocol. If the registry is conducted by a drug company as a regulatory requirement then the decision as to when to actually end the study is made jointly by the company and the regulatory authority. In other pregnancy registries, it is the Scientific Advisory Committee that decides or contributes to the decision about when to stop. Sometimes, findings from the registry might affect the decision of whether it should continue.

Criteria for possible discontinuation of a pregnancy registry include the following:

1) Sufficient information has accumulated to meet the scientific objectives of the registry (i.e., numeric targets or predetermined effect size, such as the "no evidence of risk" or "evidence of relative safety" thresholds defined in the newly created Vaccines and Medications in Pregnancy Surveillance System (VAMPSS)[20]: "Estimates of safety cannot be absolute; rather, they reflect the degree of confidence that is consistent with an observation of no increased risk between a given exposure and outcome. As more data are collected over time, power increases; for a null observation, increasing power leads to increasingly narrower confidence bounds and increasing assurance of relative safety[20]." For example, evidence of relative safety might be reached when the upper bound of the 95% confidence interval excludes a 2 fold increased risk.
2) The feasibility of collecting sufficient information diminishes to unacceptable levels because of low exposure rates, poor enrollment, and/or high rates of loss to follow-up.
3) Other methods of gathering appropriate information, such as case-control surveillance or large health care databases, become achievable or are deemed preferable.[2]

16.10. Multidrug Pregnancy Registries

A multidrug pregnancy registry actively collects information on exposure to various drug therapies in specific diseases, such as the Anti-retroviral[34] and Antiepileptic drug (NAAED)[33] pregnancy registries. In some cases, a general multidrug registry, such as that conducted by a teratogen information service, collects information on drugs for either unrelated or related indications. Multidrug registries have advantages over single drug registries with respect to both efficiency and economy and allow the examination of polytherapy. They may also have the advantage of having comparison groups of pregnant women unexposed to the specific medical product(s) of interest but with the same indication (e.g. disease registries) or exposed to other drugs for the same indication readily available.

Establishing disease pregnancy registries for common conditions in women of childbearing age (e.g. multiple sclerosis) treated with any available drug would make most sense from a logistic (avoid duplicating efforts), methodological (validity and power) and clinical (comparative safety) point of view. However, that would require collaboration among companies competing for the same therapeutic area, which may be difficult.

16.11 Multicenter and Global Registries

Like RCTs, registries can be multicenter, national, or international. Currently, registries are centralized by sponsors, CROs, government agencies or academic centers; and typically focus on a single drug,

multiple drugs within a class, multiple drugs for different indications, or are disease-based and evaluate drugs used to treat a particular condition.[1] In the future, rather than conducting a new registry for each drug, a global centralized mega pregnancy exposure registry may exist, guided and coordinated by collaborations among regulatory agencies, pharmaceutical companies, contractor organizations and academic centers.[19] Although appealing, in practice, it will be a challenge for this approach to accommodate the case-specific needs for each drug.

The setup, management, and analysis of a locally run registry represent a very different scope than a global registry. The geographic scope includes many challenges (e.g., language, culture, time zone, regulatory differences) that must be taken into consideration in the planning process. A distinct feature of a country-specific registry is that the patient population tends to be more homogeneous with respect to demographic characteristics, exposures, length of follow-up, and diagnosis of outcomes than international registries.

17. Advantages of Pregnancy Registries

Pregnancy registries are often the initial proactive step in assessing the safety of use during pregnancy of new drugs after they first are marketed because they provide a number of advantages over other approaches:[1] While many pregnant women use medications, their use of individual drugs can be quite rare.[41] By enrolling an exposure group made up only of women who took the medication(s) of interest, pregnancy registries are efficient for evaluating the effects of infrequently used drugs in the population. This is a distinct advantage over other study designs, such as case control studies and small health care utilization databases, which usually do not have sufficient power to evaluate outcomes following rare exposures.

The longitudinal nature of pregnancy registries allows the estimation of absolute risks of pregnancy outcomes. For example, registries that enroll women before any prenatal testing has been performed can estimate the risk of malformations among infants whose mothers used a drug of interest. This is in contrast to case control studies that estimate risk relative to that for a reference group. Information about the absolute risk of outcomes is particularly helpful when counseling women who are planning a pregnancy or already became pregnant while taking a drug. The prospective enrollment facilitates ascertainment of drug exposures close to the time a medication is actually used and before information about the pregnancy outcome is known. When registries interview pregnant women directly, they can obtain accurate information about the timing in relation to gestational age, dose, frequency, and duration

of medication use, as well as covariates, and can therefore reduce exposure misclassification, recall bias, and confounding.

Pregnancy registries can compare the risk of outcomes among women who have used a variety of treatments for a single condition, including different monotherapies, different polytherapy combinations, or no treatment at all. This information is useful to both women and health care providers when making decisions about whether to treat a condition during pregnancy and which of multiple alternate therapeutic strategies to use. An additional advantage is that a single registry can monitor a variety of pregnancy and infant outcomes after medication exposure, including postnatal outcomes.

18. Limitations of Pregnancy Registries

The pregnancy registry approach also has a number of limitations. While they are an efficient means to assess rare exposures, they lack the statistical power to evaluate rare outcomes. Most teratogenic exposures do not increase the prevalence of all malformations, but have a more selective effect on individual defects or distinct patterns of defects. Pregnancy registries are powered to detect common outcomes such as the total prevalence of all malformations, and can detect only very large increases in these rarer individual defects or patterns. However, many drugs associated with adverse effects in pregnancy result in only moderate increases in these rarer outcomes. Therefore, pregnancy registries are limited in their ability to detect teratogenic effects on specific malformations with statistical certainty. As such, registries can generate hypotheses that form the basis of further investigation using complimentary approaches, study designs, and data sources..[20]

Another important limitation of some registries is the lack of a comparable reference group. Ideally, a comparison group should be drawn from the same population as women with the exposure of interest using the same methods for recruitment, enrollment, and ascertainment of outcomes so that both groups have the same baseline risk for adverse pregnancy outcomes. Comparison of registry findings with data from other studies, such as population-based surveillance programs or hospital deliveries, can lead to biased results if the subjects in the reference group have characteristics different from those of the registry participants, or it the methodology for case ascertainment is different. Identifying an appropriate reference group can be particularly difficult for global registries that recruit exposed women from multiple countries with potentially different populations and backgrounds. For registries such as these, a comparable unexposed group may not exist. Even when an internal reference group is recruited,

differences between the exposed and unexposed groups in factors such as the indication for the drug or the proportion of subjects that are lost to follow-up potentially can affect the validity of the results.

An additional consideration is that findings from a pregnancy registry may not be generalizable to the broader population of all women who use a drug. Enrollment of women in pregnancy registries typically is voluntary, self-selected, and registry participants represent a small proportion of all women who have taken a drug. For these reasons, the characteristics and experience of women who participate in a registry may differ from those of nonparticipants, and those characteristics might modify the effect of the drug.

A final limitation of pregnancy registries is the length of time typically required to enroll sufficient numbers of exposed women to generate stable estimates of pregnancy outcomes. This can be affected by the frequency of exposure in the general population, and the methods and extent of recruitment efforts by the registry. Most registries continue for years before publishing final results. This extended period of evaluation before reaching conclusions regarding adverse pregnancy outcomes can be a disadvantage when there are pregnancy outcomes of concern that need to be evaluated quickly or when new therapeutic agents become available.

19. Evaluation of Reports from Pregnancy Registries

It is important to critically assess the results and conclusions of reports from pregnancy registries. Key issues to consider include the following:

Table 2. Issues to Consider when Evaluating Reports from Pregnancy Registries

Area	Issues
Objectives	What question(s) is the registry attempting to answer? Are the design and methods appropriate to do so?
Background	What condition(s) is the drug used to treat (e.g., chronic vs. episodic)? In what settings is the drug likely to be used (e.g., as primary treatment or as adjuvant therapy with other drugs)? What is the recommended therapeutic dose and duration of use? Is the drug likely to be used off-label for conditions other than the stated indication?
Study population	What is the target population from which pregnant women exposed to the

Area	Issues
	drug are drawn? Does the report describe the characteristics of women enrolled in the registry? Could the study subjects differ from women in the target population in ways that would affect the generalizability of the results?
Exposure ascertainment	What are the sources of information about drug exposure (e.g., maternal interviews, physician reports, pharmacy records)? Are these sources likely to provide valid information about how women actually used the drug? Are the exposures ascertained in sufficient detail (e.g., dose, frequency, duration, timing during gestation) to accurately assess the drug's potential effects on the outcome(s) of interests?
Outcome ascertainment	What are the sources of information about pregnancy outcomes, infant and fetal health (e.g., maternal interviews, obstetricians' reports, pediatric records)? Are these sources likely to be knowledgeable about the occurrence of the outcomes being studied? Are outcomes among stillbirths, spontaneous abortions, and elective terminations included? Are the outcomes documented in sufficient detail?
Reference group(s)	What comparison group(s) does the registry utilize? Is an internal comparison group recruited? Are there potential differences between the exposed and comparison groups that could affect the validity of the findings?
Statistical power	Does the sample size provide sufficient statistical power to meet the objective(s)?
Biases	When did subjects enroll (i.e. gestational age at enrollment for exposed and reference groups)? What proportion of registry enrollees did not complete the study (lost to follow-up)? Do their characteristics differ from those who completed the study in ways that could affect the validity of the results? Are exposed subjects comparable to the reference group? Were the same

Area	Issues
	methods for data collection and outcome definition used in the exposed and reference groups? Are there other possible sources of bias in the results? Are these biases addressed sufficiently in the analyses or sensitivity analyses?
Results	Does the report provide estimates of the absolute risk of the adverse pregnancy outcomes being studied? Are the results generalizable to the broader population of pregnant women who will use the drug?
Conclusions	Does the report explore possible alternative explanations for the findings? Does the report review and compare findings from other studies that assess the drug's effects during pregnancy, or findings for other drugs used to treat the same condition(s)? Do the registry findings provide information that will be useful to health care providers and women in making clinical decisions about use of the medication and pregnancy management?

20. Summary and Conclusions

Well-designed and executed pregnancy registries are an efficient initial approach to assess the safety of biopharmaceuticals during pregnancy and can provide data that can be used by health care providers for the treatment and counseling of patients who are pregnant or wish to become pregnant. Although pregnancy registries are more appropriate to identify or rule out large increases in the risk for malformations than to identify more modest teratogenic risks, they are a valuable tool to establish safety boundaries around risk estimates as data accumulate. Pregnancy registries have some unique characteristics that make them distinct form other registries or types of surveillance. Critical methodological issues to consider in the design include the prospective enrollment of women before the pregnancy outcome is known, inclusion of a comparable reference group, thoughtful assessment of drug exposure, ascertainment of prenatal and postnatal diagnosis, and validation of outcomes. Chance and potential biases should be considered when interpreting results from any observational study. A surveillance program should consider a combination of different sources of data so that associations detected in one study can be replicated or refuted by others.[20]

References

1. Kennedy DL, Uhl K, Kweder SL. Pregnancy Exposure Registries. Drug Safety 2004;27:215-28.
2. FDA. Guidance for industry: establishing pregnancy exposure registries. http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm071639.pdf. In. August 2002 [on-line] ed; 2002.
3. Wyszynski D. Pregnancy Exposure Registries: Academic Opportunities and Industry Responsibility. Birth Defects Res Part A Clin Mol Teratol 2009;85:93-101.
4. US Food and Drug Administration. Office of Women's Health. List of pregnancy registries [online]. Available from URL: http://www.fda.gov/womens/registries/default.htm. In.
5. EMA. Guideline on the exposure to medicinal products during pregnancy: Need for post-authorisation data, (Committee for Medicinal Products for Human Use, 2005), http://www.ema.europa.eu/pdfs/human/phvwp/31366605en.pdf. In; 2005.
6. FDAAA. US Food and Drug Administration Amendments Act (FDAAA) of 2007, Pub. L. No. 75-711, 52 Stat. 1040 (1938) as amended. . 2007.
7. Merkatz R, Temple R, Sobel S, Feiden K, Kessler D. Women in clinical trials of new drugs. A change in Food and Drug Administration Policy. N Engl J Med 1993;329:292-6.
8. Howard T, Tassinari M, Feibus K, Mathis L. Monitoring for teratogenic signals: Pregnancy registries and surveillance methods. American Journal of medical Genetics Part C: Seminars in Medical Genetics 2011;157:209-14.
9. Finer L, Henshaw S. Disparities in Rates of Unintended Pregnancy In the United States, 1994 and 2001. Perspectives on Sexual and Reproductive Health 2006;38:90-6.
10. Wilson JG. Evaluation of human teratologic risk in animals. In: Lee DH, Hewson EW, Okun D, eds. Environment and birth defects. First ed. New York and London: Academic Press; 1973:146-60.
11. Holmes L. Human teratogens: update 2010. Birth Defects Res Part A Clin Mol Teratol 2011;91:1-7.
12. McAdams M, Staffa J, Dal Pan G. Estimating the extent of reporting to FDA: a case study of statin-associated rhabdomyolysis. Pharmacoepidemiol & Drug Safety 2008;17:229-39.
13. Wysowski D, Swartz L. Adverse drug event surveillance and drug withdrawals in the United States, 1969–2002: the importance of reporting suspected reactions. Arch Intern Med 2005;165:1362-9.
14. Mathis L, Iyasu S. Safety monitoring of drugs granted exclusivity under the Best Pharmaceuticals for Children Act: what the FDA has learned. Clin Pharmacol Ther 2007;82:133-4.
15. Werler M, Louik C, Mitchell A. Case-control studies for identifying novel teratogens. Am J Med Genet 2011;in press.
16. Schneeweiss S, Avorn J. A review of uses of health care utilization databases for epidemiologic research on therapeutics. J Clin Epidemiol 2005;58:323-37.
17. Charlton RA, Cunnington M, de Vries CS, Weil JG. Data Resources for Investigating Drug Exposure during Pregnancy and Associated Outcomes. The General Practice Research Database (GPRD) as an Alternative to Pregnancy Registries. Drug Safety 2008;31:39-51.
18. Andrade S, Davis R, Cheetham T, et al. Medication Exposure in Pregnancy Risk Evaluation Program. Matern Child Health 2011;DOI 10.1007/s10995-011-0902-x.
19. Honein MA, Paulozzi LJ, Cragan JD, Correa A. Evaluation of selected characteristics of pregnancy drug registries. Teratology 1999;60:356-64.
20. Mitchell AA. Systematic Identification of Drugs That Cause Birth Defects-- A New Opportunity. N Engl J Med 2003;349:2556-9.
21. ISPE. Guidelines for good pharmacoepidemiology practices (GPP). Pharmacoepidemiol & Drug Safety 2008;17:200-8.
22. deMoor C, Covington D, Golembesky A. Differing Strategies for Analyzing Multiple Gestation Pregnancies in Pregnancy Registries: Impact on Birth Defect Rates. Pharmacoepidemiol & Drug Safety 2010;19 (Supplement 1):S124.

23. Cragan JD, Gilboa SM. Including Prenatal Diagnoses in Birth Defects Monitoring: Experience of the Metropolitan Atlanta Congenital Defects Program. Birth Defects Res Part A Clin Mol Teratol 2009;85:20-9.
24. Reiff_Eldridge R, Heffner CR, Ephross SA, Tennis PS, White A, Andrews E. Monitoring pregnancy outcomes after prenatal drug exposure through prospective pregnancy registries: A pharmaceutical company commitment. Am J Obstet Gynecol 2000;182:159-63.
25. Tabacova S, Little R, Tsong Y, Vega A, Kimmel C. Adverse pregnancy outcomes associated with maternal enalapril antihypertensive treatment. Pharmacoepidemiol & Drug Safety 2003;12:633-46.
26. Chambers CD, Johnson KA, Dick LM, Felix RJ, Jones KL. Birth outcomes in pregnant women taking fluoxetine. N Engl J Med 1996;335:1010-5.
27. Chambers C, Hernández-Díaz S, Van Marter L, et al. Selective Serotonin-Reuptake Inhibitors and Risk of Persistent Pulmonary Hypertension of the Newborn. N Engl J Med 2006;354:579-87.
28. Briggs GG, Polifka J, Committee OR. Better Data Needed from Pregnancy Registries. Birth Defects Res Part A Clin Mol Teratol 2009;85:109-11.
29. Correa A, Cragan J, Kucik J, et al. Reporting birth defects surveillance data 1968-2003 Birth Defects Res Part A Clin Mol Teratol 2007;79:65-86.
30. Rothman KJ, Ray W. Should cases with a known cause of their disease be excluded from study? Pharmacoepidemiol Drug Safety 2002;11:11-4.
31. Dolk H, Loane M, Garne E. The prevalence of congenital anomalies in Europe. Adv Exp Med Biol 2010;686:349-64.
32. Chambers C, Braddock SR, Briggs GG, et al. Postmarketing surveillance for human teratogenicity: A model approach. Teratology 2001;64:252-61.
33. Holmes L, Wyszynski D, Lieberman E. The Antiepileptic Drug Pregnancy Registry: a six year experience. Arch Neurol 2004;61:673-8.
34. Covington D, Tilson H, Elder J, Doi BS P. Assessing teratogenicity of antiretroviral drugs: monitoring and analysis plan of the Antiretroviral Pregnancy Registry. Pharmacoepidemiol & Drug Safety 2004;13:537-45.
35. Hernán MA, Hernández-Díaz S, Werler MM, Mitchell AA. Causal knowledge as a prerequisite for confounding evaluation. An application to birth defects epidemiology. Am J Epidemiol 2002;155:176-84.
36. Hernán MA, Hernández-Díaz S, Robins JM. A structural approach to selection bias. Epidemiology 2004;15:615-25.
37. Hernández-Díaz S, Schisterman EF, Hernán MA. The birth weight "paradox" uncovered? Am J Epidemiol 2006;164:1115-20.
38. Editors. The registration of observational studies-when metaphors go bad. Epidemiology 2010;21:607-20.
39. Gliklich R, Dreyer N. Registries for Evaluating Patient Outcomes: A User's Guide. AHRQ Publication No.10-EHC049. 2nd ed. Rockville, MD: Agency for Healthcare Research and Quality; 2010.
40. Roberts S, Covington D, Doi P. Pregnancy Exposure Registries: Streamlining Regulatory Requirements to Increase Enrollment. Pharmacoepidemiol & Drug Safety 2005;14 (Supplement 2).
41. Mitchell A, Gilboa S, Werler M, et al. Medication Use during Pregnancy, with Particular Focus on Prescription Drugs: 1976-2008. Am J Obstet Gynecol 2011;205:e1-8.