Published in final edited form as:
Am J Med Genet C Semin Med Genet. 2011 August 15; 157(3): 201–208. doi:10.1002/ajmg.c.30307.

Case-control studies for identifying novel teratogens

Martha M. Werler, **Carol Louik**, and **Allen A. Mitchell**
Slone Epidemiology Center at Boston University

The case control study design offers an efficient approach to measuring an association between an exposure and an outcome, especially when the outcome is rare, as is true for specific birth defects. For example, instead of following 50,000 pregnant women to have sufficient statistical power to identify a doubling in risk of oral clefts associated with a common exposure (e.g., cigarette smoking), 175 cases and 2 controls per case could be studied with equal statistical power. Examples of case sources include hospital or clinical series, or birth defect registries. For validity, control subjects should represent the population base of the cases, which can be difficult to identify for non-population-based case groups. Case-control studies typically rely on retrospective exposure measurement, which presents a major challenge and sets up the possibility of recall bias. Approaches are discussed to keep sources of bias to a minimum, including recall, non-differential information, and selection biases. Case-control studies can play an important role in this process for both hypothesis-generation and hypothesis-testing of potential teratogens. Examples of case-control studies and their contributions to the field are presented.

The case control study design offers an efficient approach to measuring an association between an exposure and an outcome, especially when the outcome is rare like specific birth defects. The case-control approach involves identifying and enrolling persons with the outcome of interest (cases) and a comparison group (controls), and then comparing the prevalence of exposure (typically retrospectively measured) between cases and controls to produce an odds ratio as the measure of association. This is in contrast to the traditional prospective, follow-up study design where cohorts of exposed and unexposed individuals are identified, enrolled, and followed to measure occurrences of the outcome of interest; the proportions of that outcome are compared between exposed and unexposed study subjects to produce a relative risk.

To understand the efficiency of the case-control design for studying specific births, maternal smoking in pregnancy and cleft lip with or without cleft palate (CLP) serves as good example. Maternal smoking is a common exposure, with an approximate prevalence of 15%, and CLP is one of the more common specific birth defects, with a prevalence of approximately 0.7 per 1000 births. In a follow-up study that enrolled 50,000 pregnant women, 7,500 cigarette-exposed pregnancies would result in 5 exposed cases under an assumption of no association. If there was an observed doubling in risk, 10 exposed cases would be expected among the 7,500 exposed pregnancies, but the 95% confidence interval around that 2-fold relative risk would be 0.98 to 4.17. In a case-control study with 75 cases and 225 controls (3 per case), a 2-fold increased odds ratio would have a 95% confidence interval of (1.09-3.83). In other words, to get to the same end point, measuring the association between maternal cigarette smoking and risk of CLP, the case-control study would need a study sample that is less than 1% of the cohort study size. Most specific birth defects are less common than CLP; thus, a cohort study would produce even fewer cases, while the case-control approach would have comparatively greater efficiency. This efficiency explains why the case-control study is the most popular design for epidemiologic studies of risk factors for birth defects.

Design and conduct

Cases

The research question will dictate what the outcome of interest is -- what specific birth defect or set of birth defects will constitute the case group. Once identified, a case definition will guide the investigator toward potential sources of cases. Clinic populations, birth defect registries, vital records, hospital discharge diagnoses, insurance claims data, and prenatal diagnosis records can all provide cases. Completeness of any these sources depends on the specific birth defect or group of birth defects of interest. For example, the neural tube defects anencephaly and spina bifida are easily identified at birth and well documented in medical records. However, birth certificates and hospital discharge diagnoses are incomplete sources of neural tube defect cases because a large fraction of affected pregnancies are terminated following prenatal diagnosis (Peller and others, 2004). Birth certificates or hospital discharge diagnoses may be poor sources for birth defects that are easily missed, misdiagnosed, or do not come to diagnosis until later in infancy or early childhood. An example of the last situation is hemifacial microsomia, in which the asymmetrically-underdeveloped craniofacial structures can be diagnosed anytime from the third trimester by ultrasound to early childhood at a first dental visit. A preferable ascertainment source would be one where cases come to diagnosis, which for hemifacial microsomia is craniofacial specialists. In choosing an ascertainment source, the investigator should take into consideration when, where, and how cases come to diagnosis.

Clinic populations are often a readily available case source. Indeed, research questions may arise from a clinician's observation of seemingly high proportions of cases being exposed to a particular agent. Depending on the specific birth defect, clinic populations may include a selection of the full range of affected individuals. For example, a highly specialized clinic might draw patients from around the world, but it may over-represent the more assertive and financially capable families. A highly selected case population is not necessarily a problem, but it would be essential that the investigator be aware of the demographics and access-to-care characteristics of a clinic population when designing a case-control study. Reducing the likelihood of selection bias is discussed below.

Birth defect registries exist in nearly every state in the U.S. and can be excellent sources of cases. The methods employed to ascertain cases vary from registry to registry, with some relying on passive reporting via birth certificates or physicians, some employing rigorous active surveillance, and others using a combination of various methods (NBDPN, 2010). The latter type of registry offers more complete ascertainment of cases by conducting reviews of birth records and adhering to detailed inclusion criteria. When considering a registry as a source of birth defect cases, one needs to determine whether prenatally diagnosed cases are included and at what age newly diagnosed cases are included.

Controls

A valid control group is essential in a case-control study, but can also be a challenging aspect in this study design. The first step is to identify the study base, which typically depends on the source of cases. Controls must represent the same population that gave rise to the cases. An advantage of identifying cases from a registry is its geographic base; controls would come from that same geographic area.

Cases that are ascertained from hospitals or clinics or a mix of sources may have a less obvious study base. When cases are ascertained from birth hospitals, controls might be births at those same institutions. However, high-risk pregnancies such as those with a prenatally diagnosed problem are often referred to tertiary care birth hospitals and therefore arise from a different background population than those pregnancies that were originally

intended to be delivered at that hospital. What is the background population for the referred cases? It can't be easily defined in simple geographic terms; rather the study base would be pregnancies that would also have been referred if the same type of condition as that of the case had been prenatally diagnosed. For cases ascertained at tertiary care birth hospitals, controls could be matched to the case according to the intended birth hospital – i.e., where the mother had originally planned to deliver. Or controls could be ascertained from the primary care provider, in this case the obstetrician or midwife, under the assumption that the decision to refer a patient to the tertiary care birth hospital would be the same for controls had they had the same condition as that of the case. The study base for cases ascertained at specialty clinics in pediatric hospitals may be defined as the population of similarly aged children who would have gone to the same pediatric hospital if they had been similarly affected. Controls could be ascertained from the cases' primary care providers, again assuming referral decisions would be the same for each primary care provider's patients.

Exposure measurement

Information on exposures can come from a variety of sources. Optimally, sources would have detailed, specific, accurate and complete documentation of exposure on all study subjects. It is particularly important that detailed Information on timing of exposure should be captured because the developmentally-relevant period for birth defects is typically a few weeks in early pregnancy, when many changes in behavior and exposure occur. Complete and accurate measurement of exposure would eliminate misclassification, though this is rarely achieved.

Measurement of ‘environmental’ exposures – meaning non-genetic factors -presents many challenges. Documentation of exposures that precede the diagnosis of a birth defect is preferred because of the possibility of reporting or diagnostic biases. Obstetric records, pharmacy records, and environmental databases are examples of such documentation, but these sources often suffer from lack of detail on timing, dose, and potential confounders. For example, an obstetric record might state that a patient was prescribed an anti-emetic to relieve severe nausea and vomiting, but it is not known whether the woman actually filled it or took the medication, or her timing or frequency of use. Biologic samples collected after birth for measurement of exposures are not subject to reporting or diagnostic biases, but may be poor indicators of exposure during the developmentally relevant time frame. Often, the most accessible source of information with sufficient detail on timing, frequency and dose of exposure is the mother herself. In fact, mother is the only source for some details. Therefore, most large-scale case-control studies of birth defects rely on mothers as the primary source of exposure information.

Evidence suggests that accuracy depends on several factors, such as the type of exposure. For example, recall of parity, age at menarche, cigarette smoking, or history of gallbladder disease has been shown to be relatively high (Hensley Alford and others, 2009; Must and others, 2002; Paganini-Hill and Ross, 1982; Sanderson and others, 1998; Tomeo and others, 1999; Yawn and others, 1998). Recall of transient, repetitive, or casual exposures is likely to be poorer (Tourangeau and others, 2000). The social or public health stigma attached to an exposure might also be important. A study of alcohol, cocaine, and marijuana use in pregnancy found that women under-reported exposures at the antenatal interview compared to the post-partum interview. The amount of under-reporting was 44% for alcohol, 57% for marijuana, and 70% for cocaine, suggesting that amount of denial varies according to its negative perception in society (Jacobson and others, 1991). Social stigma can change over time, depending on what is in the news during the period preceding data collection. Also, a longer period of time between data collection and the event that is being recalled has been shown to negatively affect accuracy of reporting (Lewis and others, 2006), as intuition would suggest. Finally, data collection methods can affect quality of responses. Open-ended

questions such as “Did you take any medications in pregnancy?” are more likely to elicit whatever is foremost on the mother's mind. More detailed and specific questions increases recall appreciably (Mitchell and others, 1986). If these factors are balanced between cases and controls, their impact is considered *random misclassification* of exposure, which tends (but not always) to result in biasing risk estimates toward no effect (Jurek and others, 2005). On the other hand, accuracy of reporting that is dependent on case–control status, can introduce bias - *recall bias* – in either direction, and deserves further consideration.

Recall Bias

Reports of positive associations between a risk factor and birth defects seem to inevitably be greeted with suspicion of recall bias. Such concerns stem, in part, from the experiences of clinicians caring for children with birth defects because mothers so often ask whether a particular event or exposure in their pregnancy was the cause. Intuitively, it makes sense that mothers whose pregnancies were affected with birth defects would search for an explanation by reviewing the course of their pregnancy, while mothers of ‘normal’ controls would be more focused on the present and future. Traditionally, recall bias is considered a possible explanation for a spurious increased risk estimate. However, differential recall could operate in the opposite direction, where mothers of cases deny exposure, resulting in a downward bias of the risk estimate.

Empirical evidence in support of this type of recall bias is scant, partially due to the difficulty of measuring it in the setting of birth defect case-control studies. A gold standard is necessary to validate retrospective reports. Data in vital records, medical records, biologic markers, and prospective studies have been used as gold standards, but each source carries its own limitations. Exposures documented in vital records, primarily birth certificates, may be less accurate than mother's report and, like maternal report, may be vulnerable to bias because the outcome of pregnancy is already known. Many exposures are not documented in medical records and even those that tend to be, such as illnesses and treatments, are not recorded with specific details on timing, severity, or dose. Also, when a notation of illness or treatment is missing, it could represent either no occurrence or unknown information. Biologic markers of exposure that are collected close to the first trimester can be an excellent gold standard when genetic variation in metabolism isn't a factor. For example, serum folate levels are a function of both folate intake and genotype of many different enzymes. A one-time biologic sample may reflect exposure status at a point in time that is etiologically irrelevant.

A superior gold standard would be prospectively collected data, but cohorts with both prospective and retrospective data collection have not been large enough to allow robust comparisons between mothers of children with birth defects and mothers of healthy children. Hence, validation studies of maternal retrospective reports have compared mothers of healthy children to those with a variety of adverse reproductive outcomes, such as prematurity, intra-uterine growth retardation, neonatal intensive care admittance, sudden infant death syndrome, miscarriages, stillbirths, and neonatal deaths (Drews and others, 1990; Klemetti and Saxen, 1967; Mackenzie and Lippman, 1989). In terms of searching for a causal exposure, the mindset of mothers whose children have a birth defect may well be different than mothers with these other experiences. Nevertheless, an upward bias of relative risk estimates was not observed for post-partum reports of most exposures (Drews and others, 1990; Mackenzie and Lippman, 1989). It is also worth noting that repetition of an interview could improve at the second time point, resulting in underestimation of true bias.

A comparison of two studies of cardiovascular birth defects in relation to the use of an anti-emetic medication (Bendectin) provides indirect evidence of recall bias (Rothman and others, 1979; Zierler and Rothman, 1985). The first study identified an increased risk based

on data derived from retrospective questionnaires in which mothers were asked a general question about drug use in pregnancy. The second study also employed a case-control design, but asked mothers standardized questions specifically about Bendectin medication use and no association was observed. The authors concluded that the results of the first study were likely due to recall bias. Thus, to reduce information bias -both *random misclassification* and *recall bias* - a standardized and detailed questionnaire should be employed. Since recall accuracy likely decreases as the length of the recall interval increases, retrospective data should be collected as soon as possible after the pregnancy to also reduce the likelihood of information bias.

Another approach to minimize the possibility of recall bias is the use of a control group comprising mothers of children with malformations other than those of the case group. In this setting, it is assumed that reporting accuracy would be similar for case and control mothers and that there is no association between the exposure of interest and the birth defects included in the control group. If the latter is not true, selection bias would be present. However, the investigator cannot be certain that such an association does not exist. A control group comprised of a wide variety of different specific malformations would dilute the impact of any unidentified associations with some specific defects on the overall control group. In support of this approach is that most teratogens are not linked to all types of malformations. On the contrary, obesity, diabetes, and heavy alcohol consumption are examples of maternal exposures that appear to affect many different developing organs and tissues in the fetus. Studies that have utilized both malformed and non-malformed control groups have shown remarkably similar prevalences of multivitamin supplementation, obesity, use of decongestants and use of analgesics (Werler and others, 1999; Werler and others, 1996; Werler and others, 2002), providing further indirect evidence against recall bias.

Measuring Associations

Case-control studies produce odds ratio estimates of the association between exposure and outcome rather than relative risks or rate ratios. Because cases and controls are identified without necessarily knowing the number of pregnancies in the background population, risks or rates of birth defects cannot be measured, which are necessary to calculate relative risks or rate ratios, respectively. Instead, the case control study measures the prevalence of exposure among cases and among controls. The odds ratio is therefore the ratio of the odds of exposure among cases to that among controls. The odds ratio is a measure of association in its own right, but its interpretation is less clear than the relative risk. Mathematically, odds ratios are good estimators of rate ratios and relative risks when the case outcome is rare, meaning occurring in than 15% of the population. Thus, odds ratios generated from case control studies of birth defects approximate relative risks, leading to interpretations that are more easily received by clinicians, patients, and the public. For example, a case control study that identified a 26% prevalence of cigarette smoke exposure among CLP cases and a corresponding 15% prevalence among controls, produces an odds ratio of 2.0 that is strictly interpreted as follows: the odds of a case being exposed to cigarette smoke is twice as high as that of controls. Because CLP is rare, we can say that the risk of having a baby born with CLP is twice as high among smokers compared to non-smokers.

In epidemiology measuring rates of outcomes takes person-time into account. In birth defects epidemiology we don't typically consider rates because we don't know the birth defect status of early pregnancy losses and therefore we can't measure time from exposure to onset of maldevelopment. If we are willing to accept that the outcome of interest is the risk of a birth defect in pregnancies that are 20 or more weeks gestation, i.e., that spontaneous abortion is a different outcome regardless of whether the conceptus was malformed or not, then time between exposure on onset of disease becomes irrelevant. Hence, risks, rather than

rates, are what we measure in prospective birth defect studies and relative risks are what we approximate with the odds ratio.

Identification of novel teratogens

The efficiency of case-control studies allows data collection to occur shortly after new exposures appear in the population, rendering them a valuable tool for uncovering teratogenic agents. However, a positive association from a single study should not be interpreted as causal. Unlike experimental studies, it is not possible to control for all potentially confounding factors or other sources of bias in case-control studies. Thus, confirmation of positive associations from additional studies is essential, and it is helpful if there is also evidence of biologic plausibility and coherence with related results from other scientific arenas. Case-control studies are also useful for estimating relative safety of a specific exposure and birth defect risks, as noted below.

The National Birth Defects Prevention Study (NBDPS) ascertains cases with selected structural malformations from population-based birth defect registries in nine states (Arkansas, California, Iowa, Georgia, Massachusetts, New York, North Carolina, Texas, and Utah) (Yoon and others, 2001). Control subjects are births without known major malformations from the same study bases that give rise to the cases. Clinical geneticists review the medical records of cases and classify defects for inclusion/exclusion and according to primary defect and the presence of associated malformations (Botto and others, 2007; Rasmussen and others, 2003). Interviews of mothers of cases and controls are conducted within two years after delivery and questions are asked on a wide range of exposures including illnesses, medications, cigarette smoking, alcohol, caffeine, and dietary intakes, and occupation. The population base of NBDPS makes it especially amenable to linkages with environmental contaminant databases. Buccal cell samples are also collected from cases and their parents (Rasmussen and others, 2002).

The NBDPS is an enormous resource for studies of risk factors for birth defects, having contributed both new and confirmatory findings to the literature. An example of a confirmatory finding is maternal obesity in relation to omphalocele in offspring; NBDPS observed an 1.6-fold increased risk for women with a pre-pregnancy body mass index ≥30 kg/m^2 (Waller and others, 2007) following two similar reports from other epidemiologic studies (Waller and others, 1994; Watkins and others, 2003). Opioid use in early pregnancy had previously been linked to cardiac malformations, but NBDPS was the first study to report greater than 3-fold increased risks of hypoplastic left heart syndrome in relation to two specific opioids - codeine and hydrocodone (Broussard and others, 2011). This new finding deserves further attention because the outcome, hypoplastic left heart syndrome, is well-defined in NBDPS and hydrocodone and codeine exposures were shown in NBDPS to be prevalent in approximately 1.5% of pregnancies. In addition, a possible mechanism was identified, based on evidence that an opioid-sensitive growth factor is expressed in developing heart tissue in rat embryos (Zagon and others, 1999).

The Slone Epidemiology Center Birth Defects Study is a long-standing, rigorous, and flexible study of risk factors for birth defects (Mitchell, in press; Mitchell and others, 1981; Werler and others, 1996). It began in the mid-1970s as a hospital-based study in greater Boston, Philadelphia and Toronto with only malformed subjects; today it enrolls both infants with a wide range of major malformations and infants without malformations. Both groups are recruited from hospitals and/or registries in Massachusetts, Rhode Island, greater Philadelphia, parts of New York State, and San Diego County. Interviews are conducted by study nurses within six months after delivery and questions are asked about demographic, reproductive, and medical factors, with a particular emphasis on medication use. BDS is especially well-suited to respond to new research questions by quickly modifying data

collection, whether it be adding a new case group, specific questions on a newly-marketed drug, or a new tool for improving reporting accuracy. Collection of buccal cell samples from babies, mothers, and fathers began in 1993 and continued until 2010; the biobank of samples from over 9500 families is an available resource for case-control studies of genetic risk factors (Hernandez-Diaz 2005).

BDS data have contributed to literature on birth defect risks in relation to numerous medications, beginning with a report on the safety of Bendectin use in relation to oral clefts and cardiac defects the early 1980s (Mitchell and others, 1981) and most recently reporting on patterns of asthma medication use in pregnancy (Louik and others, 2010). Following a report of birth outcomes in a cohort of women exposed to fluoxetine in which 2 cases of persistent pulmonary hypertension of the newborn (PPHN) were observed (Chambers and others, 1996), BDS established a collaboration with the original investigator and confirmed a positive association between the broader group of selective serotonin reuptake inhibitors and PPHN (Chambers and others, 2006). An example of a new finding from BDS is 2.5-3.9-fold increased risks of male genital malformations in association with maternal use of medications that contain phthalates (known endocrine disruptors) in the first trimester (Hernandez-Diaz and others, 2010). BDS protocols were modified to address each of these examples: For the former study, PPHN was added as a priority defect that required specific diagnostic data; for the latter study, undescended testes and primary hypospadias were added as priority defects and questions on medications were expanded to include dosage form to allow determination of phthalate content. and for the latter study, undescended testes and primary hypospadias were added as priority defects and questions were added on medications to allow determination of phthalate content.

Both NBDPS and BDS are large-scale, on-going enterprises that cover most major structural malformations. The case-control design is also ideally suited for smaller-scale studies of specific birth defects. When just one or two birth defects are the outcome of interest, methods can be tailored to maximize ascertainment and data collection efficiencies. For example, BDS data signaled a possible increased risk of gastroschisis in relation to maternal use of the decongestant pseudoephedrine (Werler and others, 1992). Because pseudoephedrine is vasoconstrictive, its use is common in pregnancy, and gastroschisis might result from vascular disruption, further study was warranted. A new case-control study was mounted that ascertained over 200 cases of gastroschisis from 15 pediatric surgeons in less than 4 years and collected detailed information on over-the-counter medications and illnesses for which decongestants are taken (Werler and others, 2002).

Because medications used in pregnancy can only be assessed for risks and safety in the post-marketing setting, this information cannot become available until some time after their approval for marketing. One formalized approach to systematically provide such assessments has been developed by Slone Epidemiology Center Birth Defect Study investigators in collaboration with investigators at the University of California San Diego, under the coordination of the American Academy of Allergy Asthma and Immunology. The program includes data collection from both case-control surveillance within the Slone Birth Defects Study and prospective registry surveillance within the Organization of Teratology Information Specialists Research Center. At present, the program is focused on surveillance of pregnancy outcomes among women who receive flu or other vaccines, take anti-viral medications for the prevention or treatment of flu, or take asthma medications during pregnancy. Details on these exposures are collected, including the type, timing, and frequency. For vaccines, the facility where it was administered is also obtained to allow for collection of additional details. The program is designed to easily expand to include other types of medication or vaccine exposures. Further, a standing independent advisory committee routinely examines the accumulating data in relation to a wide range of birth

outcomes to evaluate risks and relative safety of exposures in pregnancy (AAAAI, 2011). Although the safety of any exposure can never be considered absolute, the program investigators developed novel definitions of “relative safety”: a finding of no association with an upper 95% confidence bound of 4 or less might be termed “no evidence of risk” and a null finding with an upper bound of 2 or less might be termed “evidence of relative safety” (Schatz and others, 2011).

Another approach to routinely evaluate potential risks of medications in relation to birth defects is employed by the NBDPS. That study generates annual screens of the interview data in which all medication components (active ingredients) and products are compared to all specific birth defects and birth groups. These comparisons are in the form of odds ratios and p-values for each medication exposure in the periconceptional period (any use 1 month before through 3 months after conception). Numbers of exposed cases and controls and total numbers of case and controls groups are also included in the screen to help interpretation. Even after limiting these comparisons to those with at least five exposed cases, the screens contain over 16,000 medication–defect comparisons. A group of 10 reviewers comprising clinical geneticists and epidemiologists, is responsible for initially assessing the screen findings by taking into account the magnitude of the odds ratio, the number of exposed cases and controls, underlying pharmacology or embryology, drug indication, and patterns with other exposure-defect findings. Findings are categorized according to what action should be taken: 1) ignore due to no evidence of concern; 2) wait and watch future screens to see if association remains; 3) notify NBDPS investigator who is already conducting research on that specific medication – birth defect combination; or 4) recommend that a formal analysis be conducted, in which confounding factors, varying exposure windows, and case subgroups can be examined. This approach is targeted to identify potential risks associated with medications, but, as the NBDPS data grow in numbers of interviewed study subjects, can be modified to assess relative safety as well.

In summary, case-control studies are an efficient means for identifying novel teratogens because the number of study subjects is a small fraction of that required in a follow-up study. Case-control studies, however, are especially vulnerable to exposure information bias; extra effort is essential to reduce the potential for such bias. Regardless of study design, confirmation of positive associations is necessary from additional studies to guide interpretation. Case-control studies can play an important role in this process for both hypothesis-generation and hypothesis-testing.

References

AAAAI. American Academy of Allergy Asthma & Immunology: The Vaccines and Medications in Pregnancy Surveillance System (VAMPSS). 2011. http://www.aaaai.org/members/resources/vampss/program_information

Botto LD, Lin AE, Riehle-Colarusso T, Malik S, Correa A. Seeking causes: Classifying and evaluating congenital heart defects in etiologic studies. Birth Defects Res A Clin Mol Teratol. 2007; 79(10): 714–727. [PubMed: 17729292]

Broussard CS, Rasmussen SA, Reefhuis J, Friedman JM, Jann MW, Riehle-Colarusso T, Honein MA. Maternal treatment with opioid analgesics and risk for birth defects. Am J Obstet Gynecol. 2011

Chambers CD, Hernandez-Diaz S, Van Marter LJ, Werler MM, Louik C, Jones KL, Mitchell AA. Selective serotonin-reuptake inhibitors and risk of persistent pulmonary hypertension of the newborn. N Engl J Med. 2006; 354(6):579–587. [PubMed: 16467545]

Chambers CD, Johnson KA, Dick LM, Felix RJ, Jones KL. Birth outcomes in pregnant women taking fluoxetine. N Engl J Med. 1996; 335(14):1010–1015. [PubMed: 8793924]

Drews CD, Kraus JF, Greenland S. Recall bias in a case-control study of sudden infant death syndrome. Int J Epidemiol. 1990; 19(2):405–411. [PubMed: 2376455]

Hensley Alford SM, Lappin RE, Peterson L, Johnson CC. Pregnancy associated smoking behavior and six year postpartum recall. Matern Child Health J. 2009; 13(6):865–872. [PubMed: 18818994]

Hernandez-Diaz S, Hauser R, Mitchell AA. Phthalates in drugs and male genital malformations. Pharmacoepidemiol Drug Saf. 2010; 19:S221.

Jacobson SW, Jacobson JL, Sokol RJ, Martier SS, Ager JW, Kaplan MG. Maternal recall of alcohol, cocaine, and marijuana use during pregnancy. Neurotoxicol Teratol. 1991; 13(5):535–540. [PubMed: 1758408]

Jurek AM, Greenland S, Maldonado G, Church TR. Proper interpretation of non-differential misclassification effects: expectations vs observations. Int J Epidemiol. 2005; 34(3):680–687. [PubMed: 15802377]

Klemetti A, Saxen L. Prospective versus retrospective approach in the search for environmental causes of malformations. Am J Public Health Nations Health. 1967; 57(12):2071–2075. [PubMed: 6070247]

Lewis JD, Strom BL, Kimmel SE, Farrar J, Metz DC, Brensinger C, Nessel L, Localio AR. Predictors of recall of over-the-counter and prescription non-steroidal anti-inflammatory drug exposure. Pharmacoepidemiol Drug Saf. 2006; 15(1):39–45. [PubMed: 16136614]

Louik C, Schatz M, Hernandez-Diaz S, Werler MM, Mitchell AA. Asthma in pregnancy and its pharmacologic treatment. Ann Allergy Asthma Immunol. 2010; 105(2):110–117. [PubMed: 20674820]

Mackenzie SG, Lippman A. An investigation of report bias in a case-control study of pregnancy outcome. Am J Epidemiol. 1989; 129(1):65–75. [PubMed: 2910073]

Mitchell, AA. Studies of drug-induced birth defects. In: Strom, B., editor. Pharmacoepidemiology. 5th. Chichester, England: Wiley and Sons; in press

Mitchell AA, Cottler LB, Shapiro S. Effect of questionnaire design on recall of drug exposure in pregnancy. Am J Epidemiol. 1986; 123(4):670–676. [PubMed: 3953545]

Mitchell AA, Rosenberg L, Shapiro S, Slone D. Birth defects related to bendectin use in pregnancy. I. Oral clefts and cardiac defects. JAMA. 1981; 245(22):2311–2314. [PubMed: 7230459]

Must A, Phillips SM, Naumova EN, Blum M, Harris S, Dawson-Hughes B, Rand WM. Recall of early menstrual history and menarcheal body size: after 30 years, how well do women remember? Am J Epidemiol. 2002; 155(7):672–679. [PubMed: 11914195]

NBDPN. Selected Birth Defects Data from Population-based Birth Defects Surveillance Programs in the United States, 2003-2007. Birth Defects Res A Clin Mol Teratol. 2010; 88(12):1062–1123. [PubMed: 21157887]

Paganini-Hill A, Ross RK. Reliability of recall of drug usage and other health-related information. Am J Epidemiol. 1982; 116(1):114–122. [PubMed: 7102647]

Peller AJ, Westgate MN, Holmes LB. Trends in congenital malformations, 1974-1999: effect of prenatal diagnosis and elective termination. Obstet Gynecol. 2004; 104(5 Pt 1):957–964. [PubMed: 15516385]

Rasmussen SA, Lammer EJ, Shaw GM, Finnell RH, McGehee RE Jr, Gallagher M, Romitti PA, Murray JC. Integration of DNA sample collection into a multi-site birth defects case-control study. Teratology. 2002; 66(4):177–184. [PubMed: 12353214]

Rasmussen SA, Olney RS, Holmes LB, Lin AE, Keppler-Noreuil KM, Moore CA. Guidelines for case classification for the National Birth Defects Prevention Study. Birth Defects Res A Clin Mol Teratol. 2003; 67(3):193–201. [PubMed: 12797461]

Rothman KJ, Fyler DC, Goldblatt A, Kreidberg MB. Exogenous hormones and other drug exposures of children with congenital heart disease. Am J Epidemiol. 1979; 109(4):433–439. [PubMed: 443241]

Sanderson M, Williams MA, White E, Daling JR, Holt VL, Malone KE, Self SG, Moore DE. Validity and reliability of subject and mother reporting of perinatal factors. Am J Epidemiol. 1998; 147(2): 136–140. [PubMed: 9457002]

Schatz M, Chambers CD, Jones KL, Louik C, Mitchell AA. Safety of influenza immunizations and treatment during pregnancy: the Vaccines and Medications in Pregnancy Surveillance System. Am J Obstet Gynecol. 2011

Tomeo CA, Rich-Edwards JW, Michels KB, Berkey CS, Hunter DJ, Frazier AL, Willett WC, Buka SL. Reproducibility and validity of maternal recall of pregnancy-related events. Epidemiology. 1999; 10(6):774–777. [PubMed: 10535796]

Tourangeau, R.; Rips, LJ.; Rasinski, K. The Psychology of Survey Response. Cambridge, U.K.: Cambridge University Press; 2000. The Role of Memory in Survey Responding; p. 62-99.

Waller DK, Mills JL, Simpson JL, Cunningham GC, Conley MR, Lassman MR, Rhoads GG. Are obese women at higher risk for producing malformed offspring? Am J Obstet Gynecol. 1994; 170(2):541–548. [PubMed: 8116710]

Waller DK, Shaw GM, Rasmussen SA, Hobbs CA, Canfield MA, Siega-Riz AM, Gallaway MS, Correa A. Prepregnancy obesity as a risk factor for structural birth defects. Arch Pediatr Adolesc Med. 2007; 161(8):745–750. [PubMed: 17679655]

Watkins ML, Rasmussen SA, Honein MA, Botto LD, Moore CA. Maternal obesity and risk for birth defects. Pediatrics. 2003; 111(5 Part 2):1152–1158. [PubMed: 12728129]

Werler MM, Hayes C, Louik C, Shapiro S, Mitchell AA. Multivitamin supplementation and risk of birth defects. Am J Epidemiol. 1999; 150(7):675–682. [PubMed: 10512421]

Werler MM, Louik C, Shapiro S, Mitchell AA. Prepregnant weight in relation to risk of neural tube defects. Jama. 1996; 275(14):1089–1092. [PubMed: 8601927]

Werler MM, Mitchell AA, Shapiro S. First trimester maternal medication use in relation to gastroschisis. Teratology. 1992; 45(4):361–367. [PubMed: 1533958]

Werler MM, Sheehan JE, Mitchell AA. Maternal medication use and risks of gastroschisis and small intestinal atresia. Am J Epidemiol. 2002; 155(1):26–31. [PubMed: 11772781]

Yawn BP, Suman VJ, Jacobsen SJ. Maternal recall of distant pregnancy events. J Clin Epidemiol. 1998; 51(5):399–405. [PubMed: 9619967]

Yoon PW, Rasmussen SA, Lynberg MC, Moore CA, Anderka M, Carmichael SL, Costa P, Druschel C, Hobbs CA, Romitti PA, Langlois PH, Edmonds LD. The National Birth Defects Prevention Study. Public Health Rep. 2001; 116(1):32–40. [PubMed: 11889273]

Zagon IS, Wu Y, McLaughlin PJ. Opioid growth factor and organ development in rat and human embryos. Brain Res. 1999; 839(2):313–322. [PubMed: 10519055]

Zierler S, Rothman KJ. Congenital heart disease in relation to maternal use of Bendectin and other drugs in early pregnancy. N Engl J Med. 1985; 313(6):347–352. [PubMed: 4010751]