Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

Cassandra M. Gibbs *April 18, 2011*

Date

The Impact of Time to Maternal Interview on the Odds of a Negative Response in the National Birth Defects Prevention Study

By

Cassandra M. Gibbs

Master of Public Health

Epidemiology

Matthew J. Strickland, Ph.D.

Committee Chair

Sarah C. Tinker, Ph.D.

Committee Member

The Impact of Time to Maternal Interview on the Odds of a Negative Response in the National Birth Defects Prevention Study

By

Cassandra M. Gibbs

B.S., Duke University, 2009

Thesis Committee Chair: Matthew J. Strickland, Ph.D.

An abstract of

a thesis submitted to the Faculty of the

Rollins School of Public Health of Emory University

in partial fulfillment of the requirements for the degree of

Master of Public Health

in Epidemiology

2011

Abstract

The Impact of Time to Maternal Interview on the Odds of a Negative Response in the National Birth Defects Prevention Study

By Cassandra M. Gibbs

Objective: To assess whether the time to interview (TTI), defined as the number of months from an infant's expected date of delivery (EDD) to a maternal interview, is associated with the odds of a negative response when mothers are asked about specific exposures during pregnancy, and if so, whether these associations vary by case/control status or interview quality.

Methods: The National Birth Defects Prevention Study is an ongoing, population-based case-control study conducted in ten states. Mothers are interviewed between six weeks and two years after the infant's EDD about exposures encountered before and during pregnancy.

We stratified TTI into four categories: 2-6 months (reference group), 7-12 months, 13-18 months, and 19-24 months. We examined the following reported pregnancy experiences as dichotomous outcomes: upper respiratory infection; kidney, bladder, or urinary tract infection; morning sickness/nausea; folic acid-containing vitamin use during the periconceptional period; and assisted fertility. Covariates, selected a priori, were case status, study center, maternal education, annual family income, year of birth, parity, gestational age at delivery, birth outcome, interview quality, and language of interview.

Crude and adjusted odds ratios and 95% confidence intervals were estimated using logistic regression. Interaction was assessed for case/control status and, separately, interview quality.

Results: Overall, the adjusted odds of a negative response increased as TTI increased. For each interview item, the odds of a negative response were greatest for mothers interviewed 19-24 months after their infant's EDD. Interaction of TTI with case/control status was observed for negative reporting of upper respiratory infection, morning sickness, and folic acid-containing vitamin use. Adjusted odds of a negative response tended to be higher in controls than in cases, and the odds of a negative response increased with TTI in both cases and controls. There was no significant interaction by interview quality.

Conclusion: Results from our analysis suggest that TTI should be considered in case-control studies of infant outcome that enroll mothers at varying times after delivery. A sensitivity analysis can be a good method of assessing whether a study's conclusions might change based upon differences in reporting attributable to longer TTIs.

The Impact of Time to Maternal Interview on the Odds of a Negative Response in the National Birth Defects Prevention Study

By

Cassandra M. Gibbs

B.S., Duke University, 2009

Thesis Committee Chair: Matthew J. Strickland, Ph.D.

A thesis submitted to the Faculty of the

Rollins School of Public Health of Emory University

in partial fulfillment of the requirements for the degree of

Master of Public Health

in Epidemiology

2011

Acknowledgments

I would like to thank Sarah C. Tinker for her invaluable guidance and mentorship throughout the completion of this project.

I would also like to thank Matthew J. Strickland for his advice and assistance with manuscript preparation.

Finally, I would like to thank my family for their unending support and encouragement.

Table of Contents

Background/Literature Review

Maternal exposures during pregnancy have the potential to alter fetal development and pregnancy outcome, and it is often difficult to predict the favorable or adverse impacts that may result from a particular exposure. Epidemiologic research has yielded important information to promote healthier pregnancies. For instance, many researchers have found that smoking may increase the risk of preterm birth (1-3), caffeine may increase the risk of miscarriage (4, 5), and illicit drugs adversely impact fetal growth (6). However, there is still a lack of understanding of the causes of many adverse pregnancy outcomes, and research in this area is ongoing.

The method used to investigate pregnancy outcomes will depend on the particular outcome of interest. When researchers are examining an outcome that is relatively common, such as low birth weight or preterm birth, it may be possible to use prospective or retrospective cohorts, in which individuals are selected based on their exposure status and assessed for their subsequent outcome. Case-control studies are more often used for studies of rare outcomes, because cohort studies are not cost-effective when only a small number of individuals will eventually develop the outcome of interest. Because individual birth defects are rare, studies of these outcomes often use the case-control design.

Assessing maternal exposures during pregnancy is a significant challenge. In a prospective study, it may be difficult to establish an appropriate cohort. Some researchers may choose to follow women who are trying to get pregnant; however, this cohort would not include unplanned pregnancies, which account for approximately 50% of pregnancies in the U.S. (7). It is also difficult to establish the start of pregnancy. The date of conception can be estimated from the first day of the woman's last menstrual period (LMP), which is often self-reported. However, even with accurate self-report, bleeding during early pregnancy may be mistaken for a menstrual period; conversely, an early pregnancy loss may be mistaken for a menstrual period. Ultrasound can be used to validate the LMP estimate, but only for those pregnancies that survive to the point

of ultrasound. Enrolling women at their first prenatal visit could introduce selection bias based on maintenance of the pregnancy to first visit and access to care.

Self-reported information is subject to reporting errors and biases. Reviewing medical records may provide more reliable data than maternal self-reports, but there are concerns with this method of exposure assessment as well. Privacy regulations may hinder access to confidential patient data, and physicians may be uncooperative with the records they are willing to provide to researchers. Medical record abstraction is usually a costly and time-consuming process, and medical records often lack data on exposures of interest, such as over-the-counter drug use during pregnancy, pollutant exposures, or stress. Records may also lack information on previous pregnancy outcomes that were not reported, such as early pregnancy losses. Much of the information in the medical record, such as an individual's family medical history and cigarette and alcohol use, if present at all, is based on self-report.

Due to the difficulties of collecting information on exposures during pregnancy through prospective data collection or through medical record review, retrospective maternal interviews are often used to assess prenatal exposures in epidemiologic studies, particularly those of case-control design. Retrospective interviews are subject to reporting bias, which may be non-differential, in which the likelihood of misreport is equivalent for both cases and controls, or differential, if the accuracy of reports differs for cases and controls. Typically non-differential bias will bias results toward the null; differential bias may bias results away from or toward the null, depending on the exact nature of the reporting bias. Mothers of infants with birth defects or other disabilities may be more likely to recall exposures during their pregnancy than mothers of controls, which could lead estimates of association to be biased away from the null. Conversely, mothers of infants with a malformation may be less likely to report less socially-desirable exposures, such as drinking or smoking during pregnancy, which could lead estimates of association to be underestimated.

Given these concerns, it is important to understand the factors affecting the validity of maternal reports of exposures during pregnancy. Many researchers have found that maternal recall of *pregnancy outcomes and complications* is relatively accurate (8-14). However, in a study by Axelsson and Rylander, only 50% of the infant birth defects reported in retrospective maternal questionnaires could be verified using hospital records or central medical registers (15). Similarly, Casey et al. found poor agreement among many of the variables they examined, such as malpresentation, fetal distress, and sepsis (16).

Less research has been done on the accuracy of reporting *exposures* that occur during pregnancy. Results from some studies suggest that maternal recall of certain exposures during pregnancy is relatively accurate and that any misclassification would not alter the conclusions of a particular study (10, 17-21). However, there are many examples in the published literature in which misreporting of exposures during pregnancy was substantial and would have resulted in different study conclusions. For example, Casey et al. recruited mothers at their infant's first or second well-child visit and found that none of the ten women who had a sexually transmitted infection documented in their obstetric record reported it in a maternal interview (16).

Cartwright and Smith found that many mothers reported problems during pregnancy, such as having high blood pressure, that were not verifiable using hospital records (22). Similarly, Delgado-Rodriguez et al. found a low level of agreement between reports of maternal alcohol use documented prospectively in the obstetric record versus reports of alcohol use at the postpartum interview (kappa = 0.11 for mothers of cases and 0.03 for mothers of controls) (23). Similar results were found for anemia (kappa = 0.42 for cases and 0.21 for controls) and use of any medication during pregnancy (kappa = 0.27 among cases and 0.29 among controls); furthermore, the odds ratios for alcohol use and for the use of any drug changed direction upon comparing the postpartum reports to the obstetric records. Tilley et al. found that agreement between prenatal obstetric records and maternal questionnaires completed ten or more years after a child's birth was poor for recall of drug exposure and x-rays during pregnancy (20). Recall of

diethylstilbestrol (DES) exposure in this study was also poor. Among DES-exposed mothers who were identified via review of their prenatal records (as opposed to mothers of DES-exposed daughters who were walk-ins or referrals), 8% of the mothers said they did not take DES, and 29% of the mothers did not recall using it.

Bryant et al. compared maternal reports of short-term illnesses and medication use during the periconceptional period (up to four months after the LMP) to hospital records (24). The researchers interviewed 101 matched pairs of women, where half were interviewed prenatally (<25 weeks into pregnancy) and half were interviewed postnatally (while in the hospital after delivery). The authors found poor agreement between maternal reports and hospital records (all kappas <0.5), and most agreement was due to *failing* to report a prenatal event. Mitchell et al. found that up to 40% of women failed to report medications used during pregnancy unless the medications were queried by name (versus being asked an open-ended question or about use for selected indications) (25).

Although the potential for recall bias is a logical concern for studies of adverse birth outcomes, many studies examining this issue have observed minimal to nonexistent differential recall by exposure status (20, 21, 26, 27). However, for exposures such as urinary tract infection/yeast infection during pregnancy and use of birth control after conception, Werler et al. observed that the sensitivity of reporting among mothers of infants with malformations was significantly higher than the sensitivity among mothers of infants without malformations (28). When Delgado-Rodriguez et al. (23) compared obstetric records to postpartum maternal interviews, they observed better agreement among mothers of cases than among mothers of controls for maternal alcohol use and anemia (for alcohol use, kappa = 0.11 for mothers of cases and 0.03 for mothers of controls; for anemia, kappa = 0.42 for mothers of cases and 0.21 for mothers of controls).

In addition, while examining exposures associated with spontaneous abortions, Fenster et al. found that mothers of controls, as compared to mothers of cases, tended to omit consumption

of up to two glasses of tap or bottled water per day (29). However, the authors concluded that this level of differential reporting would not have significantly altered the conclusions of the study, and they observed no differential reporting of pregnancy history, employment, caffeine consumption, or cigarette smoking.

Many of the above studies used medical records to verify maternal self-reports. While the medical record may be a good source of validation for maternal exposures such as x-rays or prescribed medication, it may lack information on exposures such as smoking status. Another method for validating maternal reports of exposures during pregnancy is to compare reports given during pregnancy to those given after pregnancy. Maternal self-report during pregnancy is expected to be more accurate than maternal self-report after pregnancy, since the reporting would occur soon after the exposure did, and the pregnancy outcome would be unknown. These types of validation studies are not often conducted, in part due to the difficulty in measuring exposures prospectively during pregnancy, as previously described.

Klemetti and Saxen compared information from medical records with reports from two maternal interviews; the first interview was conducted during the fifth month of pregnancy, and the second was conducted up to a year after delivery (26). There was relatively good agreement *between medical records and the "prospective" interview*; when specific drugs were compared point by point, only 36 of 420 positive replies were inaccurate, and there were no discrepancies between maternal reports and the maternity welfare/hospital records of nonchronic diseases in early pregnancy. However, there was a low level of reliability *between the prospective and retrospective interviews* for positive report of medication consumption; there were only 118 identical replies, while there were 174 reports that were discrepant in time of consumption, 128 reports where the time was not mentioned or where the drug was mentioned only in the prospective study, and 98 reports that were mentioned only in the retrospective study. Out of 77 nonchronic diseases that were prospectively reported, there were only eight that were also

reported in the retrospective interview, while there were 26 additional diseases reported only in the retrospective interview.

Mackenzie et al. examined reports of 39 different prenatal exposures from interviews that were administered both before the 20th week of pregnancy and after delivery (18). The authors found good agreement among most variables (all kappas were >0.20), and the odds of reporting prenatal exposures did not systematically increase or decrease from the predelivery to postdelivery interview. However, compared to the prenatal interviews, mothers were more likely to omit illnesses and medication use in the postdelivery interview, relative to other exposures; and more likely to report nausea, poor nutrition, and coffee, wine, or liquor use during the month prior to the LMP, relative to other exposures. Pickett et al. compared prospectively reported and biologically validated smoking during pregnancy to retrospective reports more than ten years after pregnancy (19). They observed a high sensitivity and specificity of retrospective reports relative to prospective reports and biological measurements. However, among heavy smokers, the amount smoked was somewhat inaccurate in the retrospective report. Moreover, the retrospective assessment identified some smokers that had not previously been captured by the prospective interview or cotinine measurement, but most of these women reported smoking only in the first trimester or before learning of their pregnancy. These women may not have been captured in the prospective measures, since the prenatal visits during which these data were collected often occurred after the first trimester. Similarly, Tomeo et al. compared interview data collected during pregnancy to data collected more than 30 years later, and they found that maternal self-report of smoking had a sensitivity of 0.86 and a specificity of 0.94 (30). It should be noted that this method of comparison uses maternal self-reports in both cases, and there is a possibility that the prospective self-report may have been inaccurate.

Several researchers have also examined the impact of time to interview on the accuracy of maternal responses. For instances of nausea, vomiting, influenza, fever, and upper respiratory illness, Bryant et al. found that agreement between medical records and maternal interview was

higher among women who were interviewed <25 weeks into their pregnancy than among women interviewed in the hospital postdelivery. However, there were no significant differences between the groups regarding reported medication use. Klemetti and Saxen (26) interviewed a group of women about prenatal exposures both during the fifth month of pregnancy and postnatally (where the postpartum interview could be conducted up to a year after delivery). When responses were stratified by the quarter of the year in which the postnatal interview was conducted, there were no significant differences in reporting accuracy (although overall accuracy was poor in the postnatal interview when compared to the prenatal interview).

In addition, Tilley et al. interviewed women about pregnancy exposures ten or more years after delivery, and they also found no difference in reporting accuracy or missing information by time to interview (20). In a case-control study of infant leukemia, Olson et al. compared medical records to responses from maternal interviews conducted 0-8 years after delivery (27). They found that reliability of reproductive history, medical procedures, birthweight, gestational age, and postdelivery complications were only slightly better among mothers interviewed <4 years after delivery compared to mothers interviewed ≥4 years after delivery, although reliability of pregnancy complications such as anemia and toxemia did appear largely compromised with a longer time to interview. Similarly, O'Sullivan et al. examined maternal recall of infant birthweight six to fifteen years after birth and found that recall accuracy did not differ by the age of the child (11). In contrast, Seidman et al. (12) found that birthweight and gestational age were reported more accurately for the most recent births, while Rasmussen et al. (31) observed that the sensitivity of birth defect reporting was higher when there was a *longer* interval between maternal interview and delivery.

In the absence of a true gold standard for measuring the accuracy of maternal exposures during pregnancy, researchers must find alternative methods to assess validity. For instance, researchers may search for patterns in the data, such as whether subjects with a positive disease status or a longer time to interview are more likely to respond to or omit certain items on a

questionnaire, or whether these subjects are more likely to claim that they were or were not exposed to a particular hazard. These data can be used to conduct a sensitivity analysis to assess the potential impact of exposure misclassification on the estimated measures of association.

Exposure misclassification and time to interview are topics of interest for the National Birth Defects Prevention Study (NBDPS), which is the largest population-based study of birth defects in the United States. The NBDPS is an ongoing nationwide case-control study that investigates genetic and environmental factors related to a variety of non-chromosomal and non-syndromic structural birth defects and other adverse birth outcomes (32). Mothers are interviewed about prenatal exposures up to two years after the infant's estimated date of delivery. It is important to understand to what extent the time interval between estimated date of delivery and maternal interview may be associated with the accuracy of exposure reporting. The objective of my thesis study is to systematically assess whether time to interview is associated with negative exposure reports (i.e., reporting that an exposure *did not* occur) and whether this relationship differs by case/control status or quality of the interview. These results will subsequently be used in a detailed sensitivity analysis of a previously conducted study on maternal antibiotic use and selected birth defects (33).

Methods

Study Design

This analysis used data from the National Birth Defects Prevention Study. A detailed description of the NBDPS is available elsewhere (32). The NBDPS is an ongoing, population-based case-control study conducted in ten different states (Arkansas, California, Iowa, Massachusetts, New Jersey, New York, Texas, Georgia, North Carolina and Utah). Study centers identify birth defects in specified regions of their state through ongoing surveillance. Vital records and hospital birth logs are used to identify a random sample of infant controls representing the liveborn birth population of the catchment area from which the birth defects cases were identified. Mothers of cases and controls are contacted, and mothers who agree to participate provide verbal informed consent and are administered a computer-assisted telephone interview. During this interview, mothers are asked to describe certain exposures before and during pregnancy, including use of over-the-counter medications, use of vitamins, illnesses, fertility treatments, and others. Upon completing the interview, mothers are sent a kit to collect buccal cells from themselves, their infant, and the infant's father.

Variables

We defined time-to-interview (TTI) as the time interval, in months, between the infant's expected date of delivery (EDD) and the maternal interview. The NBDPS aims to enroll mothers as soon as possible after six weeks postdelivery, and women are eligible to participate up to two years after their infant's EDD. We stratified TTI into four categories: 2-6 months, 7-12 months, 13-18 months, and 19-24 months. We examined five dichotomous outcomes (yes/no): report of upper respiratory infection; kidney, bladder, or urinary tract infection (henceforth referred to as "urinary tract infection"); morning sickness/nausea; folic acid-containing vitamin use during the first month before conception through the first month of pregnancy; and assisted fertility. Mothers were classified as to whether they reported "yes" (that they had experienced the outcome) or "no"

(that they had not experienced the outcome). Mothers with missing data or who reported "Don't Know" were excluded from that particular analysis.

We selected covariates as potential confounders *a priori*: case status of the infant (case, control), study center, maternal education (less than high school, high school or equivalent, more than high school), annual family income (<$10,000, $10,000-$50,000, >$50,000), year of birth, parity (nulliparous, primiparous, multiparous), gestational age of the infant at delivery (<32 weeks, 32-36 weeks, ≥37 weeks), birth outcome (live birth, stillbirth, therapeutic abortion), and language of interview (English, Spanish). We assessed case status and interview quality as potential effect modifiers. For simplicity, mothers of cases may be referred to as "cases," and mothers of controls may be referred to as "controls," throughout the following analysis.

Data Analysis

We used two-sample t-tests or ANOVA to test whether the mean TTI differed significantly between strata. Crude odds ratios (ORs) and 95% confidence intervals (CIs) were estimated for TTI categories of 7-12 months, 13-18 months, and 19-24 months for each of the five outcomes of interest, with TTI of 2-6 months as a reference. We coded the outcomes as "1" if the mother responded "No" to a question and "0" if she responded "Yes." Thus, the ORs reflect the odds of a mother reporting no exposure when she was interviewed at a particular TTI, compared to the odds of reporting no exposure at a TTI of 2-6 months. We estimated adjusted ORs and 95% CIs using multiple logistic regression.

We assessed interaction by case status and, separately, interview quality (which was dichotomized as high quality/generally reliable versus questionable/unsatisfactory). Interaction terms were included in each model, and interaction was assessed using a likelihood ratio test. Stratum-specific estimates were calculated if interaction was detected.

Data analysis was performed using SAS version 9.1 (Cary, NC, USA). This study was approved by the Emory IRB (study IRB00046951), the CDC, and each of the study sites.

Results

There were 25,448 women with expected dates of delivery between October 1997 and December 2005 who participated in the study. We excluded 408 women (1.6%) due to missing interview date and 18 additional women (0.07%) that were interviewed in a language other than English or Spanish. Baseline demographic information is presented in **Table 1**.

Mothers of controls were generally interviewed earlier than mothers of cases; the average time to interview for controls was 9.2 months and for cases was 11.4 months ($p<0.0001$). Although many differences were statistically significant due to a large sample size, the distribution of TTI was relatively consistent across levels of maternal education, family income, parity, and language of interview. We observed some variation in TTI by study site and year of birth.

Mothers of infants born very preterm (<32 weeks) or who had a therapeutic abortion were more likely to be interviewed early, as expected, given that these subjects were identified earlier. However, the difference was minor in the case of preterm infants. Interviews rated as "high quality" were also more likely to be conducted earlier.

The mean TTI did not markedly differ among mothers reporting and not reporting upper respiratory infections; urinary tract infections; morning sickness; use of artificial reproductive technology; or folic acid use during the month before conception through the first month of pregnancy. Although there were statistically significant differences in the distribution of mean TTI by upper respiratory infection, morning sickness, and folic acid use, these differences were small in magnitude.

Crude and adjusted odds ratios for the association between TTI and reporting a negative response are presented in **Table 2**. For all five interview items, the adjusted odds of a negative response were greatest for mothers interviewed 19-24 months after their infant's EDD compared to mothers interviewed 2-6 months post-EDD.

The crude odds of reporting no upper respiratory infection (URI) were significantly higher in the 19-24 month TTI group than in the 2-6 month TTI group (crude Odds Ratio [cOR] = 1.19, 95% CI 1.09-1.32), and the magnitude of the association increased slightly with adjustment for confounders (adjusted OR [aOR] = 1.24, 95% CI 1.12-1.38). We observed no difference in the odds of a negative report of URI among mothers with 7-12 and 13-18 month TTIs compared to the reference group. The adjusted odds of not reporting a urinary tract infection (UTI) did not largely differ between the 2-6, 7-12, and 13-18 month TTI groups. The adjusted odds in the 19-24 month group were slightly higher than the adjusted odds in the reference group, with borderline statistical significance (aOR = 1.11, 95% CI 0.98-1.26). The odds of reporting no morning sickness were greater in all three TTI groups than in the reference group. There was only a small increase in the odds of a negative response among mothers with a 7-12 month TTI (aOR = 1.09, 95% CI 1.00-1.17); the increase was greater for mothers interviewed between 13 and 24 months after their infant's EDD (aOR = 1.23 [95% CI 1.12-1.34] for the 13-18 month group and aOR = 1.26 [95% CI 1.13-1.40] for the 19-24 month group). The crude odds of not reporting assisted fertility were significantly greater in the 19-24 month group than in the reference group (cOR = 1.35, 95% CI 1.10-1.66); with adjustment for potential confounders, the odds were slightly attenuated (aOR = 1.20, 95% CI 0.96-1.50). The odds of reporting no use of folic acid-containing vitamins during the periconceptional period were slightly, but significantly, higher for mothers with a TTI between 7 and 18 months (aORs [95% CIs] were 1.10 [1.02-1.19] and 1.11 [1.02-1.22] for the 7-12 month and 13-18 month TTI groups, respectively). The magnitude of the association increased among the 19-24 month TTI group (aOR = 1.29, 95% CI 1.16-1.44).

The number of individuals providing a negative response to each interview item, as well as the percent providing a negative response, are presented by case/control status and TTI in **Table 3.** In both crude (data not shown) and adjusted analyses, significant interaction between case/control status and TTI was observed for negative reporting of URI and morning sickness. The point estimates for the adjusted odds ratios of reporting no URI were higher in controls than

in cases for every TTI category, and the odds of reporting no URI increased with TTI in both cases and controls. Moreover, although the differences between cases and controls were modest in the 7-12 month and 13-18 month TTI groups, the observed difference between cases and controls was more pronounced at a TTI of 19-24 months (aOR = 1.14, 95% CI 1.02-1.29 for cases; aOR = 1.68, 95% CI 1.33-2.10 for controls).

We also observed interaction by case status for the association of TTI with the morning sickness interview item. The odds of a negative report increased with TTI in controls, but this pattern was not observed for cases, although we observed a modestly increased odds of a negative response among cases with a TTI greater than 12 months. Controls with a TTI of 19-24 months (aOR = 1.68, 95% CI 1.33-2.12) were more likely to report no morning sickness compared to controls with a TTI of 2-6 months, and this OR was much higher than that observed for cases with the longest TTI (aOR = 1.19, 95% CI 1.05-1.35).

The p-value for interaction by case status for assisted reproductive technology (ART) was not statistically significant, likely due to a small number of mothers reporting use of ART. However, the observed odds ratios differed greatly between cases and controls with a 19-24 month TTI (aOR for controls = 2.20, 95% CI 1.05-4.61; aOR for cases = 1.11, 95% CI 0.87-1.41). There was no significant interaction by case status observed between TTI and either folic acid use or urinary tract infection. Furthermore, we observed no significant interaction by interview quality (data not shown).

Discussion

There appear to be meaningful differences in the odds of a negative response between mothers with a 2-6 month TTI and mothers with a 19-24 month TTI across a variety of reported maternal exposures. Although differences were not always statistically significant, the odds of a negative response were consistently highest in the longest TTI group compared to the other TTI groups.

Among the five prenatal events examined, a urinary tract infection was the only pregnancy event for which there was no observed association with TTI. This may be because UTIs are less likely to be forgotten, since they almost always require medical treatment. The use of assisted reproductive technology would seem to be similarly difficult to forget. However, the fact that controls with a 19-24 month TTI were much less likely to report using ART than cases may suggest that these controls were systematically different in a way for which we did not adjust, making them less likely to have actually used ART.

There are multiple interpretations of the patterns that we observed in the data. Due to the longer time interval between EDD and maternal interview, mothers with a longer TTI may not remember exposures as well as mothers with a shorter TTI. Alternatively, mothers with a short TTI could have been systematically different than mothers with a longer TTI in ways that were not accounted for by controlling for our covariates. This could mean that these mothers were more likely to forget their exposures, less likely to report them if they occurred, or less likely to experience these pregnancy events. As TTI increased, the odds of a negative response increased more in controls than in cases. One reason for this observed pattern may be recall bias: mothers of children without birth defects may have been less likely to scrutinize their pregnancy exposures and therefore more likely to forget them with time. Alternatively, mothers of controls may be systematically different than mothers of cases in ways not accounted for by controlling for covariates; these mothers may be less likely to experience or report these events.

Recall bias is a potential threat to case-control studies that use interviews to assess exposures, as is using a sample of controls that is not representative of cases. As discussed earlier, many authors have found that reporting accuracy did not differ largely by case/control status, but we found some evidence to the contrary in our data.

Our study had many strengths, as well as several limitations, that should be considered when interpreting the results. The NBPDS uses a large, population-based sample. The sample size for this analysis was considerable, and few mothers were missing exposure data on TTI. However, there were several limitations to this analysis. We were not able to distinguish the reasons that control mothers or mothers with a long TTI were more likely to report a negative response. Since we relied on maternal reports, we were not able to tell whether these differences were due to reporting inaccuracy or to true differences in prenatal experiences.

Our study helps explore the impact of including mothers with a long TTI in the NBDPS. While mothers with a longer TTI may be more likely to forget pregnancy events, the NBDPS dataset would not be able to completely ascertain children affected by certain defects if it were limited only to mothers who could be interviewed within a short time after delivery. Some birth defects, such as certain heart defects, may not be identified immediately after birth and could therefore be missed if the timing of eligibility were more restrictive. Moreover, one possible interpretation of our data is that mothers with a short TTI are systematically different than mothers with a long TTI, so both should be included in order to have a representative sample of the population of infants with birth defects.

Studies of time to interview and maternal reporting are few in number. Some report time to interview in relation to pregnancy outcome (11, 12, 31). Among the studies that examine the association between pregnancy exposures and time to interview, many different TTI intervals are used. Tilley et al.'s study examined TTI only among mothers who were interviewed ten or more years after delivery (20). Bryant et al. compared interview responses collected during pregnancy

to interview responses collected in the hospital immediately following delivery (24). Other studies use distinct TTIs, as well (11, 12, 26, 27, 31).

Results from our analysis suggest that TTI needs to be considered in case-control studies of infant outcome if mothers are enrolled at varying times after delivery. A sensitivity analysis can be a good way of assessing whether a study's conclusions might change, based upon differences in reporting that are attributable to a longer TTI. However, it is unclear whether the differences between mothers with a short and long TTI are due to difficulty recalling pregnancy events or due to true differences between groups.

The results of this analysis will be used to inform a sensitivity analysis assessing the potential impact of long TTI on exposure misclassification in a previous analysis of antibiotics and birth defects (33). The association of TTI with missing exposure reports and detail of reported exposure will also be assessed and included in this comparative uncertainty analysis, which will use both Monte Carlo and Bayesian methods.

Tables

Table 1. Participant characteristics and mean time-to-interview (TTI), National Birth Defects Prevention Study, 1997-2005

	N (%)[1]	Mean TTI (SD)	p-value (Difference in mean TTI between strata)
Total			
	25,022	10.8 (5.2)	N/A
Covariates			
Case status			
Case	18,403 (73.6)	11.4 (5.3)	<0.0001
Control	6,619 (26.5)	9.2 (4.9)	
Study Center			
Arkansas	3,252 (13.0)	10.3 (6.0)	
California	3,234 (12.9)	10.7 (5.4)	
Iowa	2,575 (10.3)	11.9 (4.5)	
Massachusetts	3,365 (13.5)	10.4 (4.5)	
New Jersey	2,177 (8.7)	11.2 (5.0)	<0.0001
New York	1,908 (7.6)	9.8 (5.0)	
Texas	3,096 (12.4)	12.2 (5.7)	
CDC/Atlanta	2,873 (11.5)	9.3 (5.3)	
North Carolina	1,086 (4.3)	11.5 (4.8)	
Utah	1,456 (5.8)	11.6 (4.5)	
Maternal education			
<12 years	4,424 (17.7)	11.1 (5.4)	
12 years	6,459 (25.8)	11.1 (5.4)	<0.0001
>12 years	14,121 (56.5)	10.7 (5.1)	
Family Income			
<$10,000	4,291 (18.5)	11.3 (5.5)	
$10,000-$50,000	11,082 (47.9)	11.0 (5.3)	<0.0001
>$50,000	7,781 (33.6)	10.6 (5.0)	

Year of Birth			
1997	381 (1.5)	11.9 (5.0)	
1998	2,660 (10.6)	11.7 (4.9)	
1999	3,103 (12.4)	10.9 (5.5)	
2000	3,319 (13.3)	10.2 (5.3)	
2001	3,080 (12.3)	10.3 (5.4)	<0.0001
2002	2,667 (10.7)	10.4 (5.3)	
2003	3,017 (12.1)	11.1 (5.1)	
2004	3,515 (14.1)	11.4 (5.1)	
2005	3,280 (13.1)	10.7 (5.2)	
Parity			
0	10,542 (42.1)	10.8 (5.3)	
1	8,076 (32.3)	10.8 (5.2)	0.1238
2 or more	6,396 (25.6)	10.9 (5.3)	
Gestational age at delivery			
Very preterm (<32 wk)	1,720 (6.9)	10.3 (5.6)	
32-36 wk Preterm	3,973 (15.9)	10.6 (5.3)	<0.0001
Term (>=37 wk)	19,324 (77.2)	10.9 (5.2)	
Birth outcome			
Live Birth	24,338 (97.3)	10.9 (5.2)	
Fetal death	314 (1.3)	11.3 (5.5)	<0.0001
Termination	356 (1.4)	9.4 (5.7)	
Language of interview			
English	23,212 (92.8)	10.8 (5.2)	0.0053
Spanish	1,810 (7.2)	11.2 (5.4)	
Interview quality			
High quality	17,345 (69.5)	10.4 (5.1)	
Generally reliable	7,085 (28.4)	11.8 (5.3)	<0.0001
Questionable	501 (2.0)	12.0 (5.4)	
Unsatisfactory	45 (0.2)	11.9 (5.6)	

Reported Pregnancy Experiences			
Upper respiratory infection			
Yes, N (%)	14,470 (61.2)	10.7 (5.1)	0.0008
No, N (%)	9,171 (38.8)	10.9 (5.4)	
Kidney, bladder, or urinary tract infection			
Yes, N (%)	5,115 (20.6)	10.9 (5.3)	0.2800
No, N (%)	19,769 (79.4)	10.8 (5.2)	
Morning sickness			
Yes, N (%)	17,672 (70.7)	10.7 (5.2)	<0.0001
No, N (%)	7,333 (29.3)	11.1 (5.3)	
Assisted fertility			
Yes, N (%)	1,549 (6.4)	10.7 (5.0)	0.0585
No, N (%)	22,815 (93.6)	10.9 (5.3)	
Folic acid-containing vitamin use in the month before and month after conception			
Yes, N (%)	12,702 (50.8)	10.6 (5.1)	<0.0001
No, N (%)	12,320 (49.2)	11.1 (5.4)	

[1]Percentages do not include individuals missing data on each respective variable.

Table 2. Crude and adjusted odds ratios of a negative response[1], by time-to-interview (TTI),

National Birth Defects Prevention Study, 1997 – 2005

Reported pregnancy experience	*TTI: 2 to 6 months*	*TTI: 7 to 12 months*	*TTI: 13 to 18 months*	*TTI: 19 to 24 months*
No report of upper respiratory infection				
cOR[2]	1.0	0.90 (0.84, 0.96)	0.97 (0.89, 1.04)	1.19 (1.09, 1.32)
aOR[3]	1.0	0.98 (0.91, 1.06)	1.05 (0.96, 1.14)	1.24 (1.12, 1.38)
No report of kidney, bladder, or urinary tract infection				
cOR[2]	1.0	1.02 (0.94, 1.10)	0.98 (0.90, 1.07)	0.97 (0.87, 1.09)
aOR[3]	1.0	0.99 (0.91, 1.08)	0.99 (0.90, 1.10)	1.11 (0.98, 1.26)
No report of morning sickness				
cOR[2]	1.0	1.05 (0.98, 1.13)	1.20 (1.11, 1.30)	1.20 (1.09, 1.33)
aOR[3]	1.0	1.09 (1.00, 1.17)	1.23 (1.12, 1.34)	1.26 (1.13, 1.40)
No report of use of artificial reproductive technology				
cOR[2]	1.0	1.02 (0.89, 1.16)	0.95 (0.82, 1.11)	1.35 (1.10, 1.66)
aOR[3]	1.0	1.04 (0.90, 1.20)	0.95 (0.81, 1.12)	1.20 (0.96, 1.50)
No report of periconceptional folic acid use				
cOR[2]	1.0	1.05 (0.98, 1.12)	1.12 (1.04, 1.20)	1.43 (1.30, 1.57)
aOR[3]	1.0	1.10 (1.02, 1.19)	1.11 (1.02, 1.22)	1.29 (1.16, 1.44)

[1]A negative response is defined as a response in which the subject reports that they did NOT experience the exposure.

[2]cOR = crude odds ratio

[3]aOR = adjusted odds ratio; adjusted models included the following covariates: case status, study center, maternal education, annual family income, year of birth, parity, gestational age at delivery, birth outcome, and language of interview.

Table 3: Adjusted[1] odds ratios of a negative response[2], by time-to-interview (TTI), stratified by case/control status, National Birth Defects Prevention Study, 1997 – 2005

Reported Pregnancy Experience	*Case / control status*	*TTI: 2 to 6 months*		*TTI: 7 to 12 months*		*TTI: 13 to 18 months*		*TTI: 19 to 24 months*		*p-value for interaction*
		N (%)	OR (95% CI)	N (%)	OR (95% CI)	N (%)	OR (95% CI)	N (%)	OR (95% CI)	
No report of upper respiratory infection	Cases	1,419 (42.4)	1.0	2,883 (38.0)	0.94 (0.86, 1.03)	1,708 (39.2)	1.00 (0.91, 1.11)	884 (43.0)	1.14 (1.02, 1.29)	0.034
	Controls	806 (35.5)	1.0	929 (34.6)	1.05 (0.92, 1.19)	350 (36.8)	1.12 (0.95, 1.32)	192 (49.0)	1.68 (1.33, 2.10)	
No report of kidney, bladder, or urinary tract infection	Cases	2,765 (79.2)	1.0	6,324 (79.1)	0.96 (0.87, 1.07)	3,628 (78.6)	0.96 (0.85, 1.08)	1,728 (78.8)	1.08 (0.94, 1.24)	0.769
	Controls	1,862 (79.8)	1.0	2,288 (81.5)	1.04 (0.89, 1.21)	842 (81.7)	1.07 (0.88, 1.31)	332 (79.6)	1.20 (0.91, 1.59)	
No report of morning sickness	Cases	996 (28.5)	1.0	2,377 (29.6)	1.10 (1.00, 1.21)	1,464 (31.5)	1.21 (1.09, 1.34)	679 (30.7)	1.19 (1.05, 1.35)	0.012
	Controls	616 (26.3)	1.0	731 (25.9)	1.02 (0.90, 1.17)	323 (31.1)	1.28 (1.09, 1.52)	147 (35.1)	1.68 (1.33, 2.12)	
No reported use of assisted reproductive technology	Cases	3,162 (92.5)	1.0	7,235 (93.1)	1.03 (0.87, 1.22)	4,292 (92.8)	0.91 (0.76, 1.10)	2,091 (94.6)	1.11 (0.87, 1.41)	0.229
	Controls	2,125 (95.0)	1.0	2,519 (94.9)	1.04 (0.80, 1.37)	980 (94.9)	1.11 (0.78, 1.58)	411 (98.1)	2.20 (1.05, 4.61)	
No report of periconceptional folic acid use	Cases	1,636 (46.7)	1.0	3,884 (48.3)	1.06 (0.97, 1.17)	2,340 (50.3)	1.11 (1.00, 1.23)	1,221 (55.2)	1.24 (1.10, 1.40)	0.284
	Controls	1,119 (47.8)	1.0	1,365 (48.4)	1.18 (1.04, 1.34)	503 (48.4)	1.07 (0.91, 1.27)	252 (60.1)	1.45 (1.14, 1.84)	

[1]Adjusted models included the following covariates: case status, study center, maternal education, annual family income, year of birth, parity, gestational age at delivery, birth outcome, and language of interview.

[2]A negative response is defined as a response in which the subject reports that they did NOT experience the exposure.

References

1. Wisborg K, Henriksen TB, Hedegaard M, et al. Smoking during pregnancy and preterm birth. *Br J Obstet Gynaecol* 1996;103(8):800-5.

2. Kyrklund-Blomberg NB, Cnattingius S. Preterm birth and maternal smoking: risks related to gestational age and onset of delivery. *Am J Obstet Gynecol* 1998;179(4):1051-5.

3. Chan A, Keane RJ, Robinson JS. The contribution of maternal smoking to preterm birth, small for gestational age and low birthweight among Aboriginal and non-Aboriginal births in South Australia. *Med J Aust* 2001;174(8):389-93.

4. Cnattingius S, Signorello LB, Anneren G, et al. Caffeine intake and the risk of first-trimester spontaneous abortion. *N Engl J Med* 2000;343(25):1839-45.

5. Weng X, Odouli R, Li DK. Maternal caffeine consumption during pregnancy and the risk of miscarriage: a prospective cohort study. *Am J Obstet Gynecol* 2008;198(3):279 e1-8.

6. Zuckerman B, Frank DA, Hingson R, et al. Effects of maternal marijuana and cocaine use on fetal growth. *N Engl J Med* 1989;320(12):762-8.

7. Finer LB, Henshaw SK. Disparities in rates of unintended pregnancy in the United States, 1994 and 2001. *Perspect Sex Reprod Health* 2006;38(2):90-6.

8. Eaton-Evans J, Dugdale AE. Recall by mothers of the birth weights and feeding of their children. *Hum Nutr Appl Nutr* 1986;40(3):171-5.

9. Ekouevi K, Morgan SP. Note on the reliability and validity of mothers' retrospective reports of their children's birthweights. *Soc Biol* 1991;38(1-2):140-5.

10. Githens PB, Glass CA, Sloan FA, et al. Maternal recall and medical records: an examination of events during pregnancy, childbirth, and early infancy. *Birth* 1993;20(3):136-41.

11. O'Sullivan JJ, Pearce MS, Parker L. Parental recall of birth weight: how accurate is it? *Arch Dis Child* 2000;82(3):202-3.

12. Seidman DS, Slater PE, Ever-Hadani P, et al. Accuracy of mothers' recall of birthweight and gestational age. *Br J Obstet Gynaecol* 1987;94(8):731-5.

13. Sou SC, Chen WJ, Hsieh WS, et al. Severe obstetric complications and birth characteristics in preterm or term delivery were accurately recalled by mothers. *J Clin Epidemiol* 2006;59(4):429-35.

14. Troude P, L'Helias LF, Raison-Boulley AM, et al. Perinatal factors reported by mothers: do they agree with medical records? *Eur J Epidemiol* 2008;23(8):557-64.

15. Axelsson G, Rylander R. Validation of questionnaire reported miscarriage, malformation and birth weight. *Int J Epidemiol* 1984;13(1):94-8.

16. Casey R, Rieckhoff M, Beebe SA, et al. Obstetric and perinatal events: the accuracy of maternal report. *Clin Pediatr (Phila)* 1992;31(4):200-4.

17. Drews CD, Kraus JF, Greenland S. Recall bias in a case-control study of sudden infant death syndrome. *Int J Epidemiol* 1990;19(2):405-11.

18. 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.

19. Pickett KE, Kasza K, Biesecker G, et al. Women who remember, women who do not: a methodological study of maternal recall of smoking in pregnancy. *Nicotine Tob Res* 2009;11(10):1166-74.

20. Tilley BC, Barnes AB, Bergstralh E, et al. A comparison of pregnancy history recall and medical records. Implications for retrospective studies. *Am J Epidemiol* 1985;121(2):269-81.

21. 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):347-52.

22. Cartwright A, Smith C. Some comparisons of data from medical records and from interviews with women who had recently had a live birth or stillbirth. *J Biosoc Sci* 1979;11(1):49-64.

23. Delgado-Rodriguez M, Gomez-Olmedo M, Bueno-Cavanillas A, et al. Recall bias in a case-control study of low birth weight. *J Clin Epidemiol* 1995;48(9):1133-40.

24. Bryant HE, Visser N, Love EJ. Records, recall loss, and recall bias in pregnancy: a comparison of interview and medical records data of pregnant and postnatal women. *Am J Public Health* 1989;79(1):78-80.

25. Mitchell AA, Cottler LB, Shapiro S. Effect of questionnaire design on recall of drug exposure in pregnancy. *Am J Epidemiol* 1986;123(4):670-6.

26. 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):2071-5.

27. Olson JE, Shu XO, Ross JA, et al. Medical record validation of maternally reported birth characteristics and pregnancy-related events: a report from the Children's Cancer Group. *Am J Epidemiol* 1997;145(1):58-67.

28. Werler MM, Pober BR, Nelson K, et al. Reporting accuracy among mothers of malformed and nonmalformed infants. *Am J Epidemiol* 1989;129(2):415-21.

29. Fenster L, Swan SH, Windham GC, et al. Assessment of reporting consistency in a case-control study of spontaneous abortions. *Am J Epidemiol* 1991;133(5):477-88.

30. Tomeo CA, Rich-Edwards JW, Michels KB, et al. Reproducibility and validity of maternal recall of pregnancy-related events. *Epidemiology* 1999;10(6):774-7.

31. Rasmussen SA, Mulinare J, Khoury MJ, et al. Evaluation of birth defect histories obtained through maternal interviews. *Am J Hum Genet* 1990;46(3):478-85.

32. Yoon PW, Rasmussen SA, Lynberg MC, et al. The National Birth Defects Prevention Study. *Public Health Rep* 2001;116 Suppl 1:32-40.

33. Crider KS, Cleves MA, Reefhuis J, et al. Antibacterial medication use during pregnancy and risk of birth defects: National Birth Defects Prevention Study. *Arch Pediatr Adolesc Med* 2009;163(11):978-85.