



## Practice of Epidemiology

# Impact of Time to Maternal Interview on Interview Responses in the National Birth Defects Prevention Study

**Sarah C. Tinker\***, **Cassandra Gibbs**, **Matthew J. Strickland**, **Owen J. Devine**, **Krista S. Crider**, **Martha M. Werler**, **Marlene T. Anderka**, **Jennita Reefhuis**, for the **National Birth Defects Prevention Study**

\* Correspondence to Dr. Sarah C. Tinker, Centers for Disease Control and Prevention, 1600 Clifton Road NE, Mail-Stop E86, Atlanta, GA 30333 (e-mail: zzu9@cdc.gov).

*Initially submitted May 9, 2012; accepted for publication August 8, 2012.*

Prenatal exposures often are assessed using retrospective interviews. Time from exposure to interview may influence data accuracy. We investigated the association of time to interview (TTI) with aspects of interview responses in the National Birth Defects Prevention Study, a population-based case-control study of birth defects in 10 US states. Mothers completed a computer-assisted telephone interview 1.5–24 months after their estimated date of delivery. Proxy metrics for interview quality were whether certain exposures were reported, whether the start month of reported medication use or illness was reported, or whether responses were missing. Interaction by case status was assessed. Interviews were completed with 30,542 mothers (22,366 cases and 8,176 controls) who gave birth between 1997 and 2007. Mothers of cases were interviewed later than were mothers of controls (11.7 months vs. 9.5 months, respectively). In adjusted analyses, having a TTI that was greater than 6 months was associated with only a few aspects of interview responses (e.g., start month of pseudoephedrine use). Interaction by case-control status was observed for some exposures; mothers of controls had a greater reduction in interview quality with increased TTI in these instances (e.g., report of morning sickness, start month of acetaminophen use and ibuprofen use). The results suggest that TTI might impact interview responses; however, the impact may be minimal and specific to the type of exposure.

bias (epidemiology); case-control studies; interviews as topic

Abbreviations: EDD, estimated date of delivery; NBDPS, National Birth Defects Prevention Study; TTI, time to interview.

Maternal exposures during pregnancy have the potential to alter fetal development and pregnancy outcome, although it is often difficult to predict the impact of a particular exposure on the fetus. Epidemiologic research has yielded important information about how to promote healthier pregnancies. However, there is still a lack of knowledge regarding the causes behind most adverse pregnancy outcomes; therefore, improved methods for research in this area are under continued development.

Assessment of exposures during pregnancy poses many challenges. Pregnancy typically is not recognized until several weeks after conception, making recruitment of women in the earliest stages of pregnancy problematic (1). Prospective studies that enroll only women who are planning a pregnancy will miss the approximately 50% of pregnancies that

are unplanned in the United States (2) and will therefore be unrepresentative of the general population. To study rare outcomes, such as specific birth defects, enormous cohorts would be required to produce sufficiently powered analyses, and therefore case-control study designs are typically used for these outcomes. However, a drawback of case-control studies is that the pregnancy outcome is known at the time at which subjects are asked to self-report their exposure histories, potentially leading to recall bias.

A related concern with retrospective reports of exposures during pregnancy is the potential impact of the time between the exposure and the report in a maternal interview or survey. There has been substantial research in which investigators examined the reliability and validity of maternal recall of the

birthing experience (3–5) and birth weight (6–8) that generally has demonstrated good recall over intervals ranging beyond 50 years after delivery. More limited data are available on the assessment of exposures that occur during pregnancy, particularly those of a transient nature. In previous studies in which women were asked about exposures such as infections, medication use, or exposure to x-rays during pregnancy, the reliability of reporting decreased as the time between the exposure and interview increased. Overall, women were less likely to report an exposure that actually happened than to erroneously report an exposure that did not occur as time to interview (TTI) increased (9–11).

Validation data for exposures during pregnancy are also difficult and expensive to obtain. In the absence of validation data, an assessment of interview quality can inform the potential impact of increasing time between exposure and interview. The objective of our analysis was to assess whether the time interval from the estimated date of delivery (EDD) to the maternal interview influences the quality of the data on reported prenatal exposures in the National Birth Defects Prevention Study (NBDPS).

## MATERIALS AND METHODS

### The National Birth Defects Prevention Study

The NBDPS is a population-based case-control study of major structural birth defects conducted at 10 centers in the United States (Arkansas, California, Georgia, Iowa, Massachusetts, New Jersey, New York, North Carolina, Texas, and Utah). Institutional review boards at each center approved the study. The NBDPS methods have been described in detail previously (12). Briefly, birth defect cases are ascertained through population-based birth defect surveillance systems that collect information from hospital discharge summaries, birth logs, medical records, prenatal data from hospital records, and some specialty clinics. Possible pregnancy outcomes for mothers of case infants are live birth, stillbirth, or induced abortion. Controls were taken from a random sample of liveborn infants who did not have major birth defects and were born to mothers living in the same areas from which the case infants were drawn. They were selected from birth certificate records or hospital birth logs (13). The study population for the present analysis included case and control infants with an EDD between October 1, 1997, and December 31, 2007. During this time period, participation was 68.5% for cases and 64.9% for controls.

Participating mothers of case and control infants were administered a computer-assisted telephone interview in English or Spanish that lasted approximately 1 hour. Mothers provided informed consent before beginning the interview. During the interview, mothers were asked about their exposures before and during pregnancy in a variety of domains, including medication use and illness. Mothers were asked about their exposures in the 3 months before pregnancy and throughout their pregnancy. Mothers were eligible to be interviewed from 6 weeks to 2 years after their EDD. EDD was used to standardize the amount of time that elapsed from the time of conception for pregnancies that ended at different gestational ages.

### Time to interview

The exposure of interest in our analysis was the TTI. Mothers had the option of completing the interview in multiple sessions. We defined TTI as the time from the infant's EDD to the end of the mother's interview. For most of our analyses, we considered TTI as a categorical variable (6 weeks to 6 months, 7 months to 12 months, 13 months to 18 months, or 19 months to 24 months), with 6 weeks to 6 months as the referent category. Mothers with incomplete interviews were excluded from the analysis ( $n = 572$ ; <2%).

### Proxy metrics of interview quality

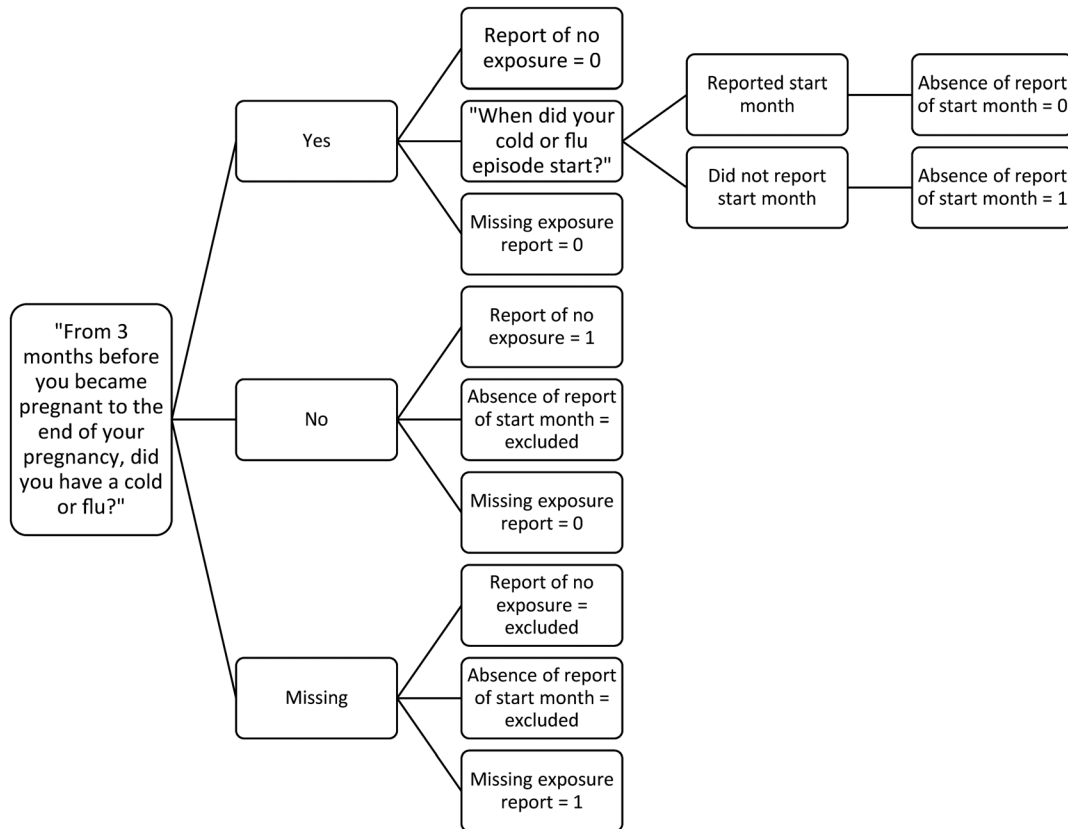
We did not have external validation data available to assess interview quality. Therefore, we used 3 different proxy metrics for recall and reporting accuracy.

*Report of no exposure.* Women were asked separate questions about whether they had experienced any for the following from 3 months before pregnancy through the end of pregnancy: upper respiratory tract infections; kidney, bladder, or urinary tract infections; injuries; morning sickness; or use of fertility medications or procedures. Response options were “yes,” “no,” “don't know,” and “refused.” We defined a report of no exposure as a mother reporting that an exposure did not occur (a “no” response). Although many negative responses are accurate, mothers may be less likely to recall an exposure and more likely to erroneously report that the exposure did not occur as the time from the exposure increases.

*Missing information on start month.* If a mother reported that she had had an illness or used a medication any time during the 3 months before pregnancy through the end of pregnancy, she was asked to provide either the exact start date or the month of pregnancy when the illness or medication use began. However, some women did not report a start month. We determined whether we were missing data on start month for infections of the upper respiratory tract or kidney, bladder, or urinary tract, as well as for use of acetaminophen, ibuprofen, pseudoephedrine, opioid analgesics, antidepressant medication, or antiepileptic medication. If a mother reported multiple instances of an exposure, she was coded as having missing information on the start month if the information was missing for any instance.

We were not able to include broad medication categories in the report of the no-exposure or missing-exposure report outcomes because mothers were not asked about specific use of most medications (i.e., “Did you take [specific drug]?”); rather, use was reported throughout the interview in response to follow-up questions after reports of specific reported illness or conditions (i.e., “For [specific condition], did you take any medications?”). Mothers were specifically questioned about their use of acetaminophen and ibuprofen and about other name-brand products that contained these ingredients. They also had the option of reporting use of these medications in response to the condition-specific questions, as described above.

*Missing responses.* Mothers whose response was “don't know” or who failed to provide a response for questions asked about injury, morning sickness, respiratory illness, or



**Figure 1.** Sample coding of interview quality variables based on response to questions on whether the mother experienced a respiratory illness during pregnancy, National Birth Defects Prevention Study, 1997–2007.

kidney, bladder, or urinary tract infection were considered to have missing data. These missing responses were evaluated as the third metric of interview quality. Figure 1 provides an example of how responses were coded based on the example of the questions mothers were asked regarding whether they experienced a respiratory illness during pregnancy.

### Statistical analyses

We conducted a descriptive analysis of the TTI data and assessed the crude and adjusted associations of TTI with our indicators of interview quality in 2 ways. We first examined mean differences in TTI by maternal characteristics and by categories of the defined metrics of interview quality. In addition, we used the categorical TTI variable as the independent variable in logistic regression models to assess potential associations between TTI and the odds of the mother's interview having the low-quality value for the given metric using the shortest TTI category (6 weeks to 6 months) as the referent. In the modeling approach, we included several covariates selected a priori as potential confounders: maternal age at EDD (<25 years, 25–34 years, or ≥35 years); maternal race/ethnicity (non-Hispanic white, non-Hispanic black, Hispanic, or other race/ethnicity); maternal educational level (less than high school graduate,

high school graduate or equivalent, or more than high school); annual family income (<\$10,000, \$10,000–\$50,000, or >\$50,000); parity (nulliparous, primiparous, or multiparous); case status (case vs. control); year of EDD; gestational age at delivery (<32 weeks, 32–36 weeks, or ≥37 weeks); birth outcome (livebirth, stillbirth, or therapeutic abortion); study center; and language of interview (English or Spanish). We also assessed whether the associations between TTI and our indicators of interview quality were different based on case status, interviewer-classified interview quality (“high quality” or “generally reliable” vs. “questionable” or “unsatisfactory”), or maternal age through stratified analysis and statistical testing for multiplicative interaction. All analyses were conducted using SAS, version 9.2 (SAS Institute, Inc., Cary, North Carolina).

### RESULTS

We analyzed data from 22,366 cases and 8,176 controls. The distributions of demographic and other descriptive characteristics are presented in Table 1. Although most of the differences in the mean TTI between groups were relatively modest, the *P* values were statistically significant for every characteristic we considered, in part because of the large sample size. The only characteristics for which there

**Table 1.** Mean Time to Interview by Selected Participant Characteristics, National Birth Defects Prevention Study, 1997–2007

Variable	No.	%	TTI in Months, <sup>a</sup> mean (SD)
Case status			
Case	22,366	73.2	11.7 (5.4)
Control	8,176	26.8	9.5 (5.0)
Maternal age, years			
<20	3,178	10.4	11.4 (5.8)
20–29	15,219	49.8	11.3 (5.4)
30–39	11,317	37.1	10.9 (5.2)
≥40	828	2.7	10.6 (5.1)
Race/ethnicity			
Non-Hispanic white	18,309	60.0	10.9 (5.2)
Non-Hispanic black	3,142	10.3	11.0 (5.9)
Hispanic	7,067	23.1	11.6 (5.6)
Other race/ethnicity	2,018	6.6	11.6 (5.5)
Maternal educational level			
Less than high school	5,371	17.6	11.4 (5.6)
High school or equivalent	7,770	25.5	11.3 (5.5)
More than high school	17,377	56.9	10.9 (5.2)
Annual household income			
<\$10,000	5,512	19.4	11.3 (5.7)
\$10,000–\$50,000	13,376	47.0	11.3 (5.4)
>\$50,000	9,587	33.7	10.9 (5.1)
Parity			
No previous live births	12,863	42.1	11.1 (5.4)
1 previous live birth	9,730	31.9	11.0 (5.3)
2 or more previous live births	7,940	26.0	11.3 (5.4)
Year of expected date of delivery			
1997	330	1.1	12.1 (5.0)
1998	2,623	8.6	11.8 (5.0)
1999	3,084	10.1	11.0 (5.6)
2000	3,333	10.9	10.3 (5.3)
2001	3,105	10.2	10.4 (5.5)
2002	2,717	8.9	10.4 (5.3)
2003	2,935	9.6	11.2 (5.1)
2004	3,490	11.4	11.5 (5.1)
2005	3,288	10.8	10.8 (5.3)
2006	2,788	9.1	12.5 (5.3)
2007	2,849	9.3	11.3 (5.6)

Table continues

was more than a 1-month difference in the mean TTI were case status, year of EDD, birth outcome, study center, and interviewer-rated quality of the interview. Cases had longer TTI than did controls, with a mean of 11.7 months compared with 9.5 months for controls. The mean TTI ranged from 10.3 months to 12.5 months across years of EDD and

**Table 1.** Continued

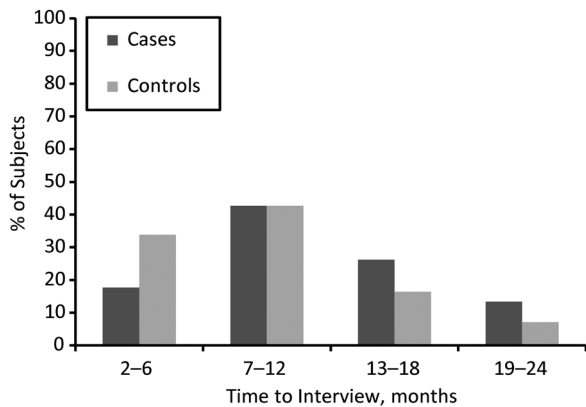
Variable	No.	%	TTI in Months, <sup>a</sup> mean (SD)
Birth outcomes			
Live birth	29,712	97.3	11.1 (5.4)
Fetal death	389	1.3	11.7 (5.8)
Therapeutic abortion	424	1.4	9.6 (5.8)
Gestational age, weeks			
<32	2,024	6.7	10.6 (5.8)
32–36	4,837	15.9	10.9 (5.5)
≥37	23,523	77.4	11.2 (5.3)
Study center			
A	3,982	13.0	11.4 (5.7)
B	3,935	12.9	10.5 (4.5)
C	3,928	12.9	10.5 (6.0)
D	3,619	11.9	12.0 (5.7)
E	3,421	11.2	9.8 (5.5)
F	3,055	10.0	12.1 (4.6)
G	2,380	7.8	12.6 (4.7)
H	2,285	7.5	10.2 (5.1)
I	2,189	7.2	11.3 (5.1)
J	1,748	5.7	11.6 (5.5)
Language of interview			
English	28,185	92.3	11.1 (5.3)
Spanish	2,357	7.7	11.6 (5.6)
Interviewer-rated quality of interview			
High	20,422	67.0	10.7 (5.3)
Generally reliable	9,379	30.8	12.0 (5.5)
Questionable	522	1.7	12.2 (5.5)
Unsatisfactory	169	0.6	12.6 (5.5)

Abbreviations: SD, standard deviation; TTI, time to interview.

<sup>a</sup> All 2-sided *P* values for differences in means were <0.005.

from 9.8 months to 12.6 months among different study centers. The mean TTI for mothers of liveborn infants was 11.1 months; for those who experienced fetal deaths, TTI was 11.7 months, and for those who underwent therapeutic abortions, it was 9.6 months. The mean TTI for interviews that the interviewer classified as high quality was 10.7 months; for interviews classified in other quality categories, the mean TTI was between 12.0–12.6 months.

Cases were about half as likely as were controls to have had their maternal interview conducted within 6 months of the EDD (17.7% and 33.8%, respectively), and they were 1.7 times more likely to have had their maternal interview conducted a year or more after the EDD (Figure 2). Because recall bias is such a big concern in studies of birth defects and because some associations did show differences between cases and controls, results are presented stratified by case status. The tests for interaction by interviewer-classified interview quality and maternal age were not



**Figure 2.** Distribution of subjects' time to interview, stratified by case status, National Birth Defects Prevention Study, 1997–2007.

statistically significant for any of the interview-quality outcome metrics we assessed.

### Report of no exposure

Some of the exposures that we considered for our analysis of no-exposure outcomes were commonly reported during pregnancy by both cases and controls, such as morning sickness, upper respiratory infections, and kidney, bladder, or urinary tract infections, which were reported by approximately 20%–70% of mothers (Table 2) depending on exposure. Injuries and use of fertility medications or procedures were rarer, with each reported by less than 10% of mothers. For all exposures and among both cases and controls, the mean TTI varied by less than 1 month among mothers who reported that these exposures occurred and mothers who reported that these exposures did not occur; however, because of the large sample sizes, some of the *P* values for these small differences were statistically significant.

Among mothers of both cases and controls, there was some suggestion that they were less likely to report an exposure as TTI increased; the associations were quite modest, although often statistically significant (Table 3). Interaction by case status was significant only for morning sickness, although the adjusted odds ratio point estimates were modest.

### Absence of report of start month of illness or medication use

The majority of mothers provided sufficient information for us to classify their start months of illness or medication use, although there was some variation (Table 2). The mean TTI for mothers who did not report a start date was longer than that for mothers who did report a start date for most of the reported exposures examined, although not all differences were statistically significant.

Although results varied by the specific reported exposure, in general we saw stronger but still modest associations between increasing TTI and decreasing interview quality, as

indicated by lack of a report of start month (Table 4). For kidney, bladder, or urinary tract infection and pseudoephedrine use, mothers of both cases and controls were generally less likely to report the start month as the TTI increased. For acetaminophen and ibuprofen use, there was little association between TTI category and lack of report of a start month among cases. Among controls, however, we observed increased odds of not reporting a start month with a TTI of greater than 12 months for acetaminophen and greater than 18 months for ibuprofen. The difference in results for cases and controls is reflected in the statistically significant tests for interaction.

Because of the small numbers of mothers who reported using opioid analgesics, antidepressant medication, or anti-epileptic medication, we were unable to conduct analyses stratified by case status or to adjust for potential confounding factors. Although the crude odds ratios suggested that mothers who reported these exposures might have been less likely to report the start month as the TTI increased, the estimates were too imprecise to allow us to draw conclusions.

### Missing exposure report

It was rare to have a missing report of a maternal exposure that we examined, with rates ranging from less than 1% for injuries, morning sickness, and kidney, bladder, or urinary tract infection to 5.6% for upper respiratory infection (Table 2). TTI tended to be longer for mothers with missing responses than for mothers who provided a response.

The results for missing exposure reports of upper respiratory infection and kidney, bladder, or urinary tract infection stratified by cases status suggested increasing odds of a missing exposure response with increasing TTI; for upper respiratory infection, this association appears slightly stronger among controls than cases (Table 5). The less-stable nonstratified results for injury and morning sickness suggested that there may also be increased odds of having a missing exposure report for these exposures as TTI increased, but the estimates were highly imprecise.

## DISCUSSION

The results of our analysis showed that the average TTI in the NBDPS varied by less than 1 month across categories for most demographic factors. We saw weak or no associations between TTI and our indicators of interview quality for most of the exposures we assessed, although we did observe modest associations between TTI and certain reported exposures (e.g., start month of pseudoephedrine use). We observed interactions by case status for no report of morning sickness, no report of start month of acetaminophen or ibuprofen use, and missing report of upper respiratory infection. In each instance, an association was observed for controls but not for cases.

### Factors contributing to TTI

There are several factors that contribute to the length of time between a mother's due date and the date she is interviewed for the NBDPS. The mean TTI for cases was about 2

**Table 2.** Indicators of Interview Quality and Mean Time to Interview Among Case and Control Mothers ( $n = 30,542$ ), National Birth Defects Prevention Study, 1997–2007

Reported Pregnancy Exposure	Cases				Controls			
	No.	%	TTI, mean (SD)	<i>P</i> Value (2-sided) for Difference in Means	No.	%	TTI, mean (SD)	<i>P</i> Value (2-sided) for Difference in Means
Injury <sup>a</sup>								
No	20,340	90.9	11.7 (5.4)	0.7422 <sup>b</sup>	7,492	91.6	9.5 (5.0)	0.03 <sup>b</sup>
Yes	2,020	9.0	11.8 (5.4)		680	8.3	9.1 (5.0)	
Missing	6	0.03	11.7 (5.4)	0.9790 <sup>c</sup>	4	0.1	11.0 (9.0)	0.54 <sup>c</sup>
Morning sickness <sup>a</sup>								
No	6,655	29.8	11.8 (5.4)	0.0418 <sup>b</sup>	2,231	27.3	9.8 (5.2)	<0.0001 <sup>b</sup>
Yes	15,691	70.2	11.7 (5.4)		5,942	72.7	9.3 (4.9)	
Missing	20	0.1	11.7 (5.4)	0.0391 <sup>c</sup>	3	0.04	8.0 (2.6)	0.61 <sup>c</sup>
Use of fertility medication or procedures <sup>d,e</sup>								
No	20,507	93.3	11.8 (5.4)	0.0061 <sup>b</sup>	7,531	95.3	9.6 (5.1)	0.09 <sup>b</sup>
Yes	1,478	6.7	11.4 (5.1)		368	4.7	9.1 (4.4)	
Upper respiratory infection <sup>f</sup>								
No	8,665	38.7	11.8 (5.5)	0.0938 <sup>b</sup>	2,945	36.0	9.7 (5.3)	<0.0001 <sup>b</sup>
Yes	12,454	55.7	11.6 (5.5)		4,865	59.5	9.2 (4.8)	
Reported start date	10,764	86.4	11.5 (5.2)	<0.0001 <sup>g</sup>	4,292	88.2	9.1 (4.8)	<0.0001 <sup>g</sup>
Didn't report start date	1,690	13.6	12.6 (5.3)		573	11.8	10.3 (5.0)	
Missing	1,247	5.6	12.5 (5.4)	<0.0001 <sup>c</sup>	366	4.5	11.0 (4.9)	<0.0001 <sup>c</sup>
Kidney, bladder, or urinary tract infection <sup>f</sup>								
No	17,504	78.3	11.7 (5.4)	0.2334 <sup>b</sup>	6,558	80.2	9.5 (5.0)	0.08 <sup>b</sup>
Yes	4,726	21.1	11.8 (5.4)		1,585	19.4	9.3 (5.1)	
Reported start date	4,516	95.6	11.7 (5.4)	0.0039 <sup>g</sup>	1,527	96.3	9.2 (5.1)	<0.005 <sup>g</sup>
Didn't report start date	210	4.4	12.9 (5.6)		58	3.7	11.2 (5.2)	
Missing	136	0.6	13.3 (5.0)	0.0007 <sup>c</sup>	33	0.4	10.7 (4.7)	0.15 <sup>c</sup>
Reported medication use <sup>h</sup>								
Acetaminophen								
No	5,799	25.9	11.9 (5.5)	0.0004 <sup>b</sup>	2,106	25.8	9.8 (5.1)	<0.005 <sup>b</sup>
Yes	16,567	74.1	11.6 (5.3)		6,070	74.2	9.4 (5.0)	
Reported start date	15,521	96.7	11.6 (5.3)	0.0038 <sup>g</sup>	5,724	94.3	9.3 (4.9)	<0.0001 <sup>g</sup>
Didn't report start date	1,046	6.3	12.1 (5.3)		346	5.7	10.7 (5.1)	
Ibuprofen								
No	15,193	67.9	11.8 (5.3)	0.2422 <sup>b</sup>	5,893	72.1	9.5 (5.1)	0.26 <sup>b</sup>
Yes	7,173	32.1	11.7 (5.4)		2,283	27.9	9.4 (4.8)	
Reported start date	6,877	95.9	11.6 (5.4)	0.1278 <sup>g</sup>	2,211	96.9	9.3 (4.8)	0.08 <sup>g</sup>
Didn't report start date	296	4.1	12.1 (5.2)		72	3.2	10.4 (5.6)	

Table continues

months longer than for controls. This is largely attributable to the fact that not all birth defects are diagnosed at birth. Although birth defects such as spina bifida or cleft lip are usually evident when an infant is born, certain heart defects and defects such as hydrocephalus or craniosynostosis may not be definitively diagnosed for weeks, months, or even years after birth; these cases cannot be ascertained by a surveillance system and assessed for eligibility in the study

until the diagnosis is finalized. Delays in birth defect surveillance procedures also contribute to a longer TTI for cases.

Controls became eligible for the study when they were born; however, a source of delay for control interviews was the availability of vital records. For both cases and controls, tracking and tracing of participants was often a multistep process that could vary by maternal demographic characteristics and could contribute substantially to the length of the

Table 2. Continued

Reported Pregnancy Exposure	Cases				Controls			
	No.	%	TTI, mean (SD)	P Value (2-sided) for Difference in Means	No.	%	TTI, mean (SD)	P Value (2-sided) for Difference in Means
Pseudoephedrine								
No	18,841	84.2	11.8 (5.4)	<0.0001 <sup>b</sup>	6,752	82.6	9.6 (5.1)	<0.0001 <sup>b</sup>
Yes	3,525	15.8	11.1 (5.1)		1,424	17.4	8.8 (4.7)	
Reported start date	3,276	92.9	11.0 (5.1)	<0.0001 <sup>g</sup>	1,341	94.2	8.7 (4.6)	0.01 <sup>g</sup>
Didn't report start date	249	7.1	12.4 (5.0)		83	5.8	10.1 (5.3)	
Opioid analgesics								
No	20,960	93.7	11.7 (5.4)	0.0093 <sup>b</sup>	7,752	94.8	9.5 (5.0)	<0.005 <sup>b</sup>
Yes	1,406	6.3	11.4 (5.4)		424	5.2	8.8 (4.8)	
Reported start date	1,382	98.3	11.4 (5.4)	0.5144 <sup>g</sup>	419	98.8	8.7 (4.8)	0.09 <sup>g</sup>
Didn't report start date	24	1.7	12.1 (6.0)		5	1.2	12.4 (4.5)	
Antidepressant medication								
No	21,038	94.1	11.7 (5.4)	0.7880 <sup>b</sup>	7,784	95.2	9.5 (5.0)	0.42 <sup>b</sup>
Yes	1,328	5.9	11.7 (5.5)		392	4.8	9.3 (4.9)	
Reported start date	1,309	98.6	11.7 (5.5)	0.0991 <sup>g</sup>	385	98.2	9.3 (4.9)	0.28 <sup>g</sup>
Didn't report start date	19	1.4	13.7 (4.9)		7	1.8	7.3 (3.7)	
Antiepileptic medication								
No	22,068	98.7	11.7 (5.4)	0.9469 <sup>b</sup>	8,097	99.0	9.5 (5.0)	0.36 <sup>b</sup>
Yes	298	1.3	11.7 (5.7)		79	1.0	9.0 (4.9)	
Reported start date	292	98.0	11.8 (5.7)	0.8026 <sup>g</sup>	78	98.7	8.9 (4.9)	0.15 <sup>g</sup>
Didn't report start date	6	2.0	11.2 (5.7)		1	1.3	16 (N/A)	

Abbreviations: SD, standard deviation; TTI, time to interview.

<sup>a</sup> Considered only in analysis of reports of no exposure and missing exposure outcome measures.

<sup>b</sup> P value for the difference in mean TTI among women who did and did not report the exposure.

<sup>c</sup> P value for the difference in mean TTI among women with and without missing exposure reports.

<sup>d</sup> Considered only in reports of the no-exposure outcome measure.

<sup>e</sup> The "gate" question regarding fertility treatment use was added to the questionnaire in 1998. Although women were asked to report fertility treatment use before that time, a missing response does not have the same interpretation.

<sup>f</sup> Considered in reports of no exposure, start month report, and missing report outcome measures.

<sup>g</sup> P value for difference in mean TTI among women who did and did not provide information on the start month of illness or medication use.

<sup>h</sup> Reported exposures for medication use were considered only in the start month report outcome measure.

TTI. Even once a mother was located, a considerable amount of time could have passed before the interviewer talked to the mother and completed the interview.

### Implications of reports of no exposure

There are 2 main reasons we could have observed an association between TTI and report of no exposure. The first is that mothers with longer TTI truly did not experience these exposures for reasons unrelated to TTI. We attempted to control for the factors that could confound the association between TTI and experiencing the exposures of interest, but there is always the possibility of residual confounding due to imprecise adjustment of measured factors or lack of adjustment for unmeasured factors.

The second is that mothers with longer TTI might be less likely to remember and thus report certain exposures. If there was recall bias such that mothers of case infants were

more likely to scrutinize their exposures during pregnancy than were mothers of control infants, we would expect to see a stronger association among controls than among cases.

We did not observe strong associations between TTI and reports of no exposure for the 5 exposures we assessed. However, we did observe modest statistically significant increased odds of a report of no exposure with increasing TTI that were larger for controls than for cases. Whether these associations are due to residual confounding or a true impact of TTI is not clear. Gestational age is an important potential confounder for this indicator of interview quality, as mothers with shorter pregnancies had less opportunity to experience the exposure of interest. Although we controlled for gestational age in our logistic regression models, we also conducted sensitivity analyses in which we stratified by gestational age ( $\geq 37$  weeks vs.  $< 37$  weeks). Results were not meaningfully different from those of our primary analysis.

**Table 3.** Adjusted<sup>a</sup> Odds Ratios for the Association Between Time to Interview and Report of No Exposure, National Birth Defects Prevention Study, 1997–2007

	Cases			Controls			P for Interaction <sup>b</sup>
	No.	aOR <sup>a</sup>	95% CI	No.	aOR <sup>a</sup>	95% CI	
No report of upper respiratory infection							
6 weeks–6 months	1,628	1.0	Referent	989	1.0	Referent	0.19
7–12 months	3,577	1.0	0.9, 1.1	1,212	1.0	0.9, 1.2	
13–18 months	2,212	1.0	0.9, 1.1	485	1.1	1.0, 1.3	
19–24 months	1,248	1.2	1.1, 1.3	259	1.4	1.1, 1.7	
No report of kidney, bladder, or urinary tract infection							
6 weeks–6 months	3,113	1.0	Referent	2,168	1.0	Referent	0.21
7–12 months	7,509	1.0	0.9, 1.1	2,820	1.0	0.9, 1.2	
13–18 months	4,554	1.0	0.9, 1.1	1,109	1.2	1.0, 1.4	
19–24 months	2,328	1.1	1.0, 1.2	461	1.2	0.9, 1.5	
No report of injury							
6 weeks–6 months	3,594	1.0	Referent	2,491	1.0	Referent	0.15
7–12 months	8,694	1.0	0.9, 1.1	3,216	1.2	1.0, 1.5	
13–18 months	5,331	1.0	0.8, 1.1	1,251	1.3	1.0, 1.8	
19–24 months	2,721	1.1	0.9, 1.3	534	1.4	1.0, 1.9	
No report of morning sickness							
6 weeks–6 months	1,122	1.0	Referent	704	1.0	Referent	0.03
7–12 months	2,838	1.1	1.0, 1.2	927	1.1	1.0, 1.2	
13–18 months	1,780	1.2	1.1, 1.3	405	1.3	1.1, 1.5	
19–24 months	915	1.2	1.1, 1.3	195	1.6	1.3, 2.0	
No report of assisted reproductive technology							
6 weeks–6 months	3,596	1.0	Referent	2,531	1.0	Referent	0.40
7–12 months	8,658	1.0	0.9, 1.2	3,161	1.1	0.8, 1.4	
13–18 months	5,415	0.9	0.7, 1.0	1,271	1.0	0.7, 1.4	
19–24 months	2,838	1.0	0.8, 1.3	568	1.7	0.9, 3.2	

Abbreviations: aOR, adjusted odds ratio; CI, confidence interval.

<sup>a</sup> Odds ratios were adjusted for maternal age, maternal race/ethnicity, maternal educational level, annual family income, maternal parity, gestational age, pregnancy outcome, study center, and language of interview.

<sup>b</sup> P value for likelihood ratio test of interaction parameters for case status and time to interview in the logistic regression model.

### Implications of absence of report of start month of exposure or a missing exposure response

Whether the start month of the reported exposure could be determined and whether the exposure response was missing are likely more indicative of interview quality than was a report of no exposure. A report of no exposure might or might not be accurate, but missing data are a direct marker of reduced interview quality.

In most NBDPS analyses, we are interested in exposures that occurred during a specific time period, such as the first trimester, which is when most organ systems are forming. If we cannot determine the start month, that subject will be excluded from the analysis because of missing data. The mothers with missing responses to the exposure questions are also not able to be included in analyses.

Missingness is almost inevitable in retrospective studies of complex exposures. It is therefore important for women to be given the option to respond “don’t know,” because forcing them to respond “yes” or “no” could lead to misclassification. However, it is also important to remember that our observation of no meaningful association between TTI and the odds of having missing information does not imply that exclusion of subjects with missing data will lead to unbiased results. It is always important that analysts consider the ramifications of implied yet unverifiable assumptions inherent in the analysis of data with missing values (14).

### Conclusion

Although an inclusion period of up to 2 years after a woman’s EDD allows for a long TTI, it also offers the



**Table 4.** Odds Ratios for the Association Between Time to Interview and Lack of Report of Start Month for Reported Illness or Medication Use, National Birth Defects Prevention Study, 1997–2007

Time to Interview	Cases			Controls			P for Interaction <sup>b</sup>
	No.	aOR <sup>a</sup>	95% CI	No.	aOR <sup>a</sup>	95% CI	
No report of start month of respiratory illness							0.91
6 weeks–6 months	207	1.0	Referent	151	1.0	Referent	
7–12 months	686	1.2	1.0, 1.4	262	1.3	1.0, 1.7	
13–18 months	536	1.2	1.0, 1.4	113	1.3	0.9, 1.8	
19–24 months	261	1.3	1.1, 1.7	47	1.3	0.8, 1.9	
No report of start month of kidney, bladder, or urinary tract infection							0.74
6 weeks–6 months	27	1.0	Referent	13	1.0	Referent	
7–12 months	86	1.4	0.9, 2.2	24	1.7	0.8, 3.6	
13–18 months	60	1.6	1.0, 2.6	16	2.8	1.2, 6.7	
19–24 months	37	1.8	1.0, 3.0	5	1.9	0.6, 6.1	
No report of start month of acetaminophen use							0.04
6 weeks–6 months	157	1.0	Referent	76	1.0	Referent	
7–12 months	446	1.1	0.9, 1.3	162	1.3	0.9, 1.9	
13–18 months	301	1.0	0.8, 1.3	73	1.9	1.3, 2.9	
19–24 months	142	1.2	0.9, 1.5	35	2.2	1.3, 3.6	
No report of start month of ibuprofen use							0.04
6 weeks–6 months	39	1.0	Referent	22	1.0	Referent	
7–12 months	135	1.5	1.0, 2.2	31	1.1	0.6, 2.0	
13–18 months	87	1.5	1.0, 2.3	9	0.8	0.4, 1.9	
19–24 months	35	1.3	0.8, 2.1	10	2.9	1.3, 6.6	
No report of start month of pseudoephedrine use							0.93
6 weeks–6 months	28	1.0	Referent	24	1.0	Referent	
7–12 months	112	1.7	1.1, 2.7	36	1.6	0.9, 2.9	
13–18 months	78	2.3	1.5, 3.7	16	2.0	1.0, 4.1	
19–24 months	31	2.2	1.3, 3.8	7	2.4	0.9, 6.4	
	<b>No.</b>	<b>cOR</b>	<b>95% CI</b>				
No report of start month of opioid analgesic use <sup>c</sup>							
6 weeks–6 months	5	1.0	Referent				
7–12 months	12	1.4	0.5, 4.1				
13–18 months	8	1.9	0.6, 5.7				
19–24 months	4	1.9	0.5, 7.2				
No report of start month of antidepressant medication use <sup>c</sup>							
6 weeks–6 months	3	1.0	Referent				
7–12 months	11	1.9	0.5, 6.8				
13–18 months	9	2.9	0.8, 10.8				
19–24 months	3	1.8	0.4, 9.1				
No report of antiepileptic medication use <sup>c</sup>							
6 weeks–6 months	1	1.0	Referent				
7–12 months	4	2.6	0.3, 23.8				
13–18 months	1	1.1	0.1, 18.2				
19–24 months	1	1.8	0.1, 29.2				

Abbreviations: aOR, adjusted odds ratio; CI, confidence interval; cOR, crude odds ratio.

<sup>a</sup> Odds ratios were adjusted for maternal age, maternal race/ethnicity, maternal educational level, annual family income, maternal parity, gestational age, pregnancy outcome, study center, and language of interview.

<sup>b</sup> P value for likelihood ratio test of interaction parameters for case status and time to interview in the logistic regression model.

<sup>c</sup> Nonstratified crude odds ratios are presented because of the small sample size.

**Table 5.** Crude Odds Ratios for the Association Between Time to Interview and Missing Exposure Reports, National Birth Defects Prevention Study, 1997–2007

	Cases			Controls			P for Interaction <sup>a</sup>
	No.	cOR	95% CI	No.	cOR	95% CI	
Missing report of respiratory illness							0.0003
6 weeks–6 months	169	1.0	Referent	75	1.0	Referent	
7–12 months	518	1.3	1.1, 1.5	162	1.7	1.3, 2.3	
13–18 months	361	1.5	1.2, 1.8	98	2.8	2.1, 3.8	
19–24 months	199	1.6	1.3, 2.0	31	2.0	1.3, 3.1	
Missing report of kidney, bladder, or urinary tract infection							0.35
6 weeks–6 months	12	1.0	Referent	6	1.0	Referent	
7–12 months	52	1.8	1.0, 3.4	16	2.1	0.8, 5.4	
13–18 months	48	2.7	1.4, 5.1	9	3.1	1.1, 8.7	
19–24 months	24	2.7	1.3, 5.3	2	1.6	0.3, 7.9	
Missing report of injury <sup>b</sup>							
6 weeks–6 months	3	1.0	Referent				
7–12 months	4	0.7	0.2, 3.1				
13–18 months	1	0.3	0.03, 3.0				
19–24 months	2	1.3	0.2, 7.5				
Missing report of morning sickness <sup>b</sup>							
6 weeks–6 months	3	1.0	Referent				
7–12 months	8	1.4	0.4, 5.2				
13–18 months	7	2.2	0.6, 8.4				
19–24 months	5	3.1	0.8, 13.1				

Abbreviations: CI, confidence interval; cOR, crude odds ratio.

<sup>a</sup> P value for likelihood ratio test of interaction parameters for case status and time to interview in the logistic regression model.

<sup>b</sup> Nonstratified odds ratios are presented because of the small sample size.

advantage of more complete case ascertainment, which is important in a population-based study such as the NBDPS. Our results suggest that increasing TTI may be associated with modest decreases in the quality of data reported during a maternal interview and that this decrease in interview quality might be more pronounced for mothers of control subjects than for mothers of case subjects. Although recall bias in birth defect studies may be a larger theoretical concern than is warranted in practice (15), it is important to continually assess this potential source of bias and to understand how it could influence study results.

A recent publication from a different birth defect case-control study (16) described a certainty algorithm that can be used to classify responses according to the timing of exposures (i.e., medications) in time windows of interest. The authors observed that when results were stratified by certainty, those with a “likely” exposure showed a stronger association than when those with “likely” and “possibly” exposures were combined. This method could potentially be modified to incorporate the TTI into the classification of exposure window certainty.

Although the present analysis focuses on data from a specific study, the NBDPS, the results should be informative to other studies of reproductive outcomes that use retrospective exposure assessment. The advantages of using NBDPS data

include the large data set covering several years and diverse regions of the United States. The NBDPS has resulted in over 120 publications on potential associations between many different exposures and birth defects and other adverse reproductive outcomes, so understanding the implication of TTI on even just this study has important implications. Despite its strengths, this analysis is severely limited by a lack of external validation data of self-reported exposures, such as medical record reviews or follow-up interviews. In addition, we examined only a small fraction of the exposures about which women were queried on the NBDPS questionnaire. Different patterns may have emerged if we had considered different exposures, although we selected several of the most important exposures assessed in NBDPS. Although we assessed only a small number of the exposures about which women were questioned in the NBDPS interview, we conducted many statistical tests, and some of our results are therefore likely due to chance. In addition, interviewers were not blinded to case status, although they are not specifically told whether an interview pertained to a case or a control. Also, a standardized interview instrument was used in an effort to reduce introduction of recall bias through interviewer probing. We combined all birth defect case groups in our stratified analyses, and it is possible that defect-specific associations may vary. Over 30 different defect categories are

included in NBDPS, and estimation of defect-specific associations for the many metrics we assessed was untenable.

Studies of potential causes of birth defects are vital for ensuring the best health for women and their infants. Because birth defects are rare outcomes and early pregnancy exposures are difficult to measure, studies with retrospective exposure assessment are typically the best design for these studies. Therefore, it is important for researchers of birth defects to have an understanding of the limitations of their data collection methods and the potential impact of these limitations on the results of their studies and interpretation of those results. According to the data presented here, the methods used for the NBDPS are likely to result in limited bias.

## ACKNOWLEDGMENTS

Author affiliations: National Center on Birth Defects and Developmental Disabilities, US Centers for Disease Control and Prevention, Atlanta, Georgia (Sarah C. Tinker, Owen J. Devine, Krista S. Crider, Jennita Reefhuis); Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, Georgia (Cassandra Gibbs); Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, Georgia (Matthew J. Strickland); Boston University School of Public Health, Slone Epidemiology Center, Boston, Massachusetts (Martha M. Werler); and Massachusetts Department of Public Health, Boston, Massachusetts (Marlene T. Anderka).

This work was supported through cooperative agreements under Program Announcement 96043, Program Announcement 02081, and Funding Opportunity Announcement DD09-001 from the Centers for Disease Control and Prevention to the Centers for Birth Defects Research and Prevention that participated in the National Birth Defects Prevention Study. Coding of drug information in the National Birth Defects Prevention Study used the Slone Drug Dictionary, under license from the Slone Epidemiology Center at Boston University, Boston, Massachusetts.

These results have been presented at the 27th International Conference on Pharmacoepidemiology and Therapeutic Risk Management, Chicago, Illinois, August 14–17, 2011; the 3rd North American Congress of Epidemiology, Montreal, Quebec, Canada, June 21–24, 2011; and the 24th Annual Meeting of the Society for Pediatric and Perinatal Epidemiologic Research, Montreal, Quebec, Canada, June 20–21, 2011.

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

Conflict of interest: none declared.

## REFERENCES

1. Ayoola AB, Nettleman MD, Stommel M, et al. Time of pregnancy recognition and prenatal care use: a population-based study in the United States. *Birth*. 2010; 37(1):37–43.
2. Finer LB, Henshaw SK. Disparities in rates of unintended pregnancy in the United States, 1994 and 2001. *Perspect Sex Reprod Health*. 2006;38(2):90–96.
3. Elkadry E, Kenton K, White P, et al. Do mothers remember key events during labor? *Am J Obstet Gynecol*. 2003;189(1): 195–200.
4. Hopkins LM, Caughey AB, Brown JS, et al. Concordance of chart abstraction and patient recall of intrapartum variables up to 53 years later. *Am J Obstet Gynecol*. 2007;196(3):233. e1–6.
5. Quigley MA, Hockley C, Davidson LL. Agreement between hospital records and maternal recall of mode of delivery: evidence from 12 391 deliveries in the UK Millennium Cohort Study. *BJOG*. 2007;114(2): 195–200.
6. Catov JM, Newman AB, Kelsey SF, et al. Accuracy and reliability of maternal recall of infant birth weight among older women. *Ann Epidemiol*. 2006;16(6):429–431.
7. Li CY, Wei JN, Lu TH, et al. Mothers tended to overreport categorical infant birth weight of their children. *J Clin Epidemiol*. 2006;59(12):1319–1325.
8. Lucia VC, Luo Z, Gardiner JC, et al. Reports of birthweight by adolescents and their mothers: comparing accuracy and identifying correlates. *Paediatr Perinat Epidemiol*. 2006; 20(6):520–527.
9. Buka SL, Goldstein JM, Spartos E, et al. The retrospective measurement of prenatal and perinatal events: accuracy of maternal recall. *Schizophr Res*. 2004;71(2–3): 417–426.
10. 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.
11. 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.
12. Yoon PW, Rasmussen SA, Lynberg MC, et al. The National Birth Defects Prevention Study. *Public Health Rep*. 2001; 116(suppl 1):32–40.
13. Cogswell ME, Bitsko RH, Anderka M, et al. Control selection and participation in an ongoing, population-based, case-control study of birth defects: the National Birth Defects Prevention Study. *Am J Epidemiol*. 2009;170(8): 975–985.
14. Daniels M, Hogan J. *Missing Data in Longitudinal Studies: Strategies for Bayesian Modeling and Sensitivity Analysis*. New York: CDC Chapman and Hall; 2008.
15. Khoury MJ, Erickson JD, James LM. Etiologic heterogeneity of neural tube defects: clues from epidemiology. *Am J Epidemiol*. 1982;115(4):538–548.
16. Yau WP, Lin KJ, Werler MM, et al. Drug certainty-response in interview-based studies. *Pharmacoepidemiol Drug Saf*. 2011;20(11):1210–1216.