Published in final edited form as:

Paediatr Perinat Epidemiol. 2011 November ; 25(6): 540–548. doi:10.1111/j.1365-3016.2011.01226.x.

Self-report versus medical record – perinatal factors in a study of infant leukaemia: a study from the Children’s Oncology Group

Anne M. Jurek^{a,*}, Sander Greenland^b, Logan G. Spector^c, Michelle A. Roesler^c, Leslie L. Robison^d, and Julie A. Ross^c

^aExponent Inc. Health Sciences, Washington, DC

^bDepartments of Epidemiology and Statistics, University of California, Los Angeles, CA

^cDivision of Pediatric Epidemiology and Clinical Research, University of Minnesota, Minneapolis, MN

^dDepartment of Epidemiology and Cancer Control, St. Jude Children’s Research Hospital, Memphis, TN, USA

Summary

In a case–control study of infant leukaemia, we assessed agreement between medical records and mother’s self-reported pregnancy-related conditions and procedures and infant treatments. Interview and medical record data were available for 234 case and 215 control mothers. Sensitivity, specificity and predictive values for maternal report were estimated for case and control mothers separately, taking the medical record as correct. For most perinatal conditions, sensitivity and specificity were over 75%. Low sensitivity was observed for maternal protein or albumin in the urine (cases: 12% [95% exact confidence interval (CI) 8%, 18%]; controls: 11% [95% CI 7%, 17%]) and infant supplemental oxygen use (cases: 25% [95% CI 11%, 43%]; controls: 24% [95% CI 13%, 37%]). Low specificity was found for peripheral oedema (cases: 47% [95% CI 37%, 58%]; controls: 54% [95% CI 43%, 64%]). Sensitivity for maternal hypertension appeared much lower for cases (cases: 46% [95% CI 28%, 66%]; controls: 90% [95% CI 70%, 99%]; $P = 0.003$). We did not detect other case–control differences in recall (differentiality), even though the average time between childbirth and interview was 2.7 years for case and 3.7 years for control mothers. Many conditions exhibited notable differences between interview and records. We recommend use of multiple measurement sources to allow both cross-checking and synthesis of results into more accurate measures.

Keywords

medical records; maternal recall; reproducibility; misclassification; recall bias

© 2011 Blackwell Publishing Ltd.

Correspondence: Dr Anne M. Jurek, Senior Scientist, 1150 Connecticut Avenue N.W., Suite 1100, Washington, DC 20036, USA. gjurek@exponent.com.

^{*}Previously at the University of Minnesota Department of Pediatrics.

The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Cancer Institute or the National Institutes of Health.

Conflict of interest: none declared.

Introduction

Case-control studies are often the only feasible design for epidemiological research on the aetiology of childhood cancer. It is unclear, however, how much mis-classification exists due to inaccurate maternal recall in these studies. In some instances, the index pregnancy occurred years ago and therefore recollecting past experiences may be difficult and subject to considerable error. To our knowledge, two case-control studies of childhood cancer have evaluated the validity of maternal recall of perinatal conditions. The first study included children aged 9 years or less diagnosed with leukaemia between 1980 and 1993;¹ the second study included children diagnosed 18 months or less between 1983 and 1988.²

Medical records are often taken as a 'gold standard' to assess the accuracy of self-reported data although such records may have errors. Here, we report results for agreement between maternal reports and medical records in a North American study of infant leukaemia.

Methods

The current study uses data from a Children's Oncology Group (COG) case-control study of infant leukaemia (COG-AE24) with a particular focus on maternal and infant characteristics as risk factors.³⁻⁵ The initial phase of the study included infants diagnosed with leukaemia (cases) during the first year of life at a COG institution between 1 January 1996 and 20 August 2002. Noncase (control) mothers were identified by random digit dialling and frequency matched on child's birth year. Of those eligible, 240 out of 348 (69%) case and 254 out of 430 (59%) noncase mothers participated in a telephone interview about events during and after pregnancy, including medical conditions, menstrual, hormone-use and contraceptive history, labour and delivery, and infant birth characteristics and hospitalisation.

Mothers were also asked for consent to obtain all hospital and health care provider records beginning 6 months prior to conception of the index child until the end of 1 week after birth or discharge from hospital (whichever occurred later). Records from all providers seen during pregnancy and delivery were requested. Providers were re-contacted if record information was incomplete. Medical records for 234 (98%) case and 215 (85%) noncase mothers who completed the telephone interview were obtained. Of those six case mothers without any medical record information, five refused to release medical records and one mother's medical record was not in English and therefore not requested. Among noncase mothers, medical records were not available for 39 (34 refused to release medical records and five consented but medical records were not received after request and follow-up or incorrect records were received). Written informed consent was obtained before the telephone interview and written release of medical records was obtained after the interview. Each of the participating COG institutional review boards approved the study.

For simplicity and brevity, we restricted maternal recall to dichotomies included in the telephone interview whose content was similar to information expected in the medical record. In calculating measures of agreement, the medical record was taken as correct. The following conditions were included: gestational diabetes, hypertension, maternal protein or albumin in urine, peripheral oedema, toxemia or pre-eclampsia, ultrasound, Caesarean section, and whether infant received phototherapy, blood products, supplemental oxygen, or was moved for special treatment. All medical conditions were identified as being mentioned in a medical record; none were based on diagnostic criteria or defined values. If a condition was not specifically noted or ruled out in a medical record, we coded it as not recorded/not abstracted and therefore not included in calculations.

Sensitivity was estimated by taking self-reported presence of the condition as a percentage of mothers having the condition indicated in the medical record. Specificity was estimated by taking the self-reported absence of the condition as a percentage of mothers not having the condition in the medical record. Both were calculated separately for mothers of cases and noncases. To test the hypothesis of nondifferentiality (which is no association of interview data with childhood cancer conditional on the medical record data), we computed an exact P -value for the association of case-control status with telephone interview data stratified on the medical record data for each perinatal condition.⁶ We also estimated positive and negative predictive values, stratified on the telephone interview data, for mothers of cases and noncases separately. Positive predictive values were estimated as the percentage showing the condition in the medical record among mothers reporting the condition in the interview. Negative predictive values were estimated as the percentage not showing the condition in the medical record among mothers not reporting the condition in the interview. Predictive values depend on exposure prevalences, which can vary across study populations.⁷ Nonetheless, we have included them because they illustrate directly the limitations of maternal report in this study.

Lastly, we calculated odds ratios (ORs) and 95% confidence intervals (CIs) for the association between mother and infant conditions and the odds of infant leukaemia for the maternal telephone interview and medical record data separately, provided each cell in the 2×2 table exceeded 3. Telephone interview ORs were also adjusted for length of time in years between childbirth and telephone interview (continuous variable); due to small cells for gestational diabetes and toxemia/pre-eclampsia, for their regressions we added two extra records (one case, one control) to stabilise the estimates, as described in Greenland (p. 526).⁸ We compared telephone interview and medical record ORs using nonparametric bootstrap methods,⁹ with 10 000 data resamplings to estimate the standard errors of the log ORs and log ratio of ORs; we added 1/2 to each cell when resampling produced a zero cell.

Results

Most mothers provided information during the telephone interview. There were 14 missing responses for the telephone interview, and one control mother responded 'don't know' for gestational diabetes. Some perinatal conditions had low prevalences in medical records. Nonetheless, we calculated sensitivities, specificities and predictive values for these conditions because they have been previously linked aetiologically to infant leukaemia.

Table 1 provides the characteristics of the mothers who completed the telephone interview and for whom at least one medical record (prenatal or labour and delivery) was obtained. Mothers for cases and non-cases were primarily married and White, with a higher proportion of case mothers being Hispanic (11% vs. 3%). Noncase mothers were somewhat more educated and reported higher household incomes than case mothers. Case mothers were younger than noncase mothers at the time of the telephone interview. Further, the average length of time between the index child's birth and telephone interview was 2.7 and 3.7 years (37% longer) for case and noncase mothers, respectively. Over 90% of case and noncase mothers consented to the release of prenatal records and labour and delivery records. The median number of prenatal visits was 12 for 220 case mothers (range 3–22) and 12 for 206 noncase mothers (range 1–18).

Table 2 provides the numbers for maternal report and medical records for the selected conditions, along with P -values testing nondifferentiality. Table 3 provides the estimated sensitivities, specificities, positive predictive values and negative predictive values comparing medical records and mothers' self-reports of maternal and infant characteristics. Caesarean section was perfectly classified. For most other perinatal factors of interest,

sensitivity and specificity were over 75%. Nonetheless, maternal protein or albumin in the urine and infant supplemental oxygen had estimated sensitivities less than 26%, while peripheral oedema had estimated specificities less than 60%. Case and control sensitivities appeared to differ for maternal hypertension (cases: 46% [95% CI 28%, 66%]; noncases: 90% [95% CI 70%, 99%]; Fisher's exact $P = 0.003$). Other systematic case-control differences in recall (differentiality, Table 2) were not detected despite an average time between childbirth and telephone interview that was 1 year longer for noncase mothers (Table 1). Nevertheless, except for maternal hypertension, the direction of differentiality remains ambiguous and nondifferentiality remains a possibility.

High predictive values were estimated for Caesarean section and ultrasound, but only four noncase mothers reported not having an ultrasound (Table 3). Negative predictive values for all pregnancy complications except protein or albumin in urine were greater than 75%. Positive predictive values were less than 75% for all conditions except protein or albumin in urine. For both case and noncase mothers, predictive values for infant received supplemental oxygen and infant moved for special treatment were greater than 60%.

Telephone interview data exhibited inverse associations with leukaemia for hypertension and infant received supplemental oxygen, whereas these perinatal conditions were positively associated using medical record data (Table 4). Protein or albumin in urine was inversely associated for medical records but positively associated for telephone interview data. Gestational diabetes was inversely associated for interview and medical record data, whereas increased odds for leukaemia were estimated for all other conditions in both data sources. Adjusting interview data for length of time between index child's birth and telephone interview made little difference, although there was a minor reversal in association for peripheral oedema and a near null result for gestational diabetes (Table 4). Comparison of unadjusted ORs for hypertension suggested that the maternal telephone interview and medical record ORs differed (bootstrap ratio of unadjusted interview OR vs. unadjusted medical record OR: 0.53 [95% CI 0.30, 0.95]; P -value = 0.03).

Discussion

We compared maternal telephone interview data with medical record data to assess maternal recall in a case-control study of infant leukaemia. Case and noncase mothers perfectly recalled having a Caesarean section. All other medical conditions we studied exhibited differences between telephone interview responses and medical records, albeit the extent of these differences was estimated imprecisely. Different data sources for maternal and infant characteristics led to qualitative differences in some instances, which appeared especially marked for hypertension.

Our analyses use slightly different sets of mothers for the record-based and interview-based estimates due to missing and 'don't know' responses. Hence, our comparisons represent the total difference between the two, combining inclusion effects (for medical record vs. telephone interview) with recall effects. Due to the minor inclusion differences, however, the inclusion effect appeared negligible; e.g. when we recomputed the hypertension ORs using only mothers with both medical record and interview data, the bootstrap ratio of unadjusted ORs (interview to medical record) became 0.58 [95% CI 0.33, 1.01] (vs. 0.53 [95% CI 0.30, 0.95] in Table 4). Therefore, our discussion focuses on recall effects.

Despite the 37% longer time between birth and interview among noncases, we did not see evidence of systematically poorer recall among noncase mothers. Hypertension was also the only condition for which a clear direction of differentiality seemed apparent, with half the recall of actual hypertension among cases compared to noncases (sensitivity = 46% for cases

vs. 90% for noncases, $P = 0.003$). Nonetheless, given the number of comparisons made, this result could easily be a chance artefact. In fact, the exact test of nondifferentiability for maternal hypertension recall (which simultaneously compares the sensitivities and specificities) gave $P = 0.11$, meaning that even the direction of differentiability is not well determined by our data.

We know of only two published validation studies of maternal risk factors for childhood leukaemia.^{1,2} Infante-Rivard and Jacques¹ used mother's self-reported responses as the gold standard for validating radiographic examinations during pregnancy, whereas our gold standard measure was the medical record. Both studies, however, estimated that sensitivity for ultrasound exceeded 95%. On the other hand, Olson *et al.*² focused on validating birthweight, reproductive history and maternal medical procedures also using medical records as the gold standard. They coded unrecorded medical record values as if the patient did not have the condition, an alternative approach to our method of coding them as not recorded/not abstracted. As a result, differences between those results and ours are not unexpected. Unfortunately, there was not sufficient information in the paper by Olson and colleagues to allow for calculation of P -values or CIs for these differences. Nonetheless, in accord with our findings, Olson *et al.* observed perfect recall of Caesarean section and very high recall for ultrasound examinations. Specificities were mostly similar to ours except for peripheral oedema in cases (63% in Olson *et al.* vs. 47% here). Sensitivities appeared most different for child received oxygen (cases: 64% vs. 25%; noncases 70% vs. 24%), maternal hypertension (cases: 75% vs. 46%; noncases: 59% vs. 90%) and peripheral oedema (cases: 69% vs. 86%; noncases: 68% vs. 89%). Among noncases alone, differences in sensitivities appeared for diabetes mentioned in the medical record (60% vs. 83%) and child moved for special treatment (67% vs. 81%), and among cases alone for toxemia or pre-eclampsia (75% vs. 47%).

Maternal recall of perinatal conditions has been explored in mothers with healthy children and other illnesses.^{10–17} In a case–control study of Sudden Infant Death Syndrome, Drews *et al.*¹⁰ found Caesarean section to have high sensitivity (cases = 100%; non-cases = 97%) and maternal hypertension to have low sensitivity (cases = 53%; noncases = 61%). Other studies^{11–17} also suggest that recall of many perinatal events is not perfect, but that method of delivery has high agreement.

The data for our analyses are from the largest case–control study of infant leukaemia to date. Cases and noncases resided throughout the United States and Canada and were diagnosed at COG institutions, providing a near population-based study.^{18,19} Nonetheless, our analysis involved only 234 out of 348 (67%) eligible case mothers and 215 out of 430 (50%) eligible noncase mothers. These limited participation rates reflect the fact that our analyses required both completion of telephone interview and release of medical records. Consequently, our findings may not reflect the results that would have been found had everyone in our target population participated. Moreover, as with all studies we cannot be sure that associations we observed (even if they are valid) generalise to other populations.

Additional data sources are rarely available to check interview responses. Perinatal medical records have the advantage of being collected prior to the diagnosis of cancer, thereby eliminating much recall bias. However, medical records can be difficult and expensive to collect. There is also unease among participants (especially healthy individuals) to permit collection of medical records due to privacy and confidentiality issues.²⁰ In addition, medical records can have errors. In fact, some maternal and infant characteristics may be more accurate using self-reported data than medical records (e.g. morning sickness).²

In addition to maternal recall bias, our agreement study may have other sources of error present. Our study depends on completeness as well as accuracy of medical records. Lack of information recorded in a medical record does not necessarily imply that a medical condition was not discussed or medically tested, and it is important to distinguish lack of recording from lack of the condition. This can be accomplished by checking each medical record for method of data capture (e.g. specific checkbox acknowledging condition vs. open form to write in the condition), a task unfortunately not feasible in this study due to the multiplicity of forms from different hospitals. Failure to record items would result in some mothers being classified as negative when actually positive (false negatives), and others being classified as positive who were actually negative (false positives).

Another possible source of error concerns the fact that the medical record data in our study were abstracted only once by two nurse abstractors. Although both nurse abstractors were trained and blinded to case-control status, we cannot rule out reporting errors that could have occurred during the abstraction process. No formal quality control process was implemented; however, data checks were performed during data cleaning. Nonetheless, we would expect these reporting errors to be sporadic rather than systematic, and more often in the form of failure to note. Despite such errors, medical records can still identify problematic recall items, especially if they exhibit conditions that were not reported by the mother (indicating that recall failures are present).

Despite such errors, sensitivity and specificity estimates from agreement studies like ours can be used as starting points for sensitivity and bias analyses of recall error.^{6,7,21} The results of our study and previous validation studies^{2,10–17} can also be useful when designing future case-control studies that will collect perinatal information. While medical procedures such as Caesarean sections and ultrasounds may be recalled accurately, maternal reports of pregnancy and neonatal complications appear to be much less reliable, especially if they are more occult conditions with technical names (e.g. proteinuria or toxemia). Hence, instead of relying on only one data source for perinatal events, multiple measurement tools may be necessary to construct more accurate perinatal measurements. Therefore, agreement studies such as ours can also guide which data source should be used for gathering exposure data for future epidemiological case-control studies, and suggest ways to combine measurements into a single more accurate summary.

Agreement and validation substudies are valuable but need some forethought to produce useful information for designing future studies and analyses. Ideally, these substudies would be included in the initial study proposals to allow their review before implementation. When designing substudies, we recommend identifying information collected by the best available measurement. Then, use this information to design the tool that is used to obtain the primary measurements (e.g. a telephone interview). For instance, one could gather medical records to understand the type of data collected by the best available measurement. One would then include interview questions that match the content of the best available measurements so that agreement can be examined whenever the latter can be obtained.

Acknowledgments

This research was supported by NIH R03 CA141403, R01 CA79940, U10 CA13539, U10 CA98543, U10 CA98413, the Children's Cancer Research Fund (Minneapolis) and NIH P30 CA77598 (University of Minnesota Masonic Cancer Center shared resource: Health Survey Research Center). The authors would like to thank the anonymous reviewers for helpful comments on a previous draft.

References

1. Infante-Rivard C, Jacques L. Empirical study of parental recall bias. *American Journal of Epidemiology*. 2000; 152:480–486. [PubMed: 10981463]
2. Olson JE, Shu XO, Ross JA, Pendergrass T, Robison LL. Medical record validation of maternally reported birth characteristics and pregnancy-related events: a report from the Children's Cancer Group. *American Journal of Epidemiology*. 1997; 145:58–67. [PubMed: 8982023]
3. Spector LG, Xie Y, Robison LL, Heerema NA, Hilden JM, Lange B, et al. Maternal diet and infant leukemia: the DNA topoisomerase II inhibitor hypothesis: a report from the Children's Oncology Group. *Cancer Epidemiology, Biomarkers & Prevention*. 2005; 14:651–655.
4. Peters AM, Blair CK, Verneris MR, Neglia JP, Robison LL, Spector LG, et al. Maternal hemoglobin concentration during pregnancy and risk of infant leukaemia: a Children's Oncology Group study. *British Journal of Cancer*. 2006; 95:1274–1276. [PubMed: 17003777]
5. Johnson KJ, Roesler MA, Linabery AM, Hilden JM, Davies SM, Ross JA. Infant leukemia and congenital abnormalities: a Children's Oncology Group study. *Pediatric Blood & Cancer*. 2010; 55:95–99. [PubMed: 20486175]
6. Greenland S. Bayesian perspectives for epidemiologic research: III. Bias analysis via missing-data methods. *International Journal of Epidemiology*. 2009; 38:1662–1673. [PubMed: 19744933]
7. Greenland, S.; Lash, TL. Bias analysis. Ch. 19. In: Rothman, KJ.; Greenland, S.; Lash, TL., editors. *Modern Epidemiology*. 3. Philadelphia, PA: Lippincott Williams & Wilkins; 2008. p. 345–380.
8. Greenland S. Invited commentary: variable selection versus shrinkage in the control of multiple confounders. *American Journal of Epidemiology*. 2008; 167:523–529. [PubMed: 18227100]
9. Carpenter J, Bithell J. Bootstrap confidence intervals: when, which, what? A practical guide for medical statisticians. *Statistics in Medicine*. 2000; 19:1141–1164. [PubMed: 10797513]
10. Drews CD, Kraus JF, Greenland S. Recall bias in a case-control study of Sudden Infant Death Syndrome. *International Journal of Epidemiology*. 1990; 19:405–411. [PubMed: 2376455]
11. Casey R, Rieckhoff M, Beebe SA, Pinto-Martin J. Obstetric and perinatal events: the accuracy of maternal report. *Clinical Pediatrics*. 1992; 31:200–204. [PubMed: 1563192]
12. Lederman SA, Paxton A. Maternal reporting of prepregnancy weight and birth outcome: consistency and completeness compared with the clinical record. *Maternal and Child Health Journal*. 1998; 2:123–126. [PubMed: 10728268]
13. Yawn BP, Suman VJ, Jacobsen SJ. Maternal recall of distant pregnancy events. *Journal of Clinical Epidemiology*. 1998; 51:399–405. [PubMed: 9619967]
14. Buka SL, Goldstein JM, Spartos E, Tsuang MT. The retrospective measurement of prenatal and perinatal events: accuracy of maternal recall. *Schizophrenia Research*. 2004; 71:417–426. [PubMed: 15474913]
15. 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. *British Journal of Obstetrics and Gynaecology*. 2007; 114:195–200. [PubMed: 17166217]
16. Rice F, Lewis A, Harold G, van den Bree M, Boivin J, Hay DF, et al. Agreement between maternal report and antenatal records for a range of pre and perinatal factors: the influence of maternal and child characteristics. *Early Human Development*. 2007; 83:497–504. [PubMed: 17071023]
17. Elliott JP, Desch C, Istwan NB, Rhea D, Collins AM, Stanziano GJ. The reliability of patient-reported pregnancy outcome data. *Population Health Management*. 2010; 13:27–32. [PubMed: 20158321]
18. Ross JA, Severson RK, Pollock BH, Robison LL. Childhood cancer in the United States. A geographical analysis of cases from the Pediatric Cooperative Clinical Trials groups. *Cancer*. 1996; 77:201–207. [PubMed: 8630931]
19. Liu L, Krailo M, Reaman GH, Bernstein L. Childhood cancer patients' access to cooperative group cancer programs: a population-based study. *Cancer*. 2003; 97:1339–1345. [PubMed: 12599243]
20. Kass NE, Natowicz MR, Hull SC, Faden RR, Plantinga L, Gostin LO, et al. The use of medical records in research: what do patients want? *The Journal of Law, Medicine & Ethics*. 2003; 31:429–433.

21. Lash, TL.; Fox, MP.; Fink, AK. Applying Quantitative Bias Analysis to Epidemiologic Data. Dordrecht: Springer; 2009.

Table 1

Characteristics of mothers from telephone interview by case status

Characteristic	Case mothers (n = 234) n (%)	Noncase mothers (n = 215) n (%)	P-value
Race			
White, not of Hispanic origin	186 (79)	187 (87)	<0.01
African American or Black, not of Hispanic origin	5 (2)	11 (5)	
Hispanic	25 (11)	6 (3)	
Native American Indian or Alaskan Native	8 (3)	2 (1)	
Asian, Asian American, or Pacific Islander	7 (3)	6 (3)	
Other	3 (1)	3 (1)	
Highest educational level			
11 years or less	24 (10)	6 (3)	0.05
High school graduate	63 (27)	58 (27)	
Post high school, vocational or business school	17 (7)	22 (10)	
Some college but no degree	52 (22)	52 (24)	
College graduate	59 (25)	57 (27)	
Advanced degree	19 (8)	20 (9)	
Combined annual household income ^a			
Up to \$10 000	18 (8)	13 (6)	0.08
>\$10 000 up to \$20 000	34 (15)	17 (8)	
>\$20 000 up to \$30 000	34 (15)	37 (17)	
>\$30 000 up to \$40 000	31 (13)	30 (14)	
>\$40 000 up to \$50 000	32 (14)	22 (10)	
>\$50 000 up to \$75 000	49 (21)	45 (21)	
>\$75 000	34 (15)	51 (24)	
Marital status			
Married	194 (83)	178 (83)	0.84
Living with someone in a marriage-like relationship	14 (6)	12 (6)	
Separated	4 (2)	7 (3)	
Divorced	5 (2)	5 (2)	
Widowed	1 (<1)	2 (1)	
Never married	16 (7)	11 (5)	
	Mean [SD]	Mean [SD]	
Age in years at index child's birth	29.1 [5.5]	30.0 [5.5]	0.08
Age in years at telephone interview completion	31.8 [5.6]	33.7 [5.9]	<0.01

SD, standard deviation.

^aHousehold income missing for two case mothers.

Note. Percentages may not add to 100% due to rounding.

Table 2

Agreement and test of nondifferentiality for maternal telephone interview vs. medical record data among case and noncase mothers

Characteristic	No. of occurrences by information in medical record, telephone interview ^a										Exact <i>P</i> -value for test of nondifferentiability ^b
	Case mothers (<i>n</i> = 234)					Noncase mothers (<i>n</i> = 215)					
	+,+	+,−	−,+	−,−	Total	+,+	+,−	−,+	−,−	Total	
Pregnancy complication ^c											
Gestational diabetes	3	1	4	203	211	5	1	6	188	200	0.56
Hypertension	13	15	15	175	218	19	2	15	171	207	0.11
Protein or albumin in urine	19	136	3	54	212	17	134	1	50	202	0.62
Peripheral oedema	67	11	49	44	171	58	7	46	53	164	0.71
Toxaemia or pre-eclampsia	7	8	5	192	212	4	3	4	194	205	1.00
Medical procedure											
Ultrasound	216	0	0	1	217	196	1	0	3	200	0.48
Caesarean section	59	0	0	170	229	44	0	0	169	213	1.00
Infant received phototherapy	1	0	0	2	3	2	0	0	0	2	_ <i>d</i>
Infant received blood products	0	0	2	56	58	0	1	0	122	123	0.10
Infant received supplemental oxygen	8	24	1	39	72	12	39	2	66	119	0.93
Infant moved for special treatment	11	3	4	59	77	13	3	7	122	145	0.92

^a +,+: yes in medical record and yes in telephone interview; +,−: yes in medical record and no in telephone interview; −,+: no in medical record and yes in telephone interview; −,−: no in medical record and no in telephone interview.

^b Exact *P*-value testing independence of maternal and infant characteristics from telephone interview data and case-control status outcome given medical record data.

^c Mentioned in medical record, not based on diagnostic criteria.

^d Conditional distribution is degenerate.

Note. Totals do not always add up to total number of cases and controls due to missing data (not recorded in medical record or during interview) and/or data not known (i.e. 'don't know' response) by mothers during telephone interview.

Table 3

Sensitivities, specificities, positive predictive values and negative predictive values for maternal telephone interview vs. medical record data among case and noncase mothers

Characteristic	Case mothers (n = 234)				Noncase mothers (n = 215)			
	Se [95% CI]	Sp [95% CI]	PPV [95% CI]	NPV [95% CI]	Se [95% CI]	Sp [95% CI]	PPV [95% CI]	NPV [95% CI]
Pregnancy complication ^a								
Gestational diabetes	0.75 [0.19, 0.99]	0.98 [0.95, 0.99]	0.43 [0.10, 0.82]	0.99 [0.97, 1.00]	0.83 [0.36, 1.00]	0.97 [0.93, 0.99]	0.45 [0.17, 0.77]	0.99 [0.97, 1.00]
Hypertension	0.46 [0.28, 0.66]	0.92 [0.87, 0.96]	0.46 [0.28, 0.66]	0.92 [0.87, 0.96]	0.90 [0.70, 0.99]	0.92 [0.87, 0.95]	0.56 [0.38, 0.73]	0.99 [0.96, 1.00]
Protein or albumin in urine	0.12 [0.08, 0.18]	0.95 [0.85, 0.99]	0.86 [0.65, 0.97]	0.28 [0.22, 0.35]	0.11 [0.07, 0.17]	0.98 [0.90, 1.00]	0.94 [0.73, 1.00]	0.27 [0.21, 0.34]
Peripheral oedema	0.86 [0.76, 0.93]	0.47 [0.37, 0.58]	0.58 [0.48, 0.67]	0.80 [0.67, 0.90]	0.89 [0.79, 0.96]	0.54 [0.43, 0.64]	0.56 [0.46, 0.66]	0.88 [0.77, 0.95]
Toxaemia or pre-eclampsia	0.47 [0.21, 0.73]	0.97 [0.94, 0.99]	0.58 [0.28, 0.85]	0.96 [0.92, 0.98]	0.57 [0.18, 0.90]	0.98 [0.95, 0.99]	0.50 [0.16, 0.84]	0.98 [0.96, 1.00]
Medical procedure								
Ultrasound	1.00 [0.98, 1.00]	1.00	1.00 [0.98, 1.00]	1.00	0.99 [0.97, 1.00]	1.00 [0.29, 1.00]	1.00 [0.98, 1.00]	0.75 [0.19, 0.99]
Caesarean section	1.00 [0.94, 1.00]	1.00 [0.98, 1.00]	1.00 [0.94, 1.00]	1.00 [0.98, 1.00]	1.00 [0.92, 1.00]	1.00 [0.98, 1.00]	1.00 [0.92, 1.00]	1.00 [0.98, 1.00]
Infant received phototherapy	1.00	1.00 [0.16, 1.00]	1.00	1.00 [0.16, 1.00]	1.00 [0.16, 1.00]	0	1.00 [0.16, 1.00]	0
Infant received blood products	0	0.97 [0.88, 1.00]	0 [0, 0.84]	1.00 [0.94, 1.00]	0	1.00 [0.97, 1.00]	0	0.99 [0.96, 1.00]
Infant received supplemental oxygen	0.25 [0.11, 0.43]	0.98 [0.87, 1.00]	0.89 [0.52, 1.00]	0.62 [0.49, 0.74]	0.24 [0.13, 0.37]	0.97 [0.90, 1.00]	0.86 [0.57, 0.98]	0.63 [0.53, 0.72]
Infant moved for special treatment	0.79 [0.49, 0.95]	0.94 [0.85, 0.98]	0.73 [0.45, 0.92]	0.95 [0.87, 0.99]	0.81 [0.54, 0.96]	0.95 [0.89, 0.98]	0.65 [0.41, 0.85]	0.98 [0.93, 1.00]

Se, Sensitivity; Sp, Specificity; PPV, positive predictive value; NPV, negative predictive value; 95% CI, exact 95% confidence interval.

^aMentioned in medical record, not based on diagnostic criteria.

Table 4

Odds ratios and 95% confidence intervals for the odds of infant leukaemia by data source

Characteristic ^a	Medical record data		Maternal telephone interview data			Bootstrap ratio		
	Cases (<i>n</i> = 234) <i>n</i> (%)	Noncases (<i>n</i> = 215) <i>n</i> (%)	Bootstrap OR ^b [95% CI]	Cases (<i>n</i> = 234) <i>n</i> (%)	Noncases (<i>n</i> = 215) <i>n</i> (%)	Bootstrap OR ^b [95% CI]	Unadjusted ORs for interview vs. medical record ^d [95% CI]; <i>P</i> -value ^e	ORs for adjusted interview vs. unadjusted medical record ^f [95% CI]; <i>P</i> -value ^e
Gestational diabetes								
Yes	4 (2)	6 (3)	0.66 [0.14, 3.02]	7 (3)	11 (5)	0.59 [0.20, 1.75]	0.89 [0.25, 3.24]; <i>P</i> = 0.86	1.49 [0.40, 5.58]; <i>P</i> = 0.55
No	207 (98)	195 (97)		227 (97)	203 (95)	0.98 [0.34, 2.81]		
Hypertension								
Yes	30 (14)	21 (10)	1.39 [0.75, 2.56]	30 (13)	36 (17)	0.74 [0.43, 1.27]	0.53 [0.30, 0.95]; <i>P</i> = 0.03	0.55 [0.30, 1.02]; <i>P</i> = 0.06
No	190 (86)	186 (90)		202 (87)	179 (83)	0.77 [0.44, 1.36]		
Protein or albumin in urine								
Yes	157 (73)	151 (74)	0.95 [0.61, 1.48]	22 (9)	18 (8)	1.13 [0.57, 2.22]	1.19 [0.56, 2.54]; <i>P</i> = 0.66	1.07 [0.48, 2.42]; <i>P</i> = 0.87
No	57 (27)	52 (26)		210 (91)	195 (92)	1.02 [0.49, 2.13]		
Peripheral oedema								
Yes	78 (46)	65 (40)	1.28 [0.82, 1.98]	148 (63)	132 (61)	1.08 [0.73, 1.60]	0.85 [0.53, 1.37]; <i>P</i> = 0.50	0.72 [0.44, 1.18]; <i>P</i> = 0.20
No	93 (54)	99 (60)		86 (37)	83 (39)	0.92 [0.61, 1.39]		
Toxaemia or pre-eclampsia								
Yes	15 (7)	7 (3)	2.05 [0.73, 5.70]	12 (5)	8 (4)	1.34 [0.49, 3.70]	0.66 [0.24, 1.81]; <i>P</i> = 0.42	0.58 [0.21, 1.61]; <i>P</i> = 0.30
No	197 (93)	198 (97)		222 (95)	207 (96)	1.20 [0.43, 3.35]		
Delivery method								
Caesarean section	59 (26)	44 (21)	1.34 [0.85, 2.09]	60 (26)	44 (20)	1.34 [0.86, 2.09]	1.01 [0.96, 1.05]; <i>P</i> = 0.82	1.05 [0.90, 1.23]; <i>P</i> = 0.50
Vaginal	170 (74)	169 (79)		174 (74)	171 (80)	1.41 [0.87, 2.27]		
Infant received supplemental oxygen								
Yes	34 (46)	51 (43)	1.14 [0.63, 2.05]	22 (10)	26 (12)	0.77 [0.41, 1.43]	0.68 [0.32, 1.45]; <i>P</i> = 0.32	0.74 [0.34, 1.65]; <i>P</i> = 0.47
No	40 (54)	68 (57)		208 (90)	189 (88)	0.84 [0.44, 1.62]		
Infant moved for special treatment								
Yes	15 (19)	16 (11)	1.92 [0.87, 4.27]	29 (13)	23 (11)	1.19 [0.65, 2.16]	0.62 [0.31, 1.23]; <i>P</i> = 0.17	0.66 [0.32, 1.37]; <i>P</i> = 0.26
No	63 (81)	129 (89)		203 (88)	192 (89)	1.27 [0.68, 2.35]		

OR, odds ratio; CI, confidence interval.

^aOdds ratios calculated for those characteristics with all cells of 2×2 tables >3 .

^bBootstrap odds ratio estimates from 10 000 resamplings.

^cBootstrap odds ratio estimate for odds ratio adjusted for length of time in years (continuous variable) between index child's birth and completion of telephone interview.

^dBootstrap ratio of unadjusted telephone interview odds ratio to unadjusted medical record odds ratio.

^eTwo-sided *P*-value testing whether the telephone interview data differ from the medical record data using a *z*-score.

^fBootstrap ratio of adjusted telephone interview odds ratio to unadjusted medical record odds ratio.

Notes. Numbers for cases and noncases are not identical for the telephone interview and medical record due to missing data (not recorded in medical record or during interview) and/or data not known (i.e. 'don't know' response) by mothers during telephone interview. Percentages may not add to 100% due to rounding.