



Published in final edited form as:

*Early Hum Dev.* 2017 February ; 105: 11–15. doi:10.1016/j.earlhumdev.2017.01.003.

## Concordance between maternal recall of birth complications and data from obstetrical records

Kate Keenan, PhD<sup>a</sup>, Alison Hipwell, PhD<sup>b</sup>, Rose McAloon, BS<sup>b</sup>, Amy Hoffmann, BA<sup>b</sup>, Arpita Mohanty, BS<sup>b</sup>, and Kelsey Magee, BS<sup>b</sup>

<sup>a</sup>Department of Psychiatry and Behavioral Neuroscience, University of Chicago, 5841 South Maryland Avenue, Chicago, IL, USA 60637

<sup>b</sup>Department of Psychiatry, University of Pittsburgh, 121 Meyran Ave, Pittsburgh, PA, USA 15213

### Abstract

**Background**—Prenatal complications are associated with poor outcomes in the offspring.

Access to medical records is limited in the United States and investigators often rely on maternal report of prenatal complications.

**Study design and aims**—We tested concordance between maternal recall and birth records in a community-based sample of mothers participating in a longitudinal study in order to determine the accuracy of maternal recall of perinatal complications.

**Subjects**—Participants were 151 biological mothers, who were interviewed about gestational age at birth, birthweight, and the most commonly occurring birth complications: nuchal cord and meconium aspiration when the female child was on average 6 years old, and for whom birth records were obtained.

**Outcome measures**—Concordance between reports was assessed using one-way random intra-class coefficients for continuous measures and kappa coefficients for dichotomous outcomes. Associations between maternal demographic and psychological factors and discrepancies also were tested.

**Results**—Concordance was excellent for continuously measured birthweight (ICC = .85,  $p < 0.001$ ) and good for gestational age (ICC = .68,  $p < 0.001$ ). Agreement was good for low birthweight (< 2500 grams) (kappa = .67,  $p < 0.001$ ), fair for preterm delivery (< 37 weeks gestation) (kappa = .44,  $p < 0.001$ ), and poor for nuchal cord or meconium aspiration. Most discrepancies were characterized by presence according to birth record and absence according to maternal recall. Receipt of public assistance was associated with a decrease in discrepancy in report of nuchal cord.

---

Address correspondence to Kate Keenan, Department of Psychiatry and Behavioral Neuroscience, University of Chicago, 5841 South Maryland Avenue, Chicago, IL 60637, [keenan@uchicago.edu], 773-702-4449.

Conflicts of interest: None

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

**Conclusions**—Concordance between maternal retrospective report and medical birth records varies across different types of perinatal events. There was little evidence that demographic or psychological factors increased the risk of discrepancies. Maternal recall based on continuous measures of perinatal factors may yield more valid data than dichotomous outcomes.

### Keywords

Maternal recall; obstetric record; perinatal complication; neurodevelopment

---

### Introduction

Numerous studies have linked obstetric complications such as low birthweight, premature delivery and cord complications with risk for a variety of poor neurodevelopmental outcomes [1]. In fact, birth complications are considered to have significant causal influences for several forms of severe mental disorders including depressive disorder, schizophrenia, and bipolar disorder [2-9]. Large community-based studies in North America that are designed to examine the prevalence and correlates of child and adolescent neurodevelopmental outcomes, frequently assess perinatal complications using maternal retrospective recall. Yet, studies that utilize this methodology have reported inconsistent findings. While some studies have reported no associations between maternal report of birth complications, and low birthweight and offspring neurodevelopmental problems [10,11], others have shown that maternal recall of birth events predicts youth behavior problems [12,13,14]. For example, in two child epidemiologic studies, maternal recall of birthweight was unrelated to offspring behavior problems [15,16]. In contrast, results from two other large-scale studies revealed significant associations between maternal recall of low birthweight and youth behavioral and emotional problems [17,18]. The variability across studies may be due in part to poor validity of the measures of birth complications.

In fact, research on the concordance between maternal recall and obstetrical record has produced varying levels of agreement. For example, a study of high-risk pregnant women (half of whom delivered twins or triplets) yielded nearly perfect agreement between data from medical record on delivery date, birthweight, cesarean delivery and neonatal intensive care unit admission and maternal reports collected by telephone interview within 2-6 weeks after delivery [22]. Buka and colleagues strategically selected subjects in which birth complications and offspring psychosis were over-represented from two larger cohorts [8]. Based on an average period of recall of 30 years, agreement with medical record ranged from good for birthweight < 5 lbs. to poor for preterm delivery. Thus, even within special subpopulations of obstetric patients, substantial variability in concordance between maternal recall and medical record is observed.

Despite the extensive literature linking obstetric issues to offspring neurodevelopment, an assessment of the validity of maternal recall of pregnancy complications has not been rigorously pursued within relevant populations. In the present study, we examine concordance between maternal recall of delivery events on average 6 years following the birth and data collected from medical records in a sample drawn from a community-based

study. In addition, we examine whether discordance in the two modes of assessment is associated with maternal sociodemographic and psychological factors.

## Materials and Methods

### Sampling of participants

Participants were 232 girls and their biological mothers recruited for the Pittsburgh Girls Study-Emotions Study (PGS-E), a sub-study of the larger Pittsburgh Girls Study (PGS), for which a stratified, random household sampling, with over-sampling of households in low-income neighborhoods, was used to identify girls who were between the ages of 5 and 8 years. Of the 2,992 eligible families 2,875 (96%) were successfully re-contacted to determine their willingness to participate in the longitudinal study, and 85% of those families agreed to participate resulting in a PGS sample of 2,450 [21].

Girls participating in the PGS-E substudy were recruited from the youngest participants in the PGS, and either screened high on measures of depressive symptoms by their self- and parent-report at age 8, or were included in a random selection from the remaining girls. The measures used to screen for depressive symptoms were the Short Moods and Feelings Questionnaire [22] and the Child Symptom Inventory [23]. Girls whose scores fell at or above the 75<sup>th</sup> percentile by their own report, their mother's report, or by both informants comprised the screen high group (n = 135). This sampling strategy was used in order to increase the base rate of depression as the girls moved into adolescence.

There were significantly more African American than Caucasian girls in the screen high group. One hundred thirty-six girls were randomly selected from those scoring below the 75<sup>th</sup> percentile and were matched to the screen high group on race. Eight families were not eligible at the time of recruitment for the PGS-E because the biological mother had died, the family had moved, or the family was no longer participating in the main study and could not be contacted. Of the 263 remaining eligible families, 232 (88.2%) agreed to participate and completed the laboratory assessment, 25 (9.5%) families refused to participate and 6 (2.3%) agreed but could not be scheduled for an assessment.

### Measures

Approval for all study procedures was obtained from the University of Pittsburgh Institutional Review Board. Written informed consent from the mother was obtained prior to any data collection. Using an interview developed by the PGS, mothers were asked to recall birthweight in pounds and ounces, whether the child was born early and if so, how many weeks early, whether the cord was wrapped around the baby's neck (nuchal cord), and whether the baby breathed in any meconium (meconium aspiration), as these perinatal complications are the most commonly occurring in this population [24]. Low birthweight was defined as < 2500 grams and preterm delivery as delivery prior to 37 weeks gestation. Any complication was defined as one or more of the following birth complications: low birthweight, preterm delivery, nuchal cord complications, and meconium aspiration. Participants with one or more missing complications were excluded from analyses. Concordance between maternal report and birth record was assessed using one-way random

intra-class coefficient (ICC) [25] for continuous measures (i.e., birthweight and gestational age) and kappa coefficients for dichotomous outcomes. Level of agreement was characterized according to accepted and comparable guidelines: .75–1.0 = excellent; .60-.74 = good; .40-.59 = fair, and 0-.39 poor [26].

The Difficult Life Circumstances Scale (DLC) is a set of 28 yes-no questions about difficult circumstances at home or work. The measure was designed to include items that would be applicable to women living in poverty such as difficulty with finances and housing. The internal consistency of the scale as measured by alpha was .64. The Perceived Stress Scale (PSS) is a 14-item scale designed to measure the degree to which situations in one's life are appraised as stressful. The internal consistency of the scale as measured by alpha was high: .85. The 21-item Beck Depression Inventory-II (BDI-II) assesses levels of depression in both clinical and non-clinical respondents. Each item requires individuals to select a statement about the degree to which they may or may not have been experiencing a specific symptom of depression in the last two weeks. Cronbach's alpha for the current sample was 0.88.

### Birth Record Methods

Trained research assistants coded data from the birth records: birthweight, gestational age, nuchal cord, and meconium aspiration. A second research assistant independently coded 25% of the birth records to assess inter-rater reliability. Agreement between trained raters was excellent: Intra-class coefficient (ICC) for birthweight = .99; ICC for gestational age = .99; kappa coefficient for nuchal cord = 1.0; and kappa coefficient for meconium aspiration = 1.0.

### Results

Of the 232 participants in the PGS-E study, 151 records were obtained (61.5%) (Figure 1). Demographic characteristics of the sample and tests of differences between those for whom a medical record was and was not obtained are presented in Table 1. Most of the mothers (67.1%) reported their race as African American, over half (52.0%) acted as single parents, and 47% percent of families received some form of public assistance (e.g., food stamps, Medicaid, or monies from public aid). No significant group differences in demographic characteristics or depression, negative life events, or perceived stress scores were observed between those for whom birth records were and were not obtained. According to the birth records, 33.3% had at least one complication, ranging from 6.9% for low birthweight and 21.2% for nuchal cord complications. Close to 20% of mothers reported at least one complication, ranging from 5.8% for nuchal cord complications and meconium aspiration and 8.6% for low birthweight (Table 2).

Concordance between maternal report and birth record was excellent for birthweight (ICC = .85, [95% CI = .80 - .89],  $p < 0.001$ ) (see Figure 2) and good for gestational age (ICC = .68, [95% CI = .60 - .77],  $p < 0.001$ ) (Table 2). Agreement was good for low birthweight (kappa = .67, [95% CI = .44 - .90],  $p < 0.001$ ), fair for preterm delivery (kappa = .44, [95% CI = .10 - .78],  $p < 0.001$ ), and poor for any complication (kappa = .34, [95% CI = .18 - .50],  $p < 0.001$ ); no significant agreement was observed for nuchal cord or meconium aspiration. The average gestational age according to the medical record for the 7 individuals discordant

on preterm delivery was 35.9 weeks, ranging from 35-37 weeks, and average birthweight for the 7 infants with discordant on low birthweight was 2781.6 grams, ranging from 2003 to 3530 grams.

The distribution of informant concordance and discrepancy varied across complication type. Discrepancies regarding preterm delivery were primarily characterized by presence according to birth record (indicated by +) and absence according to maternal report (indicated by -) (4.3%) compared to the reverse: presence according to maternal report and absence according to birth record (0.7%), whereas for low birthweight the distribution of disagreements was fairly equal. For nuchal cord, meconium aspiration or any one of the four complications, disagreements were largely characterized by presence according to birth record and absence according to maternal report (Figure 3).

Next we examined whether the odds of a discrepancy between maternal recall and the medical record were associated with maternal demographic or psychological factors. Preterm delivery and low birthweight yielded few cases of disagreements ( 5% of cases), and were therefore not considered in testing correlates of discrepancies. None of the demographic or psychological factors increased the odds of a discrepancy between maternal recall and medical record for nuchal cord, meconium aspiration or any complication. Receipt of public assistance significantly decreased the odds of discrepant reporting of nuchal cord (OR = 0.28; 95% CI = [0.11, 0.70]). Type of hospital (academic vs. non-academic) where delivery occurred was not associated with discordance between maternal report of birth complications and medical record. In addition, we tested whether birth complications varied as a function of demographic or psychological factors. The only statistically significant finding was an association between race and birthweight from the medical record (African American mean = 3136.7 versus non-African American mean = 3316.5,  $p = .032$ ). However, the magnitude of association between maternal reported birthweight and birthweight from medical record was not affected by race.

## Discussion

In this non-clinical, U.S. sample, concordance between maternal recall and data from obstetric records was good to excellent for continuous measures, but largely inadequate for dichotomous measures, with the exception of low birthweight. These results support the validity of maternal recall of obstetric issues, but suggest that validity will be higher for birthweight and gestational age compared to the presence or absence of specific delivery events. The recommendation based on the present data would be to model continuous rather than dichotomous measures in testing hypotheses regarding the association between perinatal events and offspring development. In fact, comparing the effect sizes for continuous birthweight and low birthweight (i.e., < 2500 grams) on inattention and behavior problems in the Tennessee Twin Study resulted in the authors concluding that the continuous term for birthweight was a stronger predictor than the categorical variable [12].

Our results did not reveal any systematic bias for poor agreement between informants in terms of demographic characteristics or maternal psychological factors, with the exception of receipt of public assistance, resulting in a reduction in the odds of a discrepancy in

reporting nuchal cord. A single significant finding among 18 tests of significance, however, could be attributable to chance. In a few studies, demographic characteristics, including age and race have been associated with recall of labor management, including induction, assisted rupture of membranes, and episiotomy: older age and white race were associated with better recall [27]. Our null findings with regard to maternal demographic and psychological characteristics are important, as these are factors typically associated with offspring physical and mental health.

We note several limitations that could affect the generalizability of the results. First, the sample size is relatively small and therefore the power to test discrepancies across different obstetric outcomes was limited. However, it is important to note that the sample is a community sample, enriched for depression symptoms, but unselected with regard to obstetric complications. Thus, the capacity to generalize the current findings to other community based samples is important. Second, data from medical records were collected on only 65.1% of the targeted sample. Although there were no significant differences in maternal demographic or psychological characteristics between those with and without medical record data, we cannot rule out the possibility that other biases were introduced via the reduced sample. Third, we focused on a few delivery events, and given the variability observed in our study and in others, it is possible that the concordance would be higher for delivery events not recorded, such as resuscitation, use of forceps, or admission to a neonatal intensive care nursery. Fourth, although mothers were asked in the interview to provide gestational age in weeks and days and birthweight in pounds and ounces, further probing of the responses reported only in weeks or pounds were not conducted to determine whether the mother used rounding in her responses.

Finally, although we recognize that data from the medical record is considered more reliable, such data do suffer from recording errors and missing information [28,29]. For this reason we did not focus on one source as a gold standard. Studies in which reliability within sources and validity across sources are needed to identify a gold standard. The issue is further complicated in the present study given that mothers are recalling information provided to them by medical staff as opposed to recalling information derived from their own observations.

In summary, maternal recall of some obstetric data, even up to 6 years following delivery, appears to be a valid method for data collection. As greater attention is paid to the fetal origins of health and disease [30], the results from the present study are important for obtaining information from patients for whom medical records are not available, and for testing hypotheses in large, representative samples in countries for which health records are difficult to obtain, such as the U.S.

## Acknowledgments

This study was supported by National Institute of Health grants R01 MH56630, R01 MH66167, and R01 MH093605. The funding sources listed had no involvement in the production of this manuscript.

## References

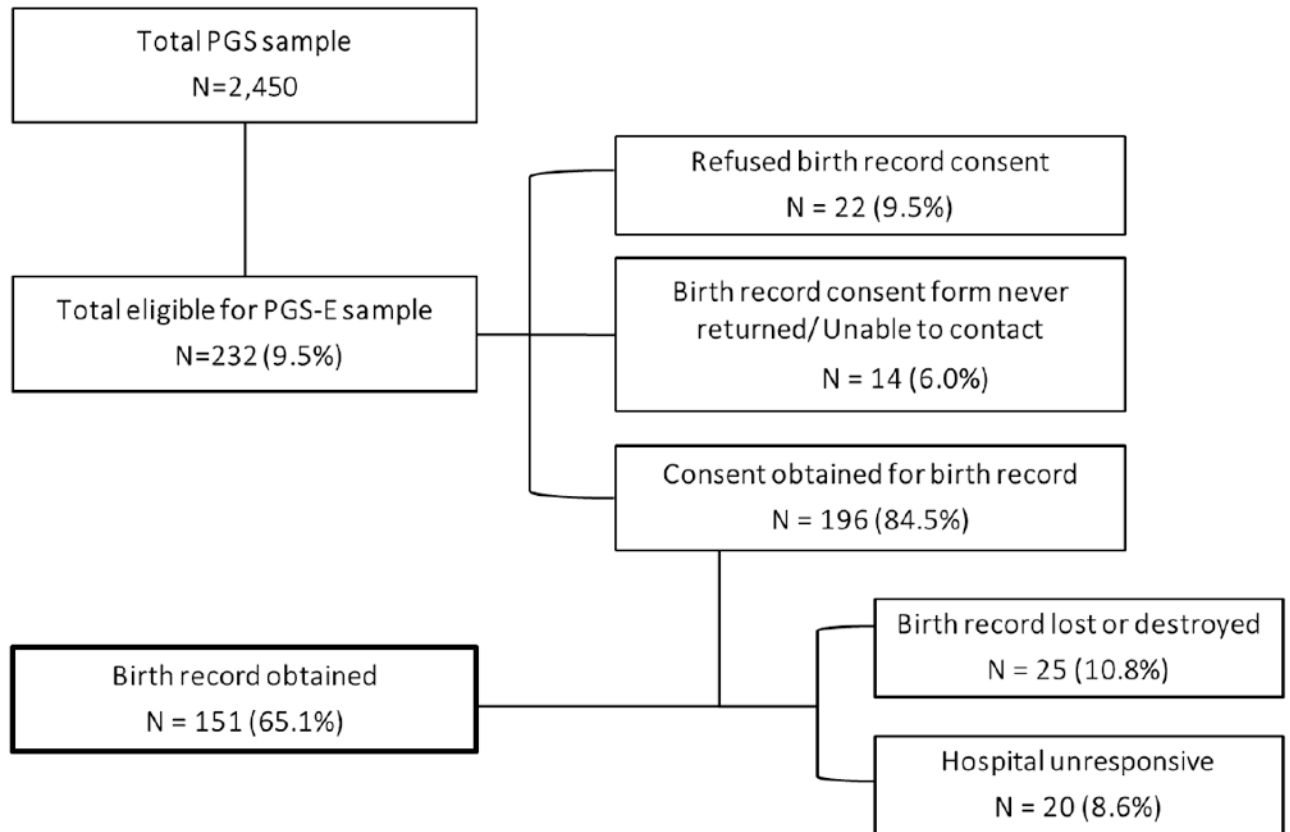
1. Dong Y, Chen S, Yu J. A systematic review and meta-analysis of long-term development of early term infants. *Neonatology*. 2012; 102:212–221. [PubMed: 22814228]
2. Nomura Y, Wickramarante P, Pilowsky D, Newcorn J, Bruder-Costello B, Davey C, et al. Low birth weight and risk of affective disorders and selected medical illness in offspring at high and low risk for depression. *Compr Psychiatry*. 2007; 48:470–478. [PubMed: 17707257]
3. Hirshfeld-Becker D, Biederman J, Faraone S, Robin J, Friedman D, Rosenthal J, et al. Pregnancy complications associated with childhood anxiety disorders. *Depress Anxiety*. 2004; 19:152–162. [PubMed: 15129417]
4. Hack M, Youngstrom E, Cartar L, Schluchter M, Taylor G, Flannery D, et al. Predictors of internalizing symptoms among very low birth weight young women. *J Dev Behav Pediatr*. 2005; 26:93–104. [PubMed: 15827460]
5. Bhutta A, Cleves M, Casey P, Cradoc M, Anand K. Cognitive and behavioral outcomes of school-aged children who were born preterm. *JAMA*. 2002; 288:728–737. [PubMed: 12169077]
6. Galera S, Zanetti A, Ferreira G, Giacon B, Cardoso L. Research with families of people with mental disorder. *Brazilian J Nur*. 2011; 64:774–778.
7. Geller D, Wieland N, Carey K, Vivas F, Petty C, Johnson J, et al. Perinatal factors affecting expression of obsessive-compulsive disorder in children and adolescents. *J Child Adol Psychopharmacol*. 2008; 18:373–379.
8. Buka S, Goldstein J, Spartos E, Tsuang M. The retrospective measurement of prenatal and perinatal events: accuracy of maternal recall. *Schizophr Res*. 2004; 71:417–426. [PubMed: 15474913]
9. Pavuluri M, Henry D, Nadimpalli S, O'Conner M, Sweeney J. Biological risk factors in pediatric bipolar disorder. *Biol Psychiat*. 2006; 60:936–941. [PubMed: 16806102]
10. Van Lieshout R, Boyle M. Canadian youth born large or small for gestational age and externalizing and internalizing problems. *Can J Psychiatry*. 2011; 56:227–234. [PubMed: 21507279]
11. Bird H, Canino G, Davies M, Zhang H, Ramirez R, Lahey B. Prevalence and correlates of antisocial behaviors among three ethnic groups. *J Abnorm Child Psychol*. 2001; 29:465–478. [PubMed: 11761281]
12. Ficks C, Lahey B, Waldman I. Does low birthweight share common genetic or environmental risk with childhood disruptive disorders? *J Abnorm Psychol*. 2013; 122:842–853. [PubMed: 23834065]
13. Costello E, Worthman C, Erkanli A, Angold A. Prediction from low birthweight to female adolescent depression: A test of competing hypotheses. *Arch Gen Psychiatry*. 2007; 64:338–344. [PubMed: 17339522]
14. Allen N, Lewinsohn P, Seeley J. Prenatal and perinatal influences on risk for psychopathology in childhood and adolescence. *Dev Psychopathol*. 1998; 10:513–529. [PubMed: 9741680]
15. Bat-Erdene U, Metcalfe A, McDonald S, Tough S. Validation of Canadian mothers' recall of events in labour and delivery with electronic health records. *BMC Pregnancy Childbirth*. 2013; 13:S3. [PubMed: 23445768]
16. Lahey B, Flagg E, Bird H, Schwab-Stone M, Canino G, Dulcan M, et al. The NIMH methods for the Epidemiology of Child and Adolescent Mental Disorders (MECA) Study: Background and methodology. *J Am Acad Child Adolesc Psychiatry*. 1996; 35:855–864. [PubMed: 8768345]
17. Lahey B, Van Hulle C, Singh A, Waldman I, Rathouz P. Higher-order genetic and environmental structure of prevalent forms of child and adolescent psychopathology. *Arch Gen Psychiatry*. 2011; 68:181–189. [PubMed: 21300945]
18. Costello E, Angold A, Burns B, Stangl D, Tweed D, Erkanli A, et al. The Great Smoky Mountains Study of Youth. Goals, design, methods, and the prevalence of DSM-III-R disorders. *Arch Gen Psychiatry*. 1996; 53:1129–1136. [PubMed: 8956679]
19. Elliott J, Desch C, Istwan N, Rhea D, Collins A, Stanziano G. The reliability of patient-reported pregnancy outcome data. *Popul Health Manag*. 2010; 13:27–32. [PubMed: 20158321]
20. Chen C, Smith L, Pierce C, Blomquist J, Handa V. Do symptoms of pelvic floor disorders bias maternal recall of obstetrical events up to 10 years after delivery? *Female Pelvic Med Recon Surg*. 2015; 21:129–134.

21. Keenan K, Hipwell A, Chung T, Stepp S, Loeber R, Stouthamer-Loeber M, et al. The Pittsburgh Girls Studies: Overview and initial findings. *J Clin Child Psychol*. 2010; 39:506–521.
22. Angold A, Costello E, Messer S, Pickles A. Development of a short questionnaire for use in epidemiological studies of depression in children and adolescents. *Int J Methods Psychiatr Res*. 1995; 5:237–249.
23. Gadow, K., Sprafkin, J. Child Symptom Inventory. Stony Brook: State University of New York at Stony Brook; 1996.
24. Centers for Disease Control and Prevention. Maternal and Infant Health. 2016. Retrieved from <http://www.cdc.gov/reproductivehealth/maternalinfanthealth/index.html>
25. Fisher, RA. Statistical methods for research workers. 12. Edinburgh: Oliver & Boyd; 1954.
26. Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychol Assess*. 1994; 6:284–90.
27. Elkadry E, Kenton K, White P, Creech S, Brubaker L. Do mothers remember key events during labor? *Am J Obstet Gynecol*. 2003; 189:195–200. [PubMed: 12861162]
28. Flood M, Small R. Researching labour and birth events using health information records: methodological challenges. *Midwifery*. 2009; 25:701–710. [PubMed: 18321619]
29. Roberts C, Ford J, Lain S, Algert C, Sparks C. The accuracy of reporting of general anaesthesia for childbirth: a validation study. *Anaesth Intensive Care*. 2008; 36:418–424. [PubMed: 18564804]
30. Barker D. The Fetal Origins of Adult Disease. *Proc R Soc Lond B Biol Sci*. 1995; 262:37–43.

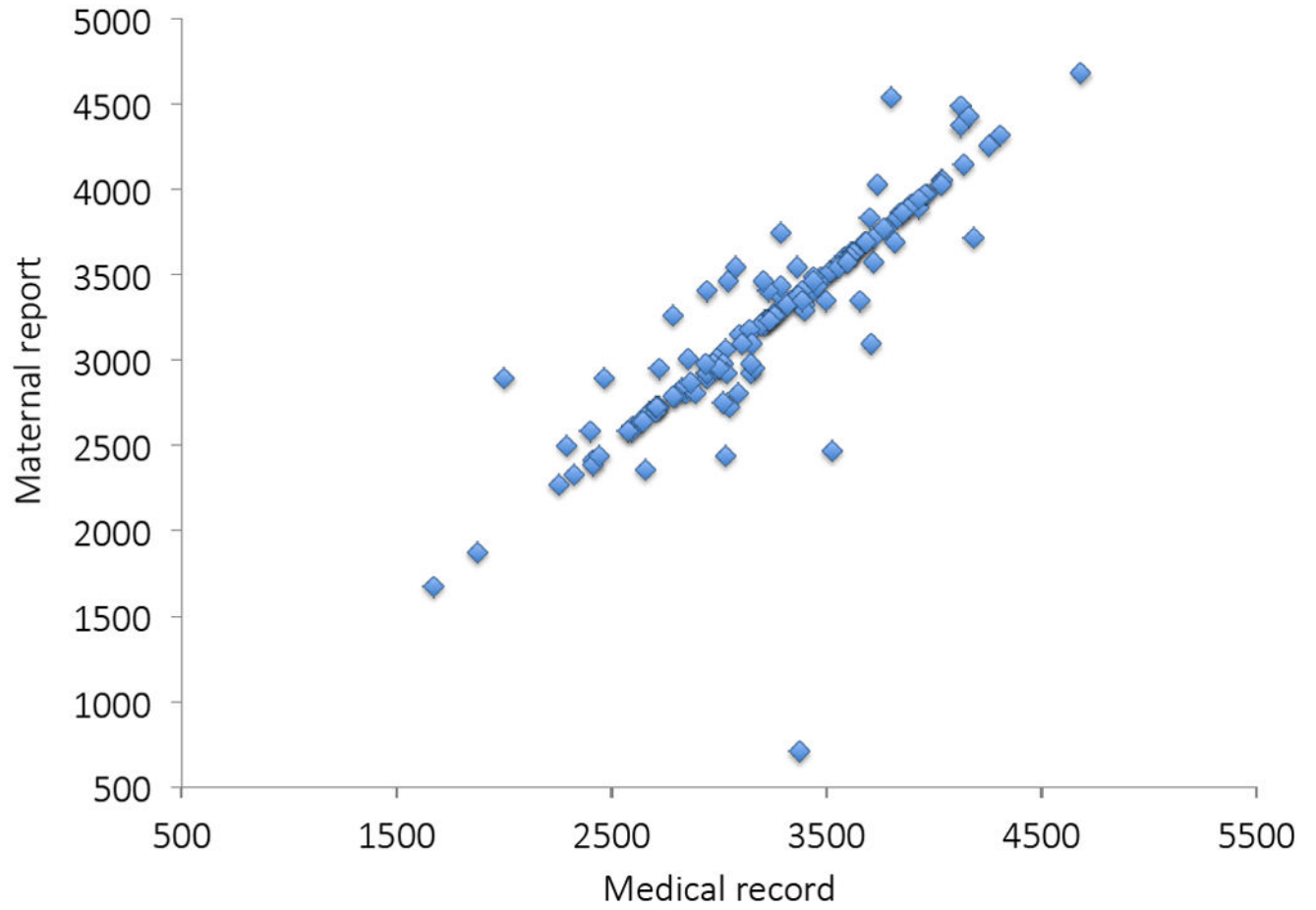


### Highlights

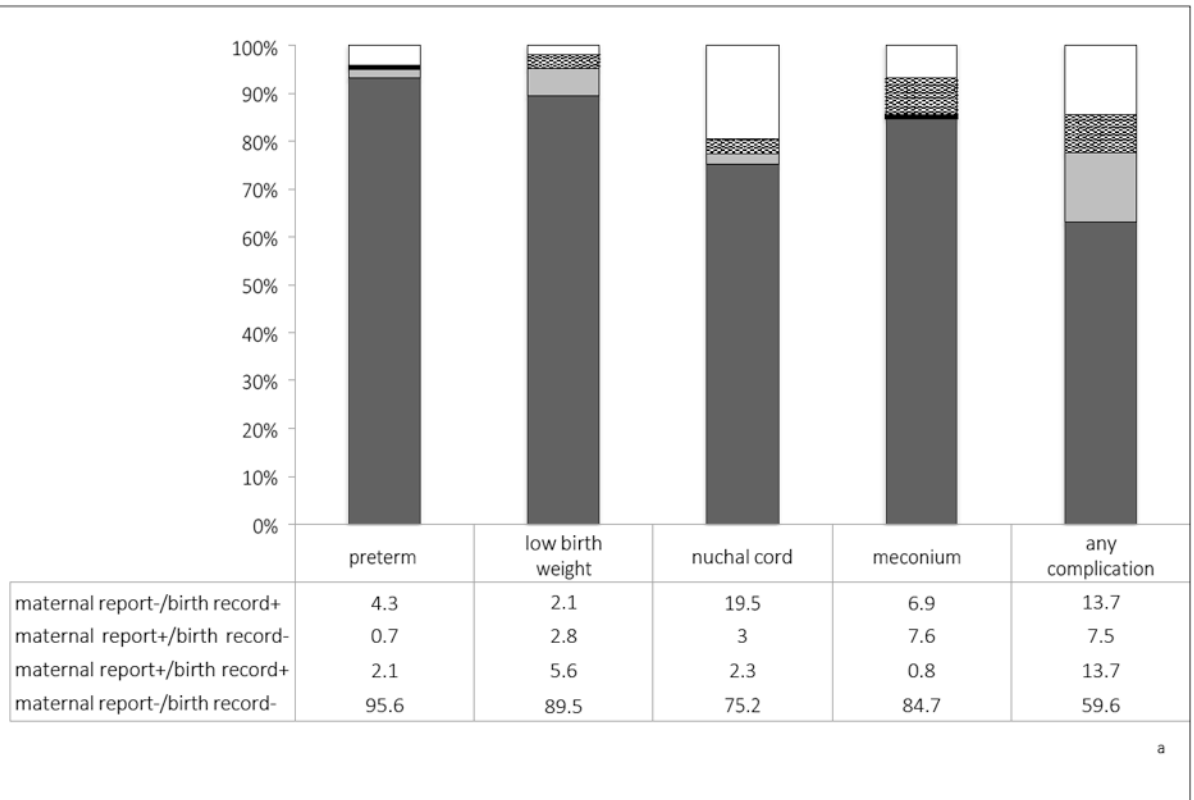
1. Correspondence between maternal retrospective report and birth record varies from strong to minimal depending on the type of prenatal information.
2. Concordance between maternal report and medical record was not affected by demographic or maternal psychological factors.
3. Maternal recall based on continuous measures of perinatal factors, such as gestational age at delivery, yields higher concordance with medical record data than dichotomous outcomes, such as preterm delivery.



**Figure 1.**  
Description of medical records obtained



**Figure 2.**  
Correspondence between maternal report and medical record on birthweight



**Figure 3.**  
 Distribution of informant agreement/disagreements  
<sup>a</sup>Any complication defined as one or more of the following complications: low birthweight, preterm delivery, nuchal cord, meconium aspiration

**Table 1**

Descriptive Statistics for demographic and maternal psychological variables

	Total Sample (n=232)		Birth Record (n=151)		No Birth record (n=81)	
	N	%	N	%	N	%
Demographic Factors						
Single Parent	114	49.1	79	52.0	35	43.8
Public Assistance <sup>a</sup>	104	45.0	71	47.0	33	41.3
African American	151	65.1	102	67.1	49	61.3
Psychological Factors	Mean	SD	Mean	SD	Mean	SD
DLC <sup>b</sup>	3.27	2.37	3.35	2.43	3.18	2.25
PSS <sup>c</sup>	2.68	6.27	23.00	5.51	22.35	5.08
BDI-II <sup>d</sup>	6.90	5.37	7.35	8.09	6.45	8.40

<sup>a</sup>Received a form of public assistance (e.g., food stamps, Medicaid, monies from public aid)<sup>b</sup>Difficult Life Circumstances Scale<sup>c</sup>Perceived Stress Scale<sup>d</sup>Beck Depression Inventory-II

Table 2

Concordance between birth record and maternal report

Infant Characteristics	Birth Record		Maternal Report		Statistic	
	Mean	SD	Mean	SD	ICC	ICC
Birthweight (grams)	3253.69	514.99	3197.72	592.99	.85	***
Gestational age (weeks)	39.06	1.66	39.33	1.59	.68	***
Perinatal Complications	N	%	N	%	kappa	
Low birthweight (< 2500 grams)	11	7.4	19	8.6	.67	***
Preterm delivery (< 37 weeks)	10	6.9	12	5.3	.44	***
Nuchal cord	29	21.2	13	5.8	.09	
Meconium aspiration	11	8.1	13	5.8	.02	
Any complication <sup>a</sup>	50	33.3	47	20.9	.34	***

Note: percentages vary slightly given incomplete data from birth record;

\*\*\*p &lt; 0.01;

\*\*\*\*p &lt; 0.001;

<sup>a</sup> Any complication defined as one or more of the following complications: low birthweight, preterm delivery, nuchal cord, meconium aspiration.