



Perinatal factors reported by mothers: do they agree with medical records?

Pénélope Troude, Laurence Foix L'Helias, Anne-Marie Raison-Boulley, Christine Castel, Christine Pichon, Jean Bouyer, Elise de La Rochebrochard

► To cite this version:

Pénélope Troude, Laurence Foix L'Helias, Anne-Marie Raison-Boulley, Christine Castel, Christine Pichon, et al.. Perinatal factors reported by mothers: do they agree with medical records?. European Journal of Epidemiology, 2008, 23 (8), pp.557-564. 10.1007/s10654-008-9268-9 . hal-02263745

HAL Id: hal-02263745

<https://hal.science/hal-02263745>

Submitted on 5 Aug 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Troude Pénélope, Foix L'Hélias Laurence, Raison-Boulley Anne-Marie, Castel Christine, Pichon Christine, Bouyer Jean, La Rochebrochard Elise (de), 2008, "Perinatal factors reported by mothers: do they agree with medical records?", **European Journal of Epidemiology**, 23(8), p. 557-564. DOI: 10.1007/s10654-008-9268-9

**Perinatal factors reported by mothers:
do they agree with medical records?**

Pénélope Troude^{1,2,3}, Laurence Foix L'Hélias^{4,5,6}, Anne-Marie Raison-Boulley⁷, Christine Castel⁶, Christine Pichon⁸, Jean Bouyer^{1,2,3}, Elise de La Rochebrochard^{1,2,3}

¹ *Ined, F-94276 Le Kremlin-Bicêtre, France*

² *Inserm U822, F-94276 Le Kremlin-Bicêtre, France*

³ *Université Paris-Sud, Faculté de Médecine Paris-Sud, F-94276 Le Kremlin-Bicêtre, France*

⁴ *Inserm U149, F-94807 Villejuif, France*

⁵ *Université Pierre et Marie Curie - Paris 6, F-75252 Paris, France*

⁶ *Service de Pédiatrie et Réanimation Néonatales, Hôpital Universitaire Antoine Bécère, F-92141 Clamart, France*

⁷ *Service de Gynécologie-Obstétrique, Hôpital Foch, F-92151 Suresnes, France*

⁸ *Maternité, Centre Hospitalier Intercommunal de Créteil, F-94010 Créteil, France*

Correspondence to:

Elise de La Rochebrochard

Unité U822

82 rue du Général Leclerc

94276 Le Kremlin-Bicêtre, France

tel: +33 1 45 21 23 33

fax: +33 1 45 21 20 75

E-mail: roche@ined.fr

Abstract

As perinatal events have been linked with diseases of later onset, epidemiological studies on child development and adult health require information on the perinatal period. When national neonatal registers do not exist, review of medical records may be impractical. However, neonatal information could be obtained by asking mothers to complete a postal questionnaire using data from the Personal Child Health Record (PCHR). We assessed agreement between medical records and mothers' reports for information on delivery and the newborn's condition at birth, based on the PCHR, a short time after delivery. Of 711 women giving birth in 3 French hospitals and invited to participate in the study, 580 (82%) completed a postal questionnaire 6 weeks after delivery, copying the data recorded in the PCHR when available. Information on pregnancy, delivery and the newborn's health at birth was independently extracted from medical records by physicians of the maternity departments. Agreement between medical records and maternal reports for a range of perinatal factors was assessed in 580 newborn-mother dyads using kappa coefficients. Agreement was excellent for first and second stages of delivery, gestational age, birth weight, birth size and head circumference (kappa coefficients 0.80 to 1.00) and good for hospitalization during pregnancy, but poor for Apgar scores. With this exception, mothers' reports appeared reliable when compared with medical records. As PCHRs exist in most developed countries, this approach could be used in epidemiological studies on child development to increase the reliability of mothers' reports of their newborn's condition at birth.

Key words: Apgar score, Birth weight, France, Infant, Medical records, Newborn.

Abbreviations: CI Confidence interval, PCHR Personal Child Health Record.

Introduction

Knowledge of events during pregnancy, delivery or the first days of life is necessary to understand health in childhood and adulthood [1, 2]. Recent studies have linked birth weight to cardiovascular diseases in adulthood [3, 4] and the association does not seem to be confounded by genetic and environmental factors, as shown in one study among twins [5]. Associations have also been found between perinatal factors and other diseases with a later onset, such as obesity [6, 7], diabetes [8], metabolic syndrome [9], cancer [10-12] and neurocognitive disabilities [13, 14]. Because events in early childhood have an impact on future health, perinatal information must be collected for studies investigating child and adult health.

When children are recruited at the time of delivery in maternity departments, information about the child and mother can be extracted from obstetric and neonatal records, but in large cohorts this is very time-consuming [15]. In the Avon Longitudinal Study of Parents and Children (ALSPAC), extraction of information related to the pregnancy and the child from written medical records took an average of four hours per record [16]. However, recruitment in maternity departments is not always possible; for example, cohorts of children conceived following in vitro fertilization (IVF) have to be recruited in assisted reproduction centers. Consequently, obstetric records are dispersed in various maternity hospitals where the delivery finally took place, making data collection complex. Similarly, if recruitment takes place some time after delivery, at 9 months as in the Millennium Cohort Study [17], medical records may have been filed, making it difficult and sometimes impossible to extract information. This is increasingly true as the child becomes older, as in the American Early Childhood Longitudinal Study-Kindergarten Cohort (ECLS-K) in which children were enrolled when they entered kindergarten [18].

In Scandinavian countries, national registers make it possible to collect perinatal information at any time in the individual's life. In Norway, a study using neonatal information obtained from linkage of 3 registries found that Apgar scores were related to subsequent death or cerebral palsy (follow-up until the age of 12), and early neonatal symptoms to minor disabilities at school age [14, 19].

When registers did not exist, questionnaires completed by parents have frequently been used to recover information related to the child's early health [18, 20-23], but the quality of this method of collection of perinatal information is still debated. Some studies have attempted to evaluate the accuracy of questionnaire data on pregnancy-related events and birth characteristics by comparing them with data extracted from medical records [15, 17, 24-31]. Most of these studies examined the accuracy of mothers' reports of their newborn's birth weight and found good agreement with medical records [24-26, 29, 30]. Conducting a medical record validation of maternally reported birth characteristics and pregnancy-related events in a case-control study of infant leukemia from United States and Canadian trial groups, Olson et al. found excellent agreement for birth weight, cesarean section and forceps delivery, and good reliability for gestational age [29]. More recently, Rice et al. examined the reliability of maternal reports and antenatal records and found excellent agreement for cesarean section, use of forceps or ventouse, and low and very low birth weight [15]. However, in this English study, women were recruited in fertility clinics and had given birth to a child conceived following IVF; this could be a population with a specifically high level of recall for antenatal and early postnatal events. Comparing birth weight obtained from maternal interviews with that documented in birth certificates for 1,432 children, a recent Taiwanese study has suggested that mothers overreported the categorical infant birth weight of their children [31]. Therefore, the accuracy of maternal reports remains a critical concern.

When considering the issue of collecting perinatal information by interviewing mothers, we hypothesized that the quality of the mothers' reports could be improved by advising them to copy data from the personal child health record (PCHR) in order to avoid recall bias. In France, the newborn's PCHR is systematically given to his/her family the day of the discharge from maternity. The PCHR contains information on health status at birth (as weight, length, head circumference and Apgar scores), the child's diseases, vaccinations and growth curves, and also advice for parents and health education pages. A specimen of the French PCHR is available at http://www.sante.gouv.fr/htm/dossiers/carnet_sante/sommaire.htm . Moreover, the PCHR is permanent and is progressively completed over time, so data are available whatever the age of the child and recall bias is minimized. In perinatal section of the PCHR, information related to pregnancy and birth is completed by midwives the day of birth in labor room; information related to health during first days of life (results of neonatal screening, first compulsory medical examination, weight at discharge from maternity) is completed by the physician who examines the newborn just before he/she leaves maternity department. As the PCHR exists in most European countries, this approach could be possible almost everywhere [32]. English, Australian and French studies that have examined use of PCHRs, although they did not specifically assess the validity of the neonatal information they contained, suggested that items on the neonatal period (birth weight, Apgar scores and information related to delivery) were satisfactorily completed and so potentially useful in epidemiological studies [33-37].

The aim of our study was to assess, in general population and a short time after delivery, the agreement between medical records and mothers' reports for information related to delivery and the newborn's condition at birth, using a postal questionnaire completed by mothers with the help of the PCHR.

Subjects and methods

Study population

The study was carried out in three maternity departments in the Paris region (France) between January 1 and February 4, 2006 at Foch and Antoine Bécère hospitals and between January 22 and March 18, 2007 at Créteil hospital. During the inclusion period, 830 mothers had a live birth. Of these 830 mothers, 51 (6%) were excluded from the study since they met the following protocol exclusion criteria: (i) age less than 18 years old ($n = 6$), (ii) difficulties with the French language ($n = 36$), (iii) major psychiatric disorder ($n = 2$), (iv) anonymous delivery ($n = 2$), or (v) transfer to intensive care unit for more than 48 hours ($n = 5$).

Of the 779 eligible women, one or two days after delivery 711 (91%) women were asked by the pediatrician to participate in the study. Sixty-eight (9%) women were not invited to take part for the following reasons: 18 for medical reasons (poor condition of the newborn ($n = 12$), death of newborn ($n = 3$) or transfer of the mother to another hospital because lack of beds ($n = 3$)), 50 women were not approached for organizational reasons (the mothers left the maternity department before being invited to take part or were not invited because their child was transferred to another care unit). Finally, 639 women agreed to participate (47 refused and 25 left the maternity unit without returning the consent form), giving an inclusion rate of 90% (639 of the 711 mothers invited to take part).

Measures

Within the week following the newborn's discharge from maternity department, physicians collected from medical records information about maternal health, pregnancy (hospitalizations, complications, treatment), delivery (gestational age, mode) and the child's health during the first days of life in the maternity department (characteristics at birth, transfer, characteristics of the child at discharge).

Six weeks after delivery, a postal questionnaire was sent to the mothers. After 3 weeks, a reminder letter was sent to those who had not replied. Three weeks later, the same postal questionnaire was again sent to those who had not replied, and ultimately, after another 3 weeks, mothers who had still not replied were contacted by telephone and asked to answer a shorter questionnaire.

The postal questionnaire included items on parents' sociodemographic characteristics, maternal health before pregnancy, obstetric complications and delivery (hospitalization for more than 24 hours, mode of delivery, gestational age), characteristics of the newborn at birth (weight, size, head circumference, Apgar scores) and conditions of return home (quality of newborn's sleep, breast-feeding). Mothers were asked to copy data from the personal child health record (PCHR) into the postal questionnaire for information reported in PCHR (gestational age, birth weight, birth size, head circumference and 1-minute and 5-minute Apgar scores). The short telephone questionnaire contained items on maternal level of education, nationality, complications and hospitalization during pregnancy, mode of delivery, gestational age, and birth characteristics reported in the PCHR. The studied items were formulated exactly in the same way as in the postal questionnaire. During the telephone interview, mothers were asked to look for the information in their child's PCHR, by providing them the page number they had to refer to.

Statistical analysis

Kappa coefficients were calculated to estimate agreement between medical records and mothers' reports for a range of perinatal factors: hospitalization during pregnancy, delivery, gestational age, birth weight, birth size, head circumference at birth and Apgar scores [38]. In multiple pregnancies ($n = 22$), analyses were conducted only among the firstborn of the twins or triplets.

Gestational age, birth weight, head circumference and Apgar scores were treated as ordered categorical variables, with three categories for gestational age (< 32 weeks, $32-37$, ≥ 37 weeks), birth weight (≤ 2500 g, $2501-3500$ g, > 3500 g), head circumference at birth (≤ 33.5 cm, $33.6-35.4$ cm, ≥ 35.5 cm), and four categories for 1- and 5-minute Apgar scores (0-4, 5-7, 8-9, 10). Beginning of labor (planned cesarean, induced labor, spontaneous labor) and mode of delivery (emergency cesarean, use of forceps or ventouse, spontaneous vaginal delivery) were also treated as ordered categorical variables. Hospitalization during pregnancy (for at least 24 hours) and birth size were treated as dichotomous variables (birth size categorized as $<$ or ≥ 50 cm).

For dichotomous variables, we calculated kappa coefficients. For variables with more than two classes, agreement was measured using weighted kappa coefficients proposed by Cohen [39] that take strength of disagreement into account.

The 95% confidence intervals for weighted kappa coefficients were calculated with the bootstrap method [40]. The lower and upper limits of the 95% CI for each kappa coefficient were defined as the 2.5th and the 97.5th percentile of the respective bootstrap distribution.

According to Landis and Koch [41], a kappa value > 0.80 represents excellent agreement, values between 0.60 and 0.80 good agreement, between 0.40 and 0.60 moderate agreement, and < 0.40 poor agreement.

Results

Table 1 shows the cumulative response rates of the postal survey among the 639 women who had agreed to participate, for each maternity department. Fifty-four percent of mothers replied to the first mailing and, after a reminder letter to non-responders, the cumulative response rate was 75%. Finally, of 639 mothers, 558 responded to the postal questionnaire and 22 responded to the telephone interview giving a total response rate of 91%. Of the respondent mothers, 93% stated they wished to receive the results of the study and 89% agreed to be contacted again for another study.

Table 2 summarizes the sociodemographic characteristics of respondent mothers (n = 580) obtained through maternal reports, for each maternity department. Most mothers were French and lived as a couple, their mean age was 32 years and 70% had an educational level higher than the baccalaureat (high school diploma). We found statistically significant differences for mother's nationality, educational level and occupation between the three hospitals, but not for maternal age, living as a couple or tobacco consumption. The proportion of French mothers, in comparison with other nationalities, was lower at Créteil hospital, while the proportion of mothers with an educational level higher than the baccalaureat was greater at Foch hospital.

Table 3 shows characteristics of pregnancy and delivery both as reported by mothers and in obstetric records, and agreement between the two sources. According to medical records, nearly 12.5% of mothers had been hospitalized for at least 24 hours at some time during pregnancy, 24% had cesarean delivery, and nearly 6.8% had delivered before term. For the first and second stages of delivery and for gestational age kappa values exceeded 0.80, corresponding to excellent agreement, whereas it was only good for hospitalization during pregnancy (kappa 0.77).

Concerning the characteristics of the newborn, according to medical records 6.2% of newborns had a birth weight equal or lower than 2500 g and nearly 10% had a 1-minute

Apgar score lower than 8 (Table 4). Values of kappa coefficients were > 0.95 for birth weight, birth size and head circumference at birth, and thus agreement between both sources was excellent. On the contrary, agreement for 1-minute and 5-minute Apgar scores was poor ($\text{kappa} < 0.40$).

Stratified analyses were conducted to assess the effect of educational level on agreement (educational level higher than baccalaureate versus equal or lower). We observed (data not shown) among both more and less educated women that agreement was excellent for stages of delivery, gestational age, birth weight, birth size and head circumference at birth with kappa ranging from 0.78 to 1.00, good for hospitalization (0.71 and 0.81 for the lower and the higher educational level, respectively) and poor for Apgar scores (kappa ranging from 0.32 to 0.40). Furthermore, agreement was not systematically higher or lower in one specific strata of educational level compared to the other, which provides arguments favoring the absence of any effect of educational level on agreement in our study.

We checked the results observed for continuous variables treated as categorical variables by estimating the intraclass correlation coefficient (ICC) which measures agreement for continuous variables [42]. This concerned gestational age, birth weight, birth size and head circumference at birth. Results were similar to those for kappa coefficients with ICC values > 0.85 . We also used the Bland and Altman graphical method for each of continuous variable [43]. We also found consistent results: Bland and Altman plots showed that the absolute differences in measurements were small and the interval of agreement for each variable was narrow compared with the mean. For instance, for birth weight, the absolute difference was 7.2g with an interval of agreement ranging from -1.4 to 15.8, whereas mean birth weight ranged from 1240 to 4900g.

Discussion

We investigated agreement between medical records and maternal reports for perinatal information, asking mothers to fill in a postal questionnaire with the help of the PCHR in order to avoid recall bias. Agreement was excellent for first and second stages of delivery, gestational age, birth weight, birth size and head circumference at birth (kappa coefficients ranged from 0.80 to 1.00), and good for hospitalization during pregnancy, whereas agreement was poor for Apgar scores at 1 and 5 minutes.

Among events related to pregnancy and delivery, we found excellent agreement for delivery characteristics (gestational age and mode of delivery), which is consistent with previous studies [15, 24, 44, 45]. For example, the UK Millennium Cohort Study found 98% of agreement between maternal reports and hospital records using three mode of delivery groups (normal, assisted, cesarean section) [17]. A Taiwanese study of 208 women found a correlation coefficient of 0.93 for gestational age [45]. Good but not excellent agreement was found in our study for pregnancy-hospitalization reports. This result is consistent with an American study of 102 mothers that showed good agreement (kappa 0.71) between medical records and maternal recall of pregnancy and hospitalizations [27]. Discrepancies between the two sources may be explained by several factors. On the one hand, if the mother was admitted to a hospital other than the maternity department where she delivered, hospitalization episodes might not be reported in obstetric records, thus leading to an underestimation of hospitalization rate based on these records. On the other hand, mothers might report short hospital stays (< 24 hours) as hospitalization, or even report hospitalization whereas they were in fact admitted for labor [24]. This would lead to an overestimation of hospitalization episodes based on mothers' reports. We examined maternal reports concerning gestational age at time of hospitalization, duration and place of hospitalization for all cases (n = 32) where

maternal and medical reports disagreed, and we considered that each of the above factors probably partially explained the discrepancies observed between the two sources.

Concerning newborn characteristics at birth, we found excellent agreement for birth weight (kappa 0.99). These findings confirmed several previous studies that found excellent agreement for birth weight reported by mothers, with a correlation coefficient > 0.90 [45] and a kappa coefficient > 0.85 [15, 29]. However, to our knowledge, the reliability of maternal reports for other characteristics of the child at birth has not previously been assessed. In our study, strength of agreement for size and head circumference at birth was similar to that for weight (kappa values 0.97 and 0.98, respectively) whereas these items of information were probably more difficult to recall than birth weight. Asking mothers to copy information from the PCHR probably increased the reliability of reports for birth size and head circumference, leading to excellent agreement as for birth weight. The few studies that have examined the completeness of neonatal data in French PCHRs found that factual information related to birth, such as weight, was nearly always recorded [35, 36]. Apgar scores, however, were more often missing in PCHRs, and when they were recorded, they were probably overestimated in comparison with the scores collected by the French Department for Mother and Infant Welfare [37]. Incorrect reporting of Apgar scores in PCHRs may explain the poor reliability of maternal reports of these scores, whereas agreement was excellent for other neonatal information. A French study has suggested that some information considered as sensitive, such as genetic diseases or low Apgar scores, might be less often recorded in PCHRs by physicians because they fear that confidentiality may be not respected [35].

In the literature, the degree of accuracy of maternal reports has been described to vary according to several factors: type of information, the way the questions are put (a list, an open-ended or a closed question), time between events and questionnaire, characteristics of respondents and health of children [27, 28, 45, 46]. Elkadry et al. found that American

mothers' recall was significantly related to demographic characteristics such as age at delivery, ethnicity or education, and also to delivery characteristics such as cesarean delivery [47]. Conducting a study on 11,890 children born in England and Wales from the Millennium Cohort Study, Tate et al. found that variance of birth weight discrepancies between mothers' reports and registration data differed according to ethnic group, ward type and socioeconomic status [46]. In our study, global educational level was higher than in the general French population, since nearly 70% of mothers had an educational level higher than the baccalaureat (high school diploma), whereas only 43% of French pregnant women had reached this level [48]. However, we found no difference in strength of agreement after stratification according to educational level. This is consistent with an English study finding that accuracy of maternal recall of infant birth weight and mode of delivery was not affected by stratification according to socio-occupational class and maternal education [15]. In the Millennium Cohort Study, Quigley et al. also found no significant association between disagreement and education, but they found higher disagreement in women who reported that English was not the only language spoken at home and in certain ethnic groups [17]. One limitation of our study was that we excluded women ($n = 36$) who had too much difficulty with the French language. The validity of maternal reports in such a population therefore needs to be investigated.

As the time elapsed between delivery and postal questionnaire was very short (6 weeks) in our study, maternal reports might be more reliable than in studies where this interval was longer. Indeed, Elkadry et al. found that the number of years since delivery was an independent predictor of poor maternal recall for key events during labor, using a multivariate logistic regression model for the prediction of one or more incorrect responses [47]. On the contrary, in their medical record validation study of maternal reports, Olson et al. observed that the range of time between delivery and interview had little effect on the reliability of maternal reports [29].

Although we found excellent agreement between mothers' reports and medical records for delivery and most of the newborn's characteristics at birth, these results should not be extrapolated to pregnancy and neonatal complications. Olson et al. found low validity and reliability for pregnancy and neonatal complications such as jaundice [29]. They hypothesized that the reliability of maternal reports for pregnancy and delivery information may differ with the nature of the factor of interest, which is coherent with our results showing poor reliability of Apgar scores.

In conclusion, we found excellent reliability of maternal reports concerning perinatal information (birth weight, birth size, head circumference at birth, gestational age, mode of delivery) when asking mothers to copy information from the PCHR. Furthermore, the postal questionnaire was well accepted by mothers since we obtained a very good response rate and most mothers agreed to participate in another study. Such a methodology could probably be transposed to other countries where PCHRs exist in order to collect reliable information on the perinatal period for epidemiological studies.

References

1. Gissler M, Järvelin M-J., Louhiala P., Rahkonen O. and Hemminki E. Can children's health be predicted by perinatal health? *Int J of Epidemiol.* 1999;28:276-80.
2. Olsen J. Prenatal exposures and long-term health effects. *Epidemiol Rev.* 2000;22(1):76-81.
3. Frontini MG, Srinivasan SR, Xu J, Berenson GS. Low birth weight and longitudinal trends of cardiovascular risk factor variables from childhood to adolescence: the bogalusa heart study. *BMC Pediatr.* 2004 Nov 3;4(1):22.
4. Mzayek F, Hassig S, Sherwin R, Hughes J, Chen W, Srinivasan S, et al. The Association of Birth Weight with Developmental Trends in Blood Pressure from Childhood through Mid-Adulthood: The Bogalusa Heart Study. *Am J Epidemiol.* 2007 May 24.
5. Bergvall N, Iliadou A, Johansson S, de Faire U, Kramer MS, Pawitan Y, et al. Genetic and Shared Environmental Factors Do Not Confound the Association Between Birth Weight and Hypertension. A Study Among Swedish Twins. *Circulation.* 2007 May 21.
6. Parsons T, Power C, Manor O. Fetal and early life growth and body mass index from birth to early adulthood in 1958 British cohort: longitudinal study. *BMJ.* 2001;323(7325):131-5.
7. Sorensen HT, Sabroe S, Rothman KJ, Gillman M, Fisher P, Sorensen TI. Relation between weight and length at birth and body mass index in young adulthood: cohort study. *BMJ.* 1997;315(7116):1137.
8. Ievins R, Roberts SE, Goldacre MJ. Perinatal factors associated with subsequent diabetes mellitus in the child: record linkage study. *Diabet Med.* 2007;24(6):664-70.
9. Boney CM, Verma A, Tucker R, Vohr BR. Metabolic syndrome in childhood: association with birthweight, maternal obesity, and gestational diabetes mellitus. *Pediatrics.* 2005;115(3):e290-6.

10. Hjalgrim LL, Rostgaard K. Birth weight and risk for childhood leukemia in Denmark, Sweden, Norway, and Iceland. *J Natl Cancer Inst.* 2004;96(20):1549-56.
11. Hodgson ME, Newman B, Millikan RC. Birthweight, parental age, birth order and breast cancer risk in African-American and white women: a population-based case-control study. *Breast Cancer Res.* 2004;6(6):R656-67.
12. Sanderson M, Daling JR, Doody DR, Malone KE. Perinatal factors and mortality from breast cancer. *Cancer Epidemiol Biomarkers Prev.* 2006;15(10):1984-7.
13. Sorensen HT, Sabroe S, Olsen J, Rothman KJ, Gillman M, Fischer P. Birth weight and cognitive function in young adult life: historical cohort study. *BMJ.* 1997;315:401-3.
14. Moster D, Lie RT, Markestad T. Joint association of Apgar scores and early neonatal symptoms with minor disabilities at school age. *Arch Dis Child Fetal Neonatal Ed.* 2002 Jan;86(1):F16-21.
15. 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 peri-natal factors: The influence of maternal and child characteristics. *Early Hum Dev.* 2006 Oct 27.
16. Golding J, Pembrey M, Jones R. ALSPAC--the Avon Longitudinal Study of Parents and Children. I. Study methodology. *Paediatr Perinat Epidemiol.* 2001 Jan;15(1):74-87.
17. 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 Feb;114(2):195-200.
18. Rose D, Bodor JN. Household food insecurity and overweight status in young school children: results from the Early Childhood Longitudinal Study. *Pediatrics.* 2006;117:464-73.

19. Moster D, Lie RT, Irgens LM, Bjerkedal T, Markestad T. The association of Apgar score with subsequent death and cerebral palsy: A population-based study in term infants. *J Pediatr*. 2001 Jun;138(6):798-803.
20. Chen J, Millar W. Birth outcome, the social environment and child health. *Health Rep*. 1999;10(4):57-67.
21. To T, Guttmann A, Dick P, Rosenfield J, Parkin P, Cao H, et al. What factors are associated with poor developmental attainment in young Canadian children? *Can J Public Health*. 2004;95(4):258-63.
22. Huddy C, Johnson A, Hope P. Educational and behavioural problems in babies of 32-35 weeks gestation. *Arch Dis Child Fetal Neonatal Ed*. 2001;85(1):F23-8.
23. Johnson A, Bowler U, Yudkin P, Hockley C, Wariyar U, Gardner F, et al. Health and school performance of teenagers born before 29 weeks gestation. *Arch Dis Child Fetal Neonatal Ed*. 2003;88(3):F190-8.
24. Cartwright A, Smith C. Some comparisons of data from medical records and from interviews with women who had recently had a live birth or stillbirth. *J Biosoc Sci*. 1979 Jan;11(1):49-64.
25. Axelsson G, Rylander R. Validation of questionnaire reported miscarriage, malformation and birth weight. *Int J Epidemiol*. 1984 Mar;13(1):94-8.
26. Ekouevi K, Morgan SP. Note on the reliability and validity of mothers' retrospective reports of their children's birthweights. *Soc Biol*. 1991 Spring-Summer;38(1-2):140-5.
27. Githens PB, Glass CA, Sloan FA, Entman SS. Maternal recall and medical records: an examination of events during pregnancy, childbirth, and early infancy. *Birth*. 1993 Sep;20(3):136-41.

28. Lumey LH, Stein AD, Ravelli AC. Maternal recall of birthweights of adult children: validation by hospital and well baby clinic records. *Int J Epidemiol.* 1994 Oct;23(5):1006-12.
29. 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. *Am J Epidemiol.* 1997 Jan 1;145(1):58-67.
30. Sanderson M, Williams MA, White E, Daling JR, Holt VL, Malone KE, et al. Validity and reliability of subject and mother reporting of perinatal factors. *Am J Epidemiol.* 1998 Jan 15;147(2):136-40.
31. Li CY, Wei JN, Lu TH, Chuang LM, Sung FC. Mothers tended to overreport categorical infant birth weight of their children. *J Clin Epidemiol.* 2006 Dec;59(12):1319-25.
32. Rollet C. Le carnet de santé pour les enfants en Europe. *Medicina e storia.* 2002;7:31-55.
33. Lakhani AD, Avery A, Gordon A, Tait N. Evaluation of a home based health record booklet. *Arch Dis Child.* 1984 Nov;59(11):1076-81.
34. Jeffs D, Nossar V, Bailey F, Smith W, Chey T. Retention and use of personal health records: a population-based study. *J Paediatr Child Health.* 1994 Jun;30(3):248-52.
35. Dommergues JP. [Personal child health record: informativity and confidentiality]. *Arch Pediatr.* 2004 Apr;11 Suppl 1:42s-5s.
36. Chalumeau M. Information recorded by maternity ward staff in permanent pediatric health records. *Arch Pediatr.* 2003;10(9):776-80.
37. Vincelet C, Tabone MD, Berthier M, Bonnefoi MC, Chevallier B, Lemaire JP, et al. [How are personal child health records completed? A multicentric evaluation study]. *Arch Pediatr.* 2003 May;10(5):403-9.
38. Cohen J. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement.* 1960;20:37-46.

39. Cohen J. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin*. 1968;70(4):213-20.
40. Efron B, Tibshirani R. Bootstrap Methods for standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy. *Statistical Science*. 1986;1(1):54-77.
41. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977 Mar;33(1):159-74.
42. Bartko JJ. The intraclass correlation coefficient as a measure of reliability. *Psychol Rep*. 1966 Aug;19(1):3-11.
43. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1:307-10.
44. Hewson D, Bennett A. Childbirth research data: medical records or women's reports? *Am J Epidemiol*. 1987 Mar;125(3):484-91.
45. Sou SC, Chen WJ, Hsieh WS, Jeng SF. Severe obstetric complications and birth characteristics in preterm or term delivery were accurately recalled by mothers. *J Clin Epidemiol*. 2006 Apr;59(4):429-35.
46. Tate AR, Dezateux C, Cole TJ, Davidson L. Factors affecting a mother's recall of her baby's birth weight. *Int J Epidemiol*. 2005 Jun;34(3):688-95.
47. Elkadry E, Kenton K, White P, Creech S, Brubaker L. Do mothers remember key events during labor? *Am J Obstet Gynecol*. 2003 Jul;189(1):195-200.
48. Blondel B, Supernant K., Du Mazaubrun C., Bréat G. Enquête périnatale 2003: situation en 2003 et évolution depuis 1998. DGS. 2003.

Table 1. Cumulative response rates of the postal survey (n = 639)

	Antoine Béclère hospital	Foch hospital	Créteil hospital	Total
Mothers included (n)	218	204	217	639
Response rates (%)				
after 1st mailing	60	57	45	54
after reminder letter	80	79	65	75
after 2nd mailing	86	93	81	87
after telephone interviews	93	93	86	91

Table 2. Mothers' self-reported characteristics (n = 580)

	Antoine Bécère hospital % (n)	Foch hospital % (n)	Créteil hospital % (n)	Total % (n)	p ⁽¹⁾
n	202	190	188	580	
Maternal age (yr)					ns
≤ 24	5% (10)	3% (6)	8% (15)	5% (31)	
25-29	21% (42)	24% (46)	29% (54)	25% (142)	
30-34	43% (85)	40% (76)	35% (66)	40% (227)	
35-39	23% (46)	26% (48)	20% (38)	23% (132)	
≥ 40	8% (16)	7% (13)	8% (14)	7% (43)	
French nationality	92% (184)	87% (166)	81% (153)	87% (503)	0.01
Educational level ≥ baccalaureat	71% (139)	82% (154)	56% (104)	70% (397)	< 1‰
Occupation					< 1‰
Managerial	24% (46)	41% (76)	15% (22)	27% (144)	
Intermediate	32% (61)	20% (37)	36% (54)	29% (152)	
Employee or worker	41% (78)	35% (66)	44% (66)	40% (210)	
Unemployed	3% (7)	4% (7)	5% (8)	4% (22)	
Living with a partner	92% (185)	96% (182)	90% (169)	93% (536)	ns
Number of cigarettes/day					ns
0	86% (165)	85% (159)	89% (154)	86% (478)	
< 10	8% (16)	9% (17)	7% (12)	8% (45)	
≥ 10	6% (11)	6% (12)	4% (7)	6% (31)	

p values for χ^2 test comparing mothers' characteristics according to maternity hospital

Table 3. Agreement between medical records and maternal reports for information related to pregnancy and delivery (n = 580)

	Medical records (%)	Maternal reports (%)	Kappa	95% CI
Pregnancy				
Hospitalization during pregnancy (n =551 ⁽¹⁾)	12.5	15.3	0.77	0.69-0.84
Delivery				
<i>First stage of labor</i> (n = 555 ⁽¹⁾)			0.83	0.76-0.88
Planned cesarean	9.6	10.1		
Induced labor	27.4	25.9		
Spontaneous labor	63.0	64.0		
<i>Second stage of labor</i> (n = 555 ⁽¹⁾)			0.99	0.98-0.99
Cesarean section	24.0	24.1		
Use of forceps/ventouse	13.1	13.2		
Spontaneous vaginal delivery	62.9	62.7		
Gestational age (n = 530 ⁽¹⁾)			0.85	0.76-0.92
< 32 weeks	0.4	0.6		
≥ 32 and < 37 weeks	6.4	7.3		
≥ 37 weeks	93.2	92.1		

(1) Number of mothers for whom information was available in both sources.

Table 4. Agreement between medical records and maternal reports for neonatal information

(n = 580)

	Medical records (%)	Mothers' reports (%)	Kappa	95% CI
Birth weight (n = 560 ⁽¹⁾)			0.99	0.98-1.00
≤ 2500 g	6.2	6.6		
> 2500 and ≤ 3500 g	58.4	58.2		
> 3500 g	35.4	35.2		
Birth head circumference (n=541 ⁽¹⁾)			0.97	0.95-0.98
≤ 33.5 cm	24.8	24.0		
> 33.5 and < 35.5 cm	43.1	42.7		
≥ 35.5 cm	32.2	33.3		
Birth length (n = 557 ⁽¹⁾)			0.98	0.97-1.00
< 50 cm	45.8	44.9		
≥ 50 cm	54.2	55.1		
1-min Apgar score (n = 497 ⁽¹⁾)			0.34	0.22-0.46
0-4	2.8	1.0		
5-7	7.2	1.0		
8-9	28	9.3		
10	62.0	88.7		
5-min Apgar score (n = 510 ⁽¹⁾)			0.38	0.18-0.57
5-7	0.4	0.4		
8-9	9.0	2.0		
10	90.6	97.6		

(1) Number of newborns for whom information was available in both sources.