

A New Method for Measuring Misclassification of Maternal Sets in Maternally Linked Birth Records: True and False Linkage Proportions

Jack K. Leiss

Published online: 26 October 2006
© Springer Science+Business Media, LLC 2006

Abstract *Objectives:* Numerous studies have used maternally linked birth records to investigate perinatal outcomes, maternal behaviors, and the quality of vital records birth data. Little attention has been given to assessing errors in the linkages and to understanding how such errors affect estimates derived from the linked data. The author developed a framework for conceptualizing maternal linkage error and measures for quantifying it, and examined the behavior of the new measures in a maternally linked file. *Methods:* Linkage errors were conceptualized as misclassification, with the classes being the maternal sets (records classified as representing different births to the same woman). The *true linkage proportion*, analogous to sensitivity, was used to capture the degree to which all of a woman's births were assigned to a single maternal set; the *false linkage proportion*, analogous to specificity, was used to capture the degree to which the assigned maternal sets combined births from different women. The behavior of the two proportions was examined by introducing increasing degrees of linkage error into a maternally linked file. *Results:* Both measures indicated greater misclassification with increasing simulated linkage errors. *Conclusions:* The new measures may be a useful tool for assessing the quality of maternally linked data, as well as other types of linked records where the linkages are within a single file. This is a necessary step towards developing methods for addressing misclassification bias in studies of maternally linked records through sensitivity analysis, adjustment, and other means.

Keywords Birth certificates · Epidemiologic methods · Medical record linkage · Sensitivity and specificity

Introduction

Maternally linked birth records datasets have emerged as a potentially powerful resource for investigating maternal and infant health in US populations [1]. These datasets consist of several consecutive years of vital birth records to which an algorithm has been applied to identify, for each woman, the set of records that represents her different births over the follow-up period [2]. These *maternal sets* are then analyzed as longitudinal data on pregnancies and births. Some maternally linked datasets include fetal death records.

Numerous studies have used maternally linked data to investigate perinatal outcomes [1, 3–16], maternal behaviors [7, 10, 11, 17, 18], and the quality of vital records birth data [19, 20]. In contrast, very little attention has been given to assessing linkage errors with respect to maternal sets and to understanding how such errors affect estimates derived from the maternally linked data. Progress in this area has been hampered by lack of a framework for conceptualizing the errors and lack of methods for quantifying them. The objectives of this study were to develop a framework for conceptualizing linkage errors in maternally linked datasets, to develop measures for quantifying the errors, and to demonstrate the application of the new measures to a maternally linked birth records dataset.

Error in assigning birth records to maternal sets was conceptualized as being analogous to misclassification in epidemiologic studies (see Table 1). This approach was chosen because the overall goal of this line of research is to develop quantitative techniques for incorporating adjustments

J. K. Leiss (✉)
Center for Health Research, Constella Group, LLC,
2605 Meridian Parkway, Suite 200, Durham,
NC 27713, USA
e-mail: jleiss@constellagroup.com

Table 1 Conceptual scheme for misclassification of maternal sets

Type of misclassification	True maternal sets		Assigned maternal sets	
	Set number	Record number	Set number	Record number
I None – assigned maternal set is equivalent to true maternal set	A	1	J	1
	A	2	J	2
	A	3	J	3
II True set is divided among different assigned sets	A	1	J	1
	A	2	J	3
	A	3	J	4
	A	4	K	2
	A	5	K	5
III No two records from the true set are included in the same assigned set (special case of II)	A	1	J	1
	A	2	K	2
	A	3	L	3
	A	4	M	4
	A	5	N	5
IV Assigned set combines different true sets	A	1	J	1
	A	2	J	2
	A	3	J	3
	B	4	J	4
	B	5	J	5
V Assigned set contains no two records from the same true set (special case of IV)	A	1	J	1
	B	2	J	2
	C	3	J	3
	D	4	J	4
	E	5	J	5
Illustration showing types II and IV together	A	1	J	1
	A	2	J	2
	A	3	J	3
	A	4	J	6
	A	5	K	7
	B	6	K	4
	B	7	K	5

for linkage errors into analysis of the linked records [21]. Two new measures, analogous to sensitivity and specificity, were developed to quantify the linkage errors. The measures were applied to a maternally linked birth records file, using a hospital birth log file as the gold standard to calculate true and false linkage rates. Varying degrees of random error were introduced into the maternal linkages and the behavior of the new measures under these known conditions was examined.

Methods

Jaro’s [22] method of record linkage, as implemented in the AutoMatch software (version 4.3, MatchWare Technologies, Inc., Kennebunk, ME), was used to construct a maternally linked dataset from North Carolina resident in-state birth and fetal death records for 1988–1997. The linkage strategy

is given in the Appendix. A typical internal validation was performed by assessing the logical consistency of selected variables across records within maternal sets. This maternally linked dataset was used as the baseline for assessing the new measures. Evaluating the quality of the maternal linkages in this file was not an objective of the present study; the results of the internal validation are given for purposes of describing the baseline data.

Five additional files were created by introducing increasing degrees of random linkage errors into the maternal sets of the original linked file. In each case, a specified proportion (1, 2, 5, 10, and 20 percent, respectively) of the records was randomly reassigned to different maternal sets, thus simulating errors in the maternal linkages. The new measures were then applied to these files, with the expectation that the measured error should increase with the increasing proportion of simulated errors.

Table 2 Two measures of misclassification of maternal sets in maternally linked birth records

True linkage proportion		False linkage proportion	
Degree to which gold standard sets are represented as sets in birth records file	Types of misclassification* counted in this cell	Degree to which birth records sets are composed of one gold standard set only	Types of misclassification* counted in this cell
No. of gold standard sets	No. of gold standard sets	No. of birth records sets	No. of birth records sets
Optimal	I	Optimal	I
Completely	I (IV)	Completely	I (II)
Partially	II (IV)	Partially	IV (III)
Not-at-all	III (IV, V)	Not-at-all	V (II, III)
Total	a + b + c	Total	d + e + f
	$a \cap d$		$a \cap d$
	$\frac{a \cap d}{a + b + c}$		$\frac{a \cap d}{d + e + f}$
	$\frac{a}{a + b + c}$		$\frac{d}{d + e + f}$
	$\frac{b}{a + b + c}$		$\frac{e}{d + e + f}$
	$\frac{c}{a + b + c}$		$\frac{f}{d + e + f}$

Note. The categories *completely*, *partially*, and *not-at-all* are mutually exclusive. *Optimal* is a subset of *completely*.

*From Table 1. The types in parentheses can exist in the cell but do not determine the counts for the proportions.

The external gold standard file consisted of the birth log for 1988–1997 of one North Carolina hospital, organized into maternal sets by the mother’s hospital ID number.

For calculating linkage error rates, AutoMatch was used to identify records in the birth file that corresponded to records in the gold standard file, matching on mother’s and infant’s names, birth dates, etc. Gold standard-birth record pairs that met the matching criteria were considered as representing the same birth. To simplify the presentation, references to birth records in the following text includes fetal death records.

Quantifying linkage errors

Errors in the composition of maternal sets were conceptualized in two dimensions: the *true linkage proportion*, analogous to sensitivity, which captures the degree to which all of a woman’s births were assigned to a single maternal set, as opposed to being divided among different sets; and the *false linkage proportion*, analogous to specificity, which captures the degree to which the assigned maternal sets combined births from different women.

The true linkage proportion was operationalized as the percent of maternal sets of size two or greater in the gold standard file that were completely, partially, or not-at-all represented as sets in the birth records file (see Table 2). *Completely* represented means that all of the births that comprised the gold standard set were assigned to the same maternal set in the birth records file (but the birth records set could include other births as well, i.e., from other gold standard sets). *Partially* represented means that at least two but not all of the births in the gold standard set were assigned to the same maternal set in the birth records file. *Not-at-all* represented means that no two births from the gold standard set were assigned to the same birth records set. In cases where there was not a birth record corresponding to the gold standard record, the gold standard record was considered assigned to a separate birth records set.

The false linkage proportion was operationalized as the percent of maternal sets of size two or greater in the birth records file that were completely, partially, or not-at-all composed of births from one gold standard set only (see Table 2). *Completely* means that all of the births in the birth records set were from a single gold standard set (but the birth records set did not necessarily encompass all of the births from that gold standard set). *Partially* means that at least two but not all of the births in the birth records set were from the same gold standard set. *Not-at-all* means that no two births in the birth records set were from the same gold standard set. In cases where there was not a gold standard record corresponding to the birth record, the birth record was considered as representing a separate gold standard set.

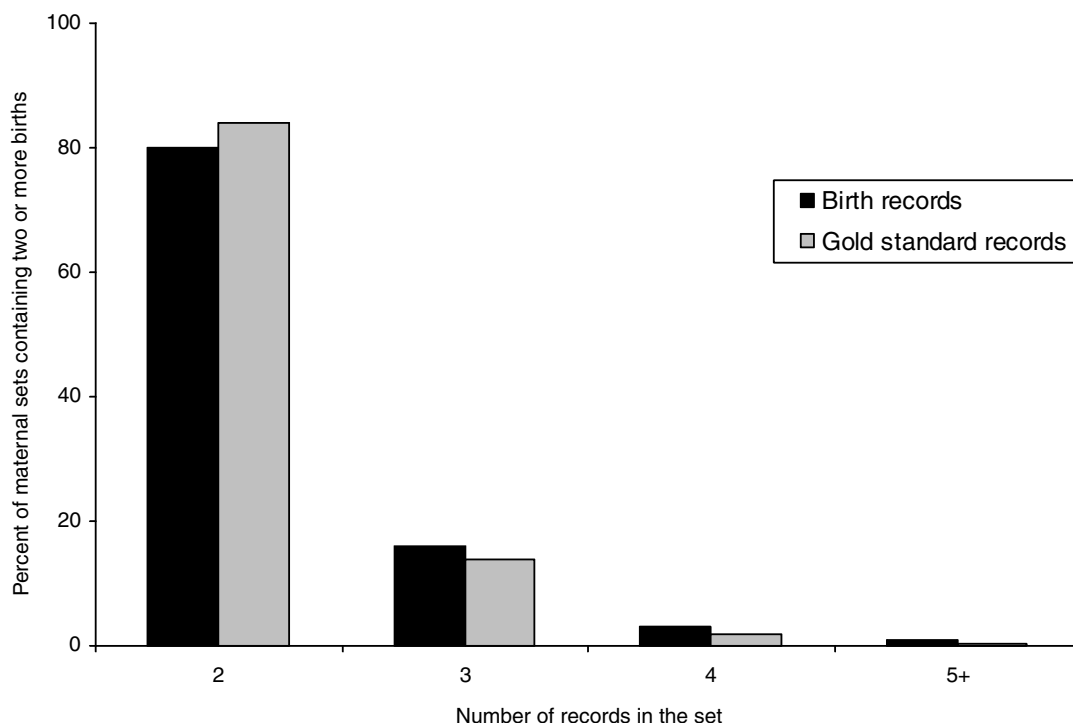


Fig. 1 Distribution of maternal sets in a maternally linked birth records file and a hospital birth log file (gold standard), by number of records in the set

The above categories are ordinal in the same direction for both the true and false linkage proportions: *completely* represents the smallest amount of misclassification and *not-at-all* represents the greatest amount of misclassification. In addition, an *optimal* category was defined for sets in which there was a one-to-one correspondence between the records in the birth set and the records in the gold standard set. Thus, *optimal* sets represent the absence of misclassification, and the *optimal* category is a subset of the *completely* category. An analogous designation for the opposite situation — total misclassification — could not be defined because any pairing of birth records sets with gold standard sets for this purpose would be arbitrary.

A variation of these measures calculated the percent of births included in sets variously categorized as above. These measures performed similarly to those using set as the unit of analysis. Only the latter are reported here.

Because the gold standard file corresponded to a small subset of the birth records file, calculation of the above measures was based on selected subsets of the birth and gold standard records. The base population for the true linkage proportion consisted of maternal sets in the gold standard file that included at least two births, along with all birth records that corresponded to those gold standard records. The base population for the false linkage proportion consisted of maternal sets in the birth records file that included at least two

births, at least one of which corresponded to a gold standard record, along with all of the gold standard records that corresponded to those birth records; birth records sets that did not have at least one corresponding gold standard record (a majority of the birth records sets) were not included in the base population for the false linkage proportion.

Results

There were 1,010,788 birth and 9,022 fetal death records in the birth file. (Due to a programming error, 10,601 (1.0%) birth records were excluded from this file.) From this, 234,235 maternal sets of two or more records were identified, 80% of which consisted of two records. The distribution of the sets by size (number of records) is shown in Fig. 1.

The results of the internal validation showed that the change in mother's age between two consecutive linked records in the maternally linked file corresponded to the difference in time between the two events in over 99% of linked record pairs; the estimated date of the beginning of gestation for a later event began after the occurrence of the previous event in 97% of pairs; the dates of occurrence of the previous event as indicated on the record for the next event and on its own record matched in over 95% of pairs; and parity increased by one between 91% of consecutive birth

Table 3 Maternal linkage error rates before introduction of simulated errors

Category	True linkage proportion (%)	False linkage proportion (%)
Optimal	72.5	35.1
Completely	87.8	36.1
Partially	9.4	7.7
Not-at-all	2.8	56.2
<i>n</i> (no. of maternal sets)	3,447	7,133

Note. The categories are in increasing order of misclassification for both true and false linkages. *Optimal* is a subset of *completely*.

records and by two or three between an additional 3% of consecutive birth records.

The gold standard file contained records for 21,875 births and 394 fetal deaths, including 3,447 maternal sets of two or more records. The distribution of the sets by size is very similar to that of the birth records. In matching the gold standard and birth files, corresponding records in the birth file were identified for 95% of the records in the gold standard file (Fig. 1).

The true and false linkage proportions for the original (i.e., without simulated errors) linkages are shown in Table 3. There were 3,447 gold standard maternal sets in the base population for the true linkage proportion, and 7,133 birth records maternal sets in the base population for the false linkage proportion. For the true linkage proportion, 87.8% of the gold standard sets were categorized as *completely*, whereas for the false linkage proportion, 36.1% of the birth records sets were categorized as *completely*. As the percentage of simulated errors increased, both linkage proportions shifted in the direction of greater misclassification (Fig. 2). With 20% of the birth records randomly re-assigned to a different maternal set, 54.5% of the gold standard sets were categorized as *completely* (true linkage proportion) and 12.9% of the birth records sets were categorized as *completely* (false linkage proportion) respectively.

Discussion

Results of epidemiologic studies using maternally linked birth records reflect bias and imprecision introduced by errors in the linkages. To date, methods for assessing the impact of those errors on the validity and reliability of the results, and for interpreting or adjusting the results as appropriate, have not been developed. As a first step in that direction, this study developed and tested new measures for quantifying maternal linkage errors. The conceptual framework and quantitative measures were guided by the general epidemiologic approach to misclassification, i.e., determining

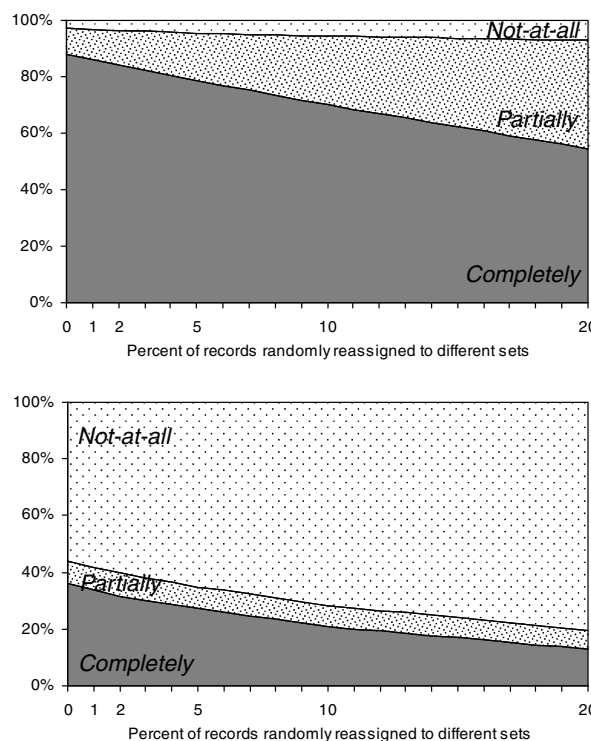


Fig. 2 True (upper graph) and false (lower graph) linkage proportions in a maternally linked birth records file with increasing levels of simulated linkage errors. For both proportions, the *completely* category represents the least degree of misclassification and the *not-at-all* category represents the greatest degree of misclassification

the sensitivity and specificity of the operationalized measure with reference to a gold standard. However, sensitivity and specificity are not directly applicable to the maternal linkage context because sensitivity and specificity are derived by comparing alternate categorizations of individuals, whereas assessing misclassification in maternal linkages involves comparing alternate compositions of groups (i.e., maternal sets).

To develop measures of misclassification for this situation, misclassification was conceptualized as a quality of the maternal sets rather than a quality of individual records. It was further conceptualized as a continuous characteristic, although it was operationalized as a categorical variable with three levels. Specifically, the true linkage proportion was proposed for capturing the notion of true positives as usually represented by sensitivity, and the false linkage proportion was proposed for capturing the notion of true negatives as usually represented by specificity.

These two measures behaved as expected when increasing degrees of error were introduced into the linkages—the distributions of the sets shifted in the direction of greater misclassification (see Fig. 2). For the true linkage proportion, the change was nearly linear and consisted primarily of a shift out of the *completely* category and into the *partially*

category. For the false linkage proportion, the rate of change decreased somewhat as the percent of errors increased, and consisted primarily of a shift out of the *completely* category and into the *not-at-all* category. These results support the conclusion that the true and false linkage proportions constitute valid measures of maternal linkage error, although further development and evaluation are needed.

In addition to serving as measures of misclassification with reference to a gold standard, the true and false linkage proportions can be used to compare different maternally linked files that are based on the same data but different linkage methods. For example, the measures can be used to compare files produced by different match specifications or different linkage software. This could provide useful information for developing final match specifications or for choosing among alternative software applications.

When suitable gold standard files are available, the true and false linkage proportions should be calculated and reported in studies analyzing the linked records, much as response rates are reported for surveys. Furthermore, reviewers and editors should expect investigators using maternally linked files to demonstrate the quality of their data by reporting the true and false linkage proportions, or other measures of linkage quality that may be developed. The current practice of reporting the percent of records that matched is meaningless as an indicator of linkage quality.

Adams et al. [2] conducted the only other published quantitative evaluation of maternal linkages. They reported the percent of sets in which the number of records in the set differed in each direction by one, two, three, or four or more from the expected number. Two such comparisons were made, one with expected numbers derived from obstetric history information on the records themselves, and the other with expected numbers obtained from interviews with a small proportion of the birth cohort mothers. The true and false linkage proportions proposed in this paper extend Adams et al.'s aggregate approach to one based on counts of maternal sets containing misclassified records. This yields additional information about the nature of the misclassification, and will facilitate the development of quantitative techniques for assessing the impact of misclassification on studies using maternally linked data.

Further development of the measures introduced in this study should examine their behavior with gold standard files of different sizes relative to the maternally linked file, as well as examining how the measures are influenced by missing data (i.e., a gold standard record that is missing a corresponding birth record, for calculating the true linkage proportion, or a birth record that is missing a corresponding gold standard record, for calculating the false linkage proportion). The gold standard file used in this study was

small relative to the birth records file. This difference in size may have produced a large proportion of missing data, which in turn may explain the different patterns shown by the true and false linkage proportions as linkage errors increased. Moreover, each missing corresponding record was counted as an additional set. Future research should determine the minimum relative size of a gold standard file necessary to obtain meaningful assessments of the linked file, and the optimal method of handling missing data under various conditions.

Future research should also investigate operationalizing the linkage proportions as continuous variables, and how the distribution of maternal set size influences the true and false linkage proportions. Although the *completely* and *not-at-all* categories capture the maximum and minimum degrees of misclassification, the *partially* category could include a wide range of intermediary degrees of misclassification. However, this is constrained by the distribution of set size. Sets of size two, which will generally account for the majority of maternally linked sets, can be categorized as *completely* or *not-at-all*, but not as *partially*.

Although the comparison files are called “gold standards,” perfect comparison files will rarely be available. Future research is needed to identify the relevant characteristics of potential comparison files, to understand how variations in these characteristics affect the true and false linkage proportions, and to identify potential sources of comparison files. Possible sources include records from maternal and child health programs, such as WIC; medical records, such as the hospital birth log used in this study; survey data, such as that used by Adams et al. [2]; and a validated subset of records, as is commonly used for internal validation studies [21].

Finally, the terminology introduced in this paper is somewhat awkward. Improvements would aid in communicating results using the new measures.

The measures developed in this paper are relevant for epidemiologic studies beyond those using maternally linked data. In general, the measures can be used to quantify errors in assignment where the unit of analysis is the group, and values for group-level variables are obtained by aggregating the values of the individual units that have been assigned to the groups. This type of design is often found in studies of institution-related populations, especially schools [23], and in follow-up studies that combine exposure measures from different sources [24]. For maternally linked and similarly structured data, the measures constitute a scheme for conceptualizing the mis-assignment and a technique for measuring it. This is a necessary step towards the ultimate goal of developing methods for assessing misclassification bias in parameter estimates derived from the linked data and for addressing such bias through sensitivity analysis [25], adjustment [21], and other means.

Appendix

Linkage strategy for the maternally linked data set

Appendix Table. Match specifications for the maternally linked data set

Matching variable	m-probability	u-probability	Type of comparison	Comparison parameter
Mother's last name	0.9	0.01	uncert	700.0
Mother's first name	0.9	0.01	uncert	700.0
Mother's maiden name	0.9	0.01	uncert	700.0
Mother's middle name	0.7	0.01	uncert	700.0
Mother's date of birth	0.99	0.01	cnt-diff	1
Mother's age	0.9	0.01	numeric	
Date of occurrence	0.9	0.01	uncert	700.0
Father's last name	0.7	0.01	uncert	700.0
Father's first name	0.7	0.01	uncert	700.0
Father's middle name	0.7	0.01	uncert	700.0
Father's date of birth	0.7	0.01	cnt-diff	1
Mother's street address*	0.9	0.01	uncert	850.0
Mother's race	0.6	0.01	uncert	700.0
Mother's Hispanic origin	0.6	0.01	uncert	700.0
Mother's place of birth	0.9	0.01	numeric	
Father's race	0.5	0.01	uncert	700.0
Father's Hispanic origin	0.5	0.01	uncert	700.0
Father's place of birth	0.5	0.01	numeric	

*Disagree weight set to zero.

The file was linked in three passes. All passes blocked on mother's month of birth. In addition, pass 1 blocked on the soundex values of mother's first and maiden names, pass 2 blocked on the soundex values of mother's first and last names, and pass 3 blocked on the soundex values of mother's last and middle names. The same value was used for the match and clerical review cutoff weights within a pass (i.e., clerical review was not performed). The values were 30.0, 20.0, and 35.0, respectively for passes 1, 2, and 3. The match specifications were identical for each pass and are given in the table above.

Acknowledgements The author thanks Drs. Russell Kirby and C. V. Ananth for comments on an earlier version of this paper. This study was funded in part by grants HD35785 from the National Institute of Child Health and Human Development, CA88757 from the National Cancer Institute, and UR6/CCU417428 from the National Center for Health Statistics.

References

- Herman AA, McCarthy BJ, Bakewell JM, Ward RH, Mueller BA, Maconochie NE, Read AW, Zadka P, Skjaerven R. Data linkage methods used in maternally-linked birth and infant death surveillance data sets from the United States (Georgia, Missouri, Utah and Washington State), Israel, Norway, Scotland and Western Australia. *Paediatr Perinat Epidemiol* 1997;11(Suppl 1):5–22.

- Adams MM, Wilson HG, Casto DL, Berg CJ, McDermott JM, Gaudino JA, McCarthy BJ. Constructing reproductive histories by linking vital records. *Am J Epidemiol* 1997;145(4):339–48.
- Adams MM, Elam-Evans LD, Wilson HG, Gilbertz DA. Rates of and factors associated with recurrence of preterm delivery. *JAMA* 2000;283(12):1591–6.
- Adams MM, Delaney KM, Stupp PW, McCarthy BJ, Rawlings JS. The relationship of interpregnancy interval to infant birthweight and length of gestation among low-risk women, Georgia. *Paediatr Perinat Epidemiol* 1997;11(Suppl 1):48–62.
- Bakewell JM, Stockbauer JW, Schramm WF. Factors associated with repetition of low birthweight: Missouri longitudinal study. *Paediatr Perinat Epidemiol* 1997;11(Suppl 1):19–29.
- Doody DR, Patterson MQ, Voigt LF, Mueller BA. Risk factors for the recurrence of premature rupture of the membranes. *Paediatr Perinat Epidemiol* 1997;11(Suppl 1):96–106.
- Holt VL, Danoff NL, Mueller BA, Swanson MW. The association of change in maternal marital status between births and adverse pregnancy outcomes in the second birth. *Paediatr Perinat Epidemiol* 1997;11(Suppl 1):31–40.
- McGuire V, Rauh MJ, Mueller BA, Hickock D. The risk of diabetes in a subsequent pregnancy associated with prior history of gestational diabetes or macrosomic infant. *Paediatr Perinat Epidemiol* 1996;10(1):64–72.
- Mueller BA, Schwartz SM. Risk of recurrence of birth defects in Washington State. *Paediatr Perinat Epidemiol* 1997;11(Suppl 1):107–18.
- Schramm WF. Smoking during pregnancy: Missouri longitudinal study. *Paediatr Perinat Epidemiol* 1997;11(Suppl 1):73–83.
- McDermott JM, Drews CD, Adams MM, Hill HA, Berg CJ, McCarthy BJ. Does inadequate prenatal care contribute to growth retardation among second-born African-American babies? *Am J Epidemiol* 1999;150(7):706–13.

12. Krulewitch CJ, Herman AA, Yu KF, Johnson YR. Does changing paternity contribute to the risk of intrauterine growth retardation? *Paediatr Perinat Epidemiol* 1997;11(Suppl 1):41–7.
13. Zhu BP, Le T. Effect of interpregnancy interval on infant low birth weight: a retrospective cohort study using the Michigan Maternally Linked Birth Database. *Matern Child Health J* 2003;7(3):169–78.
14. Beaty TH, Yang P, Munoz A, Khoury MJ. Effect of maternal and infant covariates on sibship correlation in birth weight. *Genet Epidemiol* 1988;5(4):241–53.
15. Yang P, Beaty TH, Khoury MJ, Liang KY, Connolly MA. Predicting intrauterine growth retardation in sibships while considering maternal and infant covariates. *Genet Epidemiol* 1989;6(4):525–35.
16. Cooperstock MS, Tummaru R, Bakewell J, Schramm W. Twin birth weight discordance and risk of preterm birth. *Am J Obstet Gynecol* 2000;183(1):63–7.
17. Dietz PM, Adams MM, Rochat RW, Mathis MP. Prenatal smoking in two consecutive pregnancies: Georgia, 1989–1992. *Matern Child Health J* 1997;1(1):43–51.
18. Elam-Evans LD, Adams MM, Delaney KM, Wilson HG, Rochat RW, McCarthy BJ. Patterns of prenatal care initiation in Georgia, 1980–1992. *Obstet Gynecol* 1997;90(1):71–7.
19. Green DC, Moore JM, Adams MM, Berg CJ, Wilcox LS, McCarthy BJ. Are we underestimating rates of vaginal birth after previous cesarean birth? The validity of delivery methods from birth certificates. *Am J Epidemiol* 1998;147(6):581–6.
20. Adams M. Validity of birth certificate data for the outcome of the previous pregnancy, Georgia, 1980–1995. *Am J Epidemiol* 2001;154(10):883–8.
21. Greenland S. Basic methods for sensitivity analysis and external adjustment. In: Rothman KJ, Greenland S, editors. *Modern Epidemiology*. Lippincott Williams & Wilkins, Philadelphia; 1998:347–55.
22. Jaro MA. Probabilistic linkage of large public health data files. *Stat Med* 1995;14(5–7):491–98.
23. Sirard JR, Ainsworth BE, McIver KL, Pate RR. Prevalence of active commuting at urban and suburban elementary schools in Columbia, SC. *Am J Public Health* 2005;95(2):236–7.
24. Miller AB, Howe GR, Sherman GJ, Lindsay JP, Yaffe MJ, Dinner PJ, Risch HA, Preston DL. Mortality from breast cancer after irradiation during fluoroscopic examinations in patients being treated for tuberculosis. *N Engl J Med* 1989;321(19):1285–9.
25. Lash TL, Fink AK. Semi-automated sensitivity analysis to assess systematic errors in observational data. *Epidemiology* 2003;14(4):451–8.