

Invited Commentary: Using Vital Statistics Databases for Perinatal Epidemiology: Does the Quality Go In Before the Name Goes On?

Russell S. Kirby

Most epidemiologic research using birth certificate records utilizes cross-sectional or case-control study designs. These records are typically made available to qualified researchers as computerized files of vital events for a state or the entire United States by year of occurrence, either for events occurring within the state or to residents of the state. This structure clearly lends itself to the cross-sectional approach. Occasionally, researchers step outside the box and conduct studies utilizing prospective study designs, generally with one of two approaches: 1) intergenerational studies based on record linkage between the birth certificate for the index case and the birth certificate for the mother of the index case (1–3) or 2) maternally linked pregnancy outcomes based on record linkage of the birth certificate for a firstborn infant of the index mother and her subsequent birth outcomes (4). The intergenerational approach is somewhat limited due to the smaller number of statistical data fields collected on birth certificates in the past (5). Datasets with maternally linked pregnancy outcomes can be used to study birth outcomes within sibships, especially for conditions in which a prior adverse outcome (for example, low birth weight or fetal death) may be associated with an adverse outcome in a subsequent pregnancy (6).

Over the past half-century, the US standard certificate of livebirth has gradually expanded to include numerous questions concerning reproductive history, characteristics of the infant at birth, weight gain, use of tobacco and alcohol during pregnancy, and a variety of clinical data elements ranging from risk factors associated with complications of pregnancy to intrapartum events to birth defects and neonatal therapies (5, 7, 8). This expansion of the range of data elements collected through vital registration has, unfortunately, not been accompanied by careful evaluation of the quality of the data reported on birth certificate documents.

Most published birth certificate data quality studies have limitations. First, many studies are secondary methods analyses conducted by using available data collected for another research project (9–11). The generalizability of the results of some of these analyses is limited, and it is difficult to extrapolate study results to a comprehensive statewide birth certificate data quality improvement initiative. Second, many stud-

ies are too small to yield valid and reliable results to guide epidemiologists interested in using these data; either the samples are too small (12), or the study is based on data from one or, at most, a few hospitals (13, 14). Third, the range of data elements analyzed is often narrow in focus (15–19). These studies yield useful information on the topic in question, but the results are less informative for use in designing and implementing comprehensive data-quality improvement programs. Fourth, in some studies, the comparison data used as the “gold standard” have not been subjected to an analysis of their own reliability and validity (20). Thus, there is an urgent need for careful research on the quality of data elements reported on birth certificates (21) and an equally pressing imperative for administrative health database quality improvement initiatives utilizing modern computer technologies coupled with the results of this research (22).

The study reported in this issue contributes to this literature. Adams (23) examines the quality of birth certificate reporting of information concerning prior reproductive events by using the actual birth certificate for those events as the gold standard. This study is based on a dataset created by a carefully conducted and scientifically validated record linkage project (24, 25). The analysis focuses on singleton pairs of first- and second-born infants or fetal deaths with gestational ages of 20–44 weeks. While the study period is lengthy (1980–1995), it makes possible an analysis of the effects of changes in birth certificate design on reporting of information on outcomes of prior pregnancies. Since the National Center for Health Statistics revises the national standard certificates for vital events periodically, with the next anticipated for implementation in 2003 or 2004, this examination of the effect of changes effective with the 1989 revisions serves a useful purpose.

Adams examines reporting of three specific outcomes of the prior pregnancy as reported on the vital record for the subsequent pregnancy. Owing to birth certificate form changes, the analysis of outcome of previous pregnancy was limited to the period 1980–1988. This analysis shows systematic bias toward reporting of adverse events in the previous pregnancy when an adverse event occurred in the index case, both for low birth weight and for stillbirth or neonatal death. Second, a previous history of low or high birth weight was studied for second births that occurred during 1989–1995. This analysis also shows systematic bias toward more sensitive reporting of low birth weight in the first birth if the index case was low birth weight. Finally, reporting of mortality for the first birth on the birth certificate for the second birth was compared with

Received for publication August 9, 2001, and accepted for publication August 28, 2001.

From the University of Wisconsin Medical School, Milwaukee Clinical Campus, 950 N. 12th Street, Room A-504, P.O. Box 342, Milwaukee WI 53201-0342 (e-mail: r-kirby@whin.net). (Reprint requests to Dr. Kirby at this address).

information from the linked infant death certificates. Both sensitivity and predictive value positive were high and improved during the period after birth certificate revision.

Researchers who analyze birth certificate files with cross-sectional study designs should take notice of these study findings, as should vital statisticians responsible for the production of tabulations for annual reports and state documents. If the sensitivity of these historically derived variables is only low to moderate and there is also systematic bias in the reporting of adverse events secondary to the index case, is it responsible to use these data elements in multivariate analyses or to report them without appropriate caveats in official reports?

Developing methods to improve the quality of historically derived variables reported on vital statistics documents does not require revolutionary changes in system administration but may necessitate a paradigm shift in the organization and structure of data repositories and record access. Currently, each vital record is treated as a unique document and is managed as an entity unto itself. Only in special situations such as infant mortality (and, in some states, maternal mortality) are vital records linked across persons in a time frame proximal to the occurrence of the index event. If vital records repositories were treated as resources for evaluating the quality of statistical information on incoming records, cross-validation of historically derived variables could be done as records are filed, with queries sent back to reporting sources for confirmation or correction of the items in question. Alternatively, implementation of national standards for the design and content of perinatal clinical information systems applications would ensure that information concerning maternal reproductive histories reported on vital records is derived from prenatal clinical records available during the intrapartum period rather than from maternal postpartum self-report.

Neither primary nor secondary analyses of administrative health databases should be conducted without some knowledge of the reliability and validity of the statistical data contained therein. While the creators and managers of these databases have some obligation to assure the completeness and accuracy of information reported, researchers must also engage in a symbiotic relationship with staff in hospitals and public health data agencies to support continuous data quality improvement activities. In the current political environment, state and federal health data managers will not have sufficient staff or funds to undertake this essential activity without the assistance of the research community. If additional studies such as that reported by Adams and others are not conducted, with the results utilized to create effective vital statistics quality improvement initiatives, the answer to the rhetorical question posed in the title of this commentary will be a resounding no!

REFERENCES

- Emanuel I, Filakti H, Alberman E, et al. Intergenerational studies of human birthweight from the 1958 birth cohort. I. Evidence for a multigenerational effect. *Br J Obstet Gynaecol* 1992;99:67-74.
- Hennessy E, Alberman A. Intergenerational influences affecting birth outcome. I. Birthweight for gestational age in the children of the 1958 British birth cohort. *Paediatr Perinat Epidemiol* 1998;12:S45-S60.
- Klebanoff MA, Schulsinger C, Mednick BR, et al. Preterm and small-for-gestational-age birth across generations. *Am J Obstet Gynecol* 1997;176:521-6.
- Herman AA, McCarthy BJ, Bakewell JM, et al. 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:S5-S22.
- Hetzel AM. History and organization of the vital statistics system. Hyattsville, MD: National Center for Health Statistics, 1997. (DHHS publication no. (PHS) 97-1003).
- Bakketeig LS, Hoffman HJ. Perinatal mortality by birth order within cohorts based on sibship size. *Br Med J* 1979;2:693-6.
- Tolson GC, Barnes JM, Gay GA, et al. The 1989 revision of the U.S. standard certificates and reports. *Vital Health Stat* 4 1991;28:1-34.
- Report of the Panel to Evaluate the U.S. Standard Certificates and Reports. Hyattsville, MD: Division of Vital Statistics, National Center for Health Statistics, April 2000.
- Piper JM, Mitchel EF Jr, Snowden M, et al. Validation of 1989 Tennessee birth certificates using maternal and newborn hospital records. *Am J Epidemiol* 1993;137:758-68.
- Dobie SA, Baldwin LM, Rosenblatt RA, et al. How well do birth certificates describe the pregnancies they report? The Washington State experience with low-risk pregnancies. *Matern Child Health J* 1999;2:145-54.
- Parrish KM, Holt VL, Connell FA, et al. Variations in the accuracy of obstetric procedures and diagnoses in birth records in Washington State, 1989. *Am J Epidemiol* 1993;138:119-27.
- Buescher PA, Taylor KP, Davis MH, et al. The quality of the new birth certificate data: a validation study in North Carolina. *Am J Public Health* 1993;83:1163-5.
- Snell LM, Little BB, Knoll KA, et al. Reliability of birth certificate reporting of congenital anomalies. *Am J Perinatol* 1992;9:219-22.
- Costakos DT, Love LA, Kirby RS. The computerized perinatal database: are the data reliable? *Am J Perinatol* 1998;15:453-9.
- Watkins ML, Edmonds L, McClearn A, et al. The surveillance of birth defects: the usefulness of the revised US standard birth certificate. *Am J Public Health* 1996;86:731-4.
- Clark K, Fu C-M, Burnett C. Accuracy of birth certificate data regarding the amount, timing, and adequacy of prenatal care using prenatal clinical medical records as referents. *Am J Epidemiol* 1997;145:68-71.
- McDermott J, Drews C, Green D, et al. Evaluation of prenatal care information on birth certificates. *Pediatr Perinat Epidemiol* 1997;11:105-21.
- Green DC, Moore JM, Adams MM, et al. Are we underestimating rates of vaginal birth after previous cesarean birth? The validity of delivery methods from birth certificates. *Am J Epidemiol* 1998;147:581-6.
- Baumeister L, Marchi K, Pearl M, et al. The validity of information on "race" and "Hispanic ethnicity" in California birth certificate data. *Health Serv Res* 2000;35:869-83.
- Reichman NE, Hade EM. Validation of birth certificate data. A study of women in New Jersey's HealthStart program. *Ann Epidemiol* 2001;11:186-93.
- Kirby RS. The quality of data reported on birth certificates. *Am J Public Health* 1997;87:301.
- Starr S, Starr S. Reinventing vital statistics: the impact of changes in information technology, welfare policy, and health care. *Public Health Rep* 1995;110:534-44.
- Adams MM. Validity of birth certificate data for the outcome of the previous pregnancy, Georgia, 1980-1995. *Am J Epidemiol* 2001;154:883-8.
- Adams MM, Wilson HG, Casto DL, et al. Constructing reproductive histories by linking vital records. *Am J Epidemiol* 1997;145:339-48.
- Adams MM, Berg CJ, McDermott JM, et al. Evaluation of reproductive histories constructed by linking vital records. *Paediatr Perinat Epidemiol* 1997;11:78-92.