Validity of Health Plan and Birth Certificate Data for Pregnancy Research

Susan E. Andrade[1], **Pamela E. Scott**[2], **Robert L. Davis**[3], **De-Kun Li**[4], **Darios Getahun**[5], **T. Craig Cheetham**[6], **Marsha A. Raebel**[7], **Sengwee Toh**[8], **Sascha Dublin**[9], **Pamala A. Pawloski**[10], **Tarek A. Hammad**[2], **Sarah J. Beaton**[11], **David H. Smith**[12], **Inna Dashevsky**[8], **Katherine Haffenreffer**[8], and **William O. Cooper**[13]

[1]Meyers Primary Care Institute and University of Massachusetts Medical School, Worcester, MA, USA

[2]Center for Drug Evaluation and Research, U.S. Food and Drug Administration, Silver Spring, MD, USA

[3]Center for Health Research, Kaiser Permanente Georgia, Atlanta, GA, USA

[4]Division of Research, Kaiser Permanente Northern California, Oakland, CA, USA

[5]Department of Research and Evaluation, Kaiser Permanente Southern California, Pasadena, CA, USA

[6]Pharmacy Analytical Service, Kaiser Permanente, Downey, CA, USA

[7]Institute for Health Research, Kaiser Permanente Colorado, Denver, CO, USA

[8]Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute, Boston, MA, USA

[9]Group Health Research Institute, Seattle, WA, USA

[10]HealthPartners Research Foundation, Bloomington, MN, USA

[11]LCF Research, Albuquerque, NM, USA

[12]Center for Health Research, Kaiser Permanente Northwest, Portland, OR, USA

[13]Department of Pediatrics and Preventive Medicine, Vanderbilt University School of Medicine, Nashville, TN, USA

Abstract

Purpose—To evaluate the validity of health plan and birth certificate data for pregnancy research.

Methods—A retrospective study was conducted using administrative and claims data from 11 U.S. health plans, and corresponding birth certificate data from state health departments. Diagnoses, drug dispensings, and procedure codes were used to identify infant outcomes (cardiac defects, anencephaly, preterm birth, and neonatal intensive care unit [NICU] admission) and

Corresponding author and address for reprints: Susan Andrade, Sc.D., Meyers Primary Care Institute, 630 Plantation St. Worcester, MA 01605, Telephone: (508)791-7392, Fax: (508)595-2200, sandrade@meyersprimary.org.

maternal diagnoses (asthma and systemic lupus erythematosus [SLE]) recorded in the health plan data for live born deliveries between January 2001 and December 2007. A random sample of medical charts (n = 802) was abstracted for infants and mothers identified with the specified outcomes. Information on newborn, maternal, and paternal characteristics (gestational age at birth, birth weight, previous pregnancies and live births, race/ethnicity) was also abstracted and compared to birth certificate data. Positive predictive values (PPVs) were calculated with documentation in the medical chart serving as the gold standard.

Results—PPVs were 71% for cardiac defects, 37% for anencephaly, 87% for preterm birth, and 92% for NICU admission. PPVs for algorithms to identify maternal diagnoses of asthma and SLE were   93%. Our findings indicated considerable agreement (PPVs > 90%) between birth certificate and medical record data for measures related to birth weight, gestational age, prior obstetrical history, and race/ethnicity.

Conclusions—Health plan and birth certificate data can be useful to accurately identify some infant outcomes, maternal diagnoses, and newborn, maternal, and paternal characteristics. Other outcomes and variables may require medical record review for validation.

Keywords

administrative databases; birth certificate; positive predictive value; pregnancy; validation

---

Introduction

Administrative and claims databases of health plans and birth certificate files are often used for epidemiologic studies evaluating pregnancy and infant outcomes. The usefulness of these data sources for research depends on the accuracy and completeness of the information they contain. Studies have indicated that the validity of health plan or birth certificate data for pregnancy research varies considerably depending on the specific data element or diagnosis of interest; however, most published studies in U.S. populations are based on data collected before 2000.[1–8]

The Medication Exposure in Pregnancy Risk Evaluation Program (MEPREP) is a collaborative research program developed to enable the conduct of studies of medication use and outcomes in pregnancy across participating sites.[9] Collaborators include the U.S. Food and Drug Administration and researchers at three contract sites: the HMO Research Network, Kaiser Permanente Northern and Southern California, and Vanderbilt University. In MEPREP, participating organizations (which include eleven different health plans) link administrative and claims health plan data with birth certificate data obtained from state health departments.

The aim of the present study was to evaluate the validity of health plan administrative and claims data and birth certificate files used to conduct epidemiologic studies within MEPREP. The data elements evaluated include infant outcomes (cardiac defects, anencephaly, preterm birth, admission to neonatal intensive care unit [NICU]) and maternal diagnoses (asthma, systemic lupus erythematosus [SLE]) recorded in the health plan data, as well as newborn, maternal, and paternal characteristics (gestational age at birth, birth weight, race/ethnicity, number of previous pregnancies and live births) included in the birth certificate files. These data elements were selected for evaluation based upon their importance, either for determination of medication exposure periods or as potential outcomes or confounders in medication safety research in pregnant populations, as well as researchers' specific interests for future studies. Specifically, maternal asthma was selected since uncontrolled asthma has been associated with increased maternal morbidity and mortality, and adverse pregnancy outcomes (e.g., preeclampsia, fetal growth restriction, and

preterm birth) and there are limited data on the safety of asthma medications for use during pregnancy.[10–12] An increased risk of maternal morbidity and mortality and adverse pregnancy outcomes (e.g., maternal infections, preeclampsia, stillbirth, preterm birth) have also been reported among women with SLE; a number of medications used in the treatment of SLE (e.g. methotrexate, mycophenolate, and cyclophosphamide) are contraindicated during pregnancy and there are limited data on the safety of others.[13–17] The birth certificate variables selected for validation included those that provided information not commonly found in health plan claims data (which would be expected to include maternal and infant diagnoses and procedures administered during pregnancy and the perinatal period).

Methods

Data Source

This study used data from MEPREP. Encompassed within this program are three U.S. Food and Drug Administration (FDA) contract sites that include 11 health plan-affiliated research institutions: Group Health Research Institute (Washington), Harvard Pilgrim Health Care Institute (Massachusetts), HealthPartners Research Foundation (Minnesota), Kaiser Permanente Colorado, Kaiser Permanente Georgia, Kaiser Permanente Northwest (Oregon, Washington), Meyers Primary Care Institute (Massachusetts), Lovelace Clinic Foundation (New Mexico), Kaiser Permanente Northern California, Kaiser Permanente Southern California, and Tennessee State Medicaid (through the auspices of Vanderbilt University School of Medicine). These research institutions provide access to health plan data of approximately 12 million current enrollees within nine states, covering socio-economically, geographically, and ethnically diverse populations with a broad age range receiving care within a wide array of medical care delivery models.

Researchers affiliated with the health plans extracted information on maternal and infant enrollment, demographics, outpatient pharmacy dispensings, and outpatient and inpatient health care encounters from health plan administrative and claims databases for live born deliveries. They linked the health plan data to state birth certificate data to access information on sociodemographic, medical, and reproductive factors, such as maternal race/ethnicity, parity and infant's gestational age at birth. All data were transformed into de-identified, standardized datasets. This study was approved by the Institutional Review Board of each participating organization, and the state departments of public health, where applicable.

Study Population

This retrospective study was conducted among mothers and infants enrolled in the eleven participating health plans. The source population included women delivering a live born infant between January 1, 2001 through December 31, 2007 and the infants born to these women.

At each site, potential cases of cardiac defects, anencephaly, preterm birth, or NICU admission were identified among infants in the MEPREP cohort based upon diagnoses (International Classification of Diseases, 9th revision, Clinical Modification [ICD-9-CM]) or procedure codes (Current Procedural Terminology [CPT]) recorded in health plan administrative or claims data. Diagnoses of asthma or SLE were similarly identified among mothers in the MEPREP cohort, based upon diagnosis and procedure codes and medication dispensings. Table 1 shows the codes and criteria for selection.

At each site, a random sample of patients identified with each of the diagnoses/outcomes indicated above was selected for chart review (Table 1). If a patient had more than one encounter with the diagnosis of interest during the observation period, one encounter date

was randomly selected to determine the encounter type (inpatient, outpatient, emergency department) documented in the health plan data; for infant outcomes, this encounter date also determined the source of data (infant or maternal administrative health plan records). For infant outcomes, encounters categorized as "inpatient" encounter types included data from the maternal delivery hospital admission and/or infant hospitalizations; "emergency department" and "ambulatory" encounters included only data from the infants' health plan records, not the mothers' (Table 1). The selected encounter date was also used to determine the health care provider or facility (clinic or hospital) of primary interest for medical record retrieval and review.

Medical Record Review and Adjudication

Confirmation of infant outcomes and maternal diagnoses—Medical records were located and abstracted for 802 (91%) of the 878 patients sampled for medical record review.

Chart reviews were performed using electronic or hard copy medical records. Trained chart abstractors at each site used a standard instrument to confirm the presence of a physician's diagnosis or diagnosis from a surgical, radiology, or autopsy report.

Outcomes other than preterm birth or cardiac defects were confirmed based upon documentation of the condition or history of the condition in the medical record. For preterm birth, the outcome was confirmed if the gestational age at birth documented in the medical record was < 37 completed weeks; if the gestational age at birth was not recorded, the event was confirmed based upon documentation of a diagnosis of preterm birth in the medical record. For all infants identified with a potential cardiac defect in the administrative or claims data, the chart abstraction forms and copies of de-identified medical records were reviewed and adjudicated by an investigator with expertise in birth outcomes research (RD, DG, DL, WC) to determine whether or not a cardiac defect was present. The National Birth Defects Prevention Study (NBDPS) Guidelines for *Conducting Birth Defects Surveillance* was used as a reference for the description of definitions, inclusions, and exclusions for specific malformations.[18] Documentation of a cardiologist diagnosis was sufficient evidence to confirm the diagnosis of an atrial septal defect (ASD), peripheral pulmonary stenosis (PPS), and ventricular septal defect (VSD). For all other cardiac defects of interest, an autopsy, surgical, cardiac catheterization, or echocardiography report was necessary to confirm the diagnosis, as specified in the NBDPS Guidelines.

Confirmation of data elements in birth certificate data—For infants selected for chart review based upon administrative data documentation of a cardiac defect, anencephaly, preterm birth, or NICU admission, we also assessed the validity of the following data elements from the birth certificate data through medical record review: gestational age at birth, birth weight, number of previous pregnancies and live births, and maternal and paternal race/ethnicity.

Neither the chart abstractors or adjudicators had access to birth certificate data obtained from the state departments of public health; however, they were aware of the diagnoses identified in the administrative data.

Statistical Analysis

The positive predictive values (PPVs) and 95% confidence intervals (95% CI) were calculated based upon documentation in the medical chart ("gold standard"). For each outcome or diagnosis, the PPV was calculated as the percentage of confirmed cases among patients with a code identified using the health plan administrative or claims databases. The PPV for mother's and father's race/ethnicity, as documented in birth certificate data, was

similarly determined. The proportion of infants for whom the birth weight documented in the birth certificate was within 5% of that documented in the chart was also estimated. Using the documented birth weight, we compared the proportion of infants categorized as low birth weight (LBW; < 2500 grams) in the birth certificate confirmed as LBW in the chart. The proportion of infants for whom the gestational age documented in the birth certificate was within 14 days of the gestational age documented in the medical record was evaluated; separate analyses were conducted for the gestational age values based upon date of last menstrual period (LMP), clinical assessment, or other methods included in the birth certificate. The proportion of mothers with at least one prior pregnancy or one prior live birth, as documented in the chart compared to the birth certificate, was evaluated.

For infant outcomes and maternal diagnoses, PPVs were estimated overall and according to encounter type. We also estimated PPVs based upon the source of data used to identify the outcome (infants' or the mother's health plan data). PPVs were calculated after exclusion of patients for whom charts were not reviewed. For birth certificate data, analyses were conducted for infants for whom both the chart and birth certificate were available.

Results

Infant Outcomes

Table 2 shows the PPVs of administrative health plan codes for identification of the infant outcomes of cardiac defects, anencephaly, preterm birth and NICU admission. Of the 276 infants identified with a potential cardiac defect for whom the charts were reviewed, 195 (71%) were confirmed to have at least one cardiac defect, with 269 total defects confirmed. Most defects (N=217; 81%) were confirmed based upon radiological evidence. The PPV for codes documented in the maternal hospital discharge data (35%; 95% CI 16%–57%) was substantially lower than the PPV for codes documented in the infant health plan data (74%; 95% CI 69%–79%). PPVs for specific defects varied greatly among the most commonly reported cardiac defects, with the highest PPVs found for ventricular septal defect (95%) and persistent fetal circulation (86%).

A total of 13 potential cases of anencephaly (37%) were confirmed. The PPV for codes documented in inpatient data (71%; 95% CI 44%–90%) was substantially higher than the PPV for codes documented in infant ambulatory data (6%; 95% CI 0%–27%).

The overall PPV for preterm birth was 87% (95% CI, 81%–92%). Again, the PPV for codes documented in maternal hospital discharge data was lower (76%; 95% CI 64%–88%) than that for codes in infant health plan data (92%; 95% CI 87%–98%).

The PPV for health plan data for identification of a NICU admission (92%) was similar among the most commonly used codes and different data sources.

The proportions of cases confirmed for anencephaly, preterm birth, and NICU admission were similar across FDA contract sites. However, PPVs varied across sites for overall cardiac defects, ranging from 53% (95% CI, 43%– 63%) to 87% (95% CI, 80%– 94%). For atrial septal defect, the PPVs ranged from 46% (95% CI, 29%–64%) to 88% (95% CI, 69%–97%); in contrast, the PPVs for ventricular septal defect were similar across FDA contract sites, ranging from 91% to 97%.

Maternal Diagnoses

PPVs of algorithms based upon administrative health plan data to identify asthma and SLE in the mother were excellent (95% and 93%, respectively; Table 3). The proportions of cases confirmed were similar for all encounter types and across sites.

Birth Certificate Data Elements

Table 4 reports the agreement between the birth certificate and chart data for the 598 infants for whom medical records were abstracted. For 537 infants, information on birth weight was documented in both the birth certificate and chart. The birth weight recorded in the birth certificate data was within 5% of that recorded in the chart for 497 (93%) infants. The PPVs for LBW and normal/high birth weight infants are also shown in Table 4.

For 486 infants, information on gestational age at birth was documented by at least one method (LMP, clinical assessment, other) in both the birth certificate and the chart. The gestational age estimate in the chart was based upon clinical assessment in 173 infants (36%), date of LMP in 100 infants (21%), and ultrasound in 53 infants (11%); in 160 cases (33%), the method of estimating the gestational age at birth was unknown/not documented. For 83% (95% CI, 79%–86%) of infants, the gestational age based upon the LMP documented in the birth certificate data was within 14 days of that reported in the chart; for 182 infants (41%; 95% CI, 37%–46%) the gestational age was an exact match between the birth certificate and chart. Agreement between the birth certificate and chart data was higher for gestational age based upon clinical assessment in the birth certificate, with 94% (95% CI, 92%–96%) of values within 14 days of the value reported in the chart; for 337 infants (72%; 95% CI, 66%–77%) the gestational age was an exact match between the birth certificate and chart.

Overall, agreement between birth certificate and chart data was good to excellent for the maternal and paternal characteristics examined (Table 4). For instance, 74% of mothers identified with no prior pregnancies based upon birth certificate data were also identified with no prior pregnancies based upon the chart. Agreement between the birth certificate and medical record was excellent for both maternal race/ethnicity and paternal race/ethnicity (91%), although race/ethnicity was not documented in the chart for a high proportion infants (38% were missing data on maternal race/ethnicity and 74% were missing data on paternal race/ethnicity).

Agreement between the birth certificate and medical record varied across the 3 FDA contract sites for gestational age and maternal race/ethnicity. The proportion of infants for whom the gestational age based upon LMP documented in the birth certificate was within 14 days of that reported in the chart ranged from 76% (95% CI, 69%–82%) to 88% (95% CI, 83%–93%) across sites; however, the proportion of infants for whom gestational age based upon clinical assessment documented in the birth certificate was within 14 days of that reported in the chart was more similar across sites (range: 91% to 98%). The proportion of infants for whom maternal race/ethnicity in the birth certificate data was confirmed through chart review ranged from 86% (95% CI 81%–91%) to 98% (95% CI 90–100%) across sites.

Discussion

In this study, we found that the validity of administrative health plan data for identification of infant outcomes varied by outcome, and often differed depending on the data source (infant's or mother's records), type of encounter (inpatient or outpatient), and specific diagnosis code. Overall, health plan data were good to excellent at identifying infants with ventricular septal defect (but not other cardiac defects), preterm birth, and NICU admission. Health plan data were poor at identifying infants with anencephaly. Since most infants with anencephaly die shortly after birth,[19] this outcome may be difficult to adequately assess using only health plan data, particularly if the infants die before being assigned a health plan medical record number. Linkage to data sources such as birth certificates, death certificates, and fetal death certificates may provide more complete capture of cases. Also, slightly more than half the possible cases in the present study were identified from infant ambulatory

encounters. However, on further investigation, we found that the PPV of ambulatory encounters was 6%, while the PPV of visits coded with anencephaly from the inpatient setting was 71%. The PPVs for the algorithms to identify maternal diagnoses of asthma and SLE using health plan data were high overall (93%) and differed little by encounter type. Our findings also indicate considerable agreement between birth certificate and medical record data for data elements related to birth weight, gestational age, prior obstetrical history, and race/ethnicity.

PPVs varied across study sites for some data elements, potentially due to differences in the data sources, information available in the medical chart, or differences in the study population. For example, the site with the lowest agreement for race/ethnicity between birth certificate and medical chart data had a more racially/ethnically heterogeneous population than other sites. Similarly, differences in the distribution of specific diagnosis codes for cardiac defects across sites likely influenced the findings for overall cardiac defects. Specifically, the site with the highest percentage of infants diagnosed with ventricular septal defect, and a lower percentage of infants diagnosed with atrial septal defect, had the highest overall PPV. Information available in the medical record also likely influenced findings, since some sites had more complete access to the patients' charts than others. In addition, differences in agreement for gestational age variables in the birth certificate data and charts was potentially influenced by whether LMP, clinical assessment, or ultrasound was available to determine gestational age.

Limited data are available on the validity of health plan data for identification of congenital malformations. The PPV we found for overall cardiac defects (71%) was similar to that of a prior study at one of the participating health plans.[1] In this prior study using Tennessee Medicaid data and birth certificate data for three studies conducted during the period 1985 through 2002, Cooper et al. reported that the PPV of inpatient claims compared to the medical record was 74.5% for cardiac defects overall and varied across defects from 43% for pulmonary value stenosis to 82% for atrial septal defect. It is important to acknowledge that the data analyzed by Cooper et al. were from a MEPREP participating health plan, therefore the agreement is to be expected to some extent. However, the PPV we found was far higher than that determined by Frohnert et al. using hospital discharge data from a large urban medical center in 2001, where the PPV was 36% for ICD-9-CM codes for cardiac defects based upon confirmation through medical record review.[20]

Our findings are similar to those of prior studies evaluating the validity of birth certificate data. Published studies have found birth certificate data accuracy varies considerably by data element assessed,[1–8] with the accuracy of demographic data and birth weight generally high, and the accuracy of items related to prenatal care and maternal risk factors and comorbidities generally lower. In a study assessing the validity of birth certificate data from the Ohio Department of Health from 1993 through 1995,[2] the PPV of birth certificate data was 99% for prior pregnancy and 97% for nulliparity (medical record as the gold standard). The concordance of data elements describing maternal race/ethnicity, gestational age, and birth weight was higher than 95% between the birth certificate data and medical record. Zollinger et al. similarly reported a high PPV (> 98%) for data elements describing birth weight and gestational age in Indiana birth certificate data for 1996.[8] In a prior study at one of the participating health plans among infants identified with possible congenital anomalies using Medicaid claims or birth certificate data, Cooper et al. reported that the date of the LMP in the medical record was within 14 days of the date of the LMP in the birth certificate in 94% of cases.[1]

Strengths of our study include evaluating the validity of administrative and claims codes from 11 large health plans and birth certificate data from eight states located in different

geographic regions of the U.S. In addition, the diagnoses originated from data of both HMO-owned and community medical care settings. Thus, our findings are likely generalizable to other U.S. health plans and data systems. We also were able to identify medical records for the great majority of patients (91%) selected for chart abstraction.

Limitations of this study include reliance on the medical record as the sole source of information for determining the validity of data elements from health plan administrative and claims databases and birth certificates. Often the chart review did not include evaluation of the patient's entire record. However, the PPVs reported for many data elements were high despite this limitation. The validity of data elements from the birth certificate data may have been influenced by the sample of infants for whom we reviewed charts, i.e., infants meeting our criteria for congenital malformations or other adverse conditions arising in the perinatal period. In addition, we randomly selected an encounter date when the patient had more than one encounter with the diagnosis or outcome of interest recorded in the health plan data during the observation period; thus, analyses to assess PPVs for specific encounter types and data sources did not take into account information for all encounters with the outcome of interest. Lastly, we did not determine if outcomes and diagnoses were present in the medical chart, but not the health plan administrative or claims data; thus, we could not evaluate the sensitivity and specificity of codes to identify infant outcomes and maternal diagnoses.

Conclusion

Limited evidence is available to evaluate the safety of medications administered in pregnancy. Administrative health plan databases hold promise for facilitating such safety research that otherwise could not be pursued. This study, using data from 11 diverse health plans, indicates that administrative health plan data can be used to accurately identify some important infant outcomes including specific cardiac defects, preterm birth, and NICU admission, and maternal diagnoses (asthma and SLE). Further, this study demonstrates that birth certificate data are useful to identify a number of birth (birth weight, gestational age) and maternal/paternal characteristics (race/ethnicity, prior obstetrical history). Our findings also underscore the importance of selecting both relevant settings and data sources (e.g. inpatient encounters for anencephaly) at the time of case selection. Some outcomes and variables may require medical record review for validation.

Acknowledgments

References

1. Cooper WO, Hernandez-Diaz S, Ray WA. Positive predictive value of computerized records for major congenital malformations. Pharmacoepidemiol Drug Saf. 2008; 17:455–460. [PubMed: 18081215]

2. DiGiuseppe DL, Aron DC, Ranbom L, Harper DL, Rosenthal GE. Reliability of birth certificate data: a multi-hospital comparison to medical records information. Matern Child Health J. 2002; 6(3):169–79. [PubMed: 12236664]

3. Northam S, Knapp TR. The reliability and validity of birth certificates. JOGNN. 2006; 35:3–12. [PubMed: 16466348]

4. Pearl M, Wier ML, Kharrazi M. Assessing the quality of last menstrual period data on California birth records. Paediatric Perinatal Epidemiol. 2007; 21(Suppl 2):50–61.

5. Reichman NE, Hade EM. Validation of birth certificate data. A study of women in New Jersey's HealthStart program. Ann Epidemiol. 2001; 11(3):186–93. [PubMed: 11248582]

6. Reichman NE, Schwartz-Soicher O. Accuracy of birth certificate data by risk factors and outcomes: analysis of data from New Jersey. Am J Obstet Gynecol. 2007; 197(1):32. e1–8. [PubMed: 17618747]

7. Roohan PJ, Josberger RE, Acar J, Dabir P, Feder HM, Gagliano PJ. Validation of birth certificate data in New York State. J Community Health. 2003; 28(5):335–46. [PubMed: 14535599]

8. Zollinger TW, Przybylski MJ, Gamache RE. Reliability of Indiana birth certificate data compared to medical records. Ann Epidemiol. 2006; 16(1):1–10. [PubMed: 16039875]

9. Andrade SE, Davis RL, Cheetham TC, Cooper WO, Li D, Amini T, Beaton SJ, Dublin S, Hammad TA, Pawloski PA, Raebel MA, Smith DH, Staffa JA, Toh S, Dashevsky I, Haffenreffer K, Lane K, Platt P, Scott PE. Medication Exposure in Pregnancy Risk Evaluation Program. Matern Child Health J. (in press).

10. Dombrowski MP, Schatz M. Asthma in pregnancy. Clin Obstet Gynecol. 2010; 53:301–10. [PubMed: 20436305]

11. Bakhireva LN, Schatz M, Chambers CD. Effect of maternal asthma and gestational asthma therapy on fetal growth. J Asthma. 2007; 44:71–6. [PubMed: 17454318]

12. Chambers C. Safety of asthma and allergy medications in pregnancy. Immunol Allergy Clin North Am. 2006; 26:13–28. [PubMed: 16443141]

13. Barnabe C, Faris PD, Quan H. Canadian pregnancy outcomes in rheumatoid arthritis and systemic lupus erythematosus. Int J Rheumatol. 2011; 2011:345727. [PubMed: 22028718]

14. Chakravarty EF, Nelson L, Krishnan E. Obstetric hospitalizations in the United States for women with systemic lupus erythematosus and rheumatoid arthritis. Arthritis Rheum. 2006; 54:899–907. [PubMed: 16508972]

15. Clowse ME, Jamison M, Myers E, James AH. A national study of the complications of lupus in pregnancy. Am J Obstet Gynecol. 2008; 199:127. e1–6. [PubMed: 18456233]

16. Clowse ME. Lupus activity in pregnancy. Rheum Dis Clin North Am. 2007; 33:237–52. [PubMed: 17499705]

17. Iozza I, Cianci S, Di Natale A, Garofalo G, Giacobbe AM, Giorgio E, De Oronzo MA, Politi S. Update on systemic lupus erythematosus pregnancy. J Prenat Med. 2010; 4:67–73. [PubMed: 22439065]

18. Sever, LE., editor. National Birth Defects Prevention Network (NBDPN). Guidelines for Conducting Birth Defects Surveillance. Atlanta, GA: National Birth Defects Prevention Network, Inc; Jun. 2004

19. Facts About Anencephaly. Centers for Disease Control and Prevention; http://www.cdc.gov/ncbddd/birthdefects/Anencephaly.html

20. Frohnert BK, Lussky RC, Alms MA, Mendelsohn NJ, Symonik DM, Falken MC. Validity of hospital discharge data for identifying infants with cardiac defects. J Perinatol. 2005; 25(11):737–42. [PubMed: 16163368]

Key Points

- Health plan administrative and claims data were found to accurately identify infant outcomes including specific cardiac defects, preterm birth, and admission to a neonatal intensive care unit and maternal diagnoses (asthma and systemic lupus erythematosus).

- Birth certificate data are useful to identify infant birth weight, gestational age, and maternal/paternal characteristics including race/ethnicity and prior obstetrical history.

- Our findings underscore the importance of selecting the relevant setting and data sources such as inpatient encounters for anencephaly at the time of case selection.

Table 1

Criteria for identification of infant outcomes and maternal diagnoses using the health plan administrative and claims data

Outcome	Criteria for Identification[a]	Number of Potential Cases Sampled for Chart Abstraction
Infant outcomes		
cardiac defects	ICD-9-CM 745–745.9, 746–746.9, 747.1–747.49, 747.6–.9,747.8–747.89 recorded in the infant's inpatient or outpatient records during the first year of life or in mother's record for the delivery admission	302
anencephaly	ICD-9 CM 740–740.2 recorded in the infant's inpatient or outpatient records during the first year of life or in mother's records for the delivery admission	40
preterm birth	ICD-9-CM codes 644.21, 765.0–765.28 recorded in the infant's inpatient or outpatient records during the first 30 days of life or in mother's record for the delivery admission	151
admission to neonatal intensive care unit	CPT codes 99295, 99296, 99297, 99298, 99299, 99468, 99469, 99477, 99478, 99479, 99480); at one site, neonatal intensive care admissions were identified using the site's registry and at another site, neonatal intensive care admissions were identified using birth certificate and administrative data, as well as CPT codes recorded in infant records during the first 30 days of life or in mother's record for the delivery admission	160
Maternal medical conditions		
asthma	ICD-9-CM codes 493–493.92; including women with: 1) 2 outpatient visits at least 30 days apart, 2) 1 inpatient visit with a diagnosis code, or 3) 1 outpatient visit with a diagnosis code and 1 dispensing of an asthma medication (short- or long- acting beta-agonist, inhaled corticosteroid, leukotriene receptor antagonist, mast cell stabilizer, methylxanthine) during the one year prior to pregnancy through the date of delivery	150
systemic lupus erythematosus	ICD-9-CM 710.0; including women with: 1) 2 outpatient visits at least 30 days apart or 2) 1 inpatient visit with a diagnosis code during the one year prior to pregnancy through the date of delivery	75

[a]To be selected as a potential case, the infant/mother was required to be enrolled in the health plan at least one day during the period of interest (e.g., for cardiac defects, the infant had to be enrolled at least one day during the first year of life). All cases of anencephaly meeting criteria were selected for medical chart review. Definitions of abbreviations: ICD-9-CM, International Classification of Diseases, 9th revision, Clinical Modification; CPT, Current Procedural Terminology.

Table 2

Validation of diagnosis and procedure codes in health plan administrative databases for identification of infant outcomes

	Number of charts reviewed	Number of cases confirmed	Positive predictive value (95% confidence interval)
Cardiac defects			
Total patients with 1 cardiac defect	276	195	71% (65%, 76%)
Encounter type [a]			
Inpatient	193	138	72% (65%, 78%)
Emergency department	5	1	20% (1%, 72%)
Ambulatory	78	56	72% (60%, 81%)
Source[a]			
Infant claims/administrative data	253	187	74% (69%, 79%)
Maternal hospital discharge data	23	8	35% (16%, 57%)
Most common codes observed in the population [a]			
Ostium secundum type atrial septal defect (ICD-9-CM 745.5)	96	62	65% (55%, 74%)
Ventricular septal defect(ICD-9-CM 745.4)	79	75	95% (90%, 100%)
Unspecified anomaly of the heart (ICD-9-CM 746.9)	35	24	69% (53%, 84%)
Anomalies of pulmonary artery (ICD-9-CM 747.3)	24	18	75% (53%, 90%)
Persistent fetal circulation (ICD-9-CM 747.83)	22	19	86% (65%, 97%)
Anencephaly			
Total patients	35	13	37% (21%, 55%)
Encounter type [a]			
Inpatient	17	12	71% (44%, 90%)
Ambulatory	18	1	6% (0%, 27%)
Source [a]			
Infant claims/administrative data	31	10	32% (17%, 51%)
Maternal hospital discharge data	4	3	75% (19%, 99%)
Most common code observed in the population			
Anencephaly (ICD-9-CM 740.0)	24	12	50% (29%, 71%)
Anencephaly and similar anomalies ICD-9-CM 740)	6	1	17% (0%, 64%)
Preterm birth			
Total patients	141	122	87% (81%, 92%)
Encounter type [a]			
Inpatient	136	118	87% (81%, 92%)

	Number of charts reviewed	Number of cases confirmed	Positive predictive value (95% confidence interval)
Ambulatory	5	4	80% (28%, 99%)
Source [a]			
Infant claims/administrative data	91	84	92% (87%, 98%)
Maternal hospital discharge data	50	38	76% (64%, 88%)
Most common codes observed in the population [a]			
Early onset of delivery, delivered (ICD-9-CM 644.21)	47	35	74% (62%, 87%)
Weeks of gestation, 35 to 36 weeks (ICD-9-CM 765.28)	36	36	100% (90%, 100%)
Other preterm infants, 2000 to 2499 grams (ICD-9-CM 765.18)	31	29	94% (79%, 99%)
Other preterm infants, 2500 grams and over (ICD-9-CM 765.19)	29	26	90% (73%, 98%)
Admission to neonatal intensive care unit			
Total patients	146	134	92% (87%, 96%)
Source [a]			
Infant claims/administrative data	128	119	93% (89%, 97%)
Maternal hospital discharge data	18	15	83% (59%, 96%)
Most common codes observed in the population [a]			
Neonatal critical care, initial (CPT 99295)	70	65	93% (87%, 99%)
Neonatal critical care, subsequent days (CPT 99296)	51	48	94% (84%, 99%)
Neonatal intensive care, low birth weight, 1500 to 2500 grams (CPT 99299)	31	30	97% (83%, 100%)
Health plan neonatal intensive care registry	25	25	100% (86%, 100%)

[a] If an infant had more than one encounter with the diagnosis of interest during the observation period, one encounter date was randomly selected to determine the encounter type (inpatient, outpatient, emergency department) and the source of data (infant or maternal administrative health plan records). For infant outcomes, encounters categorized as "inpatient" encounter types included data from the maternal delivery hospital admission and/or infant hospitalizations; "emergency department" and "ambulatory" encounters included only data from the infants' health plan records, not the mothers'. For some infants, more than one code for a specific outcome was recorded for the encounter date. Definitions of abbreviations: ICD-9-CM, International Classification of Diseases, 9th revision, Clinical Modification; CPT, Current Procedural Terminology.

Table 3

Validation of algorithms using health plan administrative databases for identification of maternal diagnoses

	Number of charts reviewed	Number of cases confirmed	Positive predictive value (95% confidence interval)
Asthma			
Total patients	133	126	95% (91%, 99%)
Encounter type [a]			
Inpatient	83	79	95% (91%, 100%)
Emergency department	18	17	94% (73%, 100%)
Ambulatory	32	30	95% (79%, 99%)
Systemic lupus erythematosus			
Total patients	71	66	93% (87%, 99%)
Encounter type [a]			
Inpatient	51	49	96% (87%, 100%)
Emergency department	2	2	100% (16%, 100%)
Ambulatory	18	15	83% (59%, 96%)

[a]If a women had more than one encounter with the diagnosis of interest during the observation period, one encounter date was randomly selected to determine the encounter type (inpatient, outpatient, emergency department).

Table 4

Validation of data elements in birth certificate data

Data Element From Birth Certificates	Number of charts reviewed[a]	Number confirmed	Percentage confirmed (95% confidence interval)
Birth outcome			
Birth weight[b]	537	497	93% (90%, 95%)
Low birth weight [c] (< 2500 grams	212	209	99% (96%, 100%)
Normal/high birth weight [c]	325	317	98% (96%, 99%)
Gestational age [d]			
Gestational age based upon last menstrual period	440	365	83% (79%, 86%)
Gestational age based upon clinical assessment	465	438	94% (92%, 96%)
Gestational age based upon 'other' method	20	17	85% (62%, 97%)
Prior obstetrical history			
Prior pregnancy			
No prior pregnancy	157	116	74% (67%, 81%)
1 prior pregnancy	337	323	96% (94%, 98%)
Prior live births			
No prior live birth	196	185	94% (91%, 98%)
1 prior live birth	282	259	92% (89%, 95%)
Race/ethnicity			
Maternal race/ethnicity			
Hispanic	78	67	86% (78%, 94%)
Asian	35	31	89% (78%, 99%)
Black	94	92	98% (93%, 100%)
White	133	122	92% (87%, 96%)
Other	4	3	75% (19%, 99%)
Total	344	315	92% (89%, 95%)
Paternal race/ethnicity			
Hispanic	35	31	89% (73%, 97%)
Asian	21	19	90% (70%, 99%)
Black	11	10	91% (59%, 100%)
White	64	60	94% (88%, 100%)
Other	1	0	0% (0%, 98%)
Total	132	120	91% (86%, 96%)

[a] Total charts with documentation of data element of interest with nonmissing values in birth certificate data.

[b] Weight in birth certificate data within 5% of weight reported in chart. Data for infants with values for birth weight < 454 grams or > 7262 grams are coded as missing and would be excluded from these analyses.

^{c}as determined by birth weight variable

^{d}Gestational age in birth certificate data within 14 days of that reported in chart. For MEPREP projects, data for infants with values for gestational age of < 20 completed weeks or > 45 completed weeks are coded as missing and would be excluded from these analyses.