Development of an Algorithm to Identify Pregnancy Episodes in an Integrated Health Care Delivery System

Mark C. Hornbrook, Evelyn P. Whitlock, Cynthia J. Berg, William M. Callaghan, Donald J. Bachman, Rachel Gold, F. Carol Bruce, Patricia M. Dietz, and Selvi B. Williams

Objective. To develop and validate a software algorithm to detect pregnancy episodes and maternal morbidities using automated data.

Data Sources/Study Setting. Automated records from a large integrated health care delivery system (IHDS), 1998–2001.

Study Design. Through complex linkages of multiple automated information sources, the algorithm estimated pregnancy histories. We evaluated the algorithm's accuracy by comparing selected elements of the pregnancy history obtained by the algorithm with the same elements manually abstracted from medical records by trained research staff.

Data Collection/Extraction Methods. The algorithm searched for potential pregnancy indicators within diagnosis and procedure codes, as well as laboratory tests, pharmacy dispensings, and imaging procedures associated with pregnancy.

Principal Findings. Among 32,847 women with potential pregnancy indicators, we identified 24,680 pregnancies occurring to 21,001 women. Percent agreement between the algorithm and medical records review on pregnancy outcome, gestational age, and pregnancy outcome date ranged from 91 percent to 98 percent. The validation results were used to refine the algorithm.

Conclusions. This pregnancy episode grouper algorithm takes advantage of databases readily available in IHDS, and has important applications for health system management and clinical care. It can be used in other settings for ongoing surveillance and research on pregnancy outcomes, pregnancy-related morbidities, costs, and care patterns.

Key Words. Pregnancy, maternal morbidities, research methods, episode grouper software, validity studies

While more than 6 million women in the United States become pregnant each year, assessments of the extent and nature of morbidities during pregnancy have been limited primarily to hospitalization data. These data likely capture

only the most severe morbidities, do not identify complications detected and managed in outpatient settings, and do not ascertain postpartum maternal morbidities. These shortcomings are salient because it is not possible to comprehensively assess the prevalence, severity, and economic burden of pregnancy-related morbidity by using only hospital discharge data.

Recent advances in the medical informatics of large, vertically integrated health care delivery systems (IHDS) include adopting computerized clinical and laboratory/pathology information systems. These data systems enhance opportunities for research. Using the electronic data systems of a large, nonprofit, group-model IHDS, we developed a computerized algorithm that identifies pregnancies and complications associated with these pregnancies, and estimates the extent of antepartum, intrapartum, and postpartum morbidity. This paper describes the development and validation of the algorithm to identify pregnancy episodes. A pregnancy episode grouper algorithm identifies a pregnancy and assigns an outcome, gestational age at the outcome event, estimated date of conception, and date 8 weeks after delivery or termination.

The conceptual foundation for this work was established by Hornbrook and colleagues (Hornbrook, Hurtado, and Johnson 1985; Hornbrook 1995), who defined three general types of health care episodes: (a) illness or health problem episodes—health problems requiring medical care as perceived by patients; (b) disease episodes—a medical model of a diagnosis and associated physiologic processes and related complications; and (c) care episodes—the temporally contiguous cluster of services provided for diagnosis and care of an identified disease or health problem. Pregnancy was defined as a special type of health-related episode, where an uneventful, full-term pregnancy is not defined as a disease episode, and maternal morbidities represent true disease episodes occurring simultaneously with pregnancy episodes. From a medical perspective, the pregnancy episode starts with the woman's last menstrual

Address correspondence to Mark C. Hornbrook, Ph.D., The Center for Health Research, Northwest/Hawaii/Southeast, 3800 N. Interstate Avenue, Portland, OR 97227-1110. Evelyn P. Whitlock, M.D., M.P.H., Donald J. Bachman, M.S., Rachel Gold, Ph.D., M.P.H., and Selvi B. Williams, M.D., M.P.H., are with The Center for Health Research, Northwest/Hawaii/Southeast, Portland, OR. Cynthia J. Berg, M.D., M.P.H., William M. Callaghan, M.D., M.P.H., Patricia M. Dietz, Dr.P.H., M.P.H., and F. Carol Bruce, M.P.H., are with the Division of Reproductive Health, National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control and Prevention, Atlanta, GA. Rachel Gold, Ph.D., M.P.H., and Selvi B. Williams, M.D., M.P.H., are also with the Department of Public Health and Preventive Medicine, Oregon Health & Science University, Portland, OR.

period and ends with the resolution of all potential risks and complications from the pregnancy. The pregnancy care episode begins at the date of diagnosis by a clinician. Identification and detection of early markers of pregnancy is essential for maternal epidemiology (incidence and prevalence of pregnancies in defined populations), maternal quality improvement (understanding and assuring access to and quality of prenatal care), and maternal pharmaco-epidemiology (assessing prenatal exposure to medications and other potentially risky substances that may affect fetal development [Manson, McFarland, and Weiss 2001]).

The work presented here builds on previous research at the study site by Manson, McFarland, and Weiss (2001) to identify pregnancy episodes in electronic health care databases in an IHDS. The study IHDS implemented an electronic medical records system (EMR) in 1997, largely supplanting the Manson algorithm. Specifically, the new EMR provided diagnoses and procedures on nearly all ambulatory care visits provided in IHDS medical offices, whereas Manson had diagnoses and procedures data available only from claims submitted by outside providers for care of IHDS beneficiaries, a very small proportion of the total visit volume during the 1993–1994 study period. Before 1997, ambulatory care delivered in the study IHDS medical offices was documented by a patient encounter registration system, which identified the provider, clinical department, and medical office associated with the visit and the type of visit (e.g., “new OB visit” versus “continuing OB visit”), but contained no standardized variables for diagnoses and procedures. Our algorithm integrates multiple automated information sources—ambulatory EMR (with ICD-9-CM diagnosis and CPT-4 procedure data), hospital discharge abstracts, emergency room visits, imaging procedures, laboratory test results, dispensing data, home health data, outside claims, and other databases—to enhance our ability to detect pregnancies and to construct the care episodes provided for each pregnancy in a defined population.

METHODS

Data Sources

This study was conducted using electronic data from Kaiser Permanente Northwest (KPNW), a prepaid not-for-profit group-practice IHDS serving approximately 467,000 members in the Pacific Northwest as of 2001. We developed a computerized algorithm that obtained data from all member-oriented electronic data systems, including KPNW enrollment records,

hospital discharge abstracts, ambulatory electronic medical record system (EPICCare[®]), emergency department visits, outside claims and referrals, radiology, laboratory results, home health care, and pharmacy records. To obtain data on parental race, ethnicity, and marital status for girls or women with live birth or stillbirth outcomes, we linked KPNW membership data for identified pregnancy episodes to birth certificate data from Oregon and Washington. Permissions from the IRBs of KPNW, State Health Division of Oregon, and Department of Health and Social Services of Washington State were obtained.

Study Population

The study population comprised 251,251 females aged 12–55 years who were continuously enrolled as KPNW members for a minimum of 42 days at any time during 1998–2001. The algorithm searched all databases for potential indicators of pregnancy, including: diagnosis and procedure codes from the *International Classification of Diseases, Ninth Revision, Clinical Modification* (ICD-9-CM) (CDC/NCHS 2005) and *Current Procedural Terminology, Fourth Edition* (CPT-4) (AMA 2004); *Healthcare Common Procedure Coding System* (HCPCS) (<http://www.cms.hhs.gov/MedHCPCSGenInfo/>); results of laboratory tests (e.g., pregnancy tests, maternal serum α -fetoprotein tests); pharmacy records; and imaging procedures associated with pregnancy, such as fetal ultrasound scans.

The pregnancy episode was defined as the interval between the estimated date of the last menstrual period and 8 weeks after delivery or pregnancy termination. A pregnancy episode was included in the study population only if the entire episode occurred within the study period (January 1, 1998–December 31, 2001), and if the girl or woman was enrolled as a KPNW member at the date of the pregnancy outcome. Because no ICD-9-CM code exists for last menstrual period or duration of pregnancy, we developed a series of analytic strategies using a hierarchy of pregnancy indicators to determine the beginning, outcome date, and end of pregnancy episodes. Pregnancy outcomes and outcome dates were identified, and gestational age at outcome was used to calculate the pregnancy's beginning date. The primary, and most reliable, outcome indicators were ICD-9-CM diagnostic codes (Table 1). We identified most births by the presence of a V27.0 code (single live birth delivery) in hospital discharge data (any position in the diagnosis vector). In a few instances, the algorithm was able to identify that a pregnancy had occurred but was not able to identify the outcome with precision (e.g., an early

Table 1: Logical Steps of Pregnancy Algorithm Processes

<i>Algorithm Steps</i>	<i>Data Sources and Hierarchies Referred to in Each Step</i>	<i>Population Considered</i>
Step 1: Algorithm identifies all eligible women in KPNW	KPNW membership data	KPNW population 1998–2001 (755,118) ↓ Females aged 12–55 years, KPNW members for ≥ 42 days in 1998–2001 (251,251) ↓ ... with any pregnancy indicators (32,847)
Step 2: Algorithm searches KPNW data sources for pregnancy indicators (in look-up tables)	Data sources Computerized hospital discharge records Ambulatory medical records (EPICCare [®]) Inpatient discharge files Emergency department data system Data system of outside claims and referrals Laboratory/pathology information system Imaging procedures Home health data system Pharmacy data system Preterm birth prevention data system Look-up table of ICD-9-CM, CPT-4 codes associated with pregnancy Look-up table of laboratory and imaging tests associated with pregnancy Hierarchy for data on pregnancy episode, end date, and outcome type 1. V27 code “deliver single live born” diagnoses, inpatient or emergency department 2. Procedures clearly defining end of pregnancy 3. Diagnosis related groups (DRGs) defining end of pregnancy 4. Procedures that probably ended pregnancy 5. V27, 632, or 634–637 codes for diagnoses not from phone contact, problem list 6. Outcome diagnoses, inpatient and emergency department	↓ ... with an identified pregnancy episode (21,001 women; 24,680 episodes): the study population
Step 3: Algorithm data hierarchy determines which indicators identify a pregnancy episode, episode end date, and outcome type		

7. Newborn V30-39 codes for outcome diagnoses, inpatient or emergency dept.

8. Delivery dates from the Preterm Birth Prevention Program database

9. Outcome diagnoses not listed in above categories

10. Procedures, diagnoses, DRGs before expulsion or extraction of POC

11. Procedures, diagnoses or DRGs after expulsion or extraction of POC

12. Abortion or delivery claim or referral

13. Diagnosis from emergency department record

Hierarchy for data on pregnancy beginning date

1. Subtract gestational age from discharge date in hospital discharge database

2. Subtract gestational age from date of maternal serum α -fetoprotein tests

3. Subtract 40 weeks from estimated date of delivery from Preterm Birth

Prevention Program database

4. Subtract gestational age from emergency department admission date

5. Estimate gestational age using national median gestational age by outcome

type:

Ectopic pregnancy = 8 weeks

Spontaneous abortion = 10 weeks

Therapeutic abortion = 10 weeks

Very preterm live birth = 26 weeks

Stillbirth = 28 weeks

Preterm live birth = 34 weeks

Full-term live birth = 40 weeks

Step 4: Second data
hierarchy identifies
episode beginning date

ICD-9-CM, International Classification of Diseases, Ninth Revision, Clinical Modification; CPT-4, Current Procedural Terminology, Fourth Edition;
POC, Products of conception; KPNW, Kaiser Permanente Northwest.

pregnancy loss of undetermined type); these are referred to as nondefinitive outcomes. They reflect diagnosis and procedure codes that contain only partial information about pregnancies, delivery procedures, and/or outcomes.

Gestational age was obtained from a data field in the discharge abstract for nearly all hospital births. When this field was not available, we searched for other information to estimate gestational age, including clinician orders for maternal serum α -fetoprotein testing. As a last resort, outcome-specific gestational age estimates were applied based on literature on stillbirths and live births (Copper et al. 1994; Martin et al. 2003). Gestational age data were often not available for pregnancy episodes with outcomes other than births (e.g., ectopic pregnancy, abortion, and trophoblastic disease). In these instances, outcome-specific estimates of gestational age were used (Wilcox, Treolar, and Sandler 1981; Centers for Disease Control and Prevention 2004). We calculated the beginning date of the pregnancy by subtracting gestational age from the outcome date, after gestational age had been determined, estimated, or assigned. We added 8 weeks to the outcome date to define the postpregnancy period. For this analysis, the pregnancy episode was defined as the window between the pregnancy beginning date and the end of the postpregnancy period.

During the algorithm's development and testing, pregnancy episodes were identified that had internal inconsistencies, such as pregnancy durations that were incompatible with pregnancy outcomes, overlapping or multiple pregnancy episodes, and indicators of pregnancy that occurred outside an identified pregnancy episode. The algorithm was tested and revised by comparing outcomes, beginning dates, and outcome dates with data from pregnancy-related medical records from administrative databases. This process resulted in numerous revisions to the algorithm before validation.

Validation of Algorithm

A formal validation analysis of the algorithm to detect pregnancy episodes was conducted by comparing data obtained by the algorithm with comparable data obtained from the medical chart by trained abstractors. This analysis addressed the question: If the algorithm identifies a pregnancy, do medical chart data identify the same pregnancy outcome with the same beginning and outcome dates of the pregnancy episode? To assess this, we calculated percent agreement (Hennekens, Buring, and Mayrent 1987; Gordis 1996) between the algorithm data and the medical chart data for pregnancy identification, outcome, and outcome date in a stratified sample of 511 females (see Table 2). Sampled episodes were randomly selected within pregnancy outcome strata.

Table 2: Sampling Strategy for Algorithm Validation

<i>Pregnancy Outcome Type</i>	<i>Pregnancy Episodes in Study Population</i>		<i>Pregnancy Episodes in Validation Sample</i>	
	#	%	#	%
Outcomes including live births				
Live birth, gestational age unknown	361	1.5	25	4.9
Live birth, gestational age known	16,316	66.1	225	44.0
Live birth and stillbirth	9	0.0	9	1.8
Abortions				
Therapeutic abortion	4,199	17.0	50	9.8
Spontaneous abortion	3,087	12.5	50	9.8
Morbid pregnancy outcomes				
Ectopic pregnancy	322	1.3	50	9.8
Trophoblastic disease	28*	0.1	27*	5.3
Stillbirth				
All	89	0.4	25	4.9
Nondefinitive pregnancy outcomes				
All	269	1.1	50	4.9
Live birth or stillbirth	181	0.7	25	4.9
Delivery or unknown abortion	19			
Early loss unknown outcome	58			
Outcome occurred, type unknown	8			
Unidentified abortion	3			
Total	24,680	100.0	511	100.0

*One woman in the study population had two trophoblastic outcomes during the study period; therefore, while 28 pregnancy episodes occurred during the study period, only 27 were sampled, because this woman could be selected for the validation sample only once.

We oversampled pregnancy episodes with less common outcomes, such as ectopic pregnancies, stillbirths, and therapeutic abortions. The validation sample also included all episodes identified by the algorithm as having a trophoblastic disease outcome, or a “live birth or stillbirth” outcome (a birth category with an undetermined outcome).

Using standardized data collection forms and coding instructions, trained medical record abstractors, blinded to any algorithm-determined information, obtained the beginning date, the outcome date, and the outcome of pregnancy episodes from medical records of the 511 women in the validation sample. For the outcome dates, data were considered to agree if the outcome date on the medical chart was within 30 days of the outcome date determined by the algorithm. A study physician manually reviewed and adjudicated any

pregnancy episode identified solely by the algorithm or by the medical record abstractors.

We calculated percent agreement (number agreeing divided by total number) between the outcome end date, outcome type, and gestational age determined by the algorithm and as determined by the medical record abstractors. The pregnancy outcomes of the oversampled episodes were re-weighted in proportion to the distribution of pregnancy outcomes in the KPNW study population, to estimate the algorithm's performance in the entire study population because the validation sample contained oversampled pregnancy episodes with less-common outcomes. If the algorithm and abstractors agreed precisely on the pregnancy episode outcome type, it was considered *exact agreement*. In cases where the algorithm identified a nondefinitive pregnancy outcome (such as "live birth or stillbirth") and the abstractors identified an outcome type that was a possibility within the nondefinitive outcome (such as "stillbirth"), it was considered *close agreement*.

We corrected minor systematic errors in the algorithm to create the final version after completing the validation analysis. We compared the abstractors' pregnancy parameters for the validation sample with those determined by the validation algorithm and the final revised algorithm to quantify the extent to which final changes led to improved results.

RESULTS

Of the 755,118 persons eligible for KPNW membership during the study period, 251,251 (33 percent) were females aged 12–55 years (see Table 1). Among these, 32,847 girls or women with potential pregnancy indicators were identified. Among these, 24,680 pregnancy episodes were identified in 21,001 girls or women. Most pregnancies (87 percent) were identified using diagnosis codes, including: 67 percent identified by diagnosis code V27.0 ("single live-born"); 13 percent identified with CPT-4 codes; and less than 1 percent identified by diagnosis-related group (DRG) codes, other ICD-9-CM codes, or pharmacy data.

Of the pregnancy episodes identified by the algorithm, live birth was the most frequent pregnancy outcome, as expected, with 16,677 episodes (68 percent of all episodes). There were also 4,199 therapeutic abortion episodes (17 percent); 3,087 spontaneous abortion episodes (13 percent); 322 ectopic pregnancy episodes (1 percent); 89 still birth episodes (<1 percent); 28 trophoblastic pregnancy episodes (<1 percent); and nine episodes (<1 percent)

ending in both a live birth and a stillbirth (in multiple gestations where one twin died). One hundred and eighty-one episodes (< 1 percent) were known to have ended in *either* a live birth or stillbirth; 88 pregnancy episodes (< 1 percent) had other nondefinitive outcomes. The final version of the algorithm identified 24,514 pregnancy episodes, with very similar distributions, but with only 33 episodes with nondefinitive outcomes.

Results from the validation analysis are presented in Table 3. The medical records abstractors did not identify 38 episodes identified by the algorithm, and the algorithm did not identify 24 episodes identified by the medical record abstractors (data not shown). Agreement on the pregnancy outcome date varied by the outcome type. For live births, 99 percent agreement on outcome date was achieved. As expected, agreement on outcome date was lowest when the algorithm identified a pregnancy, but failed to identify an outcome. For cases with agreement on the outcome date, close or exact agreement on the type of outcome was high for all outcomes (88–100 percent); 100 percent agreement was achieved for live births. Agreement on outcome was defined as “exact” when both the algorithm and the review of medical records identified the same outcome.

The results of the algorithm and the review of medical records were in agreement on outcome date for 640 pregnancy episodes. Eighty-one percent of the 640 episodes matching on outcome date (519 episodes) agreed on gestational age within 2 weeks, and 13 percent (84 episodes) agreed on gestational age within 15–28 days, which represents a total of 94 percent agreement on gestational age within 4 weeks or less among episodes matching on outcome type (data not shown). For pregnancy episodes with agreement on both outcome and outcome date, there was 91 percent agreement on gestational age within 4 weeks. Agreement on gestational age varied by outcome—for episodes ending in live birth, there was 98 percent agreement on gestational age within 4 weeks, both when gestational age was found in one of the databases and when it was estimated. For episodes ending in stillbirth, there was 100 percent agreement on gestational age within 4 weeks, when gestational age was found in one of the databases—no gestational ages had to be estimated for stillbirths. For episodes ending in ectopic pregnancy, there was 80 percent agreement on gestational age within 4 weeks when gestational age was found in one of the databases, compared with 70 percent agreement when gestational age was estimated. For episodes ending in spontaneous abortion, there was 100 percent agreement on gestational age within 4 weeks when gestational age was found in one of the databases, and 67 percent agreement when gestational age was estimated.

Table 3: Percent Agreement between Algorithm and Medical Record on Pregnancy Outcome (Unweighted Results)

<i>Pregnancy Outcome</i>	<i># Episodes Identified by Algorithm</i>	<i># with Agreement on Outcome Date within 30 Days</i>	<i>% Agreement on Outcome Date</i>	<i>% Exact Agreement on Outcome</i>	<i>% Close Agreement on Outcome</i>
Outcomes including live births					
Live birth	334	330	98.8	100.0	100.0
Live birth or stillbirth	22	16	72.7	18.8	100.0
Live birth and stillbirth	8	7	87.5	100.0	100.0
Abortions					
Therapeutic abortion	83	82	98.8	96.3	96.3
Spontaneous abortion	97	84	86.6	92.9	92.9
Morbid pregnancy outcomes					
Ectopic pregnancy	61	61	100.0	100.0	100.0
Trophoblastic disease	28	24	85.7	87.5	87.5
Stillbirth					
All	24	24	100.0	87.5	87.5
Nondefinitive outcomes					
All	21	12	57.1	0.0	91.7
Total	678	640	94.4	-	-

The final version of the algorithm agreed with the medical record review more closely than the validated version (data not shown). The final algorithm alone identified only 22 episodes not identified by the medical record abstractors, compared with 38 by the earlier validated version of the algorithm. Both the final algorithm and the medical records review identified 637 episodes. Of these 637 episodes agreeing on outcome date, 98 percent agreed within 14 days, and of these, 95 percent matched exactly on pregnancy outcome and 3 percent matched closely (e.g., where the algorithm identified a “live birth or stillbirth” and the medical record review identified a stillbirth). Eighty-two percent agreed on gestational age within two weeks and 13 percent within 4 weeks.

When we estimated the rates of agreement between the medical record abstractors and the final algorithm weighted to reflect the observed distribution of pregnancy outcomes in the source population, agreement on outcome date was 98 percent. For episodes with agreement on the outcome date, exact agreement on outcome type improved to 99 percent.

DISCUSSION

We developed a computerized algorithm that searched an IHDS’s readily available administrative and clinical databases to identify pregnancies occurring within a 4-year period. The algorithm’s performance was evaluated by comparing selected pregnancy parameters identified by the algorithm with comparable data abstracted from medical charts (including textual progress notes, nursing notes, surgical notes, pathology reports, radiologist interpretations, medical history, diagnostic impressions, and written treatment plans). This validation indicated a high degree of agreement between the algorithm and the medical record on pregnancy duration and outcome.

Our study builds on the algorithm developed by Manson, McFarland, and Weiss (2001) to identify markers for early pregnancy detection and pregnancy outcomes in the automated databases of the KPNW population in 1993–1994. The enhancements included in our algorithm are: (1) use of ICD-9-CM diagnosis and CPT-4 procedure codes from ambulatory visits, as documented in EMR (implemented in 1997 after Manson’s study [2001], which had only four variables to describe ambulatory visits); (2) defining a longer childbearing age range, 12–55 versus 15–44, to detect pregnancies among girls with earlier menarche and sexual activity, as well as pregnancies among older women (especially those receiving medical interventions to extend fertility);

(3) incorporating markers of spontaneous early terminations, rather than only elective early terminations; (4) use of dispensing data to identify pregnancy markers and outcomes (e.g., methotrexate for medical treatment of ectopic pregnancies and medical induction of therapeutic abortions); (5) use of the automated home health information system to detect home visits to new mothers (to detect live births delivered outside of KPNW facilities but the mother and infant are being cared for by KPNW providers); (6) use of the preterm birth prevention data system, a component of the EMR that was not available to Manson and colleagues (medical records were still hard copy in 1993–1994); (7) inclusion of ectopic pregnancy and trophoblastic disease as pregnancy outcomes; (8) incorporation of HCPCS codes for markers of pregnancy detection and outcomes (HCPCS code S0199 indicates the occurrence of a medically induced abortion by oral ingestion of medication); and (9) use of “history of” pregnancy and pregnancy-related complications ICD-9-CM codes to detect previous pregnancies that may have been managed outside the health plan. While Manson, McFarland, and Weiss (2001) recommended manual review of medical records to complement their computerized algorithm, we are attempting to avoid the need for manual review of records in the current EMR environment. Manson, McFarland, and Weiss (2001) identified only four pregnancy outcomes with their algorithm: (1) fetal death, (2) elective termination, (3) live birth, and (4) no pregnancy outcome (could be determined). In contrast, our pregnancy episode grouper algorithm identifies 13 different outcomes of pregnancy episodes (see Table 1).

We compared the distribution of pregnancy outcomes obtained by Manson, McFarland, and Weiss (2001) for 1993–1994 with our results for 1998–2001. Manson reported a rate of 23 percent for nondefinitive pregnancy outcomes, compared with our rate of 1 percent. The excess 22 percent nondefinitive outcomes are converted by our algorithm into 14 percent fetal deaths (compared with 6 percent for Manson et al.), 17 percent therapeutic abortions (versus 10 percent), and 68 percent live births (versus 61 percent). We combined our separate outcomes of spontaneous abortions (13 percent), ectopic pregnancies (1 percent), trophoblastic disease (0.1 percent), and still births (0.4 percent) to obtain an equivalent measure of fetal deaths to Manson. We believe that our more refined classification of pregnancy episodes provides much enhanced utility for clinical management and research in maternal health.

Our comprehensive, linked data systems allowed inclusion of all pregnancy outcome types, not just those for which women were hospitalized. The validated algorithm developed in this study is a useful tool for research on pregnancies and their outcomes. While the algorithm is currently structured to

KPNW, it can also be modified for application and expansion to other IHDS settings, including Federally Qualified Health Clinics, for ongoing surveillance and research. EMRs offer more information than claims and encounter data systems. Adoption of EMRs is growing as the patient safety benefits are documented by research (Feldstein 2006; Smith 2006). Because we included data elements not found in claims data, our algorithm has an advantage over commercially available episode grouper software, such as Symmetry's Episode Treatment Groups[®] (Rosen and Mayer-Oakes 1999; Symmetry Health Data Systems, Inc. [www.symmetry-health.com]). We have made the look-up tables, input file formats, and SAS code for our pregnancy grouper algorithm available for public download from the *Health Services Research* website.

Our validation approach was structured to check the accuracy of pregnancies the algorithm identified. We could not ensure that the algorithm did not miss any pregnancies among the majority of women of childbearing age who had no evidence of pregnancy-related events, because abstractors would need to review the medical charts of a prohibitively large number of females of reproductive age. Furthermore, it is possible that pregnancies occurred in the study population without evidence in the medical charts, such as miscarriages that occurred before a pregnancy was detected or women who elected to have their pregnancies managed outside the health plan (e.g., by a Planned Parenthood clinic or a neighborhood women's health clinic). Nevertheless, the rate of pregnancies in our population appears to conform to expectations. In the United States in 2000, the estimated rate of live births was 66 per 1,000 females aged 15–44 years, the rate of induced abortions was 21 per 1,000, and the rate of fetal losses was 17 per 1,000 (Ventura et al. 2004). In our data, among 131,000 girls or women aged 15–44 years during 1999–2000, the rate of live births was 85 per 1,000 girls and women, the rate of induced abortions was 17 per 1,000, and the rate of fetal losses was 13 per 1,000. Given the sociodemographic structure of our insured population, these minor differences were expected. The distribution of pregnancy outcomes was also very close to national statistics. Nationally, in 1981–1991, 63 percent of pregnancies ended in live birth, 22 percent in therapeutic abortion, and 14 percent in spontaneous abortion (Saraiya et al. 1999). Our findings were similar—68 percent, 17 percent, and 13 percent, respectively. Thus, we judge it unlikely that the algorithm “missed” substantial numbers of pregnancy episodes.

The predominance of the V27.0 code for identifying pregnancies in our study setting is likely a function of two factors: (a) the nature of the enrolled population—employed persons and Medicaid clients—who are mostly healthy, have a high proportion in the childbearing age groups, and have

health insurance; and, (b) the use of skilled medical records technicians to code diagnosis data from inpatient medical records.

The algorithm is not without limitations. We encountered considerable challenges in identifying dates of pregnancy outcomes and duration of gestation for outcomes other than live births. For example, as many early losses do not require the immediate attention of a health care provider, do not require a surgical or medical procedure or hospitalization, or have a slowly evolving clinical course preceding a conclusive diagnosis, precise information on the date the pregnancy ended and the gestational age at termination were difficult to determine. In addition, for KPNW members referred for, or receiving, outside provider care for pregnancy-related issues, such as an elective abortion, clinical records often contain little information other than the referral or claims-related date. Thus, it was necessary to estimate duration of pregnancy in these cases.

Our reliance on ICD-9-CM diagnosis and procedure codes and CPT-4 procedure codes may be another limitation to the extent that they are subject to an unknown degree of coding errors. While we are unaware of reports on the validity of pregnancy-related ICD-9-CM codes recorded during outpatient and nondelivery hospitalizations, ICD-9-CM codes recorded during delivery hospitalizations in hospital discharge databases vary in both positive predictive values and sensitivities for obstetrical procedures and diagnoses, complications of pregnancy, and preexisting maternal medical conditions (Lydon-Rochelle, Holt, Cardenas et al. 2005; Lydon-Rochelle, Holt, Nelson et al. 2005). However, the multiple sources of data in our algorithm likely attenuate the effect of relying solely on the delivery hospitalization discharge abstract. Moreover, administrative data systems of IHDSs are designed to monitor the use of resources and to estimate costs of providing health care. Therefore, it is unlikely that these data sources would overlook or omit a substantial proportion of pregnancies. In fact, they may be more likely to identify and include a pregnancy with complications because of the importance of the economic aspects of this kind of pregnancy to the IHDS. These data systems receive strong internal and external auditing. As the beginning of diagnosis-based risk adjustment of Medicare payments to HMOs, Medicare+Choice contractors (such as KPNW) have been required to submit claims data for their Medicare enrollees following the same format and documentation requirements as FFS providers to support the risk-adjusted payment models. The penalties for under- or overreporting services, and for under- and overcoding these services are substantial. CMS regularly audits the Medicare claims data submitted by KPNW to ensure regulatory compliance.

Thus, we contend that the rigors of internal KPNW and external CMS audits and the potential for severe financial penalties for incorrect data filings, serve to prevent any meaningful systematic errors of ICD-9-CM and CPT-4 coding.

Manson, McFarland, and Weiss (2001, p. 187) reported, “the procedure code for an outpatient obstetric visit frequently was not accompanied by a diagnosis code to indicate the purpose of this visit” (in the 1993–1994 KPNW data). These ambulatory diagnosis data were missing because the automated ambulatory encounter database used by Manson et al. was an appointment and visit registration system and did not include any diagnosis information. Diagnosis codes for ambulatory visits in KPNW did not become available on a routine basis until the ambulatory EMR system was installed in 1997. The EMR represents a major source of improvement of our algorithm because clinicians are responsible for documenting diagnoses and procedures at each visit before the medical record can be closed with a digital signature affixed to the visit. An “autocoder” function assists clinicians to select appropriate diagnosis and procedure codes. Thus, the quality of diagnosis and procedure data for ambulatory care is unusually high.

Manson, McFarland, and Weiss state (2001, p.186) that pregnancy outcome date was missing from substantial number of records. We had substantially less trouble with this because recording of outcome date has improved significantly. Finally, Manson, McFarland, and Weiss (2001, p. 187) state that LMP date was added to the KPNW database as of October 1999. Our review of the EMR data for 1998–2001 shows that the LMP field is available for clinicians to fill in, but that this was an elective field and is often not filled in. KPNW’s Preterm Birth Prevention Program, supported by a special user interface in the EMR, requires that the LMP field be filled in, so the collection of this information for high-risk pregnancies has improved in recent years. This, however, still leaves a gap in LMP data for low-risk pregnancies.

CONCLUSIONS

This pregnancy episode grouper algorithm advances the state-of-the-art for episode grouping related to pregnancies and maternal morbidities. While the extent of agreement between results of the algorithm and results of medical record review varies by pregnancy outcome, this algorithm is highly effective and has strong validity for establishing the presence of a pregnancy. Moreover, it performs well in identifying critical episode parameters, such as pregnancy outcome and duration.

The algorithm may be a useful tool for monitoring pregnancy outcomes in a defined population with consistent access to services. It identifies early pregnancy markers and frames a defined health-related episode of care related to pregnancy, thus creating a window through which maternal and fetal morbidities, health care utilization, and other factors can be identified, evaluated, and monitored. Consequently, this algorithm is a first step toward a comprehensive system for surveillance of pregnancy-related health and health care in an IHDS. Such surveillance could be very useful for research, for enhancing approaches to quality improvement for pregnancy-related care, and for monitoring resource use in large women's health care programs. The algorithm, for example, is sensitive to incomplete and nonspecific diagnosis and procedure information, which serve as indicators of the need to understand the reasons why pregnancy care appears to be discontinuous (e.g., apparent uncoordinated use of outside providers) and/or why providers are using nonspecific diagnosis and procedure codes rather than providing more complete documentation. Quality improvements based on the implementation of evidence-based practice recommendations at the clinician, clinic, or health-system level could be undertaken using such an algorithm. For example, the provision of recommended diabetic laboratory screening services within certain timeframes in pregnancy could be audited and prepared for feedback to clinicians, specialty chiefs, and clinic managers. Rankings of incidence rates, costs, and outcomes of pre- intra-, and postpartum maternal morbidities support strategic planning for quality improvement initiatives in obstetrical practice. Similarly, monitoring of pharmaceutical dispenses of medications that are known to be teratogenic or not sufficiently studied in pregnancy could also be accomplished and lead to quality improvement programs in many practice settings. Finally, our algorithm improves the base for identifying maternal exposures to pharmaceuticals during the pregnancy episode but before being diagnosed as pregnant. These drug exposures represent a valuable database for assessing the risks of pharmaceuticals during the early organogenesis period.

The algorithm will support a wide variety of research studies, including prospective studies, by providing a denominator (number of pregnancy episodes) for calculating rates of related morbidities. In addition, the algorithm is a powerful tool for future analyses, such as: customized quality improvement analyses; research on trends or disparities in pregnancy-related outcomes, morbidities, and patterns of care; and analyses evaluating the impact of health policy on pregnancy outcomes and related morbidities. Combining the pregnancy algorithm with obstetrics/gynecology risk management claims (White et al. 2005) may help to establish the utilization and clinical contexts associated

with malpractice claims, enabling identification of early warning signs associated with higher likelihood of such claims.

Recent advances in health information technology and the increasing rate of diffusion of EMR systems present expanding opportunities for applying and improving our algorithm. Our ongoing work will expand our pregnancy database by adding additional years to test the hypothesis that data quality in obstetrics and pediatrics in EMR systems is improving as health plans implement programs to improve revenue collection (improved coding), HEDIS scores (improved prenatal and postnatal visit compliance), and quality programs for managing maternal and neonatal morbidities.

ACKNOWLEDGMENTS

We thank Bertha Moseson, M.D., and Elizabeth Garlitz, M.D., for medical record review in support of diagnosis, procedure, pharmacy, and laboratory data coding and interpretation; reviews of episode validation data; and adjudication of differences between results of medical record reviews and results from the algorithm. We also thank Sara Gao for programming support to link data from birth certificates to the analysis database; Karen Riedlinger for programming support for the database on laboratory results; Jill Mesa and Lisa Puderbaugh for abstracting medical records for the validation sample; Lisa Fox for graphic arts support; Martha Swain, Kevin Lutz, and Margaret Sucec for technical editing; Gary Ansell for project management; and Barbara Lardy and Sharron Coleman (America's Health Insurance Plans) for technical assistance and liaison with the contracts office of the Centers for Disease Control and Prevention.

This research project was funded by Contract # CDC 200-2001-00074, Task # MC2-02, "Extent of Maternal Morbidity in a Managed Care Setting," from the Centers for Disease Control and Prevention. America's Health Insurance Plans administered this contract.

Disclosures: None.

Disclaimers: The findings and conclusions in this article are those of the authors and do not necessarily represent the views of the Centers for Disease Control and Prevention or Kaiser Permanente.

REFERENCES

American Medical Association. 2004. *Current Procedural Terminology 2005*. Chicago: American Medical Association.

- Centers for Disease Control and Prevention. 2004. "Abortion Surveillance—United States, 2001." *Morbidity and Mortality Weekly Report* 53 (9): 1–32.
- Centers for Disease Control and Prevention, National Center for Health Statistics. "International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM)" [accessed on October 25, 2005]. Available at <http://www.cdc.gov/nchs/about/otheract/icd9/abticd9.htm>
- Copper, R. L., R. L. Goldenberg, M. B. DuBard, R. O. Davis and the Collaborative Group in Preterm Prevention. 1994. "Risk Factors for Fetal Death in White, Black, and Hispanic Women." *Obstetrics and Gynecology* 84 (4): 490–5.
- Feldstein, A. 2006. "Several Interventions Improved Therapeutic Monitoring: A Randomized Trial." Poster presented at the 12th Annual Meeting of the Health Maintenance Organization Research Network, Boston.
- Gordis, L. 1996. *Epidemiology*. Philadelphia: W.B. Saunders Company.
- Hennekens, C. H., J. E. Buring, and S. L. Mayrent, Eds. 1987. *Epidemiology in Medicine*. Boston: Little, Brown & Company.
- Hornbrook, M. C. 1995. "Definition and Measurement of Episodes of Care in Clinical and Economic Studies." In *Cost Analysis Methodology for Clinical Practice Guidelines: Conference Proceedings. AHCPR Publication No. 95-0001*, edited by M. L. Grady and K. A. Weis, pp. 15–40. Rockville, MD: U.S. Department of Health and Human Services, Public Health Service, Agency for Health Care Policy and Research.
- Hornbrook, M. C., A. V. Hurtado, and R. E. Johnson. 1985. "Health Care Episodes: Definition, Measurement and Use." *Medical Care Review* 42: 163–218.
- Lydon-Rochelle, M. T., V. L. Holt, V. Cardenas, J. C. Nelson, T. R. Easterling, C. Gardella, and W. M. Callaghan. 2005. "The Reporting of Pre-Existing Maternal Medical Conditions and Complications of Pregnancy on Birth Certificates and in Hospital Discharge Data." *American Journal of Obstetrics and Gynecology* 193 (1): 125–34.
- Lydon-Rochelle, M. T., V. L. Holt, J. C. Nelson, V. Cardenas, T. R. Easterling, and W. M. Callaghan. 2005. "Accuracy of Reporting Maternal In-Hospital Diagnoses and Intrapartum Procedures in Washington State Linked Birth Records." *Paediatric and Perinatal Epidemiology* 19 (6): 460–71.
- Manson, J. M., B. McFarland, and S. Weiss. 2001. "Use of an Automated Database to Evaluate Markers for Early Detection of Pregnancy." *American Journal of Epidemiology* 154 (2): 180–7.
- Martin, J. A., B. E. Hamilton, P. D. Sutton, S. J. Ventura, F. Menacker, and M. L. Mumson. 2003. "Births: Final Data for 2002." *National Vital Statistics Reports* 52 (10): 1–113.
- Rosen, A. K., and A. Mayer-Oakes. 1999. "Episodes of Care: Theoretical Frameworks Versus Current Operational Realities." *Joint Committee Journal of Quality Improvement* 25 (3): 111–28.
- Saraiya, M., C. J. Berg, H. Shulam, C. A. Green, and H. K. Atrash. 1999. "Estimates of the Annual Number of Clinically Recognized Pregnancies in the United States, 1981–1991." *American Journal of Epidemiology* 149 (11): 1025–9.
- Smith, D. 2006. "A Cost-Effectiveness Analysis (CEA) of 3 Interventions to Enhance Laboratory Monitoring of Selected Medications" Poster Presented at the 12th Annual Meeting of the Health Maintenance Organization Research Network. Boston.

- Symmetry Health Data Systems, Inc. "Episode Treatment Groups:™ An Illness Classification and Episode Building System" [accessed on October 25, 2005]. Available at http://www.symmetry-health.com/ETGTut_Desc1.htm
- Ventura, S. J., J. C. Abma, W. C. Mosher, and S. Henshaw. 2004. "Estimated Pregnancy Rates for the United States, 1990–2000: An Update." *National Vital Statistics Reports* 52 (23): 1–12.
- White, A. A., J. W. Pichert, S. H. Bledsoe, C. Irwin, and S. S. Entman. 2005. "Cause and Effect Analysis of Closed Claims in Obstetrics and Gynecology." *Obstetrics and Gynecology* 105 (5, part 1): 1031–8.
- Wilcox, A. J., A. E. Treloar, and D. P. Sandler. 1981. "Spontaneous Abortion over Time: Comparing Occurrence in Two Cohorts of Women a Generation Apart." *American Journal of Epidemiology* 114 (4): 548–53.

SUPPLEMENTARY MATERIAL

The following supplementary material for this article is available online:

APPENDIX A. Look-up Tables for Pregnancy Episode Grouper and Maternal Morbidities

APPENDIX B. Format of Input Data Files for Pregnancy Episode Grouper

APPENDIX C. Pregnancy Episode Grouper Logic

APPENDIX D. Comparison of Pregnancy Outcomes from Manson et al. Algorithm for 1993–94 with those from Hornbrook et al. Algorithm for 1998–2001