Clinical Epidemiology

Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

Impact of the Transition from ICD–9–CM to ICD–10–CM on the Identification of Pregnancy Episodes in US Health Insurance Claims Data

This article was published in the following Dove Press journal:
Clinical Epidemiology

Amir Sarayani[1]
Xi Wang[1]
Thuy Nhu Thai[1,2,*]
Yasser Albogami[1,3,*]
Nakyung Jeon[4]
Almut G Winterstein[5,6]

[1]Department of Pharmaceutical Outcomes and Policy, College of Pharmacy, University of Florida, Gainesville, FL, USA; [2]Faculty of Pharmacy, Ho Chi Minh City University of Technology (HUTECH), Ho Chi Minh City, Vietnam; [3]Department of Clinical Pharmacy, College of Pharmacy, King Saud University, Riyadh, Saudi Arabia; [4]College of Pharmacy, Chonnam National University, Gwang-ju, Korea; [5]Department of Pharmaceutical Outcomes and Policy, College of Pharmacy and Center for Drug Evaluation and Safety (CoDES), University of Florida, Gainesville, FL, USA; [6]Department of Epidemiology, College of Medicine and College of Public Health and Health Professions, University of Florida, Gainesville, FL, USA

*These authors contributed equally to this work

Background: Before October 2015, pregnancy cohorts assembled from US health insurance claims have relied on medical encounters with International Classification of Diseases-ninth revision-clinical modification (ICD-9-CM) codes. We aimed to extend existing pregnancy identification algorithms into the ICD-10-CM era and evaluate performance.

Methods: We used national private insurance claims data (2005–2018) to develop and test a pregnancy identification algorithm. We considered validated ICD-9-CM diagnosis and procedure codes that identify medical encounters for live birth, stillbirth, ectopic pregnancy, abortions, and prenatal screening to identify pregnancies. We then mapped these codes to the ICD-10-CM system using general equivalent mapping tools and reconciled outputs with literature and expert opinion. Both versions were applied to the respective coding period to identify pregnancies. We required 45 weeks of health plan enrollment from estimated conception to ensure the capture of all pregnancy endpoints.

Results: We identified 7,060,675 pregnancy episodes, of which 50.1% met insurance enrollment requirements. Live-born deliveries comprised the majority (76.5%) of episodes, followed by abortions (20.3%). The annual prevalence for all pregnancy types was stable across the ICD transition period except for postterm pregnancies, which increased from 0.5% to 3.4%. We observed that ICD codes indicating gestational age were available for 86.8% of live-born deliveries in the ICD-10 era compared to 23.5% in the ICD-9 era. Patterns of prenatal tests remained stable across the transition period.

Conclusion: Translation of existing ICD-9-CM pregnancy algorithms into ICD-10-CM codes provided reasonable consistency in identifying pregnancy episodes across the ICD transition period. New codes for gestational age can potentially improve the precision of conception estimates and minimize measurement biases.

Keywords: pregnancy, gestational age, live birth, stillbirth, abortion, ectopic, ICD-9, ICD-10

Introduction

Health insurance claims databases are an important source of real-world data to evaluate drug effects, particularly in pediatric, geriatric, or pregnant populations where the evidence from clinical trials is limited.[1,2] Safety or effectiveness studies among pregnant women require accurate identification of pregnancy episodes (ie, start date, gestational age (GA), and pregnancy outcome) to adequately time drug exposure and avoid misclassification biases. Although claims data may ensure temporality in study of treatment effects and avoid exposure misclassification such as when relying on recall, accurate estimates of GA at birth or termination

Correspondence: Almut G Winterstein
Department of Pharmaceutical Outcomes and Policy, College of Pharmacy, University of Florida, 1225 Center Drive, HPNP 3320 Building, Gainesville, FL 32610, USA
Tel +1-352-273-6268
Email almut@cop.ufl.edu

submit your manuscript | www.dovepress.com
DovePress
http://doi.org/10.2147/CLEP.S269400

and thus, of pregnancy start are often not available.[3] Several algorithms have been proposed to identify unique pregnancy episodes using the clinical encounter information coded using the International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) system.[4–8] The validity of such algorithms has been demonstrated against birth certificate and electronic health records in different settings, including pregnancies that followed in vitro fertilization where the accurate conception date is known.[9–12] However, the transition from ICD-9 to ICD-10 codes in the US healthcare system in 2015 has imposed a new challenge to the identification of pregnancy episodes in claims databases.[13]

The ICD-10 system provides a more granular representation of clinical concepts. In total, the number of diagnosis codes has increased from ~17,000 in ICD-9 to ~71,000 in ICD-10.[14] Relevant to pregnancy research, "gestation of pregnancy" is represented by distinct maternity codes from Z3A.01 (less than 8-week gestation of pregnancy) to Z3A.49 (greater than 42-week gestation of pregnancy).[15] This representation is in contrast to the ICD-9 system, where the gestation weeks are available only for infant codes after delivery (eg, 765.28: "35–36 completed weeks of gestation"). Such improvement in detail may provide opportunities to enhance measurement of pregnancy episodes, while inconsistencies between the ICD eras may arise.

The consistency of several health outcome measurements across the ICD transition period has been evaluated in claims databases,[14,16,17] but no studies have assessed the performance of pregnancy identification algorithms (PIA) adapted to the ICD-10 era. In the present study, we aimed to use a cross-walk approach as previously applied,[14,16,17] and translate validated ICD-9-CM code to ICD-10-CM equivalents to extend a PIA into the ICD-10 era. We identified pregnancy episodes in a national private insurance claims database and explored frequencies of pregnancy endpoints and GA at birth estimates across the ICD transition period.

Materials and Methods

Data Source

We used the IBM MarketScan Commercial Claims and Encounters Databases, which include medical and pharmacy reimbursement claims incurred in the inpatient or outpatient settings. This database is a national sample of the privately insured US population and collects data from approximately 350 payers (data use agreement with the IBM company is required). An encrypted enrollee identifier allows for longitudinal follow-up across study years, and family members under the same insurance plan can be linked via a family identifier. Clinical diagnosis and medical procedures are encoded using ICD-9/10 systems, current procedural terminology (CPT) system, and healthcare common procedure coding system (HCPCS).

Study Population

We included all women of childbearing age (12–55 years) at the beginning of each study year and identified those with inpatient or outpatient claims indicating pregnancy endpoints or prenatal care visits to identify pregnancy episodes from 2006 to 2017.

Measurement of Pregnancy Episodes

(A) Pregnancy episodes with known endpoints

a. Definition of pregnancy endpoints:

We used a coding algorithm based on ICD-9, CPT, and HCPCS codes established in previous validation studies to measure various pregnancy endpoints, including livebirth, mixed birth (multiple gestations with live and non-live outcomes), ectopic pregnancy, stillbirth, spontaneous abortion, induced abortion, and unclassified delivery (delivery code with no mention of newborn status).[5,9,10]

To translate ICD-9-CM codes to the ICD-10-CM era, we first created a list of candidate ICD-10-CM codes using the conversion tool provided by the Supercoder® portal (https://www.supercoder.com) for the diagnosis codes and the MapIT tool from the Agency for Healthcare Research and Quality (https://www.qualityindicators.ahrq.gov/Resources/Toolkits.aspx) for procedure codes. We applied the Extended Bridge converter option in Supercoder® and used a simple-forward mapping approach in the MapIT tool. We then manually reviewed the candidate ICD-10 codes for each pregnancy endpoint and reconciled discrepancies for consistency with the ICD-9 system. For the live birth category, we used a validated ICD crosswalk published by the Sentinel Initiative.[18] We provide the final list of codes in eAppendix.

b. Assignment of pregnancy episode endpoint:

For each pregnancy endpoint category, we applied a clinically relevant wash-out period between consecutive endpoints as previously used for the validated algorithm:[5] 182 days for live births, 168 days for unclassified

deliveries, 168 days for stillbirths, 56 days for spontaneous abortions, 56 days for induced abortions, and 56 days for ectopic pregnancies. The date of the first observed claim describing a particular outcome was considered as the end date of a pregnancy episode. We optimized the algorithm for the live or mixed birth claims by using the date of the delivery procedure claim instead of delivery admission date and removed any outpatient claims for live, mixed, or unclassified delivery ±30 days of an inpatient live birth or delivery claim. For ectopic pregnancy episodes, we required either a procedure code for "extraction of ectopic pregnancy" or an ectopic pregnancy diagnosis code plus a confirmatory code for treatment within 14 days (methotrexate claim or general extraction procedure).[5] The list of codes is provided in eAppendix.

c. GA estimation at pregnancy endpoint

(i) Live birth/mixed birth/unclassified deliveries:

For each pregnancy episode, the last menstrual period (LMP) date was estimated as the date of pregnancy endpoint minus a GA value. The GA for live deliveries and deliveries with unknown outcome (assumed to be live) was chosen based on a previous study that established a mother-infant linkage cohort in the Medicaid database and validated claims-based GA estimation against birth certificates.[12] This algorithm uses ICD-9 codes for prematurity, weeks of gestation, low birth weight, and postterm/prolonged pregnancy on medical encounter claims for the mother or the infant within a 30-day window from delivery to assign a GA to a pregnancy episode. If no GA codes are observed within 30 days of delivery, 39 weeks is assigned to the pregnancy episodes consistent with previous algorithms. If multiple GA codes are observed, the smallest value of GA is used. In our dataset, we linked mothers and infants who had the same family plan, where infants' age was equal to zero according to insurance eligibility records, and infants' first observed medical encounter was within 30 days of mothers' delivery claim. For those episodes where linkage was not successful (eg, in instances where the infant was not added to the mother's insurance plan), we used only the information on the mother's records to estimate GA. In the ICD-10 system, the GA codes have been expanded and provide more detailed information about weeks of gestation. We used the ICD conversion tools to identify candidate ICD-10 codes, considered the Sentinel report,[18] and applied the following rules to assign numerical values to the GA codes to accommodate the changes:

1. GA type 1: These codes have specific weeks for the gestation in their description (eg, 'Z3A.31′: 31-week gestation of pregnancy). We used the GA mentioned in the code description as the numerical value for these codes.
2. GA type 2: These codes did not have specific weeks for gestation in their description. If trimester was mentioned in the code description, we used the midpoint of the trimester; For example, for "O60.13X0: Preterm labor second trimester with the preterm delivery third trimester, not applicable or unspecified", we used 35 as the numerical value for this code. For the low birth weight codes, we used the mean of the numerical value assigned to their ICD-9 counterparts. The codes for prolonged/postterm pregnancies received a similar numerical value as their ICD-9 counterparts.

We multiplied the number of weeks by 7 to convert the GA into days and estimate the LMP date. We prioritized type 1 codes over type 2 codes if we observed both types for a pregnancy episode. For those pregnancy endpoints that occurred after June 2015 and were identified in the ICD-9 era, we looked for GA codes using both ICD-9 or ICD-10 codes.

(ii) Pregnancies with non-live birth endpoints:

For non-live births, we assigned fixed GA estimates as 56 days for ectopic pregnancy, 196 days for stillbirth, and 70 days for spontaneous/induced abortions based on previous validation studies.[5] We estimated the LMP date by subtracting the GA estimate from the pregnancy endpoint date.

d. Pregnancy episode reconciliation

Because of miscoding or presence of pregnancies with multiple endpoints, identified episodes may overlap. To establish unique episodes, we used a hierarchical approach that prioritized pregnancy endpoints with superior accuracy in measuring both the endpoint itself and GA based on claims data: (1) live birth/mixed birth, (2) ectopic pregnancy, (3) stillbirth, (4) induced abortion, (5) spontaneous abortion, and (6) unclassified deliveries. We further

required a clinically relevant wash-out period for each pregnancy endpoints for the same beneficiary, using minimum GA for a given pregnancy endpoint.[5] For instance, a stillbirth endpoint may not occur 182 days before or 168 days after a live birth (see eAppendix). Also, it may not occur during a live birth episode (LMP to endpoint date).

B. Identification of pregnancy episodes with an unknown endpoint

Claims databases may not allow for observation of pregnancy outcomes because beneficiaries may lose or switch insurance. Furthermore, outcomes for unplanned pregnancies that end with elective termination may not be captured if not covered by the health plan. To increase the sensitivity of the PIA, we identified "unknown outcome" pregnancies based on ICD codes for prenatal care visits or CPT procedure codes for early screening tests, referred to as "pregnancy care" codes (see eAppendix). We only considered pregnancy care codes that did not occur during previously established pregnancy episodes, and that preceded a previously determined LMP date by at least 75 days. We further required at least 6 months of insurance enrollment before pregnancy care claims to ensure appropriate capture of previous pregnancy episodes. Finally, we applied a wash-out period of 182 days between two consecutive pregnancy episodes with an unknown outcome (similar to live birth episodes).

Each unknown outcome episode was required to have at least two pregnancy care claims 30 to 183 days apart. We determined LMP based on the first observed claim for a given episode using a fixed GA estimate of 55 days. We obtained this estimate from the median time between estimated LMP to the first pregnancy care claim for live, mixed, or unclassified delivery episodes (median=55 in the ICD-9 period).

Some pregnancy care codes in the ICD-10 system carry the trimester information in their description (eg, "Z34.01: Encounter for the supervision of normal first pregnancy, first trimester"). As with pregnancy care encounters obtained during the ICD-9 era, we retained the first encounter, but updated the assigned GA based on the median values for the time between LMP and the trimester-specific code (pregnancy care in trimester 1=55 days, prenatal screening=73, pregnancy care in trimester 2=134, trimester 3=240, pregnancy care during unknown trimester=75).

Evaluation of PIA Performance

We anchored pregnancy episodes in the calendar year of their endpoint date. For each study year, we calculated the proportion of each pregnancy endpoint among all episodes observed to evaluate possible changes in trends across the ICD transition period. We also calculated the prevalence of each pregnancy endpoint among all women of childbearing age with a minimum of 45 weeks of insurance enrollment in each respective year. We also evaluated the distribution of estimated GA for live birth, mixed birth, and unclassified delivery episodes between the ICD eras. Finally, we investigated the utilization of three prenatal screening tests (CPT coding system) as an internal validation measure since the CPT codes did not change before and after the transition in the ICD system. We evaluated the proportion of pregnancy episodes with ultrasound in the first trimester (CPT codes: 76801, 76802), nuchal translucency measurement (CPT codes: 76813, 76814), and pregnancy-associated plasma protein-A tests [PAPP-A] (CPT code: 84163). We assessed the ultrasound test for all pregnancy episodes and the two following tests for live, mixed, and unclassified delivery for the 2007–2017 study period when all the CPT codes were in effect.

Results

We identified 7,060,675 pregnancy episodes with their endpoints between 2006 and 2017. Live, mixed, and unclassified delivery episodes comprised the majority of pregnancy endpoints (70.8%), including 439,233 preterm births (8.8%). After applying the insurance enrollment requirement, 50.1% of the pregnancy episodes remained. In this sub-cohort, the distribution of pregnancy endpoints was similar to the primary cohort except for the unknown outcome episodes, which decreased from 8.7% to 1.3% (Table 1). The mean age of mothers was 30.1 years, and 40.8% resided in the South region of the US The distribution of pregnancy endpoints before imposing minimum enrollment requirements and the demographic characteristics of the sub-cohort are provided in eTable 1 and eTable 2, respectively.

Variations in the annual prevalence of preterm and full-term pregnancies were negligible. We observed a significant increase in postterm pregnancies, starting in 2015, which stabilized in 2016 and 2017. Among pregnancy episodes with non-live or unknown outcome, the trends were stable across the ICD transition period, as shown in Figure 1 and eTable-3.

Table 1 Pregnancy Episodes Identified by the Pregnancy Identification Algorithm (PIA)

Pregnancy Endpoint	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	Total
Full term (n)	124,287	150,033	184,867	243,637	245,480	262,798	283,281	225,361	221,000	177,543	169,442	151,444	2,439,173
%	67.89	68.53	65.88	69.22	68.86	68.19	71.68	69.66	71.55	69.13	68.30	66.94	69.00
Preterm (n)	11,193	14,577	19,093	24,135	24,344	25,491	25,742	20,797	20,344	16,964	16,824	15,613	235,117
%	6.11	6.66	6.80	6.86	6.83	6.61	6.51	6.43	6.59	6.61	6.78	6.90	6.65
Postterm[a]	345	584	774	1,129	1,233	1,408	1,420	1,401	1,678	3,461	8,550	7,512	29,495
%	0.19	0.27	0.28	0.32	0.35	0.37	0.36	0.43	0.54	1.35	3.45	3.32	0.83
Ectopic pregnancy	2010	2298	3261	3743	3867	4237	3536	3420	3021	2855	2682	2623	37,553
%	1.10	1.05	1.16	1.06	1.08	1.10	0.89	1.06	0.98	1.11	1.08	1.16	1.06
Stillbirth	1618	1787	2370	2902	3071	3283	3169	2744	2538	2202	2117	1973	29,774
%	0.88	0.82	0.84	0.82	0.86	0.85	0.80	0.85	0.82	0.86	0.85	0.87	0.84
Spontaneous abortion	30,727	35,212	49,462	54,171	55,546	62,197	54,804	49,070	42,932	39,342	35,637	34,743	543,843
%	16.78	16.08	17.63	15.39	15.58	16.14	13.87	15.17	13.90	15.32	14.36	15.36	15.38
Induced abortion	10,172	11,299	16,894	18,006	18,763	20,664	18,137	15,865	12,981	10,818	9,670	9,372	172,641
%	5.56	5.16	6.02	5.12	5.26	5.36	4.59	4.90	4.20	4.21	3.90	4.14	4.88
Unknown outcome	2729	3127	3870	4233	4213	5330	5134	4876	4385	3649	3169	2951	47,666
%	1.49	1.43	1.38	1.20	1.18	1.38	1.30	1.51	1.42	1.42	1.28	1.30	1.35
All pregnancies[b]	183,081	218,917	280,591	351,956	356,517	385,408	395,223	323,534	308,879	256,834	248,091	226,231	3,535,262

Notes: [a]Postterm includes pregnancies with a gestational age of more than 41 weeks. [b]Total number of pregnancy episodes observed in each year with 45 weeks of insurance enrollment following the last menstrual period.

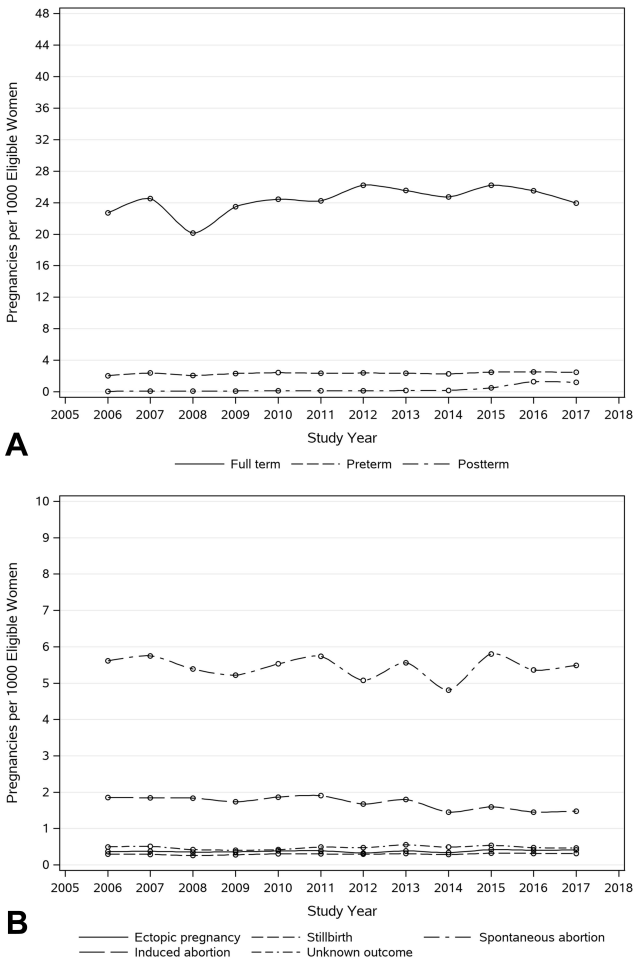

Figure 1 Secular trend of pregnancies among women of child-bearing age stratified by live, mixed, and unclassified deliveries (**A**) and non-live outcomes (**B**).
Notes: The numerator is the total number of pregnancy episodes identified by the algorithm, which had 45 weeks of continuous health plan enrollment after the estimated LMP date. The denominator includes all women aged 12–55 years with at least 45 weeks of health plan enrollment in each respective study year. Pregnancy episodes were anchored in the calendar year of their endpoint date.

We observed that 86.8% of the live, mixed, and unclassified delivery episodes in the ICD-10 era (2015–2017) had a GA code within 30 days of the endpoint, among which 97% had a GA code on the mothers' claims records. In the ICD-9 era, we observed only 23.5% of deliveries with GA codes. The distribution of GA among the live birth episodes in the ICD-9 and ICD-10 periods was not comparable. In the ICD-10 era, the observed GA ranged from 20 to 43 weeks since the GA codes in the ICD-10 system are represented on an ordinal scale with weekly increments while their ICD-9 counterparts are biweekly. Among the episodes with 39 weeks of gestation, we observed that 71.6% of the episodes had a GA code in the database, while the corresponding figure in the ICD-9 era was 5.2%. Figure 2 shows the distribution of GA among these episodes.

The assessment of the prenatal screening tests showed a consistent trend for both the proportion of episodes with the tests and the median GA. We observed the first-trimester ultrasound procedure among 45–50% of the live birth episodes across the study years with a consistent trend during the ICD transition. For other pregnancy episodes, we observe similar utilization of ultrasound procedures except for the unknown outcome episodes, which showed an increase from ~13% (2013 and 2014) to ~20% (2016 and 2017). The estimated median GA for this procedure was stable across study years except for the unknown outcome category, which showed an increase from 7 weeks to 10 weeks. The proportion of live episodes with nuchal translucency measurement and the median GA for this procedure were stable over the study years (~35% and ~12 weeks, respectively). Regarding the PAPP-A test, we observed a decreasing trend in the proportion of episodes with this test (~35% to ~20%) while the median GA for the procedure was stable across the study years (~12 weeks). The pattern of prenatal screening tests is summarized in Tables 2 and 3.

Discussion

We successfully developed the PIA to identify pregnancy episodes in the ICD-10 era. The distribution of pregnancy endpoints, including live and non-live outcomes, was comparable to the ICD-9 era. Likewise, the annual prevalence of pregnancy endpoints and the pattern of prenatal screening tests appeared to be stable across the ICD transition period. We also observed changes in LMP estimation for pregnancies with unknown outcomes, which relied on prenatal care encounters.

Accurate identification of pregnancies in claims databases can provide unique opportunities to answer clinical and policy questions. In drug safety research, observational studies using claims data have provided valuable evidence on the teratogenic effects of drugs.[19,20] In policy evaluation research, claims-based studies have quantified risk for fetal exposure associated with different risk mitigation strategies for teratogenic drugs.[21,22] Moreover, researchers may need to identify pregnancy as an essential covariate or in- or exclusion criterion in their observational studies. Our findings showed that the adaptation of the PIA to the ICD-10 system is feasible and results in a reasonably consistent measurement of pregnancy endpoints across the ICD transition. Therefore, pharmacoepidemiologic studies may utilize data from the ICD-10 era to address emerging clinical and policy questions while

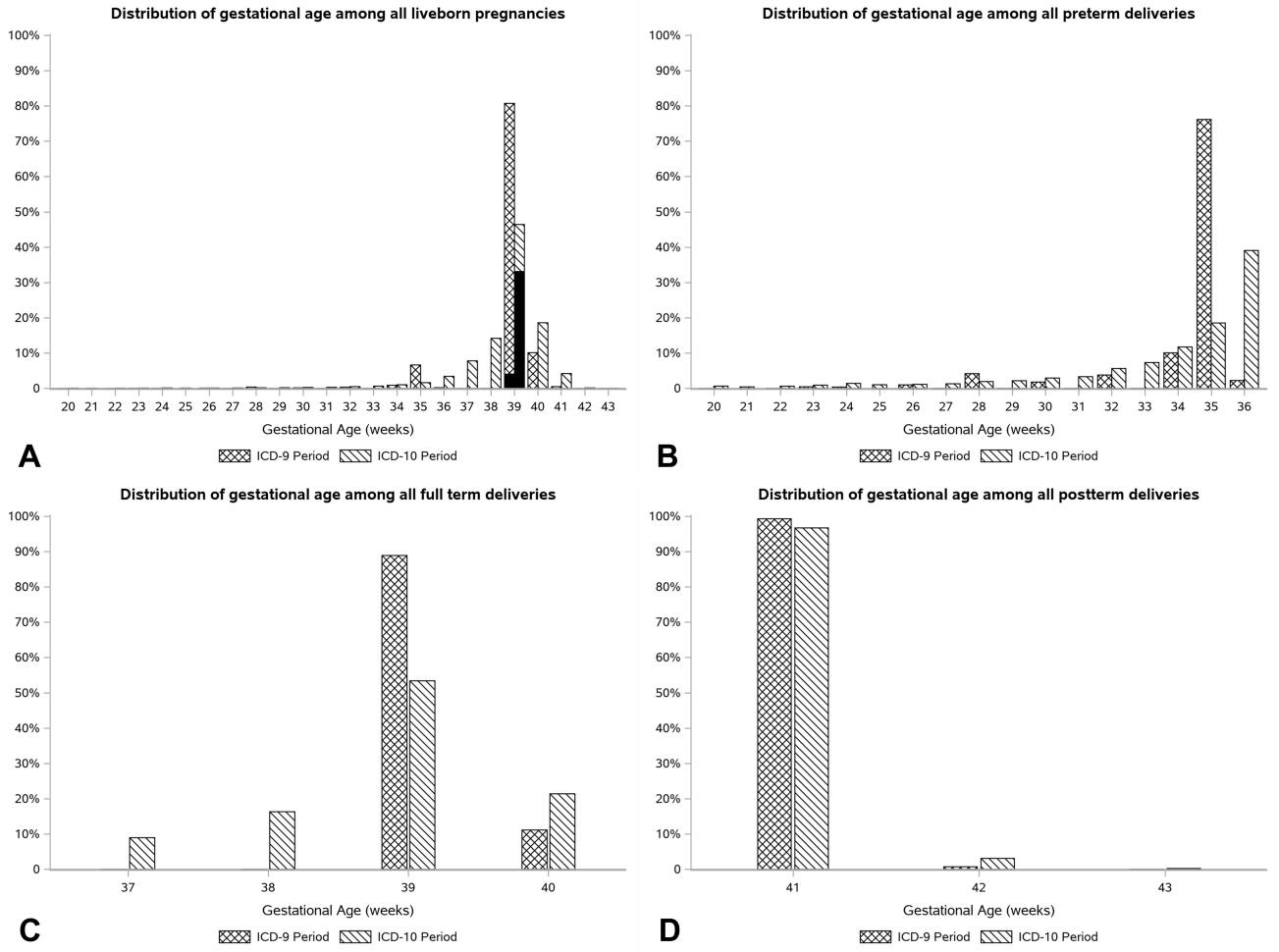

Figure 2 Distribution of gestational age assigned to all liveborn (**A**), preterm (**B**), full-term (**C**), and postterm deliveries (**D**) pregnancy episodes.
Notes: All liveborn include live, mixed, or unclassified delivery episodes. The pattern of gestational age was evaluated among pregnancies that had 45 weeks of health plan enrollment from LMP. The solid black fill pattern over the bars at 39 weeks shows the portion of episodes that had explicit GA codes on mother or infant records. (71.6% in the ICD-10 era vs 5.2% in the ICD-9).

taking into consideration the potential differences in pregnancy measurements between the two eras.

In 2017, the national estimate for late/postterm and preterm births was 6.6% and 9.9%, respectively.[23] We observed that 4.3% and 8.9% of births in our cohort were postterm and pre-term in 2017. Our study cohort was a privately insured population, and the contrast between the national estimates and our findings should be interpreted in this context. The estimated prevalence of postterm pregnancies increased sharply in 2015, which could be attributed to the new "gestation of pregnancy" codes for 41–43 weeks. Likewise, with new GA codes for GA 37 and 39 weeks, the GA distribution for full-term episodes in the ICD-10 era ranged from 37 to 41 weeks, while the majority of episodes in the ICD-9 era had a GA of 39 or 40 weeks. If we assume that the ICD-10 era provides a more accurate representation of full-term pregnancies, the LMP date for approximately 30% of the full-term episodes measured in ICD-9 is estimated up to 2 weeks earlier, which allows for potential misclassification of exposure, especially when evaluating exposure effects during the first trimester. Future studies should be cautious about exposure misclassification and particularly about differential misclassification when comparison groups are distributed unevenly across the two ICD eras (eg, because of the date of market approval or inclusion in drug formularies).

We identified pregnancies with an unknown outcome as part of our PIA to enhance the sensitivity of pregnancy measurement, where conception is the event of interest. We observed that only 10% of episodes with unknown outcomes had the insurance enrollment requirement of 45

Table 2 Proportion of Pregnancy Episodes with Prenatal Screening Procedure Codes During Pregnancy[a]

Pregnancy Endpoint	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
First-trimester ultrasound											
Full term	34.1	35.0	37.2	39.0	39.9	41.1	43.0	43.6	43.6	45.1	46.1
Preterm	41.3	42.4	44.2	45.1	46.3	48.1	50.7	50.1	50.0	51.0	52.0
Postterm[b]	37.8	37.2	38.6	39.7	42.9	44.9	43.2	45.2	44.7	43.6	46.4
Ectopic	28.2	29.7	32.0	31.2	35.0	35.3	38.6	41.2	40.5	38.7	40.8
Still birth	40.8	42.5	47.1	45.7	46.4	47.8	49.9	49.3	49.1	51.9	55.5
Induced abortion	16.8	18.0	19.2	20.3	21.4	19.5	19.2	20.0	19.2	20.8	20.6
Spontaneous abortion	26.6	28.0	28.2	28.2	28.9	28.9	29.9	30.1	29.6	29.9	29.7
Unknown outcome	11.5	12.7	11.9	13.0	13.4	13.2	13.8	13.1	14.1	20.3	19.3
Nuchal translucency measurement											
Full term	7.1	19.7	24.8	29.1	31.7	32.6	34.3	33.4	31.9	31.3	32.5
Preterm	9.8	23.2	28.5	33.1	35.6	37.2	38.8	37.6	35.8	35.4	35.0
Postterm	8.4	26.6	32.3	32.7	35.7	39.4	38.0	41.2	36.5	33.7	34.7
Pregnancy-associated plasma protein A (PAPP-A)											
Full term	12.7	17.8	22.8	26.9	29.8	31.3	30.3	26.1	20.9	18.8	17.8
Preterm	14.9	19.9	25.2	29.4	32.2	34.4	32.5	27.7	22.6	19.9	18.4
Postterm	18.2	24.4	29.7	32.0	33.0	38.1	36.1	34.3	24.0	21.4	21.1

[a]**Notes**: [a]The numbers in the table are reported as the percentage of pregnancies with the procedure during pregnancy. [b]Postterm includes pregnancies with a gestational age of more than 41 weeks.

Table 3 Median Gestational Age for Prenatal Screening Procedures During Pregnancy[a]

Pregnancy Endpoint	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
First-trimester ultrasound											
Full term	10	10	10	9	9	9	9	9	9	9	9
Preterm	9	9	9	9	9	9	9	8	8	9	8
Postterm[b]	11	11	11	11	11	10	11	10	10	10	10
Ectopic	8	8	8	8	8	8	8	8	8	8	8
Stillbirth	20	20	20	19	19	19	19	19	19	20	20
Induced abortion	9	9	9	9	9	9	9	9	9	9	9
Spontaneous abortion	9	9	9	9	9	9	9	9	9	8	8
Unknown outcome	7	7	7	7	7	7	7	7	7	10	10
Nuchal translucency measurement											
Full term	12	12	12	12	12	12	12	12	12	12	12
Preterm	11	11	11	11	11	11	11	11	11	11	11
Postterm	12	12	13	13	13	12	12	13	12	12	12
Pregnancy-associated plasma protein A (PAPP-A)											
Full term	12	12	12	12	12	12	12	12	12	12	12
Preterm	11	11	11	11	11	11	11	11	11	11	11
Postterm	13	13	13	12	13	12	12	12	12	12	12

[a]**Notes**: [a] The numbers in the table are reported as the median gestational age (weeks) at which the procedure was performed.[b]Postterm includes pregnancies with a gestational age of more than 41 weeks.

weeks after LMP. This observation implies that the outcome of such episodes was most likely missed because of health plan discontinuation. We also observed that the median GA estimate for these pregnancies increased in the ICD-10 era, which is likely attributable to the new trimester information in the ICD-10 system, allowing superior estimation of LMP. Further studies are required to validate the approach to identify pregnancy episodes based on only pregnancy care claims regardless of the pregnancy outcome.

Finally, we observed a relatively stable trend across study years for the ultrasound nuchal translucency measurement tests, while the PAPP-A test showed a downward trend since 2014. We speculate that the introduction of a new CPT code (81509; effective in 2013) might have contributed to the secular trend over the transition period. This CPT code refers to a procedure for the assessment of three serum proteins, including PAPP-A, to calculate a risk score for fetal anomalies. We also observed that the median GA for first-trimester ultrasound among stillbirth episodes is ~20 weeks. This GA estimate falls outside the first-trimester definition and is likely an artifact of GA mismeasurement for stillbirth episodes as previously discussed in the literature.[5]

Strengths and Limitations

This study assessed pregnancy outcomes for a national population of privately insured women spanning more than a decade to facilitate careful examination of the impact of ICD transition. We adapted previously developed and validated algorithms to a new coding system that is generally richer than its predecessor. In doing so, we aimed to take advantage of the new information without changing the general approach in defining unique pregnancy episodes and estimating GA. For example, we ignored the "gestation of pregnancy" codes that are attached to a variety of pregnancy-related encounters and endpoints in the ICD-10 era, because similar codes are missing for deliveries in ICD-9. The new representation of GA in the ICD-10 system could offer multiple points during pregnancy that inform about GA, thus offering superior approaches to LMP estimation, especially for non-delivery pregnancy endpoints.

Although our findings support our conclusion about consistency in the measurement of pregnancy endpoints in administrative claims databases across the ICD transition period, they do not inform about the true sensitivity and specificity of endpoint definitions, GA estimates, and the related LMP determination. Nevertheless, the performance of ICD-9 algorithms has been validated previously in such data sources, and the overall consistency in prevalence estimates suggests acceptable translation into the ICD-10 era. Finally, our findings may not be generalizable to other claims databases if patients' interaction with the healthcare system and delivery of healthcare services is expected to be different.

Conclusions

Our study showed that expansions of existing PIA into the ICD-10-CM era could provide a consistent measurement of pregnancy episodes. The availability of new codes for weeks of gestation provides an opportunity for a more precise estimation of the pregnancy start date, which should be further explored for live and non-live pregnancy endpoints.

Abbreviations

CPT, current procedural terminology; GA, gestational age; HCPCS, healthcare common procedure coding system; ICD-9-CM & ICD-9, International Classification of Diseases-9th revision-clinical modification; ICD-10-CM & ICD-10, International Classification of Diseases-10th revision-clinical modification; LMP, last menstrual period; PIA, pregnancy identification algorithm.

Ethics

The study protocol was reviewed and approved by the Institutional Review Board at the University of Florida (IRB201701362 – Exempt Research). Patient-informed consent was not required as the database was de-identified to protect privacy.

Disclosure

Dr Almut G. Winterstein reports grants from Merck, outside the submitted work. The authors report no other potential conflicts of interest for this work.

References

1. Platt R, Carnahan RM, Brown JS, et al. The U.S. Food and Drug Administration's Mini-Sentinel program: status and direction. *Pharmacoepidemiol Drug Saf*. 2012;21(Suppl 1):1–8.
2. Franklin JM, Glynn RJ, Martin D, Schneeweiss S. Evaluating the use of nonrandomized real-world data analyses for regulatory decision making. *Clin Pharmacol Ther*. 2019;105(4):867–877. doi:10.1002/cpt.1351
3. Huybrechts KF, Bateman BT, Hernandez-Diaz S. Use of real-world evidence from healthcare utilization data to evaluate drug safety during pregnancy. *Pharmacoepidemiol Drug Saf*. 2019;28(7):906–922. doi:10.1002/pds.4789

4. MacDonald SC, Cohen JM, Panchaud A, McElrath TF, Huybrechts KF, Hernandez-Diaz S. Identifying pregnancies in insurance claims data: methods and application to retinoid teratogenic surveillance. *Pharmacoepidemiol Drug Saf*. 2019;28(9):1211–1221.
5. Matcho A, Ryan P, Fife D, Gifkins D, Knoll C, Friedman A. Inferring pregnancy episodes and outcomes within a network of observational databases. *PLoS One*. 2018;13(2):e0192033. doi:10.1371/journal.pone.0192033
6. Andrade SE, Toh S, Houstoun M, et al. Surveillance of medication use during pregnancy in the mini-sentinel program. *Matern Child Health J*. 2016;20(4):895–903. doi:10.1007/s10995-015-1878-8
7. Margulis AV, Setoguchi S, Mittleman MA, Glynn RJ, Dormuth CR, Hernandez-Diaz S. Algorithms to estimate the beginning of pregnancy in administrative databases. *Pharmacoepidemiol Drug Saf*. 2013;22(1):16–24. doi:10.1002/pds.3284
8. Palmsten K, Huybrechts KF, Mogun H, et al. Harnessing the Medicaid Analytic eXtract (MAX) to evaluate medications in pregnancy: design considerations. *PLoS One*. 2013;8(6):e67405. doi:10.1371/journal.pone.0067405
9. Li Q, Andrade SE, Cooper WO, et al. Validation of an algorithm to estimate gestational age in electronic health plan databases. *Pharmacoepidemiol Drug Saf*. 2013;22(5):524–532. doi:10.1002/pds.3407
10. Hornbrook MC, Whitlock EP, Berg CJ, et al. Development of an algorithm to identify pregnancy episodes in an integrated health care delivery system. *Health Serv Res*. 2007;42(2):908–927. doi:10.1111/j.1475-6773.2006.00635.x
11. Toh S, Mitchell AA, Werler MM, Hernandez-Diaz S. Sensitivity and specificity of computerized algorithms to classify gestational periods in the absence of information on date of conception. *Am J Epidemiol*. 2008;167(6):633–640.
12. Zhu Y, Hampp C, Wang X, et al. Validation of algorithms to estimate gestational age at birth in the Medicaid Analytic eXtract-Quantifying the misclassification of maternal drug exposure during pregnancy. *Pharmacoepidemiol Drug Saf*. 2020. doi:10.1002/pds.5126. .
13. Mainor AJ, Morden NE, Smith J, Tomlin S, Skinner J. ICD-10 coding will challenge researchers: caution and collaboration may reduce measurement error and improve comparability over time. *Med Care*. 2019;57(7):e42–e46. doi:10.1097/MLR.0000000000001010
14. Panozzo CA, Woodworth TS, Welch EC, et al. Early impact of the ICD-10-CM transition on selected health outcomes in 13 electronic health care databases in the United States. *Pharmacoepidemiol Drug Saf*. 2018;27(8):839–847. doi:10.1002/pds.4563
15. Chabra S. International classification of diseases, 10th Revision, coding for prematurity: need for standardized nomenclature. *Health Care Manag (Frederick)*. 2015;34(2):123–127. doi:10.1097/HCM.0000000000000053
16. Panozzo CA, Welch EC, Woodworth TS, et al. Assessing the impact of the new ICD-10-CM coding system on pharmacoepidemiologic studies-an application to the known association between angiotensin-converting enzyme inhibitors and angioedema. *Pharmacoepidemiol Drug Saf*. 2018;27(8):829–838. doi:10.1002/pds.4550
17. Heslin KC, Owens PL, Karaca Z, Barrett ML, Moore BJ, Elixhauser A. Trends in opioid-related inpatient stays shifted after the US transitioned to ICD-10-CM diagnosis coding in 2015. *Med Care*. 2017;55(11):918–923. doi:10.1097/MLR.0000000000000805
18. Initiative S, ed.Use of Multiple Sclerosis Drugs Among Pregnant Women. 2018.
19. Hernandez-Diaz S, Huybrechts KF, Desai RJ, et al. Topiramate use early in pregnancy and the risk of oral clefts: a pregnancy cohort study. *Neurology*. 2018;90(4):e342–e351. doi:10.1212/WNL.0000000000004857
20. Huybrechts KF, Hernandez-Diaz S, Straub L, et al. Association of maternal first-trimester ondansetron use with cardiac malformations and oral clefts in offspring. *JAMA*. 2018;320(23):2429–2437.
21. Sarayani A, Albogami Y, Elkhider M, Hincapie-Castillo JM, Brumback BA, Winterstein AG. Comparative effectiveness of risk mitigation strategies to prevent fetal exposure to mycophenolate. *BMJ Qual Saf*. 2019.
22. Shin J, Cheetham TC, Wong L, et al. The impact of the iPLEDGE program on isotretinoin fetal exposure in an integrated health care system. *J Am Acad Dermatol*. 2011;65(6):1117–1125. doi:10.1016/j.jaad.2010.09.017
23. Martin JA, Hamilton BE, Osterman MJK. Births in the United States, 2018. *NCHS Data Brief*. 2019;(346):1–8.