

Identification and validation of uterine perforation, intrauterine device expulsion, and breastfeeding in four health care systems with electronic health records

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

Mary S Anthony,¹ Mary Anne Armstrong,² Darios Getahun,³ Delia Scholes,⁴ Jennifer Gatz,⁵ Renate Schulze-Rath,⁶ Debbie Postlethwaite,² Maqdooda Merchant,² Amy L Alabaster,² Giulia Chillemi,² Tina Raine-Bennett,² Fagen Xie,³ Vicki Y Chiu,³ Theresa M Im,³ Harpreet S Takhar,³ Michael Fassett,³ Jane Grafton,⁴ David Cronkite,⁴ Laura Ichikawa,⁴ Susan D Reed,^{4,7} Siu Lui Hui,⁵ Mary E Ritchey,¹ Catherine W Saltus,⁸ Elizabeth B Andrews,¹ Kenneth J Rothman,⁸ Alex Asiimwe,⁶ Richard Lynam,⁹ Juliane Schoendorf¹⁰

¹RTI Health Solutions, Research Triangle Park, NC, USA; ²Kaiser Permanente Northern California, Oakland, CA, USA; ³Kaiser Permanente Southern California, Pasadena, CA, USA; ⁴Kaiser Permanente Washington, Seattle, WA, USA; ⁵Regenstrief Institute, Indianapolis, IN, USA; ⁶Bayer AG, Berlin, Germany; ⁷University of Washington, Seattle, WA, USA; ⁸RTI Health Solutions, Waltham, MA, USA; ⁹Bayer HealthCare, Whippany, NJ, USA; ¹⁰Bayer OY, Espoo, Finland

Correspondence: Mary S Anthony
RTI Health Solutions, Research Triangle
Park, 200 Park Office Dr, NC 27709-2194
USA
Tel +1 919 485 5509
Email manthony@rti.org

Objective: To validate algorithms identifying uterine perforations and intrauterine device (IUD) expulsions and to ascertain availability of breastfeeding status at the time of IUD insertion.

Study design and setting: Four health care systems with electronic health records (EHRs) participated: Kaiser Permanente Northern California (KPNC), Kaiser Permanente Southern California (KPSC), Kaiser Permanente Washington (KPWA), and Regenstrief Institute (RI). The study included women ≤ 50 years of age with an IUD insertion. Site-specific algorithms using structured and unstructured data were developed and a sample validated by EHR review. Positive predictive values (PPVs) of the algorithms were calculated. Breastfeeding status was assessed in a random sample of 125 women at each research site with IUD placement within 52 weeks postpartum.

Results: The study population included 282,028 women with 325,582 IUD insertions. The PPVs for uterine perforation were KPNC 77%, KPSC 81%, KPWA 82%, and RI 47%; PPVs for IUD expulsion were KPNC 77%, KPSC 87%, KPWA 68%, and RI 37%. Across all research sites, breastfeeding status at the time of IUD insertion was determined for 94% of those sampled.

Conclusions: Algorithms with a high PPV for uterine perforation and IUD expulsion were developed at 3 of the 4 research sites. Breastfeeding status at the time of IUD insertion could be determined at all research sites. Our findings suggest that a study to evaluate the associations of breastfeeding and postpartum IUD insertions with risk of uterine perforation and IUD expulsion can be successfully conducted retrospectively; however, automated application of algorithms must be supplemented with chart review for some outcomes at one research site due to low PPV.

Keywords: electronic health records, intrauterine device, breastfeeding, validation study, algorithm, postpartum

Introduction

Health care databases are increasingly used for medication and device safety studies. Data sources associated with electronic health records (EHRs) have an advantage over administrative claims databases in that clinical notes can be accessed to provide additional insight into medical encounters, medical events, lifestyle factors, and health-related behaviors. In addition, results from laboratory tests, physical examination (eg, weight and physical findings such as pelvic

tenderness or visibility of intrauterine device [IUD] strings), and imaging procedures are available. When answering important clinical questions, however, methods to utilize these richer data sources must be developed and validated prior to use.

Interest in conducting a study on the safety of IUDs was prompted, in part, by information on uterine perforation reported by Heinemann and colleagues^{1,2} in the European Active Surveillance Study on Intrauterine Devices (EURAS IUD). They conducted a prospective cohort study in six European countries between 2006 and 2012. It included 61,448 women, who were followed for 12 months after IUD insertion. Breastfeeding at the time of IUD insertion and postpartum insertion (within the first 36 weeks after most recent delivery) were both associated with higher risk of IUD-related uterine perforation, which was self-reported and subsequently validated by the physician who diagnosed or treated the perforation. In another publication from this study,³ contraceptive failure was reported, with results indicating that about 25% of the unintended pregnancies occurred after an unrecognized IUD expulsion. Because of possible differences between Europe and the United States (US) in postpartum timing of IUD placement and breastfeeding practices, there is interest in conducting a similar study in the US. Retrospective cohort studies using certain data sources associated with EHRs have a potential advantage over prospective studies of being more efficient and more likely to reflect usual clinical practice.

The purpose of this validation work was to develop and validate automated algorithms that include both structured (eg, *International Classification of Diseases* [ICD] codes, Current Procedural Terminology [CPT] codes, Healthcare Common Procedure Coding System [HCPCS], National Drug Codes [NDC]), and unstructured data (eg, Natural Language Processing [NLP] terms) to identify uterine perforation and IUD expulsion involving research sites with access to EHRs. For this study, we focused on the positive predictive value (PPV) of the algorithms for uterine perforation and IUD expulsion. We also assessed the availability and completeness of breastfeeding information and the ability to determine breastfeeding status when IUD placement occurred among postpartum women. We undertook this study to determine the feasibility of conducting a multicenter cohort study utilizing health care systems with EHRs to evaluate safety and identify potential risk factors for two adverse outcomes: uterine perforation and IUD expulsion.

Materials and methods

We conducted the study at four research sites associated with US health care systems that use EHRs: Kaiser Permanente Northern California (KPNC), Kaiser Permanente Southern California (KPSC), Kaiser Permanente Washington (KPWA, formerly Group Health), and Regenstrief Institute (RI). RTI Health Solutions served as the coordinating center. As the study was intended to validate algorithms and assess feasibility, statistical hypothesis tests or measures of association between exposure and outcome are not reported. Each research site maintained their own patient-level data and provided site-specific, aggregated, de-identified results for further analysis by RTI Health Solutions. Site-specific institutional review board approvals were obtained for this study.

Study population

The source population included US women enrolled in one of four health care systems. KPNC, KPSC, and KPWA included members with health coverage through Kaiser Foundation Health Plan. These three health care systems are located on the US west coast (northern and southern California, and Washington, respectively) and include racially and ethnically diverse memberships. The RI data are from a health information exchange system that covers most of Indiana and has a larger proportion of African Americans but a lower proportion of other minorities than the other three research sites.

To be eligible for the study, women had to have evidence in their medical record of an IUD placement during the study period and could have had multiple IUD insertions during the study period. They were also required to be aged ≤ 50 years at the time of the IUD insertion and to have EHRs available for review. The earliest date for inclusion (ie, the dates at which EHR data were fully incorporated into the health plans' automated data) was January 1 of the following years: 2001 (RI), 2006 (KPWA), 2008 (KPSC), and 2009 (KPNC). The study period ended on September 30, 2015.

Variables of interest

Variables were defined with clinical input and are described below. Investigators from all four research sites collaborated in developing algorithms and approaches to identify the variables of interest. This endeavor was an iterative process, designed to improve on the PPV of the algorithms. Approaches to defining outcomes, exposures,

validation procedures, and clinical definitions for validation of outcomes were developed and standardized in research partner meetings and in documents shared across research sites. The algorithms at all research sites captured the same concepts and used a combination of structured (ICD, CPT, NDC, HCPCS) and unstructured (NLP terms) data. Tables summarizing the types of data used in the algorithms at each research site as well as the structured codes and NLP terms for uterine perforation and IUD expulsion are available in the online supplemental materials (Appendix A).

Uterine perforation included both complete (ie, IUD was documented as located in the pelvis or abdominal cavity) and partial perforation (ie, IUD was documented as embedded in the myometrium or cervix) via hysteroscopy, laparoscopy, laparotomy, or an imaging study in conjunction with evidence of a difficult IUD removal (eg, ultrasound suggesting partial embedment plus string avulsed with traction upon attempted removal).

IUD expulsion, specifically the unintended, spontaneous expulsion of the IUD through the cervix, included both partial (eg, IUD visibly extruding from external cervical os) and complete (ie, not present in the vagina, cervix, uterus, pelvic, or abdominal cavity). If the IUD was malpositioned in the uterine cavity (eg, imaging demonstrated IUD in lower uterine segment), it was not considered an expulsion.

Postpartum interval was calculated as the difference between the date of the most recent delivery before the insertion and the IUD insertion date, and it was classified into the following categories of postpartum periods: ≤ 3 days, >3 days and <4 weeks, ≥ 4 to <6 weeks, ≥ 6 to ≤ 14 weeks, >14 to ≤ 52 weeks, and >52 weeks or no prior delivery recorded.

Breastfeeding status at the time of IUD insertion was classified as “breastfeeding,” “not breastfeeding,” or “undetermined/missing.” Evidence of breastfeeding was derived from linked mother/infant records (eg, well-child care visits, infant check-ups, immunization visits) and clinical notes extracted from EHRs for the infant and/or the mother from around the time of IUD insertion with manual review of the records.

Sampling strategy

Each research site—involving project investigators, clinicians, NLP experts, programmers, and medical record analysts—identified possible cases of uterine perforation and IUD expulsion using the algorithms. Trained medical

record analysts or obstetric/gynecologic clinicians reviewed the EHRs of possible cases to determine whether there was evidence of the outcome of interest. Possible cases for which the documentation was unclear were adjudicated by clinicians.

Each research site chose a random sample of one-third of all women identified by the algorithms as possible cases of uterine perforation or IUD expulsion (maximum 100 for each outcome) to review in the EHR and determine case status (yes, no, undetermined). If the site decided to improve the algorithm, the revised algorithm was used to identify a new set of possible cases, and a random sample of one-third (maximum 100) was selected for EHR review and case status determination.

Breastfeeding status was identified among a random sample of 125 women in each health care system who had undergone IUD placement within the first 52 weeks postpartum. Where available, a random sample of 25 women within each of the 5 postpartum interval categories was selected to construct the sample of 125 women.

Analysis

We described the IUD cohort, including age, type of IUD, and categories for postpartum time interval, with summary statistics for continuous and categorical variables. Continuous variables were summarized by the mean, median, standard deviation, and quartiles; categorical variables were summarized by frequency counts and percentages.

The PPV for each outcome (ie, uterine perforation, IUD expulsion) was defined as the percentage of possible cases identified by the algorithms that were determined to be actual cases upon medical record review. PPV was calculated as $(\text{number of true positives}/\text{number sampled}) \times 100$. The exact (Clopper–Pearson) 95% confidence intervals (CIs) around the estimates were used to describe the accuracy of the algorithms in identifying uterine perforation and IUD expulsion. For breastfeeding, the number and percentage of women classified in each postpartum time interval by breastfeeding status were calculated. External validity for breastfeeding status was assessed by comparing percentages within each of the four health care systems with national surveys collecting breastfeeding status by state.⁴

Results

Table 1 includes data on the characteristics of the study population at each research site and across all research sites. The study population included 282,028 women with

Table 1 Characteristics of the study population with one or more intrauterine device insertions

Characteristic	KPNC	KPSC	KPWA	RI	All research sites
Number of women	145,004	107,148	22,683	7,193	282,028
Number of IUD insertions	168,744	122,743	26,315	7,780	325,582
Inclusion dates	2009–2015	2008–2015	2006–2015	2001–2015	
Age in years, median (Q1, Q3)	32 (26, 38)	31 (26, 38)	31 (25, 37)	29 (24, 35)	
Type of IUD, n (%)*					
LNG-releasing					246,139 (75.6)
Copper					64,652 (19.9)
Unidentified					14,791 (4.5)
Postpartum time before the IUD insertion, n (%)					
≤3 days	1,609 (1.0)	25 (<0.1)	58 (0.2)	3 (<0.1)	1,695 (0.5)
>3 days and <4 weeks	306 (0.2)	75 (0.1)	80 (0.3)	4 (<0.1)	465 (0.1)
≥4 weeks and <6 weeks	9,223 (5.5)	3,669 (3.0)	688 (2.6)	114 (1.5)	13,694 (4.2)
≥6 weeks and ≤14 weeks	28,618 (17.0)	28,166 (22.9)	5,078 (19.3)	874 (11.2)	62,736 (19.3)
>14 weeks and ≤52 weeks	10,119 (6.0)	10,495 (8.6)	1,755 (6.7)	457 (5.9)	22,826 (7.0)
>52 weeks or no prior delivery recorded	118,869 (70.4)	80,313 (65.4)	18,656 (70.9)	6,328 (81.3)	224,166 (68.9)

Notes: *For IUD type, only the data summarized across all research sites are presented to avoid releasing Kaiser Permanente proprietary information.
Abbreviations: IUD, intrauterine device; KPNC, Kaiser Permanente Northern California; KPSC, Kaiser Permanente Southern California; KPWA, Kaiser Permanente Washington; LNG, levonorgestrel; Q, quartile; RI, Regenstrief Institute.

325,582 IUD insertions. The median age of the women at the time of IUD placement at each research site varied from 29 to 32 years. Of the IUDs used across all research sites, >75% were levonorgestrel releasing and about 20% were copper; IUD type could not be determined for 4.5%.

Sixty-nine percent of IUDs were inserted in women who were >52 weeks postpartum or for whom no evidence of a delivery existed in the databases before IUD insertion. During the first-year postpartum, the largest percentage of IUD insertions occurred between 6 and 14 weeks postpartum (19.3% of all IUD insertions). Fewer than 1% of the IUDs were inserted earlier than 4 weeks postpartum.

The total number of possible uterine perforations identified for each research site using the site-specific

algorithms is provided in Table 2. The number sampled, case status based on EHR review, PPVs, and 95% CIs for the validation of potential cases of uterine perforation are also summarized in Table 2. Two of the research sites, KPWA and RI, refined their algorithms and conducted a second (RI for IUD expulsion) or third (RI for uterine perforation and KPWA for both outcomes) round of validation. The data reported here are from the final algorithm validation. Across research sites, the PPVs for the uterine perforation algorithm were 47%, 77%, 81%, and 82% for RI, KPNC, KPSC, and KPWA, respectively.

The total number of possible IUD expulsions identified for each research site using the site-specific algorithms are provided in Table 3. The number of possible cases of IUD

Table 2 Uterine perforation algorithm validation results within each research site

Research site	Potential uterine perforations identified by algorithm	Number sampled for EHR review	Case status after EHR review			PPV% (95% CI)
			Yes	No	Undetermined	
KPNC	444	100	77	16	7	77 (68–85)
KPSC	388	100	81	19	0	81 (72–88)
KPWA	121	28 [†]	23	4	1	82 (63–94)
RI	67*	30	14	14	2	47 (29–65)

Notes: *Regenstrief Institute identified only the first potential uterine perforation for each woman; the total number of potential uterine perforations was not assessed.
[†]Less than one-third of the total sample because some patients were included in prior samples.
Abbreviations: CI, confidence interval; EHR, electronic health record; KPNC, Kaiser Permanente Northern California; KPSC, Kaiser Permanente Southern California; KPWA, Kaiser Permanente Washington; PPV, positive predictive value; RI, Regenstrief Institute.

Table 3 Intrauterine device expulsion algorithm validation results within each research site

Research site	Potential IUD expulsions identified by algorithm	Number sampled for EHR review	Case status after EHR review			PPV% (95% CI)
			Yes	No	Undetermined	
KPNC	4,185	100	77	21	2	77 (68–85)
KPSC	2,376	100	87	12	1	87 (79–93)
KPWA	531	103 [†]	70	33	0	68 (58–77)
RI	268*	100	37	61	2	37 (28–46)

Notes: *RI identified only the first potential IUD expulsion for each woman; the total number of IUD expulsions was not assessed. [†]More than 100 possible cases included because all overlapping possible cases identified in the prior two samples were included.

Abbreviations: CI, confidence interval; EHR, electronic health record; IUD, intrauterine device or intrauterine system; KPNC, Kaiser Permanente Northern California; KPSC, Kaiser Permanente Southern California; KPWA, Kaiser Permanente Washington; PPV, positive predictive value; RI, Regenstrief Institute.

expulsion sampled for validation was 100 at each research site. KPWA refined its initial algorithm twice, reviewing medical records of a sample of potential cases after each revision. The case status based on EHR review, PPVs, and 95% CIs is summarized in Table 3. The PPVs were 37%, 68%, 77%, and 87% for RI, KPWA, KPNC, and KPSC, respectively.

The percentages of women in the five postpartum categories who were classified as “breastfeeding,” “not breastfeeding,” or “undetermined” at each research site and across all research sites are summarized in Table 4. Among those with an IUD insertion within 52 weeks after delivery, 92% of those sampled were classified as breastfeeding in the first 3 days postpartum, 90% in the >3 days and <4 weeks postpartum interval and by the >14–≤52-week interval only 41% of those sampled were classified as breastfeeding.

For all postpartum intervals combined, 72% of the women undergoing IUD placement at KPNC were categorized as breastfeeding at the time of insertion, as were 86% at KPSC, 80% at KPWA, and 45% at RI. Across all research sites and for all postpartum intervals, breastfeeding status at the time of IUD insertion was unable to be determined for 6% of those sampled, although this varied across research sites (range, 0–21%).

Discussion

This validation study was performed to determine the feasibility of conducting a safety study relying on automated data in these four health care systems using EHRs to assess whether risks of uterine perforation or IUD expulsion are associated with breastfeeding at the time of IUD insertion or with early postpartum IUD insertion. Although these associations have been reported in a prospective observational study, EURAS IUD,^{1,2} because of

potential differences between the European Union (EU) and US in breastfeeding practices and postpartum timing of IUD insertion, there is interest in conducting a study in the US. To our knowledge, algorithms for uterine perforation and IUD expulsion have not been validated for US EHR data, and there is little information in the literature about the feasibility of determining breastfeeding status.

While EHRs have been used in several clinical domains within the US for decades, they became widely adopted after the passage of the HITECH Act in 2009.⁵ There is consensus in the literature that wider use of EHRs has the potential to improve the quality of patient care.^{6,7} Readily accessible clinical, laboratory, and other sources of health information provide opportunities in advancing health reporting and surveillance as well as research dissemination. However, previous studies have demonstrated mixed results regarding the completeness of clinical information in the EHR and its potential impact on use for research purposes;⁸ thus, the need for this validation effort.

There were 325,582 IUD insertions identified at the four research sites included in this validation research: three Kaiser Permanente health care system-related research sites on the US west coast and one health information exchange in Indiana. Across all research sites, the mean duration of continuous enrollment was 95 months (range, 91–105 [data not shown]). This population is nearly 5 times larger than the population reported in the EURAS study, which reported 81 perforations among the 61,448 women in the first 12 months (1.3 cases/1,000 IUD insertions).¹ Therefore, with the large number of women identified in this study who received IUDs and the long duration of continuous follow-up, we anticipate the numbers will be more than adequate to conduct the planned safety study utilizing EHRs from these 4 health care systems.

Table 4 Breastfeeding status at the time of intrauterine device insertion within 52 weeks after delivery by postpartum time intervals* at each research site

Postpartum time interval	KPNC sample (n=125)	KPSC sample (n=125)	KPWA sample (n=125)	RI sample (n=125)	All research sites (N=500)
≤3 days					
Yes, n (%)	24 (96)	25 (100)	23 (92)	0 (0)	72 (92)
No, n (%)	1 (4)	0 (0)	0 (0)	0 (0)	1 (1)
Undetermined, n (%)	0 (0)	0 (0)	2 (8)	3 (100)	5 (6)
>3 days and <4 weeks					
Yes, n (%)	22 (88)	23 (92)	24 (96)	2 (50)	71 (90)
No, n (%)	3 (12)	2 (8)	1 (4)	2 (50)	8 (10)
Undetermined, n (%)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
≥4 weeks and <6 weeks					
Yes, n (%)	19 (76)	23 (92)	21 (84)	35 (52)	98 (69)
No, n (%)	6 (24)	2 (8)	2 (8)	24 (35)	34 (24)
Undetermined, n (%)	0 (0)	0 (0)	2 (8)	9 (13)	11 (8)
≥6 weeks and ≤14 weeks					
Yes, n (%)	19 (76)	19 (76)	19 (76)	14 (56)	71 (71)
No, n (%)	6 (24)	6 (24)	6 (24)	8 (32)	26 (26)
Undetermined, n (%)	0 (0)	0 (0)	0 (0)	3 (12)	3 (3)
>14 weeks and ≤52 weeks					
Yes, n (%)	6 (24)	17 (68)	13 (52)	5 (20)	41 (41)
No, n (%)	19 (76)	8 (32)	12 (48)	9 (36)	48 (48)
Undetermined, n (%)	0 (0)	0 (0)	0 (0)	11 (44)	11 (11)
All postpartum intervals					
Yes, n (%)	90 (72)	107 (86)	100 (80)	56 (45)	353 (71)
No, n (%)	35 (28)	18 (14)	21 (17)	43 (34)	117 (23)
Undetermined, n (%)	0 (0)	0 (0)	4 (3)	26 (21)	30 (6)

Notes: *Random sample of 25 selected at each research site for each postpartum time period with the exception of RI, where the sample size for ≤3 days postpartum was 3, the sample size for >3 days and <4 weeks postpartum was 4, and the sample size for ≥4 weeks and <6 weeks postpartum was 68 (oversampled to ensure total sample size was 125). **Abbreviations:** KPNC, Kaiser Permanente Northern California; KPSC, Kaiser Permanente Southern California; KPWA, Kaiser Permanente Washington; RI, Regenstrief Institute.

The main focus of this study was to determine whether algorithms with a high PPV could be developed to identify cases of uterine perforation and IUD expulsion in these 4 health care systems. The PPV at the 3 Kaiser research sites ranged from 77% to 82%; RI demonstrated a lower PPV (47%). For IUD expulsion algorithms, the PPV at the 3 Kaiser research sites ranged from 68% to 87% and was 37% at RI. The PPVs at RI were lower for both outcomes, which might be due to less availability of structured codes for uterine perforation and IUD expulsion at RI. However, given the smaller size of this research site, the RI investigators plan to conduct complete chart reviews for all possible cases of these outcomes and will not rely on automated algorithms for these outcomes. These results provide assurance that the outcomes can be identified with rather high accuracy using EHR data and algorithms

already developed and validated at the three Kaiser research sites, with some minor additional refinements.

To assess breastfeeding within this study, women who had a live birth and an IUD inserted within 52 weeks postpartum were categorized into one of five groups based on the timing of IUD insertion relative to their delivery date. Overall, 94% of the women sampled could be classified as either breastfeeding or not breastfeeding at the time of IUD insertion. Two of the research sites (KPNC and KPSC) administer a questionnaire at well-child care visits for infants up to 12 months of age that inquires whether the infant is breastfed; these data were used in this validation study and will be used in the safety study as well. Additional work will be done to develop NLP-based algorithms to determine breastfeeding status from the records that were manually reviewed for this

study. Given the high percentage of women for whom breastfeeding status can be determined, these data will be adequate for the safety study.

As would be expected, a high percentage of women at three research sites were determined to be breastfeeding in the first 4 weeks postpartum. In RI, the lower percentage of breastfeeding in the first two postpartum categories could have been a result of the small number of women with IUD insertion less than 4 weeks postpartum or from geographic differences in breastfeeding attitudes. Data from the Centers for Disease Control and Prevention (CDC) National Immunization Survey⁴ showed that for California and Washington, 92–93% of the women had ever breastfed, 63–64% were breastfeeding at 6 months, and 35–38% were breastfeeding at 12 months. For Indiana, the prevalence of breastfeeding was lower, with 74% having ever breastfed, 39% breastfeeding at 6 months, and 22% breastfeeding at 12 months. Thus, the proportions of women breastfeeding at different postpartum times in this study reflect the patterns reported by the CDC.⁴

There are several strengths and limitations to this study. Strengths include the size of the available study populations, duration of continuous enrollment at these research sites, and the multidisciplinary research teams operating at all research sites. Expertise included investigators experienced in multi-site collaborative studies, skilled programming and analytic staff, and clinicians possessing current clinical experience with IUDs and familiarity with these health care systems. These EHRs provide a rich source of clinical information, laboratory and examination results, and notes from patient/clinician interactions, all of which provide a comprehensive overview of a patient's health experiences; however, each EHR data source is different regarding the type of data available and the completeness of its capture of a patient's health care encounters. These data sources were selected because nearly all health care encounters are within the health care system and those that are not are documented through claims. Limitations include the need to develop separate algorithms at each research site since clinical practice and documentation vary; however, the algorithms at each research site mapped to the same core concepts, and many of the results appear to be consistent across research sites. The differences in coding and terminology noted among these health care systems, necessitating site-specific algorithms, suggest that other data sources would have to conduct validation studies for these types of outcomes. One research site found lower PPVs, so chart review will still be utilized, however on a relatively

small scale. To completely validate a general-purpose predictive algorithm, one would need more information than we had; specifically, we would have required a gold standard that identified a complete set of true cases. Nonetheless, we were able to assess PPV, which informs us about the potential for misclassification of the included cases. Negative predictive value of the algorithms was not done since the outcomes are so rare. In the ongoing safety study, we are focusing on relative effects (ie, relative risks, hazard ratios); therefore, we think that incomplete case ascertainment will not be a source of much bias, whereas a low PPV would be a source of bias for both relative and absolute effects. Lastly, we plan to include data through 2017 in the safety study, and this will involve use of *International Classification of Diseases, Tenth Revision* (ICD-10) codes not validated here; therefore, we will monitor whether the change in coding systems results in different rates of events.

Conclusions

This study provides evidence that a retrospective study to evaluate the associations of breastfeeding at IUD insertion and postpartum timing of IUD insertion with risk of uterine perforation or IUD expulsion can be successfully conducted in these health care systems relying largely on automated algorithms with EHR data; however, automated application of algorithms must be supplemented with chart review to adjudicate possible cases of uterine perforation and IUD expulsion outcomes at one of the research sites. Such a retrospective database study can provide larger study populations and more efficient and timely results than prospective studies that entail new data collection, such as EURAS IUD,¹ and is generally less difficult and burdensome to conduct. However, not all EHRs capture the complete health care experience for the patients, and not all variables of interest for a study are captured in routine clinical care, so each study must consider the strengths and limitations of different study approaches and data sources.

Ethics approval

Site-specific Institutional Review Board approvals were obtained for this study: Human Subjects Review Office, Group Health Research Institute #998277; Human Subjects Office, Office of Research Compliance – Indiana University #1610788954; Kaiser Permanente Northern California Institutional Review Board #CN-17-2863; Human Research Subjects Protection Office, Institutional Review Board, Kaiser Permanente Southern California #11289.

Acknowledgments

The authors would like to acknowledge the following individuals who contributed to the planning or conduct of the study: Sharon Brown (formerly of Bayer HealthCare), Kevin Filocamo (KPWA), Vina Graham (KPWA), Kelly Hansen (KPWA), Joel Martin (RI), David McSorley (RTI Health Solutions [RTI-HS]), Jeffrey Peipert (RI/Indiana University School of Medicine), Montse Soriano-Gabarro (Bayer AG), and Jane Wang (RI). Funding for this research was provided by Bayer AG, Berlin, Germany to RTI-HS, KPNC, KPSC, KPWA, and RI. RTI-HS led the design of the study and interpretation of the results in collaboration with study team members from KPNC, KPSC, KPWA, RI, and Bayer. KPNC, KPSC, KPWA, and RI conducted the analyses, which were reviewed by study team members from RTI-HS and Bayer AG. The contracts between Bayer AG and each of the other organizations (KPNC, KPSC, KPWA, RI, RTI-HS) include independent publication rights. Bayer AG was provided the opportunity to review the manuscript prior to submission and comments were advisory.

Disclosure

RS-R, AA, RL, and JS are employees of Bayer, the marketing authorization holder for three IUD brands among others that were included in this study. AA reports stocks from Bayer. MAA is the Principal Investigator at the KPNC site and employed by KPNC. She also reports grants from Bayer AG, during the conduct of the study and outside the submitted work. DG is employed by KPSC and reports grants from Bayer Pharma AG, during the conduct of the study and received grants from Bayer Pharma AG, Centers for Disease Control and Prevention, and the US National Institutes of Health/National Institute of Child Health and Human Development (NIH/NICHHD), outside the submitted work. DS is employed by KPWA and reports grants from Bayer, during the conduct of the study. JG is employed by RI and reports grants from Bayer AG, during the conduct of the study and outside the submitted work. DP is employed by KPNC and reports grants from Bayer AG, during the conduct of the study and outside the submitted work. ALA

is an employee of KPNC. GC is an employee of KPNC and reports grants from Bayer AG, during the conduct of the study and outside the submitted work. TR-B and MM are employees of KPNC and reports grants from Bayer A.G, during the conduct of the study and reports consulting fees paid to KPNC from TEVA pharmaceutical, outside the submitted work. MF is employed by KPSC and reports grants from Bayer, during the conduct of the study; grants from Bayer and NIH/NICHHD, outside the submitted work. DC is employed by KPWA and reports grants from Bayer, during the conduct of the study. SDR is employed by KPWA and reports grants from Bayer AG, during the conduct of the study. MSA, MER, CWS, EBA, and KJR are employed by RTI Health Solutions. The authors report no other conflicts of interest in this work.

References

1. Heinemann K, Reed S, Moehner S, Minh TD. Risk of uterine perforation with levonorgestrel-releasing and copper intrauterine devices in the European Active Surveillance Study on Intrauterine Devices. *Contraception*. 2015;91:274–279. doi:10.1016/j.contraception.2015.01.007
2. Heinemann K, Barnett C, Reed S, Möhner S, Minh TD. IUD use among parous women and risk of uterine perforation: a secondary analysis. *Contraception*. 2017;95:605–607. doi:10.1016/j.contraception.2017.03.007
3. Heinemann K, Reed S, Moehner S, Minh TD. Comparative contraceptive effectiveness of levonorgestrel-releasing and copper intrauterine devices: the European Active Surveillance Study on Intrauterine Devices. *Contraception*. 2015b;91:280–283. doi:10.1016/j.contraception.2015.01.011
4. Centers for Disease Control and Prevention (CDC). Breastfeeding report card United States/2014. Available from: <https://www.cdc.gov/breastfeeding/pdf/2014breastfeedingreportcard.pdf>; 2014 Accessed August 26, 2017.
5. American Recovery and Reinvestment Act, HR 1, 111th Congress, 1st Session; 2009. Available from: <https://www.congress.gov/111/plaws/publ5/PLAW-111publ5.pdf>. Accessed October 24, 2017.
6. Bell LM, Grundmeier R, Localio R, et al. Electronic health record-based decision support to improve asthma care: a cluster-randomized trial. *Pediatrics*. 2010;125(4):e770–7. doi:10.1542/peds.2009-1385
7. Jamoon E, Beatty P, Bercovitz A, Woodwell D, Palso K, Rechtsteiner E. Physician adoption of electronic health record systems: united States, 2011. NCHS Data Brief, No. 98, U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics; July 2012.
8. Weiskopf NG, Weng C. Methods and dimensions of electronic health record data quality assessment: enabling reuse for clinical research. *J Am Med Inform Assoc*. 2013;20(1):144–151. doi:10.1136/amiainl-2011-000681

Clinical Epidemiology

Dovepress

Publish your work in this journal

Clinical Epidemiology is an international, peer-reviewed, open access, online journal focusing on disease and drug epidemiology, identification of risk factors and screening procedures to develop optimal preventative initiatives and programs. Specific topics include: diagnosis, prognosis, treatment, screening, prevention, risk factor modification,

systematic reviews, risk & safety of medical interventions, epidemiology & biostatistical methods, and evaluation of guidelines, translational medicine, health policies & economic evaluations. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use.

Submit your manuscript here: <https://www.dovepress.com/clinical-epidemiology-journal>