Observational/Post Authorization Study Information

Title	Study on the Association of Uterine Perforation and IUD Expulsion With Breastfeeding Status at the Time of IUD Insertion and Postpartum Timing of IUD Insertion in Electronic Medical Record Databases – A Postmarketing Requirement for Mirena (APEX IUD)
Protocol version identifier	Version 2.0 (protocol amendment)
Date of last version of protocol	29 June 2018
IMPACT study number	19682
Study type/Study phase	☐ non-PASS ☒ PASS Joint PASS: ☐ YES ☐ NO Phase IV
Active substance	Intrauterine contraceptives: Plastic IUD with progestogen (ATC code G02BA03) and Plastic IUD with copper (ATC code G02BA02)
Medicinal product	Mirena®, Skyla®, Liletta®, Kyleena®, and ParaGard® intrauterine devices
Marketing authorization holder(s)	Bayer AG, 13342 Berlin
Research question and objectives	United States postmarketing requirement study assessing outcomes of uterine perforation and intrauterine device (IUD) expulsion in association with breastfeeding, postpartum exposures, and type of IUD. The study aims to quantify the risk of uterine perforation and IUD expulsion for the following comparisons: • Women who are breastfeeding at the time of IUD insertion versus not breastfeeding at the time of IUD insertion • Women who had a first observed IUD insertion within different time periods postpartum (i.e., ≤ 6 weeks, > 6 weeks and ≤ 14 weeks, > 14 weeks and ≤ 52 weeks) versus women who had their first observed IUD insertion more than 52 weeks postpartum, including women without a recorded delivery within the past 52 weeks This study will also assess the risk of perforation and

	expulsion (separately) by type of IUD. In addition, this study aims to assess the following interactions: • The extent to which type of IUD (LNG-releasing vs. copper IUD) modifies the association between perforation and/or expulsion and breastfeeding status and/or postpartum status • The extent to which breastfeeding status modifies the association between perforation and/or expulsion and postpartum status
Country(-ies) of study	United States
Author	[redacted]

The study will be conducted in compliance with the protocol
and any applicable regulatory requirements.

Throughout this document, symbols indicating proprietary names (®, TM) may not be displayed. Hence, the appearance of product names without these symbols does not imply that these names are not protected.

1. Table of Contents

2. List of abbreviations

ATC	Anatomical Therapeutic Chemical (Classification System)
BMI	Body Mass Index
CI	Confidence Interval
DOR	Division of Research (KPNC)
EHR	Electronic Health Record
EMA	European Medicine Agency
EU PAS Register	European Union Electronic Register of Post-Authorisation Studies
EURAS-IUD	European Active Surveillance Study for Intrauterine Devices
FDA	Food and Drug Administration
HCPCS	Healthcare Common Procedure Coding System
HIPAA	Health Insurance Portability and Accountability Act
ICD-9	*International Classification of Diseases, 9th Revision*
ICD-9-CM	*International Classification of Diseases, 9th Revision, Clinical Modification*
ID	Identification
IPTW	Inverse Probability Treatment Weighting
IRB	Institutional Review Board
IRD	Incidence Rate Difference
IRR	Incidence Rate Ratio
ISPE	International Society for Pharmacoepidemiology
IT	Information Technology
IUD	Intrauterine Device
KPNC	Kaiser Permanente Northern California
KPSC	Kaiser Permanente Southern California
KPWA	Kaiser Permanente Washington
LNG	Levonorgestrel
LNG-IUD	Levonorgestrel-Releasing Intrauterine System
NLP	natural language processing
PMR	Postmarketing Requirement
RDW	Research Data Warehouse (KPSC site)
RI	Regenstrief Institute
RR	Relative Risk
RTI-HS	RTI Health Solutions, a unit of RTI International, a nonprofit research organization
US	United States
VDW	Virtual Data Warehouse (Kaiser Permanente sites)

3. Responsible parties

Name	Role	Contact Information
Bayer – Project Sponsor		
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
RTI Health Solutions (RTI-HS) – Coordinating Center		
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
Kaiser Permanente Washington (KPWA) – Data Source Research Partner		
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
Kaiser Permanente Northern California (KPNC)– Data Source Research Partner		
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]

Name	Role	Contact Information
[redacted]	[redacted]	[redacted]
Kaiser Permanente Southern California (KPSC)– Data Source Research Partner		
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
Regenstrief Institute – Data Source Research Partner		
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]

4. Abstract

4.1 Title

Study on the Association of Uterine Perforation and IUD Expulsion With Breastfeeding Status at the Time of IUD Insertion and Postpartum Timing of IUD Insertion in Electronic Medical Record Databases – A Postmarketing Requirement for Mirena (APEX IUD)

4.2 Rationale and background

Mirena, a levonorgestrel (LNG)-releasing intrauterine system, was approved for use in the United States (US) in December 2000 (Bayer HealthCare Pharmaceuticals Inc., 2015). In August 2015, Bayer received a postmarketing requirement from the US Food and Drug Administration (FDA) to evaluate the incidence of and risk factors for uterine perforation in patients receiving Mirena intrauterine devices (IUDs) (communication from FDA to Bayer, 18 August 2015). The concerns described by the FDA, based on information from the European Active Surveillance Study for Intrauterine Devices (EURAS-IUD), were that insertion of the IUD immediately postpartum might result in a higher risk of uterine perforation and that breastfeeding at the time of IUD insertion might be associated with higher risk of uterine perforation.

In response, on October 16, 2015 [see NDA 21, 255, seq 0096], Bayer proposed to conduct a retrospective, cohort study of women with an intrauterine system or copper IUD using electronic health record (EHR) databases and included information on three potential electronic data sources to support this approach. In May 2016, the FDA responded favorably to Bayer's proposal and indicated in September 2016 that use of the retrospective approach was pending successful results

from a validation/feasibility study that should be submitted before the postmarketing requirement study, to support the proposed retrospective study within the EHR setting. A validation study has been conducted within four* EHR databases and has indicated that uterine perforation, IUD† expulsion, and breastfeeding can be adequately identified within these data sources.

The study described within this protocol is a retrospective postmarketing requirement study assessing outcomes of uterine perforation and IUD expulsion in association with breastfeeding and postpartum exposures in the EHR databases, as requested by the FDA.

4.3 Research question and objectives

The overall goal of this study is to assess the impact of breastfeeding and timing of postpartum IUD insertion on uterine perforation and IUD expulsion in a population of US women. The study aims to quantify the difference in risk of perforation and expulsion in the following groups:

- Women who are breastfeeding at the time of IUD insertion versus not breastfeeding at the time of IUD insertion
- Women who had a first observed IUD insertion within different time periods postpartum (i.e., ≤ 6 weeks, > 6 weeks and ≤ 14 weeks, > 14 weeks and ≤ 52 weeks) versus women who had their first observed IUD insertion more than 52 weeks postpartum, including women without a recorded delivery within the past 52 weeks

This study will also assess the difference in risk of perforation and expulsion by type of IUD. In addition, this study aims to assess the following interactions:

- The extent to which breastfeeding status modifies the association of perforation and/or expulsion for women with IUD insertion at different time periods postpartum (i.e., IUD insertion ≤ 14 weeks versus IUD insertion > 14 weeks postpartum) among women with recorded delivery within the past 52 weeks
- The extent to which type of IUD (LNG-releasing vs. copper IUD) modifies the association of perforation and/or expulsion with breastfeeding status and/or postpartum status

Approximately 10% of IUD insertions in the databases are expected to occur in women who have previous exposure to an IUD within the data. For most analyses, only the first IUD insertion observed in the database will be included in order to maintain independence of observations. Multiple insertions will be considered only for objectives 18 and 26 (see Section 4.3.2).

The study includes the following primary and secondary objectives. The secondary objectives are grouped by type of analysis (rates, comparative, interaction) and outcome (uterine perforation, IUD expulsion, indicators of difficult insertion).

* Fourth data source added to incorporate women not on the west coast and not in the Kaiser Permanente health care system. Inclusion of this data source increases understanding of generalizability of the study results.
† From this point forward throughout the protocol, both LNG-releasing intrauterine systems and copper IUDs will be referenced as "IUD."

4.3.1 Primary objectives

1. To evaluate whether the risk of uterine perforation among women who were breastfeeding at the time of first observed IUD insertion differs from the risk of uterine perforation among women who were not breastfeeding at the time of first observed IUD insertion
2. To evaluate whether the risk of uterine perforation among women who had a first observed IUD insertion within different time periods postpartum (i.e., ≤ 6 weeks, > 6 weeks and ≤ 14 weeks, > 14 weeks and ≤ 52 weeks) differs from the risk of uterine perforation among women who had their first observed IUD insertion more than 52 weeks postpartum, including women without a recorded delivery within the past 52 weeks

These primary objectives will include modification by the data source (i.e., interaction terms) if preliminary testing indicates statistically significant interaction between data source and breastfeeding (objective 1) or timing of postpartum insertion (objective 2). For each data source, interaction terms will be included only for the objective(s) with statistically significant interaction(s).

4.3.2 Secondary objective(s)

Rates: uterine perforation

3. To estimate the incidence rate and cumulative incidence of uterine perforation among women using IUDs
4. To estimate the incidence rate and cumulative incidence of uterine perforation among women using IUDs for the following categories:

- ≤ 6 weeks postpartum
- > 6 weeks and ≤ 14 weeks postpartum
- > 14 weeks and ≤ 52 weeks postpartum
- > 52 weeks postpartum, including women without recorded delivery within the past 52 weeks
- ≤ 14 weeks postpartum
- > 14 weeks postpartum, including women without recorded delivery within the past 52 weeks
- ≤ 36 weeks postpartum
- > 36 weeks postpartum, including women without recorded delivery within the past 52 weeks

5. To estimate the incidence rate and cumulative incidence of uterine perforation among women who were and were not breastfeeding at the time of IUD insertion
6. To estimate the incidence rate and cumulative incidence of uterine perforation among women with different types of IUD (i.e., LNG-IUD and copper IUD)
7. To estimate the incidence rate and cumulative incidence of uterine perforation among women with and without menorrhagia (heavy menstrual bleeding) in the 12 months before IUD insertion

Rates: IUD expulsion

8. To estimate the incidence rate and cumulative incidence of IUD expulsion among users of IUDs
9. To estimate the incidence rate and cumulative incidence of IUD expulsion among users of IUDs for the following categories:

- ≤ 6 weeks postpartum
- > 6 weeks and ≤ 14 weeks postpartum
- > 14 weeks and ≤ 52 weeks postpartum
- > 52 weeks postpartum, including women without recorded delivery within the past 52 weeks
- ≤ 14 weeks postpartum
- > 14 weeks postpartum, including women without recorded delivery within the past 52 weeks
- ≤ 36 weeks postpartum
- > 36 weeks postpartum, including women without recorded delivery within the past 52 weeks

10. To estimate the incidence rate and cumulative incidence of IUD expulsion among women who were and were not breastfeeding at the time of IUD insertion
11. To estimate the incidence rate and cumulative incidence of IUD expulsion among women with different types of IUD (i.e., LNG-IUD and copper IUD)
12. To estimate the incidence rate and cumulative incidence of IUD expulsion among women with and without menorrhagia in the 12 months before IUD insertion

Prevalence of difficult IUD insertion

13. To describe the prevalence of indicators of a difficult IUD insertion (e.g., need for cervical dilation or ultrasound guidance, clinician experience) among all users

Comparative: uterine perforation

14. To estimate the adjusted hazard ratio of uterine perforation among women who had a first observed IUD insertion early in the postpartum period (i.e., up to 14 weeks postpartum) versus those who had a first observed IUD insertion late in the postpartum period (i.e., more than 14 weeks postpartum, including women without recorded delivery within the past 52 weeks)
15. To estimate the adjusted hazard ratio of uterine perforation among women who had a first observed IUD insertion ≤ 36 weeks postpartum versus women who had a first observed IUD insertion > 36 weeks postpartum, including women without recorded delivery within the past 52 weeks (this objective will be performed as a sensitivity analysis; same cut point as in EURAS-IUD)
16. To estimate the adjusted hazard ratio of uterine perforation for women whose first observed IUD was a copper IUD versus women whose first observed IUD was an LNG-releasing IUD
17. To estimate the adjusted incidence rate ratio (IRR) and incidence rate difference (IRD) of uterine perforation at 1 year and 5 years of follow-up among women who had a first observed IUD insertion ≤ 36 weeks postpartum versus women who had a first observed IUD

insertion > 36 weeks postpartum, including women without recorded delivery within the past 52 weeks (same analytic approach as EURAS-IUD)

18. To estimate the adjusted hazard ratios of uterine perforation described in objectives 1, 2, and 14-16 across all subsequent insertions (i.e., not the first insertion) observed within the data. (The site-specific analyses will be performed only if there are more than 20,000 subsequent IUD insertions for that site. The pooled analysis will include all sites regardless of the number of subsequent IUD insertions at a site.)
19. To estimate the adjusted hazard ratio of uterine perforation for women using an IUD who have at least one diagnosis code indicating menorrhagia in the 12 months before IUD insertion versus IUD users who do not have this indication (this analysis will be performed only if there are more than 20,000 IUD users with an indication of menorrhagia that can be included in the analysis)

Comparative: IUD expulsion

20. To estimate the adjusted hazard ratio of IUD expulsion among women who were breastfeeding at the time of first observed IUD insertion versus those who were not breastfeeding at the time of first observed IUD insertion
21. To estimate the adjusted hazard ratio of IUD expulsion for women who had a first observed IUD insertion early in the postpartum period (i.e., up to 14 weeks postpartum) versus those who had a first observed IUD insertion late in the postpartum period (i.e., more than 14 weeks postpartum, including women without recorded delivery within the past 52 weeks)
22. To estimate the adjusted hazard ratio of IUD expulsion for women who had a first observed IUD insertion ≤ 36 weeks postpartum versus women who had a first observed IUD insertion > 36 weeks postpartum, including women without recorded delivery in the past 52 weeks
23. To estimate the adjusted hazard ratios of IUD expulsion for women who had a first observed IUD insertion in early postpartum categories versus women who had a first observed IUD insertion late in the postpartum period, using the following strata:

- ≤ 6 weeks postpartum
- > 6 weeks and ≤ 14 weeks postpartum
- > 14 weeks and ≤ 52 weeks postpartum
- > 52 weeks postpartum, including women without recorded delivery in the past 52 weeks (referent category)

24. To estimate the adjusted hazard ratio for IUD expulsion for women whose first observed IUD was an LNG-releasing IUD versus women whose first observed IUD was a copper IUD
25. To estimate the adjusted IRR and IRD of IUD expulsion at 1 year and 5 years of follow-up among women who had a first observed IUD insertion ≤ 36 weeks postpartum versus women who had a first observed IUD insertion > 36 weeks postpartum, including women without recorded delivery within the past 52 weeks
26. To estimate the adjusted hazard ratios of IUD expulsion described in objectives 20-24 across all subsequent insertions (i.e., not the first insertion) observed within the data. (The site-specific analyses will be performed only if there are more than 20,000 subsequent IUD insertions for that site. The pooled analysis will include all sites regardless of the number of subsequent IUD insertions at a site.)

27. To estimate the adjusted hazard ratio of IUD expulsion for women using an IUD who have at least one diagnosis code indicating menorrhagia in the 12 months before IUD insertion versus IUD users who do not have this indication (this analysis will be done only if there are more than 20,000 IUD users with an indication of menorrhagia that can be included in the analysis)

Interactions (effect modification)

28. To evaluate the extent to which breastfeeding status (yes vs. no) modifies the observed association of uterine perforation for women with IUD insertion at different time periods postpartum (i.e., IUD insertion ≤ 14 weeks versus IUD insertion > 14 weeks postpartum) among women with a recorded delivery within the past 52 weeks) at the time of the first observed IUD insertion
29. To evaluate the extent to which type of IUD (LNG-IUD vs. copper IUD) modifies the association between uterine perforation among women who were and were not breastfeeding at the time of first observed IUD insertion
30. To evaluate the extent to which type of IUD (LNG-IUD vs. copper IUD) modifies the association between IUD expulsion among women who were and were not breastfeeding at the time of first observed IUD insertion
31. To evaluate the extent to which type of IUD (LNG-IUD vs. copper IUD) modifies the association of uterine perforation for women with IUD insertion at different time periods postpartum (i.e., ≤ 6 weeks, > 6 and ≤ 14 weeks, > 14 and ≤ 52 weeks) versus IUD insertion more than 52 weeks postpartum, including no recorded delivery within the past 52 weeks at the time of the first observed IUD insertion
32. To evaluate the extent to which type of IUD (LNG-IUD vs. copper IUD) modifies the association of IUD expulsion for women with IUD insertion at different time periods postpartum (i.e., ≤ 6 weeks, > 6 and ≤ 14 weeks, > 14 and ≤ 52 weeks) versus IUD insertion more than 52 weeks postpartum, including no recorded delivery within the past 52 weeks at the time of the first observed IUD insertion

4.4 Study design

A retrospective cohort study design will be used to evaluate uterine perforation and IUD expulsion among women who have an IUD insertion identified within four EHR databases. The study will consider the impact of breastfeeding status at the time of IUD insertion and timing of IUD insertion during the postpartum period on the outcomes of perforation and expulsion.

The study period in each database will be from the time of the availability of electronic data capture in each data source or time of Mirena launch (01 January 2001 to 01 January 2009, depending on database) until about 30 June 2018.*

* This study will be a retrospective analysis of secondary data. This end date for the data cut will be modified based on the timing of FDA approval of the validation study results in order to obtain the maximum data available at that time.

4.5 Population

The study will utilize data from health care systems with EHR data in California, Washington state, and Indiana.

The study population will be women with evidence of insertion of an IUD during the study period who were no more than 50 years of age at the time of the insertion. Only women with available EHR records from 12 months before the day of IUD insertion and onward will be included in this study.

If women had more than one IUD insertion during the study period and were within the age limit at the time of the insertion, each postinsertion period will be considered a separate at-risk period. However, for the primary aims, the population will include only the first IUD insertion observed in the database in order to maintain independence of observations.

4.6 Variables

Intrauterine device insertion will be identified for all enrolled women via available information on IUD, medication, diagnostic, and procedure codes. The type of IUD (i.e., LNG-releasing or copper IUD) will be identified. To obtain the most information about IUD type, the clinician notes will be reviewed if sufficient information is not available from medication, diagnostic, and procedure codes. The date of the IUD insertion will be the index date.

Baseline data, including pregnancies and deliveries, will be collected across all available time in the database before IUD insertion unless otherwise stated for a specific variable, but a minimum of a 12-month look-back period before IUD insertion will be required.

Person-time at risk will be calculated from the time of insertion until the first occurrence of any of the following: uterine perforation, IUD expulsion, IUD removal, indication of IUD reinsertion, indication of new pregnancy, death, hysterectomy, expiration of IUD (e.g., 5 years from date of insertion for Mirena), disenrollment from the database, or end of study period. All person-time at risk that meets these criteria will be included, and there will be no minimum or maximum follow-up time.

4.6.1 Exposure

The postpartum interval (in days) will be calculated as the difference between the date of the most recent delivery and the IUD insertion date. Women with no evidence of delivery in the past year (52 weeks) will be classified as "> 52 weeks postpartum or no evidence of delivery."

Breastfeeding status at the time of IUD insertion will be determined from linked mother/infant records and clinical notes. Breastfeeding status will be classified as yes, no, or undetermined. Breastfeeding status will not be ascertained for women with no evidence of a live birth in the past year (52 weeks) (classified as not breastfeeding).

4.6.2 Outcomes

Uterine perforation and IUD expulsion will be determined from the EHR including physician notes using algorithms developed during the validation study. In both conditions, partial and complete cases will be considered as an outcome, and no distinction will be made between partial and complete cases. Cases will be classified as yes or no. If both perforation and expulsion have occurred on the same date, then the outcome will be classified as both perforation and expulsion, since these outcomes are evaluated separately throughout this study. No analysis will be conducted

to assess both perforation and IUD expulsion as a composite outcome or as a subgroup analysis among those with both outcomes.

4.6.3 Covariates

Demographic and clinical characteristics will be collected at baseline (i.e., before IUD insertion). Information on procedures at IUD insertion that might be indicative of a difficult insertion and provider-related characteristics will also be collected.

4.7 Data sources

Women will be identified in four EHR databases at the time of IUD insertion: Kaiser Permanente Northern California (KPNC), Kaiser Permanente Southern California (KPSC), Regenstrief Institute (RI), and Kaiser Permanente Washington (KPWA) (formerly Group Health Cooperative).

4.8 Study size

Uterine perforation is an uncommon event; 1.3 cases per 1,000 IUD insertions were reported in the EURAS-IUD at 1-year follow-up (Heinemann et al., 2015). The EURAS-IUD also showed that the incidence was higher in breastfeeding women and with early (≤ 36 weeks) postpartum insertion.

Based on results from the validation/feasibility study, the expected number of women with insertion of an IUD available through 30 June 2018, from the four health care systems is approximately N = 225,000, with breastfeeding status available for approximately 60,000 of the total. For the two primary objectives related to the risk of uterine perforation, the power of the study is based on detecting a difference in the estimated risk of uterine perforation with respect to (1) breastfeeding status at the time of IUD insertion and (2) time of IUD insertion after pregnancy delivery. The null hypothesis is no difference in risk of perforation by breastfeeding or time of IUD insertion. There will be no adjustment for multiplicity.

For a two-sided test of the hazard ratio (Schoenfeld, 1983), based on a perforation risk of 1.3 per 1,000 insertions in the 1 year after insertion, a hazard ratio of 2.0 can be detected at the two-sided $\alpha = 0.05$ significance level with the following power:

- 85% for breastfeeding versus not breastfeeding (assuming 60% of women with available breastfeeding status are breastfeeding)
- 73% for ≤ 6 weeks, > 99% for > 6 and ≤ 14 weeks, and 73% for > 14 and ≤ 52 weeks versus > 52 weeks postpartum IUD insertion (assuming a postpartum period allocation of 5:20:5:70)

4.9 Data analysis

Descriptive analyses for all variables of interest will be presented overall and stratified by database.* For categorical variables, frequencies and percentages will be presented for each level. For continuous variables, the mean, standard deviation, minimum, maximum, median, and quartiles will be presented. For estimates, two-sided 95% confidence intervals will be calculated.

* Descriptive and comparative analyses for all variables of interest will be presented overall and stratified by database, with the exception of IUD type. The variable IUD type will be analyzed by database, but only each database holder and RTI-HS will have access to the database-specific information. Data shared with Bayer will only present IUD type aggregated over all data sources. Additionally, analyses for objectives 18 and 26 will be presented by database only if there are more than 20,000 subsequent IUD insertions for the research partner that can be included in the analysis.

Crude incidence rates and cumulative incidence of the outcomes will be estimated for each exposure group (i.e., postpartum IUD insertion time category, breastfeeding status at the time of IUD insertion, type of IUD [LNG vs. copper]).

Crude hazard ratios will be calculated for each outcome without adjustment for covariates within each database. Confounding adjustment will be performed via overlap weighting. Adjusted hazard ratios will be calculated accounting for the propensity score–based weighting approach.

Effect estimates across all data sources will be estimated and presented as the main study results. Effect estimates, except IUD type, will also be presented for each database.

Analyses of the adjusted hazard ratios for the two primary objectives (uterine perforation by breastfeeding status at IUD insertion and uterine perforation by postpartum timing of IUD insertion [in four categories]) combined across all data sources will include hypothesis testing using two-sided statistical tests with a significance level of 0.05 (as requested by the FDA) to test the hypothesis of no difference in the perforation risk between the exposed and referent groups. There will be no adjustment for multiplicity. Statistical tests will be conducted on the assumption that residual confounding can be neglected. However, the validity of this assumption cannot be assessed within the framework of this study.

4.10 Milestones

This study is planned for 78 weeks from the time of FDA approval of this protocol to allow adequate time for development of the statistical analysis plan, data extraction, development of algorithms for additional variables, testing of variables for inclusion in and development of propensity scores, analysis at each site, pooled analysis of results from each site, and writing the study report.

5. Amendments and updates

Number	Date	Section of study protocol	Amendment	Reason
1	29 Jun 2018	Sections 4.3.1 and 8.1	Modified primary objectives such that postpartum exposure has four categories, and assessment of database interactions are included within analysis for primary objectives	Requested by FDA
2	29 Jun 2018	Sections 4.3.2 and 8.2	Modified secondary objectives for interactions to use (1) four categories of postpartum exposure when assessing potential effect modification of IUD type and (2) two categories of postpartum exposure when assessing potential effect modification of breastfeeding and postpartum exposure	Requested by FDA

Number	Date	Section of study protocol	Amendment	Reason
3	29 Jun 2018	Sections 4.3.2 and 8.2	Added objectives to compare outcomes among women with and without documentation of menorrhagia (heavy menstrual bleeding) up to 1 year before IUD insertion	Requested by FDA
4	29 Jun 2018	Section 9.9	Added wording regarding reliability assessment of ICD-10-CM diagnosis codes for outcomes	Requested by FDA
5	29 Jun 2018	Section 9.7.1.3.2 and 9.7.3.3.3	Clarified cohorts of interest and propensity score development for all applicable analyses	Requested by FDA
6	29 Jun 2018	Sections 4.3.2 and 8.2	Removed objectives related to assessments of IUD brand	Due to changes in data sharing expectations between research partners and Bayer

6. Milestones

Table 1. Milestones

Milestone	Planned date
Protocol and validation study results submitted to the FDA	September 2017
Submission of revised protocol and statistical analysis plan	June 2018
Last date of data in data pull	30 June 2018[a]
Study start (anticipated date of FDA approval of protocol)	30 September 2018

[a] Initiation of full study is contingent upon FDA approval of the validation study results and revised protocol.

Note: Final report submission is contingent on study start date. To complete the study, 18 months are anticipated between start date and final report.

7. Rationale and background

Mirena, a levonorgestrel (LNG)-releasing intrauterine system, was approved for use in the US in December 2000 (Bayer HealthCare Pharmaceuticals Inc., 2015). In August 2015, Bayer received a postmarketing requirement (PMR) from the US Food and Drug Administration (FDA) to conduct a "prospective, observational study of incidence and risk factors for uterine perforation among women, particularly when related to breastfeeding and timing of postpartum [LNG-IUS i.e., Mirena] insertion in US women" (communication from FDA to Bayer, 18 August 2015). The concerns described by the FDA, were further clarified in the FDA letter of 18 May 2016 entitled "Information

Request," with respect to pending labeling supplements, which indicated that based on information from the European Active Surveillance Study for Intrauterine Devices (EURAS-IUD), insertion of the intrauterine device (IUD) immediately postpartum might result in a higher risk of uterine perforation and that breastfeeding at the time of IUD insertion might be associated with higher risk of uterine perforation.

EURAS-IUD was a 12-month prospective observational study in six European countries with recruitment between 2006 and 2012 (Heinemann et al., 2015). Two cohorts were included, new users of LNG-releasing IUDs* (n = 43,078) and new users of copper IUDs (n = 18,370). During the 12 months of follow-up, there were 61 uterine perforations in the LNG-IUD group (1.4 per 1,000 insertions; 95% confidence interval [CI], 1.1-1.8) and 20 in the copper IUD group (1.1 per 1,000 insertions; 95% CI, 0.7-1.7). The authors concluded that breastfeeding at the time of IUD insertion was associated with a 6-fold increase in relative risk (RR) of uterine perforation for both groups (RR, 6.1; 95% CI, 3.9-9.6), and there was no difference between the cohorts in this elevated risk of breastfeeding: LNG-IUD (RR, 6.3; 95% CI, 3.8-10.5) and copper IUD (RR, 7.8; 95% CI, 2.8-21.4). There was also an increased risk of uterine perforation among those who had the IUD inserted within 36 weeks after the most recent delivery (Table 2).

Table 2. Perforation incidence and relative risk stratified by breastfeeding status and time since last delivery

Time since last delivery	Incidence[a] of perforation		Relative risk (95% CI) of perforation if breastfeeding
	Breastfeeding		
	Yes	No	
≤ 36 weeks	5.6 (3.9-7.9)	1.7 (0.8-3.1)	3.3 (1.6-6.7)
> 36 weeks	1.6 (0.0-9.1)	0.7 (0.5-1.1)	2.2 (0.3-16.3)
Relative risk (95% CI) of perforation if last delivery ≤ 36 weeks ago	3.4 (0.5-24.8)	2.3 (1.1-4.7)	

CI = confidence interval.
[a] Per 1,000 insertions.
Source: Heinemann et al. (2015).

Uterine perforation risk was higher for patients of clinicians who inserted fewer than 50 IUDs per year. However, there was no association between uterine perforation and other potential confounding variables including cervical dilation for IUD insertion, use of anesthesia for IUD insertion, and prior Cesarean delivery (Heinemann et al., 2015).

In the US, compared with European countries, it is more common to place IUDs immediately postpartum; therefore, the FDA is particularly interested in understanding the risk of uterine perforation in relation to the duration of time from delivery to IUD placement (communication from FDA to Bayer, 18 August 2015). The FDA indicated that breastfeeding practices are also different in the US than in Europe, so it is of interest to understand whether, in the context of US breastfeeding practices, there is an association between breastfeeding status at the time of IUD insertion and higher risk of uterine perforation.

* Note: In this protocol, the term "IUD" will be used to refer to both LNG-releasing intrauterine systems and copper IUDs.

In a submission from Bayer to the FDA (communication from Bayer to FDA, 16 October 2015 [see NDA 21, 255, seq #9]), Bayer proposed that this safety assessment study could be conducted as a retrospective, single-arm cohort study of women with an intrauterine system or copper IUD using EHR databases. Within this letter, Bayer included information on three potential EHR data sources to support this approach. The FDA responded favorably to Bayer's proposal in May 2016 to conduct the Mirena IUD uterine perforation safety assessment PMR study in these EHR databases, but indicated that successful results from a validation/feasibility study should be submitted to support the proposed retrospective study within the EHR database setting (communication from FDA to Bayer, 18 May 2016). Subsequent to Bayer's communication with the FDA (16 October 2015), a validation/feasibility study [see IND 22,697 protocol submission December 2016, seq #0096] was conducted in four EHR data sources (Kaiser Permanente Northern California [KPNC], Kaiser Permanente Southern California [KPSC], Kaiser Permanente Washington [KPWA], and Regenstrief Institute [RI]). During the validation/feasibility study, algorithms were developed and validated to assess uterine perforation and IUD expulsion, and the availability to ascertain data for breastfeeding status was reviewed. In addition, available data on difficulty of IUD insertion and continuous enrollment were assessed.

Based on the results of the validation/feasibility study, the four EHR data sources contain sufficiently reliable exposure and outcome information to address the FDA PMR (communication from FDA to Bayer, 18 August 2015). This protocol details the study design and methodology intended to provide the incidence and risk factors of uterine perforation among women using IUDs, as well as the effects of breastfeeding at the time of IUD insertion and the timing of insertion postpartum on the risk of uterine perforation.

Bayer submitted the protocol for the postmarketing requirement study to the FDA, which was formally accepted for FDA review on 13 December 2017 (PMR #3129-1). On 18 April 2018, the FDA provided comments and recommendations on the protocol that were further clarified in e-mail exchanges between Bayer and the FDA on 23 April and 01 May 2018. This version of the protocol (version 2.0) reflects these communications.

During the time frame for this study, four brands of LNG-releasing IUDs were approved for use in the US: Mirena (approved by the FDA in December 2000 for use up to 5 years before removal or replacement) (Bayer HealthCare Pharmaceuticals Inc., 2015), Liletta (approved by the FDA in February 2015 for use up to 3 years before removal or replacement) (FDA, 2015), Skyla (approved by the FDA in January 2013 for use up to 3 years) (Bayer HealthCare Pharmaceuticals Inc., 2013), and Kyleena (approved by the FDA in September 2016 for use up to 5 years) (Bayer HealthCare Pharmaceuticals Inc., 2016). A copper IUD, ParaGard, has been available in the US since 1988; in 1994, it was approved for use up to 10 years (FEI Women's Health LLC, 2005). All IUDs in use during the study time frame will be included in the PMR study.

8. Research questions and objectives

The overall goal of this study is to assess the impact of breastfeeding and timing of postpartum IUD insertion on uterine perforation and IUD expulsion (evaluated separately) in a population of US women.

In response to the FDA PMR, the study aims to quantify the difference in risk of perforation and expulsion in (1) women who were breastfeeding at the time of IUD insertion versus women who were not breastfeeding at the time of IUD insertion and (2) women who had IUD insertion within different postpartum time periods versus women who had IUD insertion with no recorded delivery

in the past 52 weeks. In addition, the study intends to assess the effect of breastfeeding on the difference in risk of perforation and expulsion associated with the time period of postpartum IUD insertion. The study will also assess whether the risks of uterine perforation and/or IUD expulsion differ by the type of IUD or by whether documentation of menorrhagia appears in the year before IUD insertion, and whether there are interactions with breastfeeding status and postpartum timing of IUD insertion. The study intends to estimate the prevalence of indicators of difficult insertion.

Approximately 10% of IUD insertions in the databases are expected to occur in women who have previous exposure to an IUD within the data. For most analyses, only the first IUD insertion observed in the database will be included in order to maintain independence of observations. Multiple insertions will be considered only for objectives 18 and 26 (see Section 8.2). The objectives listed below are numbered sequentially to enable referencing of specific objectives within the analysis section.

8.1 Primary objective

The primary objectives in this study are as follows:

1. To evaluate whether the risk of uterine perforation among women who were breastfeeding at the time of first observed IUD insertion differs from the risk of uterine perforation among women who were not breastfeeding at the time of first observed IUD insertion
2. To evaluate whether the risk of uterine perforation among women who had a first observed IUD insertion within different time periods postpartum (i.e., ≤ 6 weeks, > 6 weeks and ≤ 14 weeks, > 14 weeks and ≤ 52 weeks) differs from the risk of uterine perforation among women who had their first observed IUD insertion more than 52 weeks postpartum, including women without a recorded delivery within the past 52 weeks

Both primary objectives will include modification by the data source (i.e., interaction terms) if preliminary testing indicates statistically significant interaction between data source and breastfeeding (objective 1) or timing of postpartum insertion (objective 2). Interaction terms for data source will be included only for the objective(s) with statistically significant interaction(s).

8.2 Secondary objective(s)

There are descriptive and comparative secondary objectives in this study. The secondary objectives are grouped by type of analysis (rates, comparative, interaction) and outcome (uterine perforation, IUD expulsion, indicators of difficult insertion).

Rates: uterine perforation

3. To estimate the incidence rate and cumulative incidence of uterine perforation among women using IUDs
4. To estimate the incidence rate and cumulative incidence of uterine perforation among women using IUDs for the following categories:

    - ≤ 6 weeks postpartum
    - > 6 weeks and ≤ 14 weeks postpartum
    - > 14 weeks and ≤ 52 weeks postpartum
    - > 52 weeks postpartum, including women without recorded delivery within the past 52 weeks

- ≤ 14 weeks postpartum
- > 14 weeks postpartum, including women without recorded delivery within the past 52 weeks
- ≤ 36 weeks postpartum
- > 36 weeks postpartum, including women without recorded delivery within the past 52 weeks

5. To estimate the incidence rate and cumulative incidence of uterine perforation among women who were and were not breastfeeding at the time of IUD insertion
6. To estimate the incidence rate and cumulative incidence of uterine perforation among women with different types of IUD (i.e., LNG-IUD and copper IUD)
7. To estimate the incidence rate and cumulative incidence of uterine perforation among women with and without menorrhagia in the 12 months before IUD insertion

Rates: IUD expulsion

8. To estimate the incidence rate and cumulative incidence of IUD expulsion among users of IUDs
9. To estimate the incidence rate and cumulative incidence of IUD expulsion among users of IUDs for the following categories:

    - ≤ 6 weeks postpartum
    - > 6 weeks and ≤ 14 weeks postpartum
    - > 14 weeks and ≤ 52 weeks postpartum
    - > 52 weeks postpartum, including women without recorded delivery within the past 52 weeks
    - ≤ 14 weeks postpartum
    - > 14 weeks postpartum, including women without recorded delivery within the past 52 weeks
    - ≤ 36 weeks postpartum
    - > 36 weeks postpartum, including women without recorded delivery within the past 52 weeks

10. To estimate the incidence rate and cumulative incidence of IUD expulsion among women who were and were not breastfeeding at the time of IUD insertion
11. To estimate the incidence rate and cumulative incidence of IUD expulsion among women with different types of IUD (i.e., LNG-IUD and copper IUD)
12. To estimate the incidence rate and cumulative incidence IUD expulsion among women with and without menorrhagia in the 12 months before IUD insertion

Prevalence of difficult IUD insertion

13. To describe the prevalence of indicators of a difficult IUD insertion (e.g., need for cervical dilation or ultrasound guidance, clinician experience) among all users

Comparative: uterine perforation

14. To estimate the adjusted hazard ratio of uterine perforation among women who had a first observed IUD insertion early in the postpartum period (i.e., up to 14 weeks postpartum) versus those who had a first observed IUD insertion late in the postpartum period (i.e., more than 14 weeks postpartum, including women without recorded delivery within the past 52 weeks)
15. To estimate the adjusted hazard ratio of uterine perforation among women who had a first observed IUD insertion ≤ 36 weeks postpartum versus women who had a first observed IUD insertion > 36 weeks postpartum, including women without recorded delivery within the past 52 weeks (this objective will be performed as a sensitivity analysis; same cut point as in EURAS-IUD)
16. To estimate the adjusted hazard ratio of uterine perforation for women whose first observed IUD was a copper IUD versus women whose first observed IUD was an LNG-releasing IUD
17. To estimate the adjusted IRR and IRD of uterine perforation at 1 year and 5 years of follow-up among women who had a first observed IUD insertion ≤ 36 weeks postpartum versus women who had a first observed IUD insertion > 36 weeks postpartum, including women without recorded delivery within the past 52 weeks (same analytic approach as EURAS-IUD)
18. To estimate the adjusted hazard ratios of uterine perforation described in objectives 1, 2, and 14-16 across all subsequent insertions (i.e., not the first insertion) observed within the data. (The site-specific analyses will be performed only if there are more than 20,000 subsequent IUD insertions for that site. The pooled analysis will include all sites regardless of the number of subsequent IUD insertions at a site.
19. To estimate the adjusted hazard ratio of uterine perforation for women using an IUD who have at least one diagnosis code indicating menorrhagia in the 12 months before IUD insertion versus IUD users who do not have this indication (this analysis will be performed only if there are more than 20,000 IUD users with an indication of menorrhagia that can be included in the analysis)

Comparative: IUD expulsion

20. To estimate the adjusted hazard ratio of IUD expulsion among women who were breastfeeding at the time of first observed IUD insertion versus those who were not breastfeeding at the time of first observed IUD insertion
21. To estimate the adjusted hazard ratio of IUD expulsion for women who had a first observed IUD insertion early in the postpartum period (i.e., up to 14 weeks postpartum) versus those who had a first observed IUD insertion late in the postpartum period (i.e., more than 14 weeks postpartum, including women without recorded delivery within the past 52 weeks)
22. To estimate the adjusted hazard ratio of IUD expulsion for women who had a first observed IUD insertion ≤ 36 weeks postpartum versus women who had a first observed IUD insertion > 36 weeks postpartum, including women without recorded delivery in the past 52 weeks

23. To estimate the adjusted hazard ratios of IUD expulsion for women who had a first observed IUD insertion in early postpartum categories versus women who had a first observed IUD insertion late in the postpartum period, using the following strata:
    - ≤ 6 weeks postpartum
    - > 6 weeks and ≤ 14 weeks postpartum
    - > 14 weeks and ≤ 52 weeks postpartum
    - > 52 weeks postpartum, including women without recorded delivery in the past 52 weeks (referent category)
24. To estimate the adjusted hazard ratio for IUD expulsion for women whose first observed IUD was an LNG-releasing IUD versus women whose first observed IUD was a copper IUD
25. To estimate the adjusted IRR and IRD of IUD expulsion at 1 year and 5 years of follow-up among women who had a first observed IUD insertion ≤ 36 weeks postpartum versus women who had a first observed IUD insertion > 36 weeks postpartum, including women without recorded delivery within the past 52 weeks
26. To estimate the adjusted hazard ratios of IUD expulsion described in objectives 20-24 across all subsequent insertions (i.e., not the first insertion) observed within the data. (The site-specific analyses will be performed only if there are more than 20,000 subsequent IUD insertions for that site. The pooled analysis will include all sites regardless of the number of subsequent IUD insertions at a site.)
27. To estimate the adjusted hazard ratio of IUD expulsion for women using an IUD who have at least one diagnosis code indicating menorrhagia in the 12 months before IUD insertion versus IUD users who do not have this indication (this analysis will be done only if there are more than 20,000 IUD users with an indication of menorrhagia that can be included in the analysis)

Interactions (effect modification)

28. To evaluate the extent to which breastfeeding status (yes vs. no) modifies the association of uterine perforation for women with IUD insertion at different time periods postpartum (i.e., IUD insertion ≤ 14 weeks versus IUD insertion > 14 weeks postpartum) among women with a recorded delivery within the past 52 weeks at the time of the first observed IUD insertion
29. To evaluate the extent to which type of IUD (LNG-IUD vs. copper IUD) modifies the association between uterine perforation among women who were and were not breastfeeding at the time of first observed IUD insertion
30. To evaluate the extent to which type of IUD (LNG-IUD vs. copper IUD) modifies the association between IUD expulsion among women who were and were not breastfeeding at the time of first observed IUD insertion
31. To evaluate the extent to which type of IUD (LNG-IUD vs. copper IUD) modifies the association of uterine perforation for women with IUD insertion at different time periods postpartum (i.e., ≤ 6 weeks, > 6 and ≤ 14 weeks, > 14 and ≤ 52 weeks) versus IUD insertion more than 52 weeks postpartum, including no recorded delivery within the past 52 weeks, at the time of the first observed IUD insertion

32. To evaluate the extent to which type of IUD (LNG-IUD vs. copper IUD) modifies the association of IUD expulsion for women with IUD insertion at different time periods postpartum (i.e., ≤ 6 weeks, > 6 and ≤ 14 weeks, > 14 and ≤ 52 weeks) versus IUD insertion more than 52 weeks postpartum, including no recorded delivery within the past 52 weeks, at the time of the first observed IUD insertion

9. Research methods

9.1 Study design

A retrospective cohort study design will be used to evaluate uterine perforation and IUD expulsion among women with an IUD insertion identified within EHR data. The study will consider the impact of breastfeeding status at the time of IUD insertion and timing of IUD insertion during the postpartum period on the outcomes of perforation and expulsion.

This study will include all women with evidence of an IUD insertion that had at least 12 months of enrollment history preceding IUD insertion. (RI, which does not have enrollment dates, will require a clinical visit at least 12 months before IUD insertion.) The 12-month enrollment before inclusion of IUD insertions will be used to gather baseline data, including data on the exposures, time postpartum, and breastfeeding.

Baseline data—such as patient demographics, patient characteristics (e.g., personal history of gynecologic conditions such as endometriosis), procedure characteristics, medications, and comorbid conditions (e.g., diabetes)—will be collected from all time in the database before the index date (which will be defined as the day of IUD insertion).

Patients will be followed from the time of IUD insertion until the first occurrence of any of the following: uterine perforation, IUD expulsion, IUD removal, indication of IUD reinsertion, indication of new pregnancy, hysterectomy, death, expiration of IUD (e.g., 5 years for Mirena), disenrollment from the database, or end of the study period. All person-time at risk that meets these criteria will be included, and there will be no requirement for minimum or maximum follow-up time. All IUD insertions occurring with at least 12 months of enrollment before the insertion that are noted within the data sources will be included in the study. The index date will be captured for each insertion, and baseline data will be collected for each index date. The main analyses for the study will assess only the first observed IUD insertion for each woman in the database. Secondary analyses will be conducted assessing all subsequent IUD insertions (i.e., after the first observed IUD insertion), as recorded in the database. The sequential number of each insertion as captured in the data for each woman will be collected and included as a baseline covariate within these secondary analyses.

9.2 Setting

This study will be conducted using data from four health care systems with EHRs: KPNC, KPSC, KPWA, and RI. The investigators at these sites (research partners) are working collaboratively to develop a common approach to study design and implementation as outlined in the following sections and detailed within the statistical analysis plan.

The exposure and outcome algorithms at all sites were developed collaboratively to capture the same concepts, but differ in specific terminology to the extent that there are differences across sites. Site investigators affiliated with each data system will be responsible for implementation of the

study protocol at their sites. Results will be summarized in a final study report by RTI Health Solutions (RTI-HS) in collaboration with the site investigators and Bayer AG.

9.2.1 Study time frame

Time windows

The earliest possible start for a patient to be eligible for the study population of women with IUD insertion will be 01 January 2001 (after approval of Mirena), and the latest date for a patient to be included in the study population will be 2 months before the end date of the data pull (anticipated to be about 30 June 2018[*] to coincide with anticipated approval of the protocol). The study start date at each site will be dictated by when EHRs were implemented or the time when Mirena was launched (RI). Further, the start date at each site for inclusion in the breastfeeding assessment will be dictated by the date at which breastfeeding data became available. The end date was chosen to coincide with the expected availability of complete data at the time of the data cut for the analysis. The start and end dates at each site are listed in Table 3.

Table 3. Start dates of EHR data and breastfeeding data and end date of study period, by data source

Site	Start date: EHR data	Start date: breastfeeding data	End date (date of data pull, based on anticipated approval of protocol)
KPWA	01 January 2006	01 January 2006	~30 June 2018
KPNC	01 January 2009	01 January 2009	~30 June 2018
KPSC	01 January 2008	01 January 2010	~30 June 2018
Regenstrief Institute	01 January 2001	01 January 2001	~30 June 2018

EHR = electronic health record; KPWA = Kaiser Permanente Washington; KPNC = Kaiser Permanente Northern California; KPSC = Kaiser Permanente Southern California.

Index date

The index date will be the date of IUD insertion.

9.2.2 Selection criteria

Inclusion criteria for the source population

Each IUD insertion is eligible for inclusion in the study if it meets *all* of the following criteria:

- Evidence in the database of insertion of an IUD (LNG-releasing, copper, or unidentified type) during the study time window for each site (through approximately 30 April 2018).
- Patient enrolled in the database with electronic medical records available for review for at least 12 months before the IUD insertion to ensure identification of any deliveries in the 12 months before IUD insertion and to provide a minimum time for capture of baseline data among IUD users.

[*] This study will be a retrospective analysis of secondary data. This end date for the data cut will be modified based on the timing of FDA approval of the study protocol in order to obtain the maximum data available at that time.

Exclusion criterion for the source population

IUD insertions will be excluded from the study if a patient meets the following criterion *at the time of the IUD insertion*:

- Aged more than 50 years at the time of the IUD insertion (IUD insertions that occur in eligible patients at younger ages will be included)

9.2.3 Study population

Source

This study will be conducted using EHR data from four health care systems: KPNC, KPSC, KPWA, and RI. The source population will be women with evidence, in their medical record, of insertion of an LNG-releasing IUD, a copper IUD, or an unidentified type of IUD during the study period and who were aged less than 50 years at the time of the IUD insertion. Only those with electronic medical records available for review beginning 12 months before the day of IUD insertion will be included in this study.

If women had more than one IUD insertion during the study period and were within the age limit at the time of the insertions, only the first observed IUD insertion will be included in the primary analyses. In a secondary analysis, each postinsertion period will be considered a separate at-risk period. Based on data from the validation study, this is anticipated to occur in 6% to 15% of the study population.

Sampling strategy

The first observed IUD insertion for each woman that meets study inclusion/exclusion criteria will be included in the primary analyses. All eligible IUD insertions occurring during the study period will be included in the study for secondary analysis.

Study population characteristics

The source population will include women in the US. Three of the data sources include individuals with health maintenance organization insurance coverage on the west coast of the US (Washington state, northern Idaho, and northern and southern California) and are ethnically diverse. The fourth data source is a health information exchange located in the midwest (Indiana), which includes all patients regardless of health insurance status and has a larger proportion of African Americans than the other sites but a lower proportion of other minorities.

9.3 Variables

9.3.1 Baseline characteristics

Baseline characteristics will be assessed before the index date for each eligible IUD insertion. The look-back time, all available data before the index date (unless otherwise specified for a particular variable), will be used to evaluate patient characteristics and the potential for confounding. Because all patients in the study are required to have at least 12 months of data before the first index date, there will be a minimum of 12 months of data from which to evaluate baseline characteristic values. For some patients, more information will be available, and all information within the database (see time frame for each database in Table 3) will be considered to reduce misclassification of baseline information (Brunelli et al., 2013).

Demographic characteristics

The following demographic variables will be assessed as potential confounders.

- **Age:** age in years as of the index date.
  - Three categories divided at tertiles (or closest integer cut point [i.e., in years]) for descriptive tables.
  - Continuous variable in propensity score models (including higher order terms age^2 and age^3, if appropriate).
- **Race/ethnicity:** categorical variable with nine categories: non-Hispanic white, Hispanic white, non-Hispanic black, Hispanic black, other Hispanic, Asian/Pacific Islander, multiple races/ethnicities, other race/ethnicity, unknown
- **Smoking status:** indicator variable (0 = no recent smoking, 1 = recent smoking [active smoker within 365 days before index date]) for smoking status as of the index date, where available (i.e., KPSC; partial KPWA, RI).
- **Calendar year of the index date**: year of IUD insertion.
- **Month of the index date:** 12-level categorical variable corresponding to month of the IUD insertion.
- **Duration of the look-back period at the index date:** continuous variable with a minimum of 365 days may be categorized after examining frequency distribution.

Clinical characteristics

- **Body mass index (BMI):** continuous variable assessed at the index date or the closest date before or after the index date.
  - If BMI is not recorded within the EHR, then weight and height closest to the index date (before or after the index date) will be used to calculate BMI (weight [in kg] divided by height [in meters] squared, kg/m^2).
- **Dysmenorrhea:** four-level categorical variable for whether the patient was diagnosed with dysmenorrhea
  - Diagnosed in the year prior to the index date, but not diagnosed before that time
  - Not diagnosed in the year prior to the index date, but was diagnosed before that time
  - Diagnosis recorded both within year prior to the index date and before that time
  - No diagnosis of dysmenorrhea within data
- **Fibroids:** indicator variable (0 = No, 1 = Yes) for whether the patient was ever diagnosed with or reported diagnosis of uterine fibroids prior to the index date.
- **Parity:** cumulative number of viable pregnancies (i.e., carried to at least 20 weeks gestation) prior to the index date..

The following baseline characteristics will be captured only among women who had at least one delivery prior to the index date:

- **Cesarean delivery** will be captured in two variables:
  - Indicator variable (0 = No, 1 = Yes) for whether the patient ever had a Cesarean delivery prior to the index date.
  - Indicator variable (0 = No, 1 = Yes) for whether the patient had a Cesarean delivery for the most recent delivery that is within 52 weeks prior to the index date.

Procedure-related characteristics

- **Concomitant gynecological procedure:** indicator variable (0 = No, 1 = Yes) for whether the IUD insertion was performed during the same visit as another gynecological procedure or surgery.
  - The following will be considered concomitant gynecological procedures: abortion, aspiration and curettage, dilation and curettage, excision/biopsy of cervix or uterus, ablation, colposcopy, hysteroscopy, laminaria, laparoscopy, lysis adhesions, myomectomy, nerve procedure, salpingectomy/oophorectomy. If insufficient data are available to assess (RI only), then concomitant gynecological procedures will be missing.
- **IUD insertion count:** count of the number of IUD insertions for this woman, including the current insertion, that was identified before or on the index date within the data source.
- **Initial IUD insertion**: indicator variable (0 = No, 1= Yes) for first insertion seen within data.

Indicators of a difficult IUD insertion

- **Difficult insertion:** indicator variable (0 = No, 1 = Yes) for whether any of the following occurred on the index date (or in the 7 days before the index date for misoprostol): *cervical dilation, ultrasound guidance*, *paracervical block*, *provider note*, *use of misoprostol*
- **Cervical dilation** will be identified on the day of the IUD insertion and will be classified as yes or no. If there is no information in the record that cervical dilation was done, then the classification will be "no."
- **Ultrasound guidance** for the placement of the IUD will be identified on the day of the IUD insertion and will be classified as yes or no. If there is no information in the record that ultrasound guidance was used, then the classification will be "no."
- **Paracervical block:** indicator variable (0 = No, 1 = Yes) for whether the patient received a paracervical block during the IUD insertion procedure. If there is no information in the record that a paracervical block was used, then the classification will be "no."
- **Provider note indicating a difficult insertion or complicated procedure:** indicator variable (0 = No, 1 = Yes) for whether the patient record includes a notation from the provider regarding a difficult insertion or complicated procedure. If there is no notation of this in the record, then the classification will be "no."
- **Use of misoprostol:** indicator variable (0 = No, 1 = Yes) for whether the patient received misoprostol during the 7 days before the IUD insertion procedure.

Provider-related characteristics, where available (i.e., KPNC, KPSC, KPWA)

- **Provider number of IUD insertions in the previous year**: number of IUD insertions the provider performed in the previous year.
- **Provider annualized number of IUD insertions in the previous year:** provider number of IUD insertions in previous year divided by the number of months provider was employed by health care system represented in data source.
- **Categorical indicator for number of IUD insertions in the previous year:** 0 = fewer than 50 IUD insertions in the previous year, 1 = 50 or more IUD insertions in the previous year.
- **Provider length of employment in the previous year:** continuous variable of the number of days employed within the health care system in the past year.

9.3.2 Exposure

- **Pregnancy delivery date** is the date on which delivery occurred.
- The **days postpartum** will be calculated as the difference between the date of the most recent delivery and the IUD insertion date and will be expressed in days. Data for days postpartum will not be captured for women with no evidence of delivery in the past year (52 weeks).
- **Postpartum status** will consist of three variables: two dichotomous variables and a four-level categorical variable. In all of these variables, women with no evidence of delivery in the past year (52 weeks) will be classified as "no delivery in the past 52 weeks."
    - **Postpartum status categories for primary objective** will consist of the following four categories (as proposed by the FDA).
        - ≤ 6 weeks postpartum
        - > 6 weeks and ≤ 14 weeks postpartum
        - > 14 weeks and ≤ 52 weeks postpartum
        - > 52 weeks postpartum (including women without recorded delivery in the past 52 weeks)
    - **Postpartum status at IUD insertion for sensitivity analysis** will be a dichotomous comparison of ≤ 14 weeks postpartum versus > 14 weeks postpartum (including women with no evidence of delivery in the past year [52 weeks]).
    - **Postpartum status similar to EURAS-IUD** will be a dichotomous comparison of IUD insertions occurring ≤ 36 weeks postpartum versus > 36 weeks postpartum (including women with no evidence of delivery in the past year [52 weeks]).
- **Breastfeeding status** at the time of IUD insertion: any evidence of breastfeeding (i.e., any breastfeeding or pumping across a 24-hour period) at the time of IUD insertion will be determined based on linked mother/infant records (e.g., well-child visits, infant check-ups, and immunization visits) and clinical notes for the woman and infant. Breastfeeding status will be classified as yes (last breastfeeding date within 14 days before IUD insertion or after IUD insertion), no (last breastfeeding date more than 14 days before IUD insertion; first non-breastfeeding date before IUD insertion), or undetermined. Breastfeeding status will not be ascertained for women with no evidence of a live birth in the past year (52 weeks) and those women will be classified as not breastfeeding.

- **IUD type**: three-level categorical variable of inserted IUD type:
  - LNG-IUD: Mirena, Liletta, Skyla, Kyleena
  - Copper IUD: ParaGard, other copper
  - Unknown IUD type
- **Menorrhagia: diagnosis of menorrhagia will be assessed in two variables**
  - As an exposure: indicator variable (0 = No, 1 = Yes) for whether the patient was diagnosed with menorrhagia in the year (365 days) prior to the index date
  - As a covariate within other exposure assessments: four-level categorical variable for whether the patient was diagnosed with menorrhagia
    - Diagnosed in the year prior to the index date, but not diagnosed before that time
    - Not diagnosed in the year prior to the index date, but was diagnosed before that time
    - Diagnosis recorded both within the year prior to the index date and before that time
    - No diagnosis of menorrhagia within data

9.3.3 Outcome measures

- **Person-time at risk** will be calculated from the IUD insertion date until the first occurrence of any of the following: uterine perforation, IUD expulsion, or censoring date.
- **Date uterine perforation confirmed** is the date on which uterine perforation is documented. This may be complete perforation, with IUD migration into the pelvis or abdominal cavity, or partial perforation (i.e., incomplete, with IUD embedded in the myometrium). Cases of both partial and complete perforation will be considered under the umbrella term "perforation." Cases will be classified as yes, no, or undetermined perforation.
- **Date IUD expulsion confirmed** is the date on which IUD expulsion is documented. IUD expulsion, which is the unintended, spontaneous expulsion of the IUD, will be determined from the EHR, including clinical notes, using algorithms developed during the validation study. Both partial and complete expulsions will be considered under the umbrella term "expulsion." Cases will be classified as yes, no, or undetermined expulsion.

If both perforation and expulsion (e.g., complete or partial perforation of the *vagina or cervix* by the IUD) occurred and were documented on the same date, then the outcome will be classified as both perforation and IUD expulsion, since these outcomes are evaluated separately throughout this study. If both perforation and IUD expulsion occurred for the same IUD insertion but on different dates, then the earlier date will constitute a stopping date for assessment of all objectives. No analysis will be conducted to assess both perforation and IUD expulsion as a composite outcome or as a subgroup analysis among those with both outcomes.

9.3.4 Additional parameters

Start and stop dates

- **IUD insertion date** is the date on which IUD insertion is documented. This is the starting date (index date) for person-time at risk.
- **Beginning date of study period:** the first date EHR data are available from the data source for this study (as listed within Table 3)
- **End date of study period:** the last date on which EHR data are available from the data source for this study (as listed within Table 3)

- **Date of start of enrollment** (KPNC, KPSC, and KPWA only) is the earliest date of enrollment in the database for the woman (will be used to calculate look-back period)
- **Date of first clinical encounter** (RI only) is the earliest in-person visit in the database for the woman (will be used to calculate look-back period)
- **Date of disenrollment** (KPNC, KPSC, and KPWA only) is the date, after the index date, on which the woman is no longer enrolled in an eligible insurance plan (one gap of ≤ 31 days per year will be allowed)
- **Date of last clinical encounter** (RI only) is the last date on which a woman has an in-person encounter that is recorded in the database
- **Censoring date** is the earliest of the following dates: date of removal of IUD, date of IUD reinsertion, date of start of new pregnancy, hysterectomy date, date of bilateral oophorectomy and other types of sterilization, expiration of IUD (5 years after insertion of Mirena, Kyleena, or unknown IUD, 10 years after insertion for ParaGard and other copper IUDs, 3 years after insertion of Skyla and Liletta), death date, date of disenrollment from the database, or date of last clinical encounter in database.

Other parameters

- **Database**: categorical variable of the four data sources included in the study
- **Continuous enrollment**, in days, that each individual is in the data source will be calculated starting on the earliest date of date of enrollment in database or date of first clinical encounter in database and ending on the earliest of date of disenrollment from database, date of last clinical encounter in database, or end date of study period, allowing up to one 31-day gap in enrollment each year to be considered continuously enrolled. The individual's start and end dates in the enrollment files will be used for KPNC, KPSC, and KPWA. For RI, the start date will be designated as the patient's first clinical encounter within the dataset, and the end date will be the patient's last observation within the dataset. There may be multiple continuous enrollment periods for a woman who moves out of then back into the database.
- **Live birth at most recent delivery:** indicator variable (0 = No, 1 = Yes) for whether the patient had pregnancy ending in live birth within the past 52 weeks.

9.4 Data sources

There will be four EHR data sources for this study: KPNC, KPSC, KPWA, and RI. Data in different files within each data source will be linked by the patient's identification number. Descriptions of each data source follow (Sections 9.4.1 through 9.4.4).

9.4.1 Kaiser Permanente Northern California

The KPNC region in California extends from Santa Rosa and Sacramento in the north, to Modesto in the east, and south to San Jose and Fresno and includes the entire San Francisco Bay Area. It covers 21 hospitals and 238 medical offices. KPNC covers approximately 4 million patients, representing half of the commercially insured patients and one quarter of the Medicare patients in the area.

Data for KPNC are housed within a comprehensive EHR system that captures every patient encounter in every department, including hospital, emergency, ambulatory surgical, specialist, and generalist care encounters; clinic visits and telephone encounters; physiological measures; procedures; laboratory and radiology testing; and diagnoses. The comprehensive EHR system was fully implemented in 2009. Standardized research datasets—including enrollment,

sociodemographics, pharmacy, encounters, diagnoses, procedures, vital signs, census, and laboratory results—are maintained for the purposes of research. Data are linked across all databases via a unique identifier. Infant records are maintained and can be linked to the mother's delivery record data.

In the validation study, continuous enrollment in KPNC was measured via enrollment files. Of all IUD insertions in this data source, 67% (more than 100,000) were in women with at least 12 months of continuous enrollment before the date of insertion.

9.4.2 Kaiser Permanente Southern California

KPSC is Kaiser Permanente's largest region, with 4.6 million members who broadly represent the diversity of age, sex, and race/ethnicity in the southern California population. KPSC covers 14 hospitals and over 220 medical offices.

The KPSC EHR system was fully implemented in 2008 and integrates all aspects of care, including pharmacy and laboratory services, appointments, registration, and billing. Standardized research datasets are maintained similar to those in KPNC, including date and site of care, diagnosis codes, procedure codes, vaccinations, prescription medications, vital sign, radiology, clinical reports, and laboratory results, as well as member demographics and enrollment information.

Each KPSC member is assigned a unique medical record number upon joining the health plan. This number is retained for life, irrespective of leaving and rejoining the health plan. This unique number allows for the linkage across all databases (both clinical and administrative). The prenatal database includes data on live births, and infant records can be linked to the mother's data.

In the validation study, continuous enrollment in KPSC was measured via enrollment files. Of all IUD insertions in this data source, 67% (more than 80,000) were in women with at least 12 months of continuous enrollment before the date of insertion.

9.4.3 Kaiser Permanente Washington

Based in Washington state, KPWA (formerly Group Health Cooperative) is a nonprofit health system that currently serves over 650,000 members and provides primary, specialty, home health, and inpatient skilled nursing care. Members reside in 22 counties in Washington and northern Idaho. Approximately 70% of patients receive comprehensive care in KPWA-owned facilities, including 25 primary care medical centers and six specialty medical centers. The remaining 30% receive care from contracted provider networks in geographic areas not served by KPWA medical centers but reimbursed by KPWA.

The EHR system was fully implemented in 2006 and includes datasets on enrollment, encounters, diagnoses, procedures, vital signs, radiology, pathology, laboratory tests, and pharmacy dispensings. Data are linked across all databases via a unique member identifier. The mother-infant database (used to collect breastfeeding data) includes data on women with live births and linked infant records and is currently up to date through 2015.

In the validation study, continuous enrollment in KPWA was measured via enrollment files. Of all IUD insertions in this data source, 64% (more than 15,000) were in women with at least 12 months of continuous enrollment before the date of insertion.

9.4.4 Regenstrief Institute

Regenstrief has research access to the Indiana Health Information Exchange, which serves over 17 million patients and includes clinical data from 103 Indiana hospitals, 41 core hospital systems, 60 community clinics, and the state and local public health departments of Indiana. Data from health care encounters are available for this study since 2001 and are captured in a standardized fashion for inpatient admission/discharge information; outpatient visit information; laboratory values; microbiology, pathology, radiology, and cardiology reports; and clinical notes. Data from the datasets are linked via a unique identifier across institutions.

In the validation study, continuous enrollment was measured via health care encounters. Of all IUD insertions in this data source, 74% (~5,700) were in women with at least one clinical encounter 12 months or more before the date of insertion.

9.5 Study size

Uterine perforation is an uncommon event, with 81 uterine perforations reported among 61,448 women over 12 months after insertion of either an LNG-releasing or copper IUD (1.3 cases per 1,000 IUD insertions) in EURAS-IUD (Heinemann et al., 2015). EURAS-IUD also found that the risk was higher in breastfeeding women (5.3 per 1,000 insertions) than in those not breastfeeding (0.9 per 1,000 insertions) and in women with early (≤ 36 weeks) postpartum insertion (5.6 per 1,000 insertions) than later (> 36 weeks) postpartum insertion (1.6 per 1,000 insertions) (Heinemann, 2013). About 11% of the women in EURAS-IUD were breastfeeding, and approximately 20% had an early postpartum insertion (Heinemann, 2013).

There were approximately 325,000 IUD insertions identified during the time frame of the validation/feasibility study (i.e., end date, 30 September 2015), and approximately 65% of those occurred after at least 12 months within the data source. In approximately 90% of IUD insertions, this was the first observed insertion for each woman. Approximately 30% of women were identified as having an IUD inserted within the first 52 weeks postpartum. Based on these results and assumptions—approximately three additional years of data more than the validation study and allowing for a loss of approximately 15% of the insertions due to missing data and the propensity score trimming process (Section 9.7.1.3.2)—the number of first insertions of an IUD available through June 2018 from the four health care systems is anticipated to be approximately 225,000, with breastfeeding status available for approximately 60,000. For the two primary objectives related to the risk of uterine perforation, the power of the study is calculated for differences in the estimated risk of uterine perforation with respect to (1) breastfeeding status at the time of IUD insertion and (2) time interval of the IUD insertion following delivery. The null hypothesis to be tested for primary objective 1 is that the natural logarithm of the adjusted summary perforation hazard ratio for breastfeeding versus not breastfeeding at the time of IUD insertion is equal to 0. Using information from validation study, the expected allocation ratio of breastfeeding versus not breastfeeding is 60:40 for this objective. Three null hypotheses to be tested for primary objective 2 are that the natural logarithm of the adjusted summary perforation hazard ratio for early (i.e., ≤ 6 weeks postpartum, > 6 and ≤ 14 weeks, > 14 and < 52 weeks) versus later (i.e., > 52 weeks) postpartum IUD insertion is equal to 0. Based on information from the validation study, the expected allocation ratio of the corresponding postpartum period categories is 5:20:5:70 for this objective.

Power calculations for the expected number of IUD insertions were performed using PASS 14 software (NCSS, LLC, Kaysville, Utah) for a two-sided test of the hazard ratio (Schoenfeld, 1983). Table 4 indicates the power to detect various hazard ratios at the two-sided $\alpha = 0.05$ significance

level based on a perforation risk of 1.3 per 1,000 insertions and the percentage of insertions expected for the exposure groups of interest, i.e., breastfeeding versus not breastfeeding and early versus late postpartum insertion for the two cutpoints of interest.

Table 4. Power to detect hazard ratio for uterine perforation based on anticipated number of intrauterine device insertions and exposure group allocation

Exposure groups	Number of insertions expected	Allocation % (exposed: unexposed)	Hazard ratio that can be detected (% power)				
			1.5	1.75	2.0	2.25	2.5
Primary objective, 1: Breastfeeding vs. not breastfeeding	60,000	60:40	42	68	85	94	98
Primary objective, 2: Categories of postpartum insertion timing: ≤ 6 weeks vs. > 52 weeks, including women without a recorded delivery within the past 52 weeks	168,750	5:70	32	54	73	85	92
Primary objective, 2: Categories of postpartum insertion timing: > 6 and ≤ 14 weeks vs. > 52 weeks, including women without a recorded delivery within the past 52 weeks	202,500	20:70	78	97	> 99	> 99	> 99
Primary objective, 2: Categories of postpartum insertion timing: > 14 weeks and ≤ 52 weeks vs. > 52 weeks, including women without a recorded delivery within the past 52 weeks	168,750	5:70	32	54	73	85	92

Table 5 displays the power to detect various hazard ratios at the two-sided $\alpha = 0.05$ significance level for the risk of perforation among LNG-releasing versus copper IUDs based on a perforation risk of 1.3 per 1,000 insertions and the percentage of insertions expected for the exposure groups of interest, i.e., timing of postpartum insertion and breastfeeding versus not breastfeeding.*

* These numbers are provided per the FDA request in the letter of 18 May 2016 to address the request to account for "sufficient number of users of both types of IUDs to allow secondary analyses of the comparative safety of LNG IUSs vs. copper IUDs with regard to early postpartum insertion and breastfeeding status at the time of insertion." These power calculations are not necessarily directly linked to objectives, but rather indicate that adequate data are likely to be available to characterize comparative safety of LNG IUSs and copper IUDs.

Table 5. Power to detect hazard ratio for uterine perforation for LNG versus copper IUDs based on anticipated number of intrauterine device insertions and an 80% (LNG) versus 20% (copper) exposure group allocation

Exposure groups	Number of insertions expected	Hazard ratio that can be detected (% power)				
		1.5	2.0	3.0	4.0	5.0
LNG vs. copper	225,000	79%	> 99%	> 99%	> 99%	> 99%
LNG vs. copper for postpartum insertion ≤ 6 weeks	11,250	9%	18%	39%	56%	69%
LNG vs. copper for postpartum insertion > 6 weeks and ≤ 14 weeks	45,000	24	56	92	99	> 99
LNG vs. copper for postpartum insertion > 14 weeks and ≤ 52 weeks	11,250	9	18	39	56	69
LNG vs. copper for postpartum insertion > 52 weeks	157,500	64	98	> 99	> 99	> 99
LNG vs. copper for breastfeeding at the time of insertion	36,000	20%	47%	85%	97%	> 99%
LNG vs. copper for not breastfeeding at the time of insertion	24,000	15%	34%	69%	87%	95%

LNG = levonorgestrel.

9.6 Data management

This study will use data previously collected in EHRs and other electronic administrative and clinical databases at the four study sites. The data will be deidentified and sent from each site to RTI-HS for analysis.

Data management will be conducted in accordance with standard operating procedures developed for the study and used across all sites. Routine procedures include checking electronic files, maintaining security and data confidentiality, following the statistical analysis plan, and performing quality-control checks of all programs. All analyses, including conversion of the original data to analysis variables at each site, will be performed using SAS software, version 9.3 or higher (SAS Institute, Inc., Cary, North Carolina).

Specifics from each data source are described in Sections 9.6.1 through 9.6.4.

9.6.1 Kaiser Permanente Northern California

The Kaiser Permanente health plan maintains comprehensive electronic administrative and clinical databases that are linked to the individual member through a unique medical record number assigned at enrollment. Medical record numbers are not re-issued after a member leaves the health plan; therefore, linkage is assumed to be 100%.

At KPNC, deployment of the EHR system (called HealthConnect from Epic) began in 2005, with complete deployment across all sites by 2009 (2008 for outpatient and 2009 for inpatient). Data are housed in a Clarity database, which is a relational database residing on a Teradata platform and consisting of thousands of tables that can be linked by various primary keys, such as medical record number, patient identification (ID), encounter ID, medication ID, diagnosis ID, procedure ID.

Teradata mr (structured query language) is used to extract data from these various tables; further data manipulation is done either in SQL (Structured Query Language) or in SAS version 9.3.

In addition, data are also managed within the Division of Research (DOR) Virtual Data Warehouse (VDW). The VDW resides on an Oracle platform and is also available as SAS datasets on a secure UNIX server. The VDW pools together data from various sources to bring together both Epic and pre-Epic data and includes clinical, demographic, enrollment, census, and mortality information.

Data extraction will be from Clarity-based tables, as well as the in-house DOR VDW and archived databases. As inclusion and exclusion criteria are applied to build the cohort, all relevant cohort-defining databases will be saved on the secure DOR servers that are backed up every day by the DOR information technology (IT) department. Every analyst in DOR is assigned a secure space by DOR IT security on several servers that can be accessed only by that analyst. In addition, analysts have shared space on these secure servers that can be accessed only by relevant project members. DOR IT security is responsible for providing the governance, guidance, and tools to protect confidential and nonpublic Kaiser Permanente information. DOR IT partners with the National Compliance Organization, Technology Risk Organization, and The Permanente Medical Group to lay the foundation for operational strategies and programs that meet Kaiser Permanente's security obligations and position DOR to become an industry leader in research information security.

Access to the EHR requires authorization from KPNC IT to conduct medical record review validation of electronically extracted data. Each clinician, DOR programmer analyst, and medical record analyst is required to enter a unique password assigned to them to access the EHR. Access to EHR records expires in 90 days if unused. Reapplication to DOR IT or KPNC IT is required to regain EHR access.

9.6.2 Kaiser Permanente Southern California

The KPSC EHR system (HealthConnect) was fully implemented in 2008, and the back-end database, Clarity, was the primary source of patient encounter data beginning in 2005. The research database team at KPSC extracts Clarity, legacy, and claim data and integrates them with historical data prior to HealthConnect into a comprehensive Research Data Warehouse (RDW). The RDW contains information on all utilizations within the KPSC system, including date and site of care, diagnosis codes, procedure codes, vaccinations, vital signs, prescription medications, radiology, clinical reports, laboratory results, as well as member demographics and enrollment information. The RDW is updated weekly.

KPSC also builds and maintains a VDW based on the RDW to support collaborative studies across various research networks. The RDW and VDW are stored on a secure UNIX server. This server is kept in a secure facility with multiple power sources and backup power provision. All data stored on this server are backed up nightly. Access to the RDW and VDW is limited to authorized programmers and statisticians within the Department of Research and Evaluation.

The research database team at KPSC will manage the study databases and provide the support needed to meet study objectives. The analyst/programmer will perform routine range and consistency checks.

This study will be based on administrative databases and electronic medical records. Procedures mandated by the institutional review board (IRB) and the Health Insurance Portability and Accountability Act (HIPAA) for the protection of confidentiality for patient data and will be carefully followed. The analysis datasets created by KPSC will be stored and archived at KPSC as

per the applicable requirements and retention policies. Computer files associated with this project will be kept in a password-protected environment. If hard copies of the data are generated for the study, they will be stored in locked file cabinets accessible only to the investigators and KPSC study staff. All reports and published results from this study will be limited to statistical compilations of the data that do not identify individual patients. Only aggregate data and summary tables, which will not contain patient-level information, will be reported and shared with the sponsor.

The KPSC principal investigator will be responsible for ensuring that KPSC policies and procedures for confidentiality and security are followed for this project.

9.6.3 Kaiser Permanente Washington

In collaboration with RTI-HS, Bayer, and the other participating research sites, the KPWA study team will identify the study variables of interest (exposure, outcomes, additional covariates) to accomplish the study aims and perform requisite analyses. The project team, led by the analyst and programmer, will develop the programming specifications. To create the study database, the programmer will develop the code to extract data from the existing health plan administrative databases, which access the health plan's EHR Clarity database and the KPWA VDW. No direct contact with health plan members will occur. The programming and creation of raw and analytic data files will be done in SAS version 9.4. It is also anticipated that a data collection form (or forms) for review/validation of outcomes, and possibly breastfeeding status, from the EHR may be developed using standard software such as Access or Redcap.

Data security: All study data will be stored in secure computing locations within KPWA Health Research Institute that are backed up nightly. Prior to any data collection, all study protocols will be submitted for review and will receive approval from the KPWA IRB. Once the project datasets are created, the analytic files will be deidentified at the individual level, and each woman will be identified only by a study ID.

Data quality: The quality of KPWA data are assessed and improved through two mechanisms: dedicated quality-assurance programming and crowdsourcing via the VDW user base. Workgroups are responsible for authoring quality-assurance programs that assess adherence to the VDW data model and identify anomalies in the data. These quality checks range from verifying the existence of variables and assuring that they contain permissible values to more sophisticated analyses requiring clinical or scientific knowledge that compare rates and trends of events across institutions. Due to the now long-term use of the VDW, this crowdsourced quality-assurance approach effectively identifies data anomalies. Site data managers investigate these anomalies and report resolutions in the issue tracker.

The programs that will select the study data from the health plan data sources will be reviewed by the analyst and study team once the programs have been created and again as data become available. As they are created, the datasets will be checked by the programmer for range values, consistency, and completeness.

9.6.4 Regenstrief Institute

Data for this study will come from both structured data (ICD-9,* Current Procedural Terminology, and Healthcare Common Procedure Coding System [HCPCS] codes and National Drug Codes) and

* ICD-9 = *International Classification of Diseases, 9th Revision.*

unstructured data from clinical notes. The natural language processing (NLP) pipeline developed by Regenstrief to analyze unstructured data has tools to pull relevant notes for specified cohorts, techniques to find all related terms/synonyms, an approach to reducing "false-positive" hits through exclusion of negation (e.g., not uterine perforation) and family history, and a validation tool built in to add structured data back to the dataset. Data from the various institutions that contribute to the database are stored separately, but patient records are linkable across all sources via a global medical record number. The data are updated nightly. The data manager has access to the source data once IRB approval is granted. Chart reviewers can see identified data, if necessary.

Regenstrief Institute's personnel have been trained in methods to protect patient confidentiality, and efforts will be made to minimize the risk to patients as data are extracted and analyzed. Regenstrief's secure servers are protected by a firewall, and only deidentified data will be shared by the data analyst with the study team.

9.7 Data analysis

An overview of the data analysis can be found below. General statistical analyses and methodology for this study will be presented first, followed by specific data analyses related to each objective. A detailed description of variable definitions, planned analyses, and display specifications will be included within the statistical analysis plan. Statistical tests will be conducted on the assumption that residual confounding can be neglected. However, the validity of this assumption cannot be assessed within the framework of this study.

Research partners at each site will create a deidentified analytic dataset that will be shared with RTI-HS. Analyses related to primary and secondary objectives will be performed at the coordinating center (RTI-HS) on the patient-level data from all four health care systems. In addition to the pooled results, the results for each objective will be presented separately for each database, with one exception (for IUD type) described below.

9.7.1 General analytic approach

9.7.1.1 Descriptive analyses

Descriptive analysis of each variable will be conducted before other analyses. Each health care system will apply the inclusion and exclusion criteria to its data to obtain the study population.

Descriptive analyses for all variables of interest (Section 9.3, as appropriate) will be presented overall and within each database for the study cohort.* For categorical variables, frequencies and percentages will be presented for each level. Continuous variables will be summarized by the mean, standard deviation, minimum, maximum, median, and quartiles. The proportion of missing data will be captured for each variable.

* Descriptive and comparative analyses for all variables of interest will be presented overall and stratified by database, with the exception of IUD type. The variable IUD type will be analyzed by database, but only each database holder and RTI-HS will have access to the database-specific information. Data shared with Bayer will only present IUD type aggregated over all data sources. Additionally, analyses for objectives 18 and 26 will be presented by database only if there are more than 20,000 subsequent IUD insertions for that research partner which can be included in the analysis.

Study cohort at baseline

Descriptive statistics will be obtained for study cohorts at baseline, including overall and within each of the following exposure groups:

- Breastfeeding status (yes, no, undetermined)
- Postpartum period (using the four postpartum status variables defined in Section 9.3.2)
- IUD type (i.e., LNG or copper)

Outcomes

Characteristics of patients experiencing outcomes will be presented for each study outcome. Characteristics will include frequencies and percentages for each level of each outcome (including not experiencing the outcome) and by demographics and clinical characteristics of patients at the time of IUD insertion.

9.7.1.2 Crude incidence rates and crude cumulative incidence

9.7.1.2.1 Crude incidence rates

While the main study analyses will account for the anticipated underlying change in risk of outcomes across the time women are exposed to inserted IUDs, constant incidence rates will also be calculated across the person-time women contribute to the study. Following characterization of variables, person-time at risk and crude incidence rates of outcomes will be calculated. Crude incidence rates will be assessed rather than incidence proportions, since patients will contribute variable time at risk to the study. Crude incidence rates will be calculated for all study cohorts and within levels of exposure variables.

Crude incidence rates will be calculated as the number of outcomes occurring during the person-time at risk divided by the total person-time at risk (in person-years). Crude incidence rates will be reported as point estimates (number of cases per 1,000 person-years) and 95% CIs.

9.7.1.2.2 Crude cumulative incidence

Crude estimates of the cumulative incidence, defined as number of outcomes occurring up to a timepoint out of the number of IUD insertions, will be estimated using the Kaplan-Meier method. The corresponding curve over time, also known as the failure function (i.e., 1-survival function), will be plotted. Crude cumulative incidence will be estimated and plotted for all study cohorts and within levels of exposure variables.

9.7.1.3 Crude and adjusted hazard ratios

9.7.1.3.1 Crude hazard ratios

In binary comparisons, crude hazard ratios will be estimated for the exposed group (e.g., breastfeeding at the time of IUD insertion) relative to the referent group (e.g., not breastfeeding at the time of IUD insertion) using Cox regression models. In categorical comparisons, crude hazard ratios will be estimated for each exposure group (e.g., IUD insertion ≤ 6 weeks postpartum) relative to the referent group (IUD insertion > 52 weeks postpartum) using Cox regression models. These crude hazard ratios will be calculated for each outcome without adjustment for covariates. All crude hazard ratios will be reported as point estimates with 95% CIs.

The proportional hazards assumption between each exposure and outcome pairing will be assessed using visual examination of hazard functions, log-log survival curves, and goodness-of-fit testing

using Schoenfeld residuals (Kleinbaum and Klein, 2012). For violations of the proportional hazards assumption, time-dependent exposure covariates will be included in crude and adjusted hazard ratio models by fitting interaction terms with continuous or categorical time. Additional details are included in the statistical analysis plan.

9.7.1.3.2 Control for confounding effects

Confounding will be controlled through the use of propensity scores, based on the values of covariates at the time of IUD insertion. Propensity scores estimate the probability that a given patient will be exposed conditional on measured covariates and can serve as a summary confounder variable. Propensity scores can perform better than conventional regression methods when the number of events relative to the number of potential confounders is small, because rather than having to model the events with many variables, which may lead to overfitting of the outcome model, one can instead model the exposure, for which the larger number of exposed people provides sufficient data to accommodate a rich model (Cepeda et al., 2003). This advantage may be important in this study, given the low number of expected events, particularly for uterine perforation, within this study.

Separate propensity score models will be developed for exposure-outcome pairings related to the primary objectives, IUD type, and menorrhagia. Additionally, separate propensity score models will be developed for assessment of first observed IUD insertions and for assessment of subsequent IUD insertions. Thus, 16 propensity score models will be developed to assess exposures: 8 for models including first observed IUD insertions and 8 for models including subsequent IUD insertions (Table 6). One propensity score model will also be developed to assess the interaction between breastfeeding and early versus late postpartum IUD insertion. This will yield a total of 17 propensity score models for this study.

Table 6. Propensity score models for postmarketing requirement study defined by exposure and outcomes of interest

Model number	Exposure (dependent variable of propensity score model)	Outcome (not included in propensity score model)
1 (primary objective)	Breastfeeding status (yes vs. no)	Uterine perforation
2 (primary objective)	Postpartum insertion (4 categories)[a]	Uterine perforation
3	IUD type (LNG vs. copper)	Uterine perforation
4	Menorrhagia (yes vs. no)	Uterine perforation
5	Breastfeeding status	IUD expulsion
6	Postpartum insertion (4 categories)[a]	IUD expulsion
7	IUD type (LNG vs. copper)	IUD expulsion
8	Menorrhagia (yes vs. no)	IUD expulsion
9	Interaction of breastfeeding and early vs. late postpartum (breastfeeding/≤ 14 weeks; breastfeeding/> 14 weeks; no breastfeeding/≤ 14 weeks; no breastfeeding/> 14 weeks [referent])	Uterine perforation

[a] Secondary objectives include dichotomization of “early” and “late” postpartum categories. Separate propensity score models will not be *developed* for these objectives. Rather, the propensity scores calculated with the four-category variable will be used, and the distribution of scores will be collapsed into the categories for each secondary objective. Assessment of distributions of propensity scores will be performed within these collapsed categories.

Propensity scores for dichotomous exposure variables will be estimated by fitting a logistic regression model that incorporates data source (i.e., KPNC, KPSC, KPWA, or RI) and measured potential predictors of exposure as independent variables (all baseline variables in Section 9.3.1 will be considered). The dependent variable in the propensity score model is exposure status (e.g., women breastfeeding at the time of IUD insertion vs. not breastfeeding at the time of IUD insertion).

Propensity scores for the categorical variable (i.e., timing of postpartum insertion) will be estimated by fitting a multinomial logistic regression model that incorporates data source (i.e., KPNC, KPSC, KPWA, or RI) and measured potential predictors of exposure as independent variables (all baseline variables in Section 9.3.1 will be considered). The dependent variable in the propensity score model is exposure category (i.e., ≤ 6 weeks postpartum; > 6 and ≤ 14 weeks postpartum; > 14 and ≤ 52 weeks postpartum; > 52 weeks postpartum [referent]).

Covariates will be assessed for inclusion in propensity score models based on association with the study outcome (Brookhart et al., 2006) and thus will not be outcome blinded. Categorical variables will be assessed for inclusion based on indicator coding of the categories. Continuous variables (including integer count variables) will be assessed for inclusion as continuous, dichotomous, and categorical (i.e., indicator coded) variables, as appropriate. Covariates will be included in the propensity score model if the crude hazard ratio is greater than 1.11 or less than 0.90. Additional confounders will be selected for inclusion within propensity score models if a 10% change in the hazard ratio of the *exposure*-outcome relationship occurs when adjusting for that variable, including a 10% change in any level of a categorical exposure variable.

From the fitted logistic regression models, propensity scores will be estimated for each IUD insertion. The distribution of propensity scores among categories (e.g., breastfeeding at the time of IUD insertion versus not breastfeeding at the time of IUD insertion) will be examined.

The propensity scores will be used to calculate weights for each IUD insertion within each exposure group. The weights will be the "overlap weights" and will use data from the IUD insertions included in the propensity score models after reweighting (Li et al., 2018).[*] To assess whether covariates are balanced across exposure groups after weighting, the distribution of each variable will be compared between categories of the exposure variable, and balance parameters (e.g., standardized differences) (Austin and Stuart, 2015) will be calculated. Pairwise balance parameters (e.g., pairwise standardized differences) will be used for the categorical exposure variable (postpartum timing) in which each category will be compared to the referent group. The balance between exposure groups will be assessed overall and within each data source. If the groups are unbalanced on key covariates

[*] Overlap weighting was chosen due to the versatility of the methodology. While the method has been subject to simulations and has been used in practice, further testing is underway. If concerns about this methodology arise due to articles published while this study is underway, then inverse probability treatment weighting (IPTW) will be used as the weighting method. If this occurs, IUD insertions will be excluded when there are no comparable patients in one group versus the other (i.e., nonoverlapping propensity scores in the upper and lower tails of the propensity score distribution); patients with propensity scores close to that boundary will be excluded from further analysis. Weighing will then be based on the inverse probability of treatment, and extreme weights (e.g., > 10^4) will be excluded (Kurth et al., 2006; Sturmer et al., 2010). This process is called "trimming" and may affect one group more than the other. From the weighted analysis, the average treatment effect within each exposure group will be obtained (Austin, 2013; Austin, 2014). If the exposed and unexposed groups are unbalanced after trimming and application of IPTW, then the logistic regression model will be revised by including interaction terms or higher order terms, and the covariate balance between the two groups will be re-evaluated based on the revised model (Austin, 2011; Rosenbaum and Rubin, 1984).

after application of overlap weighting, then the logistic regression model will be revised by including interaction terms (e.g., with data source), higher order terms, or transformation of variables, and the covariate balance between the groups overall and within each data source will be re-evaluated based on the revised model (Austin, 2011; Rosenbaum and Rubin, 1984). When satisfactory balance between the exposed and unexposed groups is achieved (e.g., standardized difference < 0.1), the weighting will be incorporated in modeling for confounder-adjusted outcome assessments (Section 9.7.1.3.3). If satisfactory balance is difficult to achieve, especially if there are disparate allocation percentages between groups, matching of patients between groups may be considered.

9.7.1.3.3 Estimation of confounder-adjusted effect measures

The adjusted hazard ratios and 95% CIs for outcomes between exposure groups will be estimated using weighted Cox regression models with effects for exposure status and interaction with site (as appropriate). Time-dependent exposure covariates will be included if violation of the proportional hazard assumption is identified in the unweighted Cox model (described in Section 9.7.1.3.1). Hazard ratios will be adjusted for possible confounding effects using overlap weighting (Section 9.7.1.3.2). If breastfeeding status, postpartum timing, or IUD type are not included within a propensity score (as independent variables when not the dependent variable), then a separate Cox model also including these variables as covariates will be developed. Adjusted hazard ratios will be reported as point estimates with 95% CIs. For any models including time-dependent exposure covariates, separate adjusted hazard ratios will be reported for the estimates of the effect of the exposure over time.

9.7.1.4 Crude and adjusted IRR and IRD

9.7.1.4.1 Crude IRR

Crude IRR will be estimated for the exposed group(s) (e.g., breastfeeding at the time of IUD insertion) relative to the referent group (e.g., not breastfeeding at the time of IUD insertion) from measures obtained in Section 9.7.1.2.1. Crude IRR will be calculated as the crude incidence rate in the exposed divided by the crude incidence rate in the unexposed. These crude IRR will be calculated for each outcome without adjustment for covariates. All crude IRR will be reported as point estimates with 95% CIs.

9.7.1.4.2 Crude IRD

Crude IRD will be estimated for the exposed group(s) (e.g., breastfeeding at the time of IUD insertion) relative to the referent group (e.g., not breastfeeding at the time of IUD insertion) from measures obtained in Section 9.7.1.2.1. Crude IRD will be calculated as the crude incidence rate in the exposed minus the crude incidence rate in the unexposed. These crude IRDs will be calculated for each outcome without adjustment for covariates. All crude IRDs will be reported as point estimates with 95% CIs.

9.7.1.4.3 Adjusted IRR

The IRR will be adjusted for possible confounding effects via weighted estimation of the rates using overlap weights (Section 9.7.1.3.2) derived from the same propensity score models as those developed for adjustment of the hazard ratios. Adjusted IRR will be calculated as the weighted incidence rate in the exposed divided by the weighted incidence rate in the unexposed referent. If breastfeeding status, postpartum timing, or IUD type are not included within a propensity score model (as independent variables when not the dependent variable), then the adjusted IRR will

include adjustment for these variables as strata and will be calculated using the Mantel-Haenszel approach outlined in Rothman et al. (2008). Adjusted IRR will be reported as point estimates with 95% CIs.

9.7.1.4.4 Adjusted IRD

The IRD will be adjusted for possible confounding effects via weighted estimation of the rates using overlap weights (Section 9.7.1.3.2) derived from the same propensity score models as those developed for adjustment of the hazard ratios. Adjusted IRD will be calculated as the weighted incidence rate in the exposed minus the weighted incidence rate in the unexposed referent. If breastfeeding status, postpartum timing, or IUD type are not included within a propensity score model (as independent variables when not the dependent variable), then the adjusted IRD will include adjustment for these variables as strata and will be calculated using the Mantel-Haenszel approach outlined in Rothman et al. (2008). Adjusted IRD will be reported as point estimates with 95% CIs.

9.7.1.5 Missing data

Missing data will be treated as missing, and no imputations will be performed. Where appropriate, variables will include a "missing" category for analyses. Consequently, data analyses will be conducted using all women and all insertions to the extent possible with respect to their observed available data (i.e., the IUD insertion will not be included in an analysis if missing data for any variable in that analysis, except where "missing" is a separate category for the variable), and the percentage of women or insertions with missing data will be provided for key variables of interest.

Counts of missingness will be reported in descriptive analysis of categorical variables, and percentages for the nonmissing categories will be based on the number of nonmissing values. For continuous variables, the number of nonmissing values will be reported, and descriptive summaries will be based on the number of nonmissing values.

9.7.2 Study cohorts of interest

The complete study population will be defined by the study inclusion and exclusion criteria. However, subgroups of the complete study cohort will be used to address certain study objectives. The study cohorts of interest are outlined below.

9.7.2.1 Study cohorts defined by exposure variables

Only women with eligible IUD insertions and nonmissing exposures will be included in analyses reliant on exposure status. Three cohorts of interest are defined based on need for nonmissing data:

- **Complete study population**: all first IUD insertions included in the data[*] based on study inclusion and exclusion criteria outlined in Section 9.2.2. This study cohort will include women with missing data for breastfeeding status, and/or IUD type.[†] The missing data within breastfeeding status and IUD type may be considered a separate group within the analyses.

[*] Objectives 18 and 23 are the only objectives that assess <u>all</u> IUD insertions. All other objectives assess only the first IUD insertion within the data for each woman.

[†] No missing data are anticipated for postpartum/no delivery status or menorrhagia. For both of these variables, lack of documentation associated with the variable would default the status for that IUD insertion to the unexposed or referent category of the variable.

- **Breastfeeding status**: all IUD insertions among women who were less than 52 weeks postpartum at the time of IUD insertion and had either "yes" or "no" breastfeeding status at the time of IUD insertion will be included in this cohort. IUD insertions among women who have undetermined breastfeeding status will not be included.
- **IUD type**: all IUD insertions with a known type of IUD (either LNG-IUD or copper IUD) will be included in this cohort. IUD insertions among women who had an undetermined IUD type will not be included.

9.7.3 Study objective–specific data analysis

9.7.3.1 Analyses of primary objectives

In order to address the primary objectives 1 and 2 in Section 8.2, the primary endpoints in the study are the adjusted hazard ratios for uterine perforation among the following groups of women (Table 7):

1. Women who were breastfeeding at the time of first observed IUD insertion versus those who were not breastfeeding at the time of first observed IUD insertion.
2. Women who had a first observed IUD insertion within different time periods postpartum (i.e., ≤ 6 weeks, > 6 weeks and ≤ 14 weeks, > 14 weeks and ≤ 52 weeks) versus those who had a first observed IUD insertion late in the postpartum period (more than 52 weeks postpartum, including those without recorded delivery in the past 52 weeks).

Table 7. Study cohorts relevant primary objectives

Objective number	Brief description	Complete study population	Breastfeeding status available	IUD type available
1	By breastfeeding status		X	
2	By postpartum category	X		

IUD = intrauterine device.

Adjusted hazard ratios will be developed for each group of women as described in Section 9.7.1.3.3. Additionally, the interaction effect between the data source and the exposure will be assessed after confounding adjustment. The interaction will be assessed by including terms for exposure, database, and the interaction between the database and the exposure in the weighted Cox models. A type 3 group test for the interaction terms will be conducted. If the test is statistically significant ($P < 0.05$), then the interaction terms will be retained in the final model, and the adjusted hazard ratios will be reported for each data source. If the interaction terms are not deemed significant, then the interaction terms will be removed, and the overall adjusted hazard ratios (the main effect) will be reported as the results of the primary analyses.

Two-sided 95% CIs of the adjusted hazard ratios for uterine perforation will be presented, and a two-sided overall test of the null hypothesis that the natural logarithm of the adjusted hazard ratio equals to 0 will be performed for each of the primary objectives.

The null hypothesis to be tested for primary objective 1 is that the natural logarithm of the adjusted perforation hazard ratio for breastfeeding women versus women who were not breastfeeding at the time of first observed IUD insertion is equal to 0 (i.e., hazard ratio is equal to 1). A P value ≤ 0.05 for this test will reject the null hypothesis of no difference in the adjusted hazard ratio among breastfeeding women versus women who were not breastfeeding at the time of first observed IUD

insertion, indicating there is a difference in the risk of uterine perforation among breastfeeding women versus women who were not breastfeeding at the time of first observed IUD insertion; a *P* value > 0.05 would indicate there is insufficient evidence of a difference in the risk of uterine perforation among breastfeeding women versus women who were not breastfeeding at the time of first observed IUD insertion.

The null hypothesis to be tested for primary objective 2 is that the natural logarithm of the adjusted perforation hazard ratio for women with IUD insertion within different time periods postpartum (i.e., ≤ 6 weeks, > 6 weeks and ≤ 14 weeks, > 14 weeks and ≤ 52 weeks) versus late (i.e., > 52 weeks postpartum, or no recorded delivery within the past 52 weeks) postpartum IUD insertion is equal to 0 (i.e., hazard ratios are equal to 1). A *P* value ≤ 0.05 for these tests will reject the null hypothesis of no difference in the adjusted summary hazard ratios for an early category versus later postpartum IUD insertion, indicating there is a difference in the risk of uterine perforation for the early category versus later postpartum IUD insertion; a *P* value > 0.05 would indicate there is insufficient evidence of a difference in the risk of uterine perforation for the early category versus later postpartum IUD insertion. There will be no adjustment for multiplicity.

As a sensitivity analysis, confounding effects will be accounted for by including selected key covariates in the unweighted Cox models. Due to the sparse outcomes, limited covariates will be selected for inclusion based on their association with the study outcome.

In addition, crude hazard ratios (Section 9.7.1.3.1) will be reported overall and by data source (with the exception of the hazard ratio for IUD type, which will be reported only overall).

9.7.3.2 Analyses of secondary objectives for incidence rates and cumulative incidence

Estimation of crude incidence rates and crude cumulative incidence will be conducted for secondary objectives 3-12 in Section 8.2.

First, crude incidence rates and crude cumulative incidence will be calculated as described in Section 9.7.1.2. These measures will be assessed within relevant cohorts of interest as indicated by "X" in Table 8. The crude cumulative incidence will be plotted overall for each outcome (objectives 3 and 8).

Table 8. Study cohorts relevant to objectives for incidence rates and cumulative incidence

Objective number	Brief description	Complete study population	Breastfeeding status available	IUD type available
3	Overall	X	X	
4	By postpartum category	X		
5	By breastfeeding status		X	
6	By IUD type			X
7	By menorrhagia	X		
8	Overall	X	X	X
9	By postpartum category	X		
10	By breastfeeding status		X	
11	By IUD type			X
12	By menorrhagia	X		

IUD = intrauterine device.

9.7.3.3 Analyses of difficult insertion

The prevalence of indicators of a difficult insertion (objective 13) will be presented via contingency tables (Difficulty = yes, no) including frequencies and percentages of each level of each exposure and outcome variable.

9.7.3.4 Analyses of comparative secondary objectives

9.7.3.4.1 Comparing adjusted hazard ratios for uterine perforation and IUD expulsion among first observed IUD insertions

Estimation of adjusted hazard ratios associated with the first observed IUD insertions will be conducted for secondary objectives 14-16, 19-24, and 27 in Section 8.2 using the same analysis approach as that used for primary objectives.

Crude hazard ratios (Section 9.7.1.3.1) will also be reported. These measures will be assessed within relevant cohorts of interest as indicated by "X" in Table 9.

Table 9. Study cohorts relevant to comparative objectives

Objective number	Brief description	Complete study population	Breastfeeding status available	IUD type available
Adjusted hazard ratio for uterine perforation				
14	By early (< 14 weeks) vs. late (≥ 14 weeks) postpartum	X		
15	By ≤ or > 36 weeks postpartum	X		
16	By IUD type			X
19	By menorrhagia	X		
Adjusted hazard ratio for IUD expulsion				
20	By breastfeeding status		X	
21	By early (≤ 14 weeks) vs. late (> 14 weeks) postpartum	X		
22	By ≤ or > 36 weeks postpartum	X		
23	By postpartum category	X		
24	By IUD type			X
27	By menorrhagia	X		

IUD = intrauterine device.

9.7.3.4.2 Comparing adjusted IRR and adjusted IRD for uterine perforation and IUD expulsion among first observed IUD insertions

Estimation of adjusted IRR and IRD associated with the first observed IUD insertions within the data will be conducted for secondary objectives 17 and 25 in Section 8.2 to assess the change in rate of uterine perforation and IUD expulsion associated with early versus late postpartum IUD insertion (with a cut point at 36 weeks). For these analyses, follow-up data will be truncated at 1 and 5 years to provide an analytic approach similar to that seen in EURAS-IUD.

Crude IRR and IRD will be generated from the crude incidence rates obtained in objectives 4 and 9 and will be developed as described in Sections 9.7.1.4.1 and 9.7.1.4.2. After adjustment via weighting, adjusted IRR and IRD with associated 95% CIs will be calculated for each outcome as described in Sections 9.7.1.4.3 and 9.7.1.4.4. These measures will be assessed within relevant cohorts of interest as indicated by "X" in Table 10.

Table 10. Study cohorts relevant to IRR and IRD in comparative objectives

Objective number	Brief description	Complete study population	Breastfeeding status available	IUD type available
Adjusted hazard ratio for uterine perforation				
17	By ≤ or > 36 weeks postpartum	X		
Adjusted hazard ratio for IUD expulsion				
25	By ≤ or > 36 weeks postpartum	X		

9.7.3.4.3 Comparing adjusted hazard ratios for uterine perforation and IUD expulsion among subsequent IUD insertions

Estimation of adjusted hazard ratios associated with subsequent (i.e., not the first) IUD insertions will be conducted for secondary objectives 18 and 26 in Section 8.2.

These analyses will be conducted similarly to those described in Sections 9.7.3.4.1 and 9.7.3.4.2. Options to account for correlation within women with multiple IUD insertions will be explored for use with models for subsequent insertions to account for the inclusion of multiple IUD insertions per woman. If correlation among insertions within a woman cannot be adequately addressed, then descriptive analyses will be conducted for assessment of subsequent IUD insertions within a woman.

Pooled analysis will include all sites regardless of the number of subsequent IUD insertions at a site. Site-specific analyses will be performed only if there are more than 20,000 subsequent IUD insertions for that site.

9.7.3.4.4 Assessing effect modification

Estimation of effect modification of the adjusted hazard ratios will be conducted for secondary objectives 28-31 in Section 8.2.

The crude and adjusted hazard ratios will be estimated as described in Sections 9.7.1.3.1 and 9.7.1.3.3 within each level of the potential modifier. Cohort(s) of interest for each objective are indicated by "X" in Table 11.

Table 11. Study cohorts relevant to effect modification objectives

Objective number	Brief description	Complete study population	Breastfeeding status available	IUD type available
Uterine perforation outcome				
28	Breastfeeding modifies early (≤ 14 weeks) vs. late (> 14 weeks) postpartum and vice versa		X	
29	IUD type modifies breastfeeding		X	
31	IUD type modifies postpartum categories			X
30	IUD type modifies breastfeeding		X	
32	IUD type modifies postpartum categories			X

IUD = intrauterine device.

For objective 28, the Cox models will include breastfeeding status, early (≤ 14 weeks) versus late (> 14 weeks) postpartum, and their interaction. The *P* value of the type 3 group test for interaction will be reported. For reporting purposes, the group of no breastfeeding and > 14 weeks postpartum will be considered the referent, and hazard ratios will be reported for postpartum period ≤ 14 weeks and breastfeeding, ≤ 14 weeks and no breastfeeding, and > 14 weeks and breastfeeding. The adjusted hazard ratio will be obtained using the weighted Cox model (Section 9.7.1.3.3). One propensity score model (using these four categories) will be developed. Balance on baseline covariates among the four categories in the weighted sample will be assessed.

For objectives 29-32, the Cox models will include the exposure of interest (breastfeeding status or postpartum categories), the outcome of interest (uterine perforation or IUD expulsion), IUD type, and the interaction between exposure and IUD type. The *P* value of the type 3 group test for interaction will be reported. The hazard ratio for exposure of interest will be reported within each level of the IUD type. The adjusted hazard ratio will be obtained using the weighted Cox model (Section 9.7.1.3.3). The weights will be estimated using the same propensity score models developed for the exposure-outcome pairing (Table 6).

9.8 Quality assurance and quality control

Standard operating procedures at RTI-HS will guide the conduct of the study. For data analyses at each site, the standard operating procedures for the site will be used to ensure data quality and security. Specifically, these procedures include internal quality audits, rules for secure and confidential data storage, methods to maintain and archive project documents, quality-control procedures for programming, standards for writing analysis plans, and requirements for senior scientific review. Range checks and general frequency tables will be produced such that missing values, outliers, and inappropriate or abnormal values will be identified. All data will be checked for duplicate records (e.g., two records for one individual, two records for one procedure on same day). A record of data quality problems and resolutions will be kept at each site conducting the data analysis. All inconsistencies and/or data quality issues will be documented. A senior-level data

analyst at each site will review all SAS and data extraction code prior to study completion to ensure that the data extractions and case identifications are accurate and complete.

To ensure consistency across study sites, information on methods and approaches to ascertainment of exposure and outcome information will be shared. All key study documents, such as the statistical analysis plan, abstraction forms, and study reports will undergo quality-control review, senior scientific review, editorial review, and review by all site investigators.

Procedures will be consistent with the FDA's *Best Practices for Conducting and Reporting Pharmacoepidemiologic Safety Studies Using Electronic Healthcare Data* (FDA, 2013) and International Society for Pharmacoepidemiology (ISPE) *Guidelines for Good Pharmacoepidemiology Practices (GPP)* (ISPE, 2015). The European Medicines Agency (EMA) *Guideline on Good Pharmacovigilance Practices (GVP), Module VIII – Post-Authorisation Safety Studies*, echoes this approach (EMA, 2016). At RTI-HS, an independent Office of Quality Assurance will perform audits and assessments of the RTI-HS activities that involve various aspects of the project, including but not limited to education and training documentation and IRB documentation. Such audits will be conducted by the Office of Quality Assurance according to established criteria in standard operating procedures and other applicable procedures.

9.9 Limitations of the research methods

As with any observational study, there will be a potential for unmeasured differences between the treatment groups that affect their risk of outcomes. Utilization of propensity scores and use of overlap weights for the relative comparison and adjusting the hazard ratios for potential confounders will help to reduce this, but there is always the possibility of residual confounding, which would affect calculated point estimates, 95% CIs, and *P* values. An unmeasured confounder would have to be very unbalanced between cohorts to have a large impact on the outcomes. Unmeasured confounding could result in incorrect findings in the comparison of defined cohorts.

As with any health care database study used for secondary data analysis, data are not available prior to the start date of database enrollment for the individual. Thus, data are not available regarding use of an IUD, pregnancy, or baseline covariates prior to enrollment within the database. A minimum 12-month look-back period prior to IUD insertion will be required for inclusion in the study population, but all available time in the database prior to IUD insertion will be used to improve the assessment of potential confounders (Brunelli et al., 2013).

Hypothesis testing is planned for the effects of breastfeeding at the time of IUD insertion and timing of postpartum IUD insertion on the outcome of uterine perforation. No adjustment for multiplicity is planned. Adjustment for multiplicity either recalculates the probabilities or adjusts the interpretation from a statistical test to control against type I error (i.e., false-positive, the statistical test is "significant" when the null hypothesis is true). However, adjustment for multiplicity can increase the type II error (i.e., false-negative, the statistical test is "significant" when the null hypothesis is false). A balance between type I and type II error is particularly important when the research question addresses a safety outcome. In this case, we will not adjust for multiplicity because we do not want to increase the possibility of *finding* no increased risk of uterine perforation *if* there were in fact an actual increased risk of uterine perforation associated with either breastfeeding at the time of IUD insertion and timing of postpartum IUD insertion.

Propensity scores will be used to measure the probability of being "exposed" given specified covariates. The propensity scores will be developed with respect to the outcomes being assessed

within this study and thus is not outcome blinded. This is a variable selection technique that elicits good results for propensity score models (Brookhart et al., 2006).

The results of this study are dependent on accurate capture of data and definitions of variables. Since variables will be determined from diagnosis codes (ICD-9-CM, ICD-10-CM*), Current Procedural Terminology codes, medication codes (National Drug Codes), and clinical notes (i.e., via NLP) there is a possibility of misclassification. Algorithms for the outcome variables, uterine perforation and IUD expulsion, have been validated in these four databases prior to use of ICD-10-CM coding. No formal validation of the algorithms with ICD-10-CM codes to identify uterine perforation or IUD expulsion will be done. However, the rates of these outcomes will be reviewed prior to and after the implementation of ICD-10-CM coding to ensure consistency over time. For variables that have not been validated in these databases, algorithms validated in other data sources (e.g., administrative claims) will be used to identify conditions and medication dispensing, when available. In addition, the study team will develop and share conceptual definitions across data sources to standardize approaches to data capture.

There is potential for underreporting of outcomes within the data sources since women will need to seek treatment in order to have an outcome diagnosed. Asymptomatic perforation or IUD expulsion may not be captured. In addition, there may be a lag time between occurrence of perforation or IUD expulsion and the time that the woman seeks treatment. So there is the possibility of missing outcome occurrence. However, this approach is representative of the way that these outcomes would appear within clinical practice, is similar to EURAS-IUD, and there is no reason to believe that this would be differential by exposure group (e.g., breastfeeding vs. not breastfeeding at the time of IUD insertion). So the absolute incidence rates may not be entirely accurate because of potential underestimation due to underreporting or inaccurate coding; or potential overestimation if the algorithm identifies outcomes (e.g., uterine perforation) that are not true cases. However, the ratios should provide an unbiased estimate of any differential risk between exposure groups. We will also assess rates of uterine perforation and IUD expulsion in the literature to provide an external context for the incidence rates.

There is potential for misclassification and missingness of breastfeeding status at the time of IUD insertion. Data sources obtain information for this variable in different ways, with some utilizing linkage to well-child visits for the infant and others using data from the woman's chart either for postpartum visits or on the date of IUD insertion. The complexity of identifying the information within the charts is somewhat mitigated through the involvement of both clinicians and seasoned data informaticists working within each health care system who have experience with identifying such information. Further, the classification of breastfeeding as yes or no is a crude dichotomy and does not follow the potential biological mechanism for breastfeeding to affect uterine perforation or IUD expulsion. For example, breastfeeding once per day differs from eight times per day (the recommendation from one research partner), and thus combining all breastfeeding into one category combines a heterogeneous experience.

The complexity of the US health care environment and the changes in treatment patterns over calendar time mean that there is potential for differences across data sources to occur due to the

* ICD-9-CM = *International Classification of Diseases, 9th Revision, Clinical Modification;*
ICD-10-CM = *International Classification of Diseases, 10th Revision, Clinical Modification.*

different starting times for each data source. Calendar time-specific analyses can be done if there is any evidence for this.

Most of the data will be from west coast health care systems with just one data source from central US, but there is considerable diversity in factors such as race/ethnicity within each of the data sources. Further, in the validation study, there were no significant differences in the study population characteristics or outcome prevalence across these data sources.

9.10 Other aspects

RTI-HS will serve as a coordinating center for this project and will coordinate the activities of the four research partners and Bayer (e.g., ensure that timelines are being met and facilitate communications). RTI-HS will receive deidentified person-level data from research partners and perform the pooled analyses. Person-level data will not be shared with Bayer. Data on IUD *brand* may be assessed at each site but will not be included in the datasets transferred to RTI-HS for data analysis. Overall data on IUD type, combined across all sites, will be shared with Bayer. Deidentified research partner data on IUD type (not IUD brand) may be shared between the respective research partner and RTI-HS, but will not be provided to Bayer.

RTI-HS will be the primary author for the protocol, statistical analysis plan, and a study report for this PMR study. The study report will be submitted to Bayer, and Bayer will submit the report to the FDA. RTI-HS will solicit and incorporate input from the research partners at each site and the Bayer study team for the protocol, statistical analysis plan, and study report.

10. Protection of human subjects

This is a noninterventional study using secondary data collection and poses only minimal risk for patients (e.g., potential for breach of confidentiality within health plan due to extraction of data from records). All data collected in the study will be deidentified, minimizing risk of breach of confidentiality with regard to personal identifiers or health information.

RTI-HS will obtain approval or exemption from the RTI International IRB before conducting the study. KPNC, KPSC, KPWA, and RI will also obtain approval or exemption from their IRBs before conducting the study. Data protection and privacy regulations will be observed in all aspects of data storage and analysis.

11. Management and reporting of adverse events/adverse reactions

This research study will use only data that have already been collected at the time the research was performed (i.e., secondary data analysis). The outcomes of uterine perforation and IUD expulsion will be reported in aggregate with the study results. Based on current guidelines from ISPE (ISPE, 2015) and the EMA (EMA, 2014), noninterventional studies such as the one described in this protocol, conducted using medical chart reviews or electronic claims and health care records, do not require reporting of adverse events or reactions.

12. Plans for disseminating and communicating study results

In Section V of *Guidelines for Good Pharmacoepidemiology Practices (GPP)*, the ISPE contends that “there is an ethical obligation to disseminate findings of potential scientific or public health importance” (ISPE, 2015); for example, results pertaining to the safety of a marketed medication; “…the marketing authorization holder should communicate to the Agency and the competent

authorities of the Member States in which the product is authorized the final manuscript of the article within 2 weeks after first acceptance for publication."

Results of this study will be published following guidelines, including those for authorship, established by the International Committee of Medical Journal Editors (ICMJE, 2016). When reporting results of this study, the appropriate Strengthening the Reporting of Observational Studies in Epidemiology checklist will be followed (von Elm et al., 2008). The Consolidated Standards of Reporting Trials (CONSORT) statement (Moher et al., 2001) refers to randomized studies, but provides useful guidance applicable to nonrandomized studies.

The publication committee, which includes representatives from Bayer, investigators of each partner institution, and RTI-HS, will agree upon a publication plan that will result in collaborative and inclusive publication(s) and presentation(s) based on these study results—following publication guidelines of the International Committee of Medical Journal Editors (ICMJE, 2016). Bayer will ensure that the FDA is informed prior to any publication. Communication via appropriate scientific venues, e.g., International Society for Pharmacoepidemiology, will be considered.

13. List of references

Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. Multivariate Behav Res. 2011 May;46(3):399-424.

Austin PC. The performance of different propensity score methods for estimating marginal hazard ratios. Stat Med. 2013 Jul 20;32(16):2837-49.

Austin PC. The use of propensity score methods with survival or time-to-event outcomes: reporting measures of effect similar to those used in randomized experiments. Stat Med. 2014 Mar 30;33(7):1242-58.

Austin PC, Stuart EA. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. Stat Med. 2015 Dec 10;34(28):3661-79.

Bayer HealthCare Pharmaceuticals Inc. Kyleena (levonorgestrel-releasing intrauterine system) [prescribing information]. 2016. Available at: https://www.accessdata.fda.gov/drugsatfda_docs/label/2016/208224s000lbl.pdf. Accessed 22 May 2017.

Bayer HealthCare Pharmaceuticals Inc. Mirena (levonorgestrel-releasing intrauterine system) [prescribing information]. October 2015. Available at: http://www.accessdata.fda.gov/drugsatfda_docs/label/2015/021225s031lbl.pdf. Accessed October 14, 2016.

Bayer HealthCare Pharmaceuticals Inc. Skyla (levonorgestrel-releasing intrauterine system) [prescribing information]. September 2013. Available at: http://www.accessdata.fda.gov/drugsatfda_docs/label/2013/203159s002lbledt1.pdf. Accessed October 14, 2016.

Brookhart MA, Schneeweiss S, Rothman KJ, Glynn RJ, Avorn J, Sturmer T. Variable selection for propensity score models. Am J Epidemiol. 2006 Jun 15;163(12):1149-56.

Brunelli SM, Gagne JJ, Huybrechts KF, Wang SV, Patrick AR, Rothman KJ, et al. Estimation using all available covariate information versus a fixed look-back window for dichotomous covariates. Pharmacoepidemiol Drug Saf. 2013 May;22(5):542-50.

Cepeda MS, Boston R, Farrar JT, Strom BL. Comparison of logistic regression versus propensity score when the number of events is low and there are multiple confounders. Am J Epidemiol. 2003 Aug 1;158(3):280-7.

EMA. Guideline on good pharmacovigilance practices (GVP). Module VI – Management and reporting of adverse reactions to medicinal products. European Medicines Agency; 8 September 2014. Available at: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2014/09/WC500172402.pdf. Accessed 06 April 2015.

EMA. Guideline on good pharmacovigilance practices (GVP). Module VIII – Post-authorisation safety studies (Rev 2). European Medicines Agency; 04 August 2016. Available at:

http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2012/06/WC500129137.pdf. Accessed 06 August 2016.

FDA. Guidance for industry and FDA staff. Best practices for conducting and reporting pharmacoepidemiologic safety studies using electronic healthcare data. US Department of Health and Human Services, Food and Drug Administration; May 2013. Available at: http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM243537.pdf. Accessed 05 January 2015.

FDA. Liletta (levonorgestrel-releasing intrauterine system): FDA approval package (application number 206229Orig1s000). Center for Drug Evaluation and Research, United States Food and Drug Administration; 2015. Available at: http://www.accessdata.fda.gov/drugsatfda_docs/nda/2015/206229Orig1s000Approv.pdf. Accessed October 14, 2016.

FEI Women's Health LLC. ParaGard T 380A intrauterine copper contraceptive [prescribing information]. September 2005. Available at: http://www.accessdata.fda.gov/drugsatfda_docs/label/2005/018680s060lbl.pdf. Accessed October 14, 2016.

Heinemann K. European Active Surveillance Study for Intrauterine Devices (EURAS-IUD). Final study report. Data on file. 2013.

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 Apr;91(4):274-9.

ICMJE. Recommendations for the conduct, reporting, editing, and publication of scholarly work in medical journals. International Committee of Medical Journal Editors; December 2016. Available at: http://www.icmje.org/urm_main.html. Accessed 11 January 2017.

ISPE. Guidelines for good pharmacoepidemiology practices (GPP). Revision 3. International Society for Pharmacoepidemiology; June 2015. Available at: http://www.pharmacoepi.org/resources/guidelines_08027.cfm. Accessed 24 May 2016.

Kleinbaum DG, Klein M. Evaluating the proportional hazards assumption. In: Kleinbaum DG, Klein M, editors. Survival analysis: a self-learning text, 3rd edition. New York: Springer-Verlag; 2012. p. 161-200.

Kurth T, Walker AM, Glynn RJ, Chan KA, Gaziano JM, Berger K, et al. Results of multivariable logistic regression, propensity matching, propensity adjustment, and propensity-based weighting under conditions of nonuniform effect. Am J Epidemiol. 2006 Feb 1;163(3):262-70.

Li F, Morgan KL, Zaslavsky AM. Balancing covariates via propensity score weighting. J Am Stat Assoc. 2018 2018/01/02;113(521):390-400.

Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. Lancet. 2001 Apr 14;357(9263):1191-4.

Rosenbaum P, Rubin D. Reducing bias in observational studies using subclassification on the propensity score. J Am Stat Assoc. 1984;79(387):516-24.

Rothman KJ, Greenland S, Lash TL, editors. Modern epidemiology. 3rd ed. Philadelphia: Lippincott Williams & Wilkins; 2008.

Schoenfeld DA. Sample-size formula for the proportional-hazards regression model. Biometrics. 1983 Jun;39(2):499-503.

Sturmer T, Rothman KJ, Avorn J, Glynn RJ. Treatment effects in the presence of unmeasured confounding: dealing with observations in the tails of the propensity score distribution—a simulation study. Am J Epidemiol. 2010 Oct 1;172(7):843-54.

von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. J Clin Epidemiol. 2008 Apr;61(4):344-9.

Annex 1. List of stand-alone documents

None.