Observational Study Information

Acronym/Title	**Incidence and Trend of Ectopic Pregnancy 2009-2018 - A population-based Study**
Protocol version and date	Version 2.0, 28 October 2020
IMPACT study number	20257
Study type / Study phase	N/A
Medicinal product	Hormonal (Levonorgestrel-releasing intrauterine systems) and non-hormonal intrauterine devices
Comparator / Reference therapy	Oral contraception, Depot medroxyprogesterone Acetate
Study Initiator and Funder	Bayer AG, 13342 Berlin
Research question and objectives	
Country(-ies) of study	United States
Author	PPD PPD PPD PPD PPD USA

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

1 Table of contents 2
2 List of abbreviations 4
3 Responsible parties 5
4 Abstract 6
4.1 Title 6
4.2 Rationale and background 6
4.3 Research question and objectives 6
4.4 Study design 7
4.5 Population 7
4.6 Variables 7
4.7 Data sources 7
4.8 Study size 7
4.9 Data analysis 8
4.10 Milestones 8
5 Amendments 9
6 Milestones 9
7 Rationale and background 10
8 Research questions and objectives 11
8.1 Primary objectives 11
8.2 Secondary objectives 12
9 Research methods 12
9.1 Study design 12
9.2 Setting 13
9.2.1 Study population 13
9.2.2 Study time frame 13
9.2.3 Selection criteria 14
9.2.4 Representativeness 14
9.3 Variables 14
9.3.1 Exposure definition 14
9.3.2 Outcomes definition 16
9.3.3 Covariate definition 16
9.4 Data sources 18
9.4.1 PPD 19
9.4.2 PPD 19
9.5 Study size 20
9.6 Data management 20
9.6.1 PPD 20
9.6.2 PPD 21
9.6.3 Validation of exposure and outcome ascertainment 21
9.7 Data analysis 21

9.7.1 Descriptive analyses 22
9.7.2 Specific Aim 1a. 22
9.7.3 Specific aim 1b. 22
9.7.4 Specific aim 1c. 23
9.7.5 Specific aim 2a. 23
9.7.6 Specific aim 2b. 24
9.7.7 Specific aim 2c. 25
9.7.8 Specific aim 2d ..... **Error! Bookmark not defined.**
9.8 Quality control 25
9.9 Limitations of the research methods 25
9.10 Other aspects 26

10 Protection of human subjects 26

11 Management and reporting of adverse events/adverse reactions 27

12 Plans for disseminating and communicating study results 27

13 References 27

Annex 1: List of stand-alone documents 29

2 List of abbreviations

AE	Adverse Event
BMI	Body Mass Index
CI	Confidence Interval
COC	Combined OCP (estrogen and progestin)
PPD	PPD
CPT	Current Procedural Terminology
DMPA	Depot medroxyprogesterone
EDC	Estimated Date of Confinement
EGA	Estimated Gestational Age
EHR	Electronic Health Record
EMA	European Medicine Agency
EURAS	European Active Surveillance Study for Intrauterine Devices
FDA	Food and Drug Administration
ICD	International Classification of Diseases
HIPPA	Health Insurance Portability and Accountability Act
HMO	Health Maintenance Organization
IRB	Institutional Review Board
ISPE	International Society for Pharmacoepidemiology
IT	Information Technology
IUD	Intrauterine Device
PP	PPD
PPD	PPD
PPD	PPD
PPD	PPD
N/A	Not Applicable
NDC	National Drug Code
NLP	Natural Language Processing
OCP	Oral Contraceptive Pill (all types)
OS	Observational Study
PID	Pelvic Inflammatory Disease
PMR	Post Market Research
POP	Progestin-only OCP
RDW	Research Data Warehouse
R&E	Department of Research and Evaluation (PPD
SAP	Statistical Analysis Plan
STI	Sexually Transmitted Infection
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
VDW	Virtual Data Warehouse

3 Responsible parties

Name	Role	Contact Information
Bayer – Project Sponsor		
PPD	Project Leader - Epidemiology	E-mail: PPD Telephone: PPD Location: Berlin, Germany
PPD	Epidemiologist	E-mail: PPD Telephone: PPD Location: Berlin, Germany
PPD	Study Safety Leader	E-mail: PPD Telephone: PPD Location: Espoo, Finland
PPD	Study Medical Expert	E-mail: PPD Telephone: PPD Location: Basel, Switzerland
PPD	Statistician	E-mail: PPD Telephone: PPD Location: Wuppertal, Germany
PPD PPD **Principal data source**		
PPD	Principal Investigator, Clinician Scientist	E-mail: PPD Telephone: PPD Location: PPD USA
PPD	Co-Investigator, Biostatistician	E-mail: PPD Telephone: PPD Location: PPD USA
PPD	Co-Investigator, Biostatistician	E-mail: PPD Telephone: PPD Location: PPD USA
PPD PPD **Research Partner, additional data source**		
PPD	Principal Investigator, Research Investigator PPD	E-mail: PPD Telephone: PPD Location: PPD USA

Name	Role	Contact Information
PPD	Co-Investigator Clinician Scientist	E-mail: PPD Telephone: PPD Location: PPD USA

4 Abstract

4.1 Title

Incidence and Trend of Ectopic Pregnancy 2009-2018 - A population-based Study

4.2 Rationale and background

Incidence, diagnosis, and management of ectopic pregnancy underwent significant increases during the 1980's and '90s, and rates appeared to stabilize from 2000 to 2007.[1-3] There is a consensus that available effective contraceptive methods reduce the absolute risk of ectopic pregnancy by lowering the risk of pregnancy overall. However, in the case of method failure, pregnancies in women using intrauterine devices (IUDs), some progestin-only contraceptives and tubal ligation are more likely to be ectopic than pregnancies in women in the general population and women using combined oral contraceptives or barrier methods.[4-6]. There has been increased use of long acting reversible contraceptives in the last decade, particularly intrauterine devices (IUDs), but the information on the incidence rate of ectopic pregnancy in women using different contraceptive methods is limited over this time. The goal of this project is to assess the feasibility of and generate data using electronic health records on ectopic pregnancy incidence trends and risk factors in women of reproductive age of the two largest regions that PPD serves, PPD and PPD. We propose to conduct a population-based cross-sectional and cohort study to evaluate the incidence rate of ectopic pregnancy over the 10-year period from 2009-2018. We will assess incidence rates in the subpopulation of women with current hormonal and non-hormonal IUDs, combined (COC) and progestin-only (POP) oral contraceptive pills (OCP), and depot medroxyprogesterone (DMPA) use, and assess potential risk factors for ectopic pregnancy.

4.3 Research question and objectives

The overall goal of this study is to assess the incidence rate of ectopic pregnancy over the last decade in a representative population of US women and assess potential risk factors associated with ectopic pregnancy. The research will utilize electronic health records to assess whether ectopic pregnancy incidence rates have changed over the last decade overall as well as in women using prescription contraceptive methods of interest including IUDs, OCPs, and DMPA. The study has the following objectives :

Aim 1: To describe the incidence and temporal trends of ectopic pregnancy during the past decade in women of reproductive age at PPD and PPD and in subpopulations with select prescriptive contraceptive use, in particular hormonal and non-hormonal IUDs.

Aim 2: To describe demographic and clinical risk factors associated with ectopic pregnancy and the temporal trend in the proportion of ectopic pregnancies that are managed surgically vs. medically..

4.4 Study design

To achieve the aims of this project, we propose to conduct a population-based cross-sectional and cohort study of women of reproductive age at PPD and PPD using data abstracted from PPD 's electronic health record (EHR) and regional claims systems and administrative databases.

4.5 Population

The study will take place within health care systems with EHR data in California at PPD and PPD .

The study population will include women enrolled in the health plans of the two systems who were age 15 to 44 years from January 1, 2009 to December 31, 2018.

4.6 Variables

Ectopic pregnancies, defined as extra-uterine pregnancies will be identified using a combination of (ICD-9-CM, ICD-10-CM), CPT codes, medication codes (NDC and PP specific codes).

Person-time at risk will be based on membership enrollment months from the time of eligibility for inclusion (age 15-44 after January 1, 2009) using the membership databases. For sub-aims person-time at risk will be based on evidence of pregnancy (live birth or induced abortion) or evidence of time from start to discontinuation ("end") of IUD, OCP, or DMPA use. Type of IUD (hormonal or non-hormonal) and type of OCP (combined or progestin-only) will be identified.

To assess risk factors for ectopic pregnancy baseline characteristics including demographic variables (age, race/ethnicity, socioeconomic status, and smoking), clinical (reproductive, prior adnexal surgeries, sexually transmitted infections (STIs), endometriosis, infertility, and sterilization) will be assessed for at least one year before study inclusion. Women are required to have at least 12 months of continuous membership (no more than a 30-day gap) prior to the date of inclusion in the study. The look-back time will be all available data to 2008. Clinical characteristics will be assessed before the index date for the analysis of risk factors (Aim 2a) and censoring events which make a woman not at-risk for pregnancy will be identified over the follow-up period.

4.7 Data sources

Women will be identified from EHR databases at PPD and PPD .

4.8 Study size

Ectopic pregnancy is a relatively uncommon event. Van Den Eeden et al reported an annual rate of 1.03 per 1,000 women 15-44 years old from 1997 – 2000 at PPD .[1] Preliminary query of PPD diagnosis and procedure codes for ectopic pregnancy for the time period of 2007-2016 revealed 10,693 episodes of ectopic pregnancy, of which 2,617 were associated with a surgical procedure code. Preliminary query of PPD diagnosis and procedure codes for ectopic pregnancy for the time period of 2007-2016 revealed 10,203 episodes of ectopic pregnancy, of which 3,758 were associated with a surgical procedure code. Given the large cohorts of women of reproductive age at PPD and PPD the study has more than adequate power to detect a difference in linear trend of ectopic pregnancy over time. For the IUD cohort, with a total of 1,725,000 PY, there is 83% power to detect a linear decreasing trend over the 10-year study period from an ectopic pregnancy incidence

rate of 0.71/1000PY to 0.55/1000PY using a two-sided Z test with continuity correction and a significance level of 0.05. Since the IUD cohort is the smallest, we therefore have more than adequate power to detect similar trends for the OCP cohort and the overall cohort of women of reproductive age.

4.9 Data analysis

Descriptive analysis of relevant variables will be conducted initially for PPD and PPD Both will apply the inclusion and exclusion criteria to its data to obtain the study population. 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. 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 index date.

Ectopic pregnancy rates will be calculated based on two denominators: (1) woman-years or the length of enrollment among women aged 15–44 years during the study period, and (2) number of pregnancies in this group of women during the same study period. We will include live births, ectopic pregnancies, and induced abortions in the second denominator. Poisson regression models adjusting for overdispersion, with calendar-year fitted as a continuous variable will be used to evaluate any linear trends in rates over the study period between 2009 and 2018.

Ectopic pregnancy incidence rates and trends will be calculated for women with current hormonal and non-hormonal IUDs, combined and progestin-only OCPs, and DMPA use. Poisson regression models adjusting for overdispersion, with calendar-year fitted as a continuous variable will be used to evaluate any linear trends in rates over the study period between 2009-2018. We will estimate the ectopic pregnancy incidence rates of women using each of the five contraceptive methods of interest and the ectopic pregnancy incidence rate for the overall population of women of reproductive age using incidence rate ratios and 95% confidence intervals.

To analyze risk factors for ectopic pregnancy, total woman-years of hormonal and non-hormonal IUD, combined and progestin-only OCP, DMPA and other hormonal contraceptive exposure will be calculated; periods in which women are not at-risk for pregnancy will be excluded. Poisson regression with robust variance will be used to estimate risk ratios and 95% confidence intervals of ectopic pregnancy across age groups, race/ethnicity groups, contraceptive exposure groups, and covariates while controlling for potential confounders.

4.10 Milestones

The study is planned for 18 months including time for receipt of IRB approval. It is anticipated that it will take 3-4 months to execute the contract and to develop a Data Use Agreement with PPD for an anticipated start date of July 1, 2018. We will submit the IRB application as soon as the contract is executed. Expected time for IRB approval to be granted at PPD is 3 months therefore start of data collection is planned for October 1, 2018. For PPD information on maternal socioeconomic factors is based on data entered on the infant birth certificate collected by Vital Records of the California State; therefore, PPD will obtain California State IRB approval. On average, it takes up to 4 months for California State IRB approval. We plan to provide the statistical analysis plan by November 30, 2018. Validated coding algorithms used to ascertain outcomes (ectopic pregnancy) , exposures (contraceptive use , pregnancies, abortions) and important covariates will be summarized

in standard operating procedures. Once data is abstracted and data checks have been performed we will conduct statistical analyses and produce a preliminary report by December 2019. We will prepare the final report and manuscript draft in the last 3 months.

5 Amendments

28 October 2020 - Post-hoc analyses aimed at 1. Contraceptive identification and validation, 2. Additional analysis to augment Aim 1b and 2a analyses, and 3. Study timeframe extension to include 2019 data

The purpose of the post-hoc analyses is to:

1. Improve the contraception identification algorithm (specifically IUD type identification)
2. Conduct a formal validation of electronic abstraction of contraceptive use patterns
3. Examine incidence of ectopic pregnancy and risk factors for ectopic pregnancy using all observation time for the Aim 1b cohort (i.e. contraceptive exposure time and time of non-use of contraceptives, without censoring observation time at the end of contraceptive use) with non-contraceptive use as a reference category.
4. The study timeframe will also be extended to include 2019 data (limited to these post-hoc analyses).

6 Milestones

Table 1: Milestones

Milestone	**Planned date**
Start of data collection	1 October 2018
End of data collection	30 September 2019
Anticipated contract execution	*1 July 2018*
Anticipated IRB approval	*1 October 2018*
Delivery of the statistical analysis plan	*30 November 2018*
Preliminary report of study results	31 December 2019
Delivery of Study Report	30 March 2020
Delivery of final study report including post-hoc analysis module	July 2021

7 Rationale and background

Ectopic pregnancies account for 2% of all pregnancies and are the leading cause of pregnancy related maternal mortality in the first trimester.[7] Trends in ectopic pregnancy are difficult to examine for the following several reasons. The rate of ectopic pregnancy is relatively low and is usually expressed as the number of cases per reported pregnancies. Data on abortions or pregnancies that end in early miscarriage may not be as complete as pregnancies that result in livebirths. For this reason, rates can also be expressed as the proportion of all women of reproductive age in the population. In addition, women with ectopic pregnancies are increasingly managed as outpatients and are not necessarily included in hospital statistics.[8,9] As inpatient hospital treatment has decreased, the number of health care visits for a single ectopic pregnancy has increased leading to the potential for multiple counting of cases.[10] As a result, surveillance and efforts to monitor national trends is difficult.

In the U.S. the rate of ectopic pregnancy appeared to rise from 1970 to 1989 from 4.5 to 16.0 per 1,000 reported pregnancies.[11] The U.S. Centers for Disease Control and Prevention (CDC) reported that the estimated total number of ectopic pregnancies in 1992 was 108,800 for a rate of 19.7 per 1,000 reported pregnancies.[12] This is the latest year for which nationwide data from the CDC are available. Hoover et al. reported ectopic rates from 2002 to 2013 using data from the Truven Health MarketScan Commercial and Medicaid Claims Database; however, capture of pregnant women for the denominator using claims data was problematic.[13,14] Using computerized data from PPD, Van Den Eeden et al. found that the rate of ectopic pregnancy in 1997-2000 was similar to the national rate at 20.7 per 1,000 reported pregnancies and 1.03 per 1,000 women 15–44 years old.[1] Trabert et al reported ectopic pregnancy rates using computerized data from Group Health Cooperative in Washington State from 1993-2007. The crude ectopic pregnancy rate appeared to slightly decrease from 17.1 per 1,000 pregnancies in 1993-1995 to 13.6 per 1,000 pregnancies in 2005-2007 (p-value for trend test <0.05).[2,15] Hoover et al. had similar findings, reporting stabilization in incidence from 2002 to 2007 using a commercial claims database.[14]

Increases in ectopic pregnancy rates may be explained by a combination of factors including improvement in diagnosis capabilities and increasing rates of chlamydia infection. In contrast, changes in use of effective contraceptive methods should lead to decreases in ectopic pregnancy rates. Because ectopic pregnancies are relatively uncommon events, the vast majority of published studies assessing risk factors for ectopic pregnancies are case-control studies. These studies have generated a conflicting array of results for current contraceptive use because of methodologic differences in the choice of controls.[16] The selection of women with live births as controls is appropriate if the hypothesis does not relate to exposures that prevent pregnancy (e.g., current contraceptive use). In case-control studies where cases with ectopic pregnancies are compared to women with live births as controls, current contraceptive use appears to be a risk factor for ectopic pregnancy; however when cases with ectopic pregnancies are compared to non-pregnant controls, contraceptive use appears protective.[4,17-19] For a number of other possible risk factors for ectopic pregnancy, with case-control studies, whether pregnant or nonpregnant control are selected , it is difficult to minimize bias introduced by factors related to contraceptive practice.

There is a consensus that available effective contraceptive methods reduce the absolute risk of ectopic pregnancy by lowering the risk of pregnancy overall. However, when there is method failure, pregnancies in women using intrauterine devices (IUDs), some progestin-only contraceptives and tubal ligation are more likely to be ectopic than pregnancies in women in the

general population and women using combined oral contraceptives or barrier methods.[4-6] The results of one of the largest prospective cohort studies suggest that current contraceptive users have a lower rate of ectopic pregnancy than non-contraceptive users. In an epidemiological surveillance in a population of approximately 2.7 million women of reproductive age living in Beijing, the incidence was 1.80 per 1,000 married women using no contraceptives compared to 0.54 per 1,000 married women using contraceptives. The rates were lowest for women with female sterilization at 0.18 per 1,000 married women. The rates were 2.43 per 1,000 married women using natural family planning, 0.65 for IUD users, 0.21 for OCP users, and 0.57 condoms/spermicides.[20] The ectopic pregnancy incidence rate is well-documented in women who use IUDs due to the availability of data from randomized clinical trials and prospective cohort studies in these populations.[21-23] In the EURAS study, Heinemann et al demonstrated an ectopic pregnancy incidence of 0.02 per 100 woman-years (95% CI: 0.01–0.03) for the levonorgestrel IUD and 0.08 per 100 woman-years (95% CI: 0.04–0.13) for the copper IUD.[21]

The retrospective cohort studies by Van Den Eden et al. and Trabert et al. are dated and the investigators did not assess current contraceptive use or other risk factors for ectopic pregnancy. In the last decade there has been a change in the mix of contraceptive methods used, with greater emphasis on long acting reversible contraceptives, particularly hormonal IUDs. Comparative data on various methods is lacking; generation of evidence on ectopic pregnancy incidence and trends, and risk factors, including current use of various prescription contraceptives in the last decade is of interest. Electronic health records provide a unique opportunity to conduct pharmacoepidemiologic studies as well as to investigate the natural history of selected disorders such as ectopic pregnancy while accounting for several potential confounding variables. Once algorithms are established for abstracting data and analyzing outcomes, the data can be tracked over time and used as a benchmark for assessing the effectiveness of contemporary IUDs in reducing the ectopic pregnancy rate.

8 Research questions and objectives

The overall goal of this study is to assess the incidence rate of ectopic pregnancy over the last decade in a representative population of US women and assess potential risk factors associated with ectopic pregnancy. The study is designed to address the following research questions:

- What is the incidence rate of ectopic pregnancy among women of reproductive age at PPD and PPD?
- What is the incidence rate of ectopic pregnancy among women with current IUD, OCP, and DMPA use compared to the overall population of women of reproductive age?
- What are the temporal trends in ectopic pregnancy incidence rates over the last decade overall and in women with current prescriptive contraceptive use (hormonal and non-hormonal IUDs, combined OCPs (COC) and progestin-only OCPs (POP), and DMPA)?
- What are the potential risk factors associated with ectopic pregnancy in women of reproductive age at PPD and PPD?
- What are the trends in management of ectopic pregnancy over the last decade?

8.1 Primary objectives

The primary aims of the study are:

1a. To describe the incidence and temporal trends of ectopic pregnancy during the past decade in women of reproductive age at PPD and PPD.

1b. To describe the incidence and temporal trends of ectopic pregnancy in women with current hormonal and non-hormonal IUDs, COCs and POPs, and DMPA use.

1c. To describe the incidence of ectopic pregnancy during current contraceptive use and non-contraceptive use (i.e. including contraceptive exposure time and exposure time of non-use of contraceptives) (Objective 1c is added as post-hoc analysis with amendment dated 28 October 2020).

8.2 Secondary objectives

The secondary aims of the study are:

2a. To describe potential risk factors associated with ectopic pregnancy in women of reproductive age at PPD and PPD, including demographic risk factors (e.g. age, race), current contraceptive use (hormonal and non-hormonal IUDs, COCs and POPs, and DMPA), infectious (e.g. STIs, Pelvic Inflammatory Disease [PID]), and reproductive (e.g. previous ectopic, endometriosis, and infertility diagnosis or treatment).

2b. To describe the temporal trend in the proportion of ectopic pregnancies that are managed surgically vs. medically.

2c. Enhanced identification of contraception type and validation of contraceptive use patterns (Objective 2c is added as post-hoc analysis with amendment dated 28 October 2020).

2d. To describe potential risk factors associated with ectopic pregnancy in women with current contraceptive use and non-contraceptive use including all observation time for the Aim 1a cohort (i.e. all potentially at-risk exposure time will be categorized as contraceptive exposure time or time of non-use of contraceptives) (Objective 2d is added as post-hoc analysis with amendment dated 28 October 2020).

9 Research methods

9.1 Study design

To achieve the aims of this project we propose to conduct a population-based cross-sectional and cohort study of women of reproductive age at PPD and PPD using data abstracted from PPD electronic health record (EHR), regional claims systems, and administrative databases. For the primary study aim, to describe the incidence and temporal trend of ectopic pregnancy, the study will identify women with the outcome of interest - ectopic pregnancy in women of reproductive age enrolled in the health plan from January 1, 2009 to December 31, 2018. For the post-hoc analyses with the amendments dated 28 October 2020 we will extend the eligible enrollment and observation period to December 31, 2019. All person-time at-risk will be counted in the denominator, whether continuous or not, until death, disenrollment from the database, or end of the study period. First and all subsequent ectopic pregnancies in the same women will be counted. We will also describe the incidence and temporal trends in ectopic pregnancy using women with live births and induced abortions as the denominator. To assess incidence rates and trends in women using select prescription contraceptives, we will repeat the analyses with person-time at risk for the women exposed to the methods of interest as the denominator.

To capture person-time at risk of contraceptive use (Aim 1b and 1c) and to assess potential risk factors associated with ectopic pregnancy (Aim 2a and 2d), we will include women with at least 12 months of continuous enrollment before the study inclusion date. The 12-month enrollment period before study inclusion will be used to gather baseline data, including data on the exposures and covariates. Baseline data on clinical characteristics (e.g., history of ectopic pregnancy and surgeries)—will be collected from all time in the database back to January 1, 2008. For Objective 2c (added as post-hoc analysis with amendment dated 28 October 2020), we will also identify diagnostic and procedure codes for IUD use extending back to 1998 for additional information on IUD type for IUDs that were initiated before 2008.

9.2 Setting

PPD and PPD represent the two largest of PPD nine regional entities nationwide. PPD is an integrated health care system with a service area that encompasses the PPD and the PPD from the PPD area in the north to Fresno in the south. PPD provides care to approximately 4.1 million racial/ethnically diverse members – over 30% of the insured population in its service area. PPD operates 21 hospitals and over 200 outpatient clinics and utilizes an EHR based on an EPIC® platform. PPD is the largest PP system, it also has a highly racial/ethnically diverse population of approximately 4.3 million members in PPD and counties throughout PPD. It has 15 hospitals and over 227 outpatient clinics and provides comprehensive health care. PPD also utilizes an EHR based on an EPIC® platform. Both PP entities provide comprehensive care; members receive their care essentially exclusively from PPD physicians and allied staff in the medical centers and medical office buildings owned or operated by the health plan.

9.2.1 Study population

Source and sampling strategy

This study will be conducted using EHR data from PPD and PPD The source population will include all women who were age 15 to 44 years from January 1, 2009 to December 31, 2018 who were enrolled in the health plan for at least one month over the study period. For the post-hoc analyses with the amendments dated 28 October 2020 we will extend the eligible enrollment and observation period to December 31, 2019. PPD had approximately 940,000 women and PPD had approximately 1,020,000 women age 15-44 years enrolled in the health plan in 2017.

9.2.2 Study time frame

Time windows

The earliest possible start for a woman to be eligible for the study population January 1, 2009 and the latest date for a patient to be included in the study population will be December 31, 2018. For the post-hoc analyses with the amendments dated 28 October 2020 we will extend the eligible enrollment and observation period to December 31, 2019. The study start date is based on when EHRs were fully implemented in PPD and PPD and corresponds closely with the latest timeframe for which data on ectopic pregnancy incidence rates are available. The end date was chosen to coincide with an approximately 10-year window. We will conduct preliminary data pulls and data validation in the last quarter of 2018; however, full data abstraction will occur in 2019 which corresponds with the expected availability of complete data and allowance for data cleaning

prior to the time of the data cutoff for the analysis. For the post-hoc analyses data abstraction will be completed in March 2021.

Index date

9.2.3 The index date for all study aims including the post-hoc analyses is the first date of eligible health plan enrollment during the study period. Selection criteria

Inclusion criteria

To estimate incidence of ectopic pregnancy in women of reproductive age (Aim 1a), all women will be included in the study if they were at least age 15 years and not older than 44 years and had at least 1 month of health plan enrollment during the study period. To estimate incidence of ectopic pregnancy in select contraceptive users (Aim 1b) and to assess risk factors for ectopic pregnancy (Aim 2a) women must have at least 12 months of continuous membership (no more than a 30-day gap) prior to the date of inclusion in the study to determine baseline contraceptive use and risk factor information. For the post-hoc analyses Aim 1c and 2d, women must also have at least 12 months of continuous membership (no more than a 30-day gap) prior to the date of inclusion in the study to determine baseline contraceptive use and risk factor information.

Exclusion criteria

For Aim 1a, no exclusions will be applied to mirror methodology of prior studies.[1,2] For Aim 2a. women considered not at-risk for pregnancy will be identified using available information on exclusionary diagnoses or procedures (natural or surgical menopause, ovarian failure, or hysterectomy) at or before respective index. The same exclusion criteria will be applied for the post-hoc analyses in Aim 1c and 2d. Information on sexual activity and orientation is not reliably collected and varies over time; therefore, analyses will be based on person-years of health plan enrollment but not sexual exposure.

9.2.4 Representativeness

PPD and PPD provide health coverage for about 8.5 million patients, representing roughly a third of the commercially insured patients and one quarter of the Medicare patients in the state. They are broadly representative of the population of California with the exception of extremes of income.

9.3 Variables

9.3.1 Exposure definition

For Aim 1a. person-time at risk will be based on membership enrollment months from the time of eligibility for inclusion (age 15-44 after January 1, 2009) using the membership databases. All person-time will be included, whether continuous or not, until death, disenrollment from the database, or end of the study period (December 31, 2018). For the post-hoc analyses with the amendments dated 28 October 2020 we will extend the eligible enrollment and observation period to December 31, 2019. For Aim 1b, 1c, 2a and 2d. pregnancies resulting in live birth will be identified using perinatal databases described in data sources (section 9.4). Induced abortions will be identified using a combination of (ICD-9-CM, ICD-10-CM), Current Procedural Terminology (CPT) codes, and medication codes (National Drug Codes [NDC] and PP specific codes). To estimate the number of pregnancies in a study year, women with a pregnancy that begins and ends in

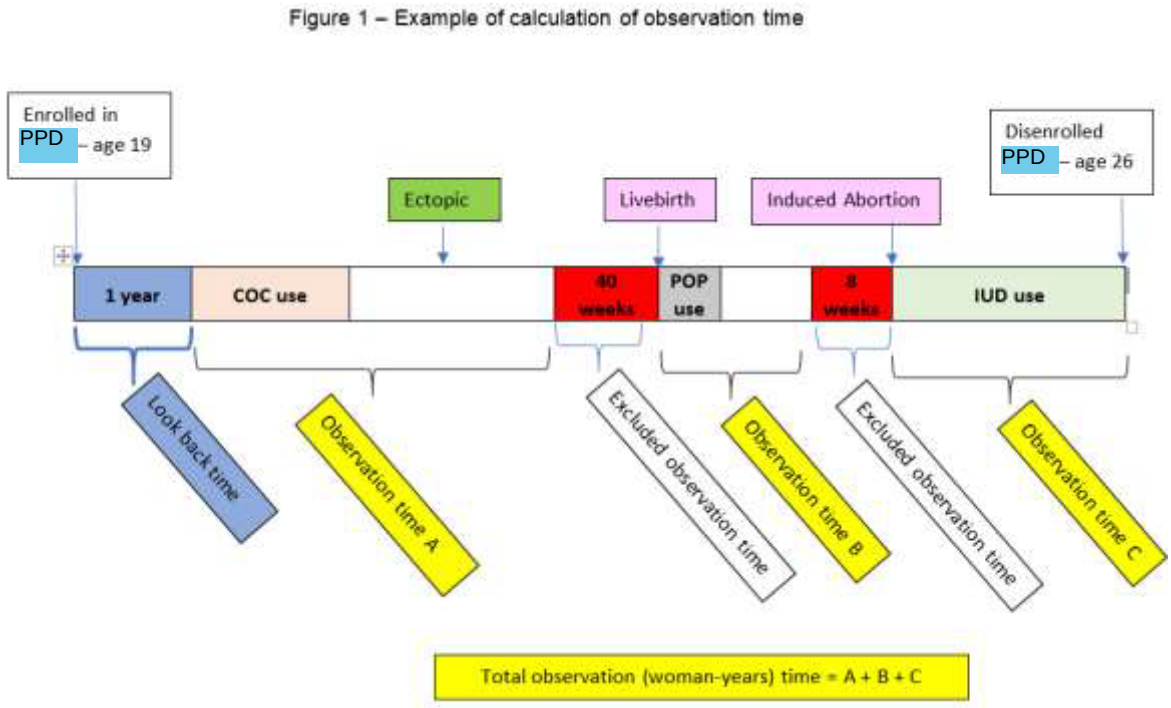
different years will be counted in the year of the delivery. Induced abortions will be counted based on the procedure date. For Aim 2a. women will not contribute person-time during pregnancy resulting in a live birth (time calculated from the date of delivery and estimated gestational age at delivery) or induced abortion (time estimated at 60 days from procedure date).

Contraceptive exposure time (Aim 1b. and 2a) will be based on evidence of time from start to discontinuation ("end") of methods in the EHR databases prior to disenrollment. We will use a combination of relevant (ICD-9-CM, ICD-10-CM), CPT codes, medication codes (NDC and PPD specific) to determine start and end dates. Over time a woman may use different methods and have many episodes of method use (Figure 1). For Aim 2a we will extract data on other hormonal contraceptives including implants, patches and rings, however, these methods will be grouped together as "other methods" and will not be analyzed separately. To capture exposure time:

- IUD start date will be based on evidence of insertion procedure and medication codes. End dates will be based on evidence of removal diagnosis and procedure codes, evidence of insertion of a subsequent IUD or implant, or subsequent DMPA injection or contraceptive prescription, livebirth, or induced abortion. Since women may have an IUD inserted prior to enrollment at PPD or PPD and since IUD removals may not be reliably coded, data on IUD use may also be augmented using surveillance codes and information extracted from clinical notes using NLP algorithms, as needed, specifically developed for this project.
- OCP (COC and POP) exposure time will be based on number of packs dispensed in the pharmacy databases using relevant medication codes. Days covered per pack will be assumed to be 28. All durations of prescriptions for OCPs will be extended by up to 14 days to account for delays in initiating use after the method is dispensed.[24]
- DMPA exposure time will based on injection date. Days covered per injection will be assumed to be 90.
- For Aim 2a. other prescription contraceptive exposure (contraceptive patch and vaginal ring and implants) will be identified. Exposure time for patches and rings will be based on number of patches and rings dispensed. Days covered per box of patches and each ring will be assumed to be 28. Implant start date will be based on evidence of insertion procedure and medication codes. End dates will be based on evidence of removal diagnosis and procedure codes, start of another method, live birth, or induced abortion. Implant surveillance codes will be used to capture use that initiated prior to study inclusion.
- For Objective 1 c and 2d (added as post-hoc analysis with amendment dated 28 October 2020), implant start date will be based on evidence of insertion procedure and medication codes. End dates will be based on evidence of removal diagnosis and procedure codes, evidence of insertion of a subsequent IUD or implant, or subsequent DMPA injection, livebirth, or induced abortion. Since women may have an Implant inserted prior to enrollment at PPD or PPD, data on Implant use may also be augmented using surveillance codes.

We will develop algorithms to adjudicate instances where there are overlapping episodes of different contraceptive methods (i.e. OCPs dispensed without evidence of an IUD removal). Final diagnosis, procedure, and medication codes, and NLP algorithms utilized, as well as method exposure definitions will be modified based on review of sample data pulls and included the standard operating procedures.

Figure 1. Determining Contraceptive Exposure



9.3.2 Outcomes definition

Ectopic pregnancies, defined as extra-uterine pregnancy, will be identified using a combination of (ICD-9-CM, ICD-10-CM), CPT codes, medication codes (NDC and PP specific codes). Information extraction from clinical notes using NLP algorithms specifically developed for this project will also be used to ascertain the outcome of interest as needed per study site. Ectopic pregnancies will be categorized as surgical or medical using a combination of diagnosis, procedure, and medication codes. Multiple ectopic pregnancy visits occurring within a 90-day period will be considered part of the same pregnancy episode, with the diagnosis date defined as the date of the first ectopic pregnancy visit in an episode. For women who have more than one successive ectopic pregnancy, each ectopic pregnancy will be counted for aim 1a; however women will be censored after the first ectopic pregnancy for aim 2a. To obtain information on medical treatment of ectopic pregnancy, the pharmacy database will be used to identify methotrexate injections for cases between 7 days before and 90 days after the diagnosis date. An ectopic pregnancy episode will be classified as involving a surgical treatment if any procedure code for the removal of an ectopic pregnancy is found during an episode, regardless of the presence of medical treatment.

9.3.3 Covariate definition

For Specific aim 2a., the analysis of risk factors for ectopic pregnancy, baseline demographic and clinical characteristics will be assessed before the index date for each woman. Women will be excluded based on identification of variables consistent with relevant exclusionary conditions (section 9.2.3). For Specific aim 2a. women are required to have at least 12 months of continuous membership (no more than a 30-day gap) prior to the date of inclusion in the study. The look-back time will be all available data to 2008. Demographic and clinical characteristics will be assessed

before the index date for each woman for this analysis. Data will be extracted using relevant diagnostic and procedure codes as well as information documented as past medical and surgical history. For some patients, more information will be available, and all information within the database will be considered to reduce misclassification of baseline information.[25] In addition censoring events which make a woman not at-risk for pregnancy will be identified over the follow-up period.

Demographic variables:

- **Date of birth**
- **Age:** age in years as of the index date.
  - Six categories (15-19, 20-24, 25-29, 30-34, 35-39, 40-44) for descriptive tables.
  - Continuous variable for Poisson regression.
- **Race/ethnicity:** categorical variable with five categories: White, Black, Hispanic, Asian/Pacific Islander, other race/ethnicity
- **Socioeconomic status:** census tract data on population living below the poverty line
- **Smoking status:** (0 = Never smoked, 1 = Ever smoked) for smoking status as of the index date.
- **Date of death**

Clinical characteristics:

- **Parity:** any births prior to the index date (0 = No, 1 = Yes)
- **Cesarean delivery:** ever had prior to the index date (0 = No, 1 = Yes)
- **History of spontaneous abortion** (0 = No, 1 = Yes)
- **Ectopic pregnancy** (treated medically or surgically): ever had prior to the index date (0 = No, 1 = Yes)
- **Appendicitis/appendectomy:** ever had prior to the index date (0 = No, 1 = Yes)
- **Chlamydia or gonorrhea:** ever had prior to the index date (0 = No, 1 = Yes)
- **Pelvic inflammatory disease:** ever had prior to the index date (0 = No, 1 = Yes)
- **Endometriosis:** ever had prior to the index date (0 = No, 1 = Yes)
- **Anatomic abnormalities of the uterus** (congenital malformation): ever had prior to the index date (0 = No, 1 = Yes)
- **Sterilization (any):** ever had prior to the index date (0 = No, 1 = Yes)
  - Tubal ligation: (0 = No, 1 = Yes) or date
  - Essure: (0 = No, 1 = Yes) or date
  - Salpingectomy (0 = No, 1 = Yes) or date
  - non-specified: (0 = No, 1 = Yes) or date
- **Bilateral oophorectomy:**
  - ever had prior to the index date (0 = No, 1 = Yes) (exclusionary)
  - Date of procedure after index date (censoring event)
- **Adnexal surgery** (including surgery for ectopic pregnancy): ever had prior to the index date (0 = No, 1 = Yes)
- **Hysterectomy:**
  - ever had prior to the index date (0 = No, 1 = Yes) (exclusionary)
  - Date of procedure after index date (censoring event)
- **Menopause or ovarian failure:**

  - ever had prior to the index date (0 = No, 1 = Yes) (exclusionary)
  - Date of diagnosis after index date (censoring event)
- **Infertility diagnosis:** ever had prior to the index date (0 = No, 1 = Yes)

Other variables:

- **IUD type**: three-level categorical variable of inserted IUD type:
  - Homonal IUD
  - Non-hormonal IUD
  - Unknown IUD type
- **OCP type**: three-level categorical variable of OCP type:
  - COC (various)
  - POP
  - Unknown OCP type
- Other hormonal contraception (Implant, Patch, Ring)
- **Duration of the look-back period at index date:** continuous variable with a minimum of 52 weeks; may be categorized after examining frequency distribution.
- **Database**: categorical variable of the two databases included in the study: PPD or PPD

Diagnosis and procedure codes along with information extraction from clinical notes using NLP algorithms specifically developed for this project and ongoing Bayer-PMR study will also be used to ascertain covariates of interest as needed. Codes and algorithms will be based on review of sample data pulls and included within the standard operating procedures.

9.4 Data sources

There will be two separate EHR data sources for this study: PPD and PPD. PPD and PPD computerized health databases contain health plan enrollment information, inpatient and outpatient clinical visits, external claims and pharmacy records. The inpatient database captures all inpatient hospitalization visits, recording admission and discharge dates as well as up to ten ICD-9/10 discharge diagnoses and procedures and up to ten CPT codes. The outpatient database captures all primary care outpatient clinic visits, urgent care visits, and ER visits recording ICD-9/10 diagnoses and procedures as well as CPT-4 codes. The external claims database captures all outpatient (clinic, urgent care, and ER) and inpatient visits by PP enrollees to non-PP facilities where PP is financially responsible for the care. The pharmacy database captures medications dispensed to PP enrollees with a pharmacy benefit plan at PP -owned pharmacies. All databases are linked through a unique medical record number assigned to each enrollee, precluding multiple counting of the same health event for individuals across sources.

Both PPD and PPD access the Virtual Data Warehouse (VDW) which was created to facilitate multi-site research projects. Local variables are standardized using consistent naming, definitions, and formats. It is "virtual" in the sense that physical data remains at each of the participating sites. During a project, the programmer at the data-coordinating center (or leading site) writes a single SAS program that can be distributed to all the participating sites throughout the Health Care Systems Research Network, formerly known as the HMO Research Network. This program can be executed at each site with minor modifications.

9.4.1 PPD

The PPD Division of Research (DOR) has developed and implemented a research database on an Oracle platform that serves as a repository for PPD clinical and administrative databases in PP HealthConnect® (PPD) data sources. The DOR research database aggregates into a single data warehouse the hundreds of separate regional data sources that researchers have typically accessed to extract data elements needed for specific research projects. It conforms PPD data into standard database structures that are optimized for research query and retrieval. It uses a security model that meets all HIPAA, IRB and PP requirements. The database is now used by PPD researchers to support virtually all research projects, including epidemiological studies, health services research, and informatics projects. The database was developed in parallel with the deployment of PPD within all PP medical facilities. It receives daily updates from PPD 's clinical databases, in addition to other PP clinical data sources not available through PPD .

PPD will leverage existing perinatal data resources; specifically, the Infant and Maternal Cohort files maintained by the PPD under the PPD , a practicing PPD neonatologist and researcher. The files are prospective datasets indexed on the cohort of live births in PPD The files are compiled monthly and are maintained within a proprietary DOR PRU SAS application. (Note: this resource made available to research teams is not an IRB-approved research project. Variables in the datasets include maternal demographics, inpatient utilization at the delivery/birth encounter, and diagnoses and procedures from the delivery/birth encounter. PPD will also use a hysterectomy database [IRBNet ID:1272878-20], which includes all hysterectomies performed at PPD from the year 2000 onward.

9.4.2 PPD

The Research Database Team at PPD 's Research and Evaluation extracts data from Clarity (the back-end database to HealthConnect®), legacy, and claims 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 healthcare 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.

PPD also maintains a specialized Perinatal Data Mart, which contains medical and obstetrical information on all pregnancies and births at PPD Hospitals. The database includes data on many potential prenatal risk factors, prenatal visit information, complications of pregnancy and delivery, birth outcomes including but not limited to gestational age, birth weight, plurality, birth order, and obstetrical procedures, maternal age, race/ethnicity, education, pre-pregnancy and delivery weight and height as well as lifestyle factors such as tobacco smoking and alcohol use during pregnancy.

In addition, PPD maintains a pregnancy file with pregnancy outcomes contributed by two ongoing PPD large scale research projects: The Vaccine Safety Datalink and The Center for Safety and Effectiveness Research. The pregnancy file is created by a validated pregnancy episode algorithm combining all potentially relevant diagnosis codes, procedure codes, laboratory tests, pharmacy records, imaging procedures to assign or determine pregnancy outcome type to each pregnancy episode. Furthermore, there is also a designed flowsheet to capture the unintended pregnancy information collected during obstetric visits in the EHR system.

9.5 Study size

Ectopic pregnancy is a relatively uncommon event. Van Den Eeden et al reported an annual rate of 1.03 per 1,000 women 15-44 years old from 1997 – 2000 at PPD .[1] Preliminary query of PPD diagnosis and procedure codes for ectopic pregnancy for the study period of 2007-2016 revealed 10,693 episodes of ectopic pregnancy, of which 2,617 were associated with a surgical procedure code. Review of a random sample of remaining episodes revealed evidence of medical treatment for an ectopic pregnancy in 90% of charts. Preliminary query of PPD diagnosis and procedure codes for ectopic pregnancy for the time period of 2007-2016 revealed 10,203 episodes of ectopic pregnancy, of which 3,758 were associated with a surgical procedure code. Aim 1 for the study is descriptive; however, we estimate that over the time period there will be approximately 240,000 IUD insertions (1,725,000 person-years [PY]) and over 1 million OCP users (8,625,000 person-years). For the IUD cohort, with a total of 1,725,000 PY, there is 83% power to detect a linear decreasing trend over the 10-year study period from an ectopic pregnancy incidence rate of 0.71/1000PY to 0.55/1000PY using a two-sided Z test with continuity correction and a significance level of 0.05. Since the IUD cohort is the smallest, we therefore have more than adequate power to detect similar trends for the OCP cohort and the overall cohort of women of reproductive age.

9.6 Data management

This study will use data previously collected in EHRs and other electronic administrative and clinical databases described in section 9.4. 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 limited dataset (stripped of all identifiers except for dates) will be sent from PPD to PPD for analysis. In order to complete analyses of Aims 2a and 2d, a limited dataset will be sent from PPD to PPD . The PPD and PPD principal investigators will be responsible for ensuring that policies and procedures for confidentiality and security are followed for this project. Data management will be conducted in accordance with standard operating procedures developed for the study and used across the two 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. This includes performing range checks and producing general frequency tables such that missing values, outliers, and inappropriate or abnormal values will be identified. We will also review preliminary data across the two sites for consistency (i.e. significant differences in baseline frequencies of exposure and outcome data across sites will be queried to identify sources of discrepancies). Description of specific data management plans for each site follow.

9.6.1 PPD

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 or later. 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 PPD information. DOR IT partners with the National Compliance Organization, Technology Risk Organization, and The PPD Medical Group to lay the foundation for operational strategies and programs that meet PPD 's security obligations. Access to the EHR requires authorization from PPD 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.

9.6.2 PPD

The research database team at PPD will manage the study databases and provide the support needed to meet study objectives. Data is extracted from Clarity (the back-end database to HealthConnect®), legacy, and claims and is integrated with historical data prior to HealthConnect® into a comprehensive Research Data Warehouse (RDW). 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 analysis datasets created by PPD will be stored and archived at PPD as per the applicable requirements and retention policies. Computer files associated with this project will be kept in a password-protected environment.

9.6.3 Validation of exposure and outcome ascertainment

To assess the quality of available electronic information on ectopic pregnancies and prescribed contraceptive method, each site will conduct a review of a random sample of ectopic pregnancies stratified by type (surgical or methotrexate) and year up to 100 total (50 PPD an d50 PPD) ~ 5% of ectopic pregnancies). Physician notes and prescribed medications will be reviewed to determine ectopic pregnancy status (yes, no, undetermined/missing) and contraceptive method as electronically abstracted (yes, no, undetermined/missing) Chart reviews will be done in an iterative process; algorithms will be modified based on finding of the chart review. Patient identifiers (except for service dates) will be replaced with unique study IDs; the identifiers and tables linking patient identifiers to the study ID will be destroyed at the end of the study.

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 data analyses related to each objective. A detailed description of variable definitions, null hypotheses, planned analyses, and display specifications will be included within the statistical analysis plan.

Limited analytic datasets (dates of birth and death, dates for services, diagnoses, procedures, and prescriptions are the only PHI included) will be shared (via secure file transfer protocol) between sites for analysis. Analyses related to primary objectives, and secondary objectives 2b, 2c, and 2d will be performed by PPD on the patient-level data from both sites; analyses related to secondary objectives 2a and 2d will be performed by PPD on the patient-level data from both sites. In addition to the aggregated results, the results for each objective will be presented separately for each site.

9.7.1 Descriptive analyses

Descriptive analysis of relevant variables will be conducted initially for PPD and PPD separately. Both will apply the inclusion and exclusion criteria to its data to obtain the study population.

Descriptive analyses for all variables of interest (Section **9.33**) 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. For estimates, two-sided 95% confidence intervals will be calculated. The proportion of missing data will be captured for each variable.

Demographic and clinical characteristics of patients experiencing ectopic pregnancies will be presented stratified by:

- Surgically treated
- Medically treated
- Unknown/other (includes expectant management)

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 index date.

9.7.2 Specific Aim 1a.

Ectopic pregnancy rates will be calculated based on two denominators: (1) woman-years or the length of enrollment among women aged 15–44 years during the study period, and (2) number of pregnancies in this group of women during the same study period. We will include live births, ectopic pregnancies, and induced abortions in the second denominator. Both denominators will be stratified by 1–calendar year group and 5-year age group. Rates per 1,000 woman-years and per 1,000 pregnancies with 95% confidence intervals will be calculated as the number of ectopic pregnancy cases divided by the total woman-years and the total number of pregnancies, respectively. We will assess for changes in the age distribution of PP enrollees over the 10-year study period and calculate age-adjusted rates per 1,000 woman-years, standardized to the PP population distribution as appropriate. Poisson regression models adjusting for overdispersion, with calendar-year fitted as a continuous variable will be used to evaluate any linear trends in rates over the study period between 2009 and 2018. Ectopic pregnancy rates and trend tests will be calculated for all age groups combined as well as stratified by 5-year age group.

9.7.3 Specific aim 1b.

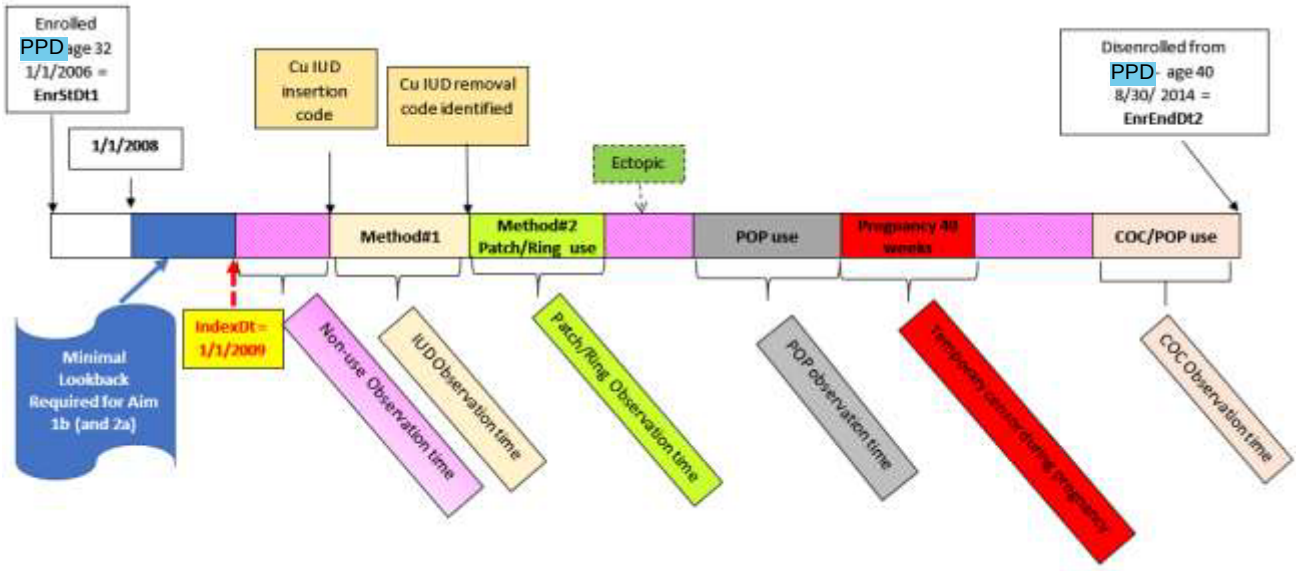
Contraceptive exposure episodes for the contraceptives of interest for each woman will be identified overtime; women may have multiple episodes of IUD, OCP and DMPA exposure. Use of hormonal contraception will be updated throughout the follow-up period, and the status of women will change when they discontinue or change the type of hormonal contraception used. Total woman-years of IUD, OCP and DMPA exposure will be calculated. Ectopic pregnancy incidence rates and trends will be calculated based on 5 denominators: hormonal and non-hormonal IUDs, COCs and POPs, and DMPA use. Incidence rates will be expressed per 1,000 woman-years of contraceptive method exposure with 95% confidence intervals. Poisson regression models adjusting for overdispersion, with calendar-year fitted as a continuous variable will be used to evaluate any linear trends in rates

over the study period between 2009 and 2018. We will calculate the ectopic pregnancy incidence rates and 95% confidence intervals of women using each of the three contraceptive methods of interest and the ectopic pregnancy incidence rate for the overall population of women of reproductive age.

9.7.4 Specific aim 1c.

This aim is added as post-hoc analysis with amendment dated 28 October 2020. For Aim 1c we will consider all observation time, including periods of contraceptive non-use. In this analysis, patch, ring, and implant exposure time will not be excluded (Figure 2). Ectopic pregnancy incidence rates and trends will be calculated based on hormonal and non-hormonal IUDs, IUD type unknown, COCs and POPs, DMPA use, other methods (patch, ring, and implant) and contraceptive non-use. We will exclude observation time during pregnancies that end in a live birth or induced abortion. Incidence rates will be expressed per 10,000 woman-years of contraceptive method exposure. Poisson regression models adjusting for overdispersion, with calendar-year fitted as a continuous variable will be used to evaluate any linear trends in rates by method and for non-use over the study period between 2009 and 2019. We will do a sensitivity analysis in which we will exclude ongoing IUD use (IUD use episodes identified by isolated removal or surveillance codes).

Figure 2. Contraceptive Exposure for Aim 1c



9.7.5 Specific aim 2a.

For this analysis, women who are not at-risk for pregnancy during the study period will be excluded. Women will also be censored at the time of natural or surgical menopause, ovarian failure, or hysterectomy. Women will not contribute time during pregnancy resulting in a live birth, or induced abortion. Excluded time will be calculated based on delivery date and estimated gestational age

(EGA) at delivery or procedure date for induced abortion. Similar to aim 1b. woman-years of IUD, COC, POP, and DMPA contraceptive exposure will be calculated. In addition, use of hormonal contraceptives will be categorized as current use or previous use (discontinuation more than 1 month previously). The outcome of interest is the first ectopic pregnancy (yes/no) during the study period. Poisson regression with robust variance will be used to estimate risk ratios and 95% confidence intervals of ectopic pregnancy across age groups, race/ethnicity groups, contraceptive exposure groups, and relevant covariates while controlling for potential confounders. Thirty-day periods will be used as a time scale in the Poisson regression. The study population will be followed until the diagnosis of ectopic pregnancy, end of health plan enrollment, age 45 years, death, or the end of follow-up on December 31, 2018. Simple adjusted models controlling for age, calendar year, and hormonal contraceptive use and fully adjusted models including relevant risk factors will be developed.

9.7.6 Specific aim 2b.

Ectopic pregnancies will be categorized to the degree possible as surgical treatment or ectopic pregnancies treated with methotrexate. Cases which treatment type cannot be determined (including expectant management) will be classified as unknown/other. The proportion of cases with medical and surgical treatment each year over the study period will be calculated and the trend will be evaluated using the Cochran-Armitage test for a linear trend in proportions.

9.7.7 Specific aim 2c

This aim is added as post-hoc analysis with amendment dated 28 October 2020. In order to validate the contraceptive use algorithms, we will look at an additional cohort of 400 women with contraceptive use irrespective of ectopic pregnancy status (200 at PPD and 200 at PPD). Since observation time for non-contraceptive use is greater than observation time for all methods combined, we will over sample women who used a method at the index date. Trained abstractors will conduct chart reviews to determine contraceptive use over the study period status using a structured abstraction form. We will assess the diagnostic validity of the contraceptive algorithms against the gold-standard "true case" as determined by chart review. Condom use will be considered as using no method. Cases where contraceptive use status is unclear will be identified and adjudicated by a clinician on an ongoing basis. We will exclude cases with insufficient information to determine contraceptive use status.

We will summarize the performance by method and site calculating:

- sensitivity - percentage of cases where a contraceptive use episode identified by chart review was correctly classified by the algorithm
- false positive rate - percentage of cases where a contraceptive use episode identified by chart review was incorrectly classified by the algorithm (i.e. use of a different method or non-use identified)
- PPV (positive predictive value) - percentage of cases where a contraceptive use episode determined by the algorithm was confirmed by chart review
- Overall accuracy – percentage of cases where all episodes of contraceptive use for a women were correctly identified.
- F-score (the weighted harmonic mean of the test's precision and recall): 2 x (PPV x Sensitivity) / (PPV + Sensitivity)

9.7.8 Specific aim 2c.

This aim is added as post-hoc analysis with amendment dated 28 October 2020. We will repeat the Aim 2a analysis, risk factors for ectopic pregnancies, using data from Aim 1c; however, for this analysis the reference category for contraceptive exposure group will be non-use of contraception.

9.8 Quality control

Standard operating procedures at PPD and PPD will guide the conduct of the study and will be used to ensure data quality and security. Procedures will be consistent with the FDA's Best Practices for Conducting and Reporting Pharmacoepidemiologic Safety Studies Using Electronic Healthcare Data.[26] 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. 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 SAP, data abstraction protocol, and study reports will undergo quality-control review, senior scientific review, editorial review, and review by PPD and PPD investigators.

9.9 Limitations of the research methods

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), CPT codes, medication codes (NDC and PP specific), and clinical notes (i.e., via NLP) there is a possibility of misclassification. Algorithms for the outcome variables and exposures will be validated in the databases. While we attempted to identify all potential EP cases for the EP validation by using cases with either an EP-related diagnostic or procedure code, it is possible that EP cases that were incorrectly or not coded were not captured which would have falsely increased our assessment of the sensitivity of the algorithms. However, in our health care system that relies on insurance reimbursement it is unlikely for an EP case to not have documentation with either a diagnosis or procedural code. For variables that have not been validated in these databases, algorithms validated in other data sources (e.g., PP VDW, PMR study, and administrative claims) will be used to identify conditions and medication dispensing, when available. In addition, the study team will develop and share variable definitions across data sources to standardize approaches to data capture.

As with all health care database studies used for secondary data analysis, data are limited to information documented by the health care system and are not necessarily available prior to the start date of database enrollment for the individual. While we will collect information recorded as past medical and surgical history, this is information may not be consistently documented. A minimum 12-month look-back period prior to the study period will be required for inclusion in Aim 1b and 2a, but all available time in the database back to January 1, 2008 will be used to improve the assessment of potential confounders. [25]

All of the data are from two large health care systems in California, both within PPD, but there is considerable diversity in factors such as race/ethnicity within each of the data sources. As with all observational studies, there will be a potential for unmeasured bias that will influence

results. Utilization of Poisson regression for the risk factor analysis (aim 2a) and adjusting the risk ratios for potential confounders will help to reduce this, but there is always the possibility of residual confounding, which would affect calculated point estimates and 95% CIs. An unmeasured confounder would have to be very unbalanced between cohorts to have a large impact on the outcomes.

We do not have data on sexual activity hence we will over estimate the number of women at risk for pregnancy. Assessing sexual activity is beyond the scope of this EHR database study. Risk of sexual activity will be assumed across all contraceptive exposure groups; however there may be bias by contraceptive method. Adjusted hazard ratios should provide a less biased estimate of any differential risk between exposure groups. There is the potential for underreporting of induced abortions within the data sources since women may not report abortions or seek abortions outside the health care system. However underreporting of induced abortions is well known and our ascertainment is representative of the way that these outcomes would appear within clinical practice and prior studies of this outcome.[27] History of spontaneous abortions is not reliable reported and documented. We do not capture spontaneous abortions that occur during the study period therefore we may misclassify some person-time at risk. Many women with spontaneous abortions recognized them after the event, do not seek medical attention, are managed expectantly, and may have multiple visits over time for one event.[28,29] Therefore documentation is incomplete and difficult to interpret. There may also be some underreporting of ectopic pregnancies due to misdiagnosis.

9.10 Other aspects

PPD will serve as a coordinating center for this project and will coordinate the activities of the two sites and Bayer (e.g., ensure that timelines are being met and facilitate communications). PPD will receive the person-level limited dataset from PPD and perform the pooled analyses on the primary objectives and secondary objectives 2b and 2c. PPD will receive the person-level limited dataset from PPD and perform the pooled analyses on secondary objectives 2a, and 2d.

PPD will be the primary author for the protocol, statistical analysis plan, standard operating procedures, and a study report for this study. PPD will solicit and incorporate input from PPD and the Bayer study team for the protocol, statistical analysis plan, and study report. To ensure consistency across study sites, information on methods and approaches to ascertainment of exposure and outcome information will be shared across sites.

10 Protection of human subjects

This study uses secondary data collection; there is no patient contact or intervention and therefore poses only minimal risk for patients (e.g., potential for breach of confidentiality within health plan due to extraction of data from records). We plan to conduct the study with Waiver of Informed Consent and HIPAA Privacy rule Authorization because the study will use only data that have already been collected at the time the research and it was not possible to know which patients would have been eligible for inclusion. All data collected in the study will be stripped of personal identifiers except for dates of services after the dataset has been collected and validated, minimizing risk of breach of confidentiality of protected health information. PPD will obtain IRB approval before conducting the study. PPD 's IRB will be asked to cede approval to the PPD IRB. For PPD information on maternal socioeconomic factors is based on data entered on the infant birth certificate collected by Vital Records of the California State; therefore, PPD will obtain California State IRB approval. 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.

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). Based on current guidelines from ISPE [30] and the EMA [31], 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

Results of this study will be published following guidelines, including those for authorship, established by the International Committee of Medical Journal Editors [32]. When reporting results of this study, the appropriate Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) checklist will be followed [33]. Bayer, and the investigators at each data site will agree upon a publication plan that will result in collaborative publication(s) based on these study results. Communication via appropriate scientific venues, e.g., Society of Family Planning, American Congress of Obstetrics and Gynecology and the Association of Reproductive for Reproductive Medicine will be considered.

13 References

1. Van Den Eeden SK, Shan J, Bruce C, Glasser M. Ectopic pregnancy rate and treatment utilization in a large managed care organization. Obstet Gynecol 2005;105:1052-7.

2. Trabert B, Holt VL, Yu O, Van Den Eeden SK, Scholes D. Population-based ectopic pregnancy trends, 1993-2007. Am J Prev Med 2011;40:556-60.

3. Farquhar CM. Ectopic pregnancy. Lancet 2005;366:583-91.

4. Xiong X, Buekens P, Wollast E. IUD use and the risk of ectopic pregnancy: a meta-analysis of case-control studies. Contraception 1995;52:23-34.

5. McCann MF, Potter LS. Progestin-only oral contraception: a comprehensive review. Contraception 1994;50:S1-195.

6. Peterson HB, Xia Z, Hughes JM, Wilcox LS, Tylor LR, Trussell J. The risk of ectopic pregnancy after tubal sterilization. U.S. Collaborative Review of Sterilization Working Group. N Engl J Med 1997;336:762-7.

7. Mikhail E, Salemi JL, Schickler R, Salihu HM, Plosker S, Imudia AN. National rates, trends and determinants of inpatient surgical management of tubal ectopic pregnancy in the United States, 1998-2011. J Obstet Gynaecol Res 2018.

8. Loffer FD. Outpatient management of ectopic pregnancies. Am J Obstet Gynecol 1987;156:1467-72.

9. Ory SJ. New options for diagnosis and treatment of ectopic pregnancy. JAMA 1992;267:534-7.

10. Zane SB, Kieke BA, Jr., Kendrick JS, Bruce C. Surveillance in a time of changing health care practices: estimating ectopic pregnancy incidence in the United States. Matern Child Health J 2002;6:227-36.

11. Goldner TE, Lawson HW, Xia Z, Atrash HK. Surveillance for ectopic pregnancy--United States, 1970-1989. MMWR CDC Surveill Summ 1993;42:73-85.

12. Centers for Disease C, Prevention. Ectopic pregnancy--United States, 1990-1992. MMWR Morb Mortal Wkly Rep 1995;44:46-8.

13. Hoover KW, Tao G, Kent CK. Trends in the diagnosis and treatment of ectopic pregnancy in the United States. Obstet Gynecol 2010;115:495-502.

14. Tao G, Patel C, Hoover KW. Updated Estimates of Ectopic Pregnancy among Commercially and Medicaid-Insured Women in the United States, 2002-2013. South Med J 2017;110:18-24.

15. Correction. Am J Prev Med;42:107.

16. Weiss NS, Daling JR, Chow WH. Control definition in case-control studies of ectopic pregnancy. Am J Public Health 1985;75:67-8.

17. Mol BW, Ankum WM, Bossuyt PM, Van der Veen F. Contraception and the risk of ectopic pregnancy: a meta-analysis. Contraception 1995;52:337-41.

18. Furlong LA. Ectopic pregnancy risk when contraception fails. A review. J Reprod Med 2002;47:881-5.

19. Li C, Zhao WH, Meng CX, et al. Contraceptive Use and the Risk of Ectopic Pregnancy: A Multi-Center Case-Control Study. PLoS One 2014;9:e115031.

20. Zhang Z, Weng L, Zhang Z, et al. An epidemiological study on the relationship of ectopic pregnancy and the use of contraceptives in Beijing--the incidence of ectopic pregnancy in the Beijing area. Beijing Collaborating Study Group for Ectopic Pregnancy. Contraception 1994;50:253-62.

21. Heinemann K, Reed S, Moehner S, Minh TD. Comparative contraceptive effectiveness of levonorgestrel-releasing and copper intrauterine devices: the European Active Surveillance Study for Intrauterine Devices. Contraception 2015;91:280-3.

22. Rowe P, Farley T, Peregoudov A, et al. Safety and efficacy in parous women of a 52-mg levonorgestrel-medicated intrauterine device: a 7-year randomized comparative study with the TCu380A. Contraception 2016;93:498-506.

23. Franks AL, Beral V, Cates W, Jr., Hogue CJ. Contraception and ectopic pregnancy risk. Am J Obstet Gynecol 1990;163:1120-3.

24. Nielsen LH, Lokkegaard E, Andreasen AH, Keiding N. Using prescription registries to define continuous drug use: how to fill gaps between prescriptions. Pharmacoepidemiol Drug Saf 2008;17:384-8.

25. Brunelli SM, Gagne JJ, Huybrechts KF, et al. Estimation using all available covariate information versus a fixed look-back window for dichotomous covariates. Pharmacoepidemiol Drug Saf 2013;22:542-50.

26. U.S. Department of Health and Human Services FaDA. Best Practices for Conducting and Reporting Pharmacoepidemiologic Safety Studies Using Electronic Healthcare Data 2013.

27. Fu H, Darroch JE, Henshaw SK, Kolb E. Measuring the extent of abortion underreporting in the 1995 National Survey of Family Growth. Fam Plann Perspect 1998;30:128-33, 38.

28. Wilcox AJ, Weinberg CR, O'Connor JF, et al. Incidence of early loss of pregnancy. N Engl J Med 1988;319:189-94.

29. Rafi J, Khalil H. Expectant management of miscarriage in view of NICE Guideline 154. J Pregnancy 2014;2014:824527.

30. ISPE. Guidelines for good pharmacoepidemiology practices (GPP). Revision 3: International Society for Pharmacoepidemiology; 2015 June.

31. EMA. Guideline on good pharmacovigilance practices (GVP). Module VI – Management and reporting of adverse reactions to medicinal products: European Medicines Agency; 2014 8 September.

32. ICMJE. Recommendations for the conduct, reporting, editing, and publication of scholarly work in medical journals: International Committee of Medical Journal Editors; 2016 December.

33. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Journal of clinical epidemiology 2008;61:344-9.

Annex 1: List of stand-alone documents

N/A