Use of The Tree-Based Scan Statistic for Surveillance of Maternal Outcomes Following Medication Use During Gestation

Sentinel Methods

Thuy N Thai[1], Almut G Winterstein [1,2,3], Elizabeth A Suarez[4], Michael Nguyen[5], Danijela Stojanovic[5], Jane Liedtka[6], Abby Anderson[7], Di Zhang[8], Yueqin Zhao[8], Monica Munoz[9], Wei Liu[10], Steven Bird[10], Inna Dashevsky[4], David Cole[4], Talia Menzin[4], Sandra DeLuccia[4], Jennifer Noble[4], Judith C Maro[4]

1. Department of Pharmaceutical Outcomes and Policy, College of Pharmacy, University of Florida, Gainesville, FL; 2. Center for Drug Evaluation and Safety, University of Florida, Gainesville, FL; 3. Department of Epidemiology, College of Medicine and College of Public Health and Health Professions, University of Florida, Gainesville, FL; 4. Department of Population Medicine, Harvard Pilgrim Health Care Institute and Harvard Medical School, Boston, MA; 5. Office of Surveillance and Epidemiology, Center for Drug Evaluation and Research, US Food and Drug Administration, Silver Spring, MD; 6. Division of Pediatric and Maternal Health, Center for Drug and Evaluation Research, US Food and Drug Administration, Silver Spring, MD; 7. Division of Urology, Obstetrics and Gynecology, Center for Drug and Evaluation Research, US Food and Drug Administration, Silver Spring, MD; 8. Office of Biostatistics, Center for Drug Evaluation and Research, FDA, Silver Spring, MD; 9. Division of Pharmacovigilance, Center for Drug Evaluation and Research, US Food and Drug Administration, Silver Spring, MD 10. Division of Epidemiology, Center for Drug and Evaluation Research, US Food and Drug Administration, Silver Spring, MD

Version 1.0

June 17, 2021

The Sentinel System is sponsored by the U.S. Food and Drug Administration (FDA) to proactively monitor the safety of FDA-regulated medical products and complements other existing FDA safety surveillance capabilities. The Sentinel System is one piece of FDA's Sentinel Initiative, a long-term, multi-faceted effort to develop a national electronic system. Sentinel Collaborators include Data and Academic Partners that provide access to healthcare data and ongoing scientific, technical, methodological, and organizational expertise. The Sentinel Operations Center is funded by the FDA through the Department of Health and Human Services (HHS) Contract number HHSF223201400030I.

Use of the Tree-Based Scan Statistic for Surveillance of Maternal Outcomes Following Medication Use during Gestation

Sentinel Methods

Table of Contents

History of Modifications

Version	Date	Modification	Author
1.0	06/17/2021	Original Version	Sentinel Operations Center

1 Introduction

Medication use during pregnancy is common. In the United States, about 70% of pregnant women use at least one prescription medication during pregnancy and four medications on average.[1] Using medications during pregnancy imposes potential risk on pregnancy outcomes, including teratogenic risk for fetal development as well as risk for maternal adverse outcomes such as gestational diabetes. Almost 98% of medications approved from 2000 to 2010 have an undetermined teratogenic risk.[2] For pregnant women with acute or chronic diseases unrelated to pregnancy, there is also risk if treatment is avoided due to unknown teratogenic effects.[3] More evidence is needed to evaluate these competing risks and to guide women and clinicians in making decisions.

Pregnant women are rarely enrolled in clinical trials, and so there is limited human data on medication safety during pregnancy at the time of product approval. To fill this gap, there is an important need for post-marketing pregnancy safety studies (e.g., pregnancy registries, pregnancy surveillance programs or epidemiologic studies using data collected from routine clinical care).[4] The advantages of registry data include the prospective nature of the data collection, which provides detailed patient information.[5] However, registry data have limited sample size, lack of appropriate control groups, and loss of generalizability due to selection bias.[6]

Real-world data from routine clinical care can address some of these limitations and can be a useful complement to pregnancy registry data.[6] Administrative claims data capture routinely reimbursed healthcare utilization and do not require active recruitment of patients, allowing larger and more heterogenous cohorts for analysis. Incidence rates and risk estimates associated with drug exposure during pregnancy can be calculated with these data and be compared with women exposed to an active comparator drug or unexposed women from the same source population. One limitation of administrative claims data is that evidence of pregnancy is typically indicated by livebirth delivery.[7,8] Gestational age (and thus start of pregnancy) are more challenging to determine for pregnancies that do not result in livebirth, although there is active research in this area.[9,10]

TreeScan™ (http://www.treescan.org) is a signal identification method that evaluates thousands of outcomes simultaneously to identify potential adverse events after adjusting for multiple testing.[11] TreeScan can be used to screen for maternal complications during pregnancy and to identify unusual elevated frequencies of these complications for further evaluation. Moreover, because potential maternal complications are classified into a hierarchical tree structure, TreeScan can screen for specific potential adverse events as well as more composite clinical concepts. TreeScan is also compatible with multiple study designs and can be used with appropriate methods to control potential confounding in observational studies.[12] TreeScan has been tested in the general adult and adolescent population[12-17] and has been used to assess birth outcomes,[18,19] but has not been used to assess maternal complications in the pregnant population. Simulation studies using TreeScan have shown consistent ability to maintain overall random Type 1 error over multiple hypothesis tests, and have quantified the effects of imperfect sensitivity on Type II error (i.e., statistical power).[14,15,20] FDA initiated this project to develop new methods to conduct surveillance using real-world data for maternal and obstetric outcomes following medication use during pregnancy.

2 Specific Aims

We will assess the performance of the TreeScan method to identify signals for maternal and obstetric adverse outcomes occurring from 20 weeks of gestation to 30 days after delivery among women with livebirths exposed to oral macrolides compared to oral penicillins.

3 Case Study: Use of Macrolides and Penicillins

The purpose of this methods project is to evaluate the performance of TreeScan rather than find potential unexpected maternal complications. Therefore, we selected a case study of drug exposure that is expected to yield a large enough sample size to detect outcomes with smaller baseline prevalence or smaller increased risk. We also considered the availability of an appropriate active comparator to limit the need to control for potential confounders. Finally, we prioritized exposure and control drugs with known safety profiles so that the TreeScan results could be interpreted in the context of a robust body of published existing safety data.

Antibiotics are among the most common medications used during pregnancy.[1,21] The percentage of women having livebirths with ≥ 1 dispensing of macrolides or penicillins during pregnancy was 17% and 18%, respectively.[21] Macrolides and penicillins share similar indications such as upper and lower respiratory tract infections, gastrointestinal infections, and sexually transmitted infections. Both macrolides and penicillins are generally considered safe antibiotics during pregnancy, with limited safety concerns regarding maternal outcomes.

Although not the focus of our analysis, there is a more complicated safety profile for fetal and neonatal outcomes following maternal exposures than maternal outcomes. Some studies have suggested that women with macrolide exposure would have a small increased risk of miscarriage and fetus with congenital malformation, cerebral palsy or epilepsy.[22-24] Penicillins were not associated with congenital malformations in most studies;[25-30] however, some studies have reported a small increased risk of oral clefts with exposure during the first trimester and a higher risk of necrotizing enterocolitis with amoxicillin and clavulanic acid with exposure during the third trimester.[31-34] These well-documented safety profiles make macrolides and penicillins ideal candidates for our exposure and control drugs for this case study.

4 Study Design

4.1 Data and Study Period

We will use the IBM MarketScan® Research Database, one of the largest samples of employer-sponsored health insurance enrollees in the U.S. It provides longitudinal detail for patient-level healthcare utilization. Our study period is from October 1, 2015 to February 29, 2020, which confines disease codes to the International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM) codes and excludes healthcare utilization after the start of the COVID-19 pandemic. Based on our study parameters, the first valid livebirth delivery date is October 26, 2016 and the final valid livebirth delivery date is January 30, 2020.

4.2 Defining Livebirth Pregnancy Episodes and Exclusion Criteria

We will include single livebirth deliveries using a previously validated code set of ICD-10-CM diagnosis, ICD-10 procedure Coding System (ICD-10-PCS) and Current Procedural Terminology, Fourth Edition (CPT-4) codes recorded in any care setting and without any restriction in code position.[35] Pregnant women aged 10 to 54 years with livebirth delivery are required to have continuous medical and drug coverage (with a 45-day gap allowance) for at least 391 days before to 30 days after the delivery date. The pregnancy start date is calculated from the estimated gestational age at delivery using a validated algorithm.[7,35] We will exclude livebirth pregnancy episodes that also had at least one livebirth delivery code within 301 days before the defined delivery date.

We will exclude multiple livebirth or pregnancies with both livebirth and stillbirth outcomes. Both macrolides and penicillins are used as a prophylaxis for pregnant women with preterm premature rupture of membranes (PPROM). Therefore, maternal complications that manifest after PPROM diagnosis may be related to PPROM and not to the antibiotic exposure. We will

therefore exclude deliveries with a PPROM-related diagnosis (ICD-10-CM codes O42.xx) from pregnancy start to the date of antibiotic initiation. Any pregnancy episodes with at least one dispensing or one outpatient procedure related to any oral or injectable teratogenic drug from pregnancy start date to index date are also excluded. Exposure to these drugs may complicate a pregnancy and manifest consequent maternal complications.

Figure 1 shows a design diagram for livebirth delivery selection including related exclusion criteria. Code lists to describe the concepts shown in the design diagram and in this protocol are located in the Appendix organized as described in Table 1. Appendix A includes the list of generic and brand names to identify teratogenic drugs through National Drug Codes (NDC). Appendix B is a list of Healthcare Common Procedure Coding System (HCPCS) codes of those teratogenic drugs.

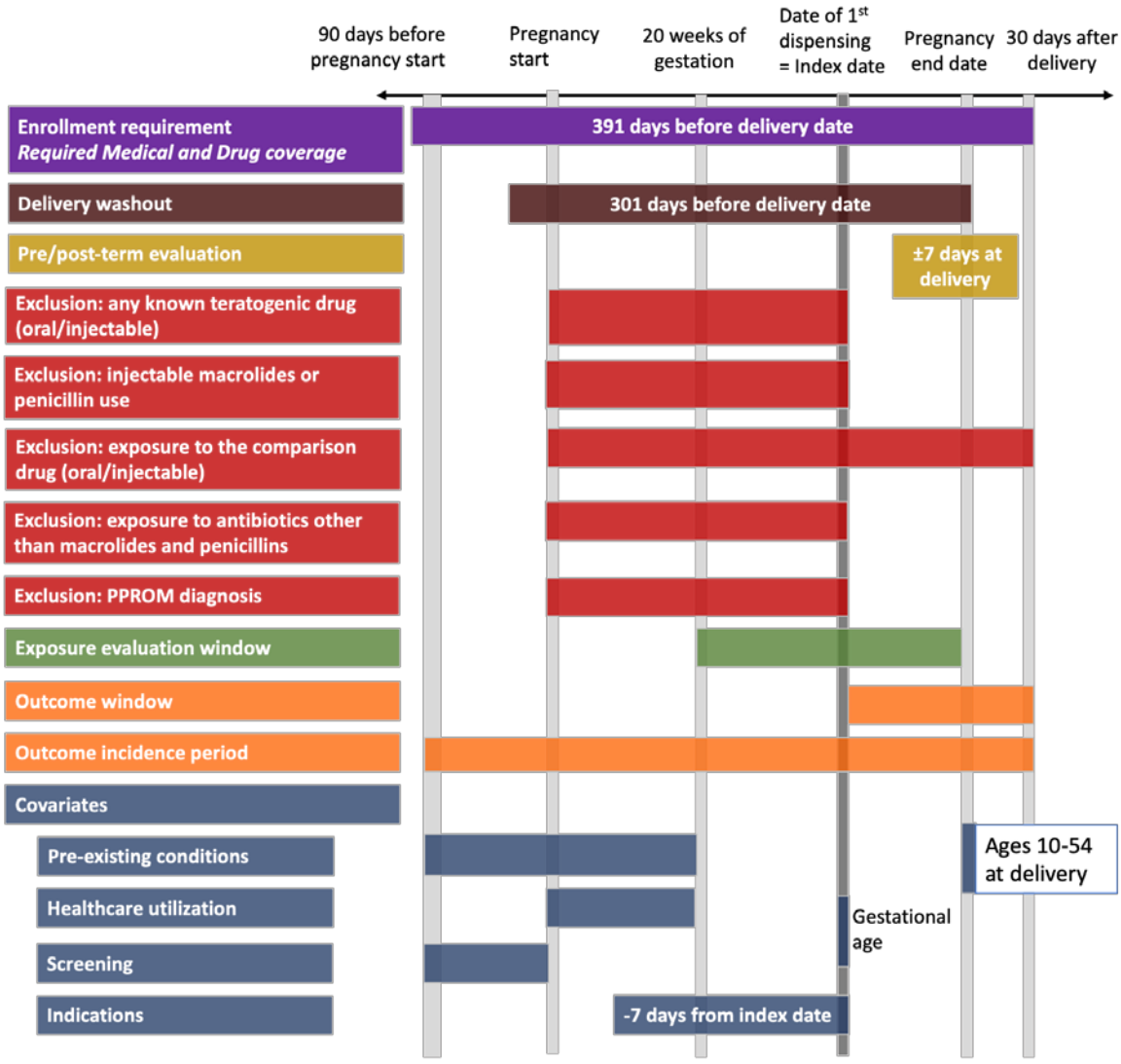

Figure 1. Design diagram for the macrolides and penicillins case study

Table 1. Summary of appendix tables and their content

Appendix	Appendix name	Description	Role
A	Generic and Brand Names of Medical Products to Identify Teratogenic Drugs with Oral and Injectable Routes for Exclusion	List of teratogenic drugs	Exclusion criterion
B	List of Healthcare Common Procedure Coding System (HCPCS) Codes to Identify Teratogenic Drugs with Oral and Injectable Routes for Exclusion	HCPCS codes of teratogenic drugs	Exclusion criterion
C	Generic and Brand Names of Medical Products to Identify Oral Macrolides and Penicillins	List of oral macrolides and penicillin	Study drug exposure definition and exclusion criterion
D	Generic and Brand Names of Medical Products to Identify Injectable Macrolides and Penicillins	List of injectable macrolides and penicillin	Exclusion criterion
E	List of Healthcare Common Procedure Coding System (HCPCS) Codes to Identify Injectable Macrolides and Penicillins	HCPCS codes of injectable macrolides and penicillin	Exclusion criterion
F	Generic and Brand Names of Medical Products to Identify Other Antibiotics with Oral and Injectable Routes for Exclusion	List of other antibiotics with both oral and injectable routes	Exclusion criterion
G	List of Healthcare Common Procedure Coding System (HCPCS) Codes to Identify Other Antibiotics with Oral and Injectable Routes for Exclusion	HCPCS codes of other antibiotics with both oral and injectable routes	Exclusion criterion
H	List of International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM) Diagnosis Codes to Identify Preexisting Conditions	List of preexisting conditions	Covariates in propensity score
I	List of International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM), Healthcare Common Procedure Coding System (HCPCS), and Current Procedural Terminology, Fourth Edition (CPT-4) Diagnosis and Procedure Codes to Identify Screening	List of screening tests	Covariates in propensity score
J	List of International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM) Codes to Identify Exposure Indications	List of indications	Covariates in propensity score
K	List of International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM) Codes Included in the Maternal Outcome Tree	List of maternal outcome tree	Outcome definition

Appendix	Appendix name	Description	Role
L	List of International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM) Codes to Identify Maternal Infection Outcomes Related to Antibiotic Indications	List of maternal outcomes related to infections	Alert classification

4.3 Defining Exposure

We will use NDC codes to identify pregnant women with oral macrolide or penicillin exposure. The list of generic and brand name medical products used to identify oral macrolide and penicillin exposure is in Appendix C. A pregnant woman will be defined as a macrolide user if she had at least one dispensing for an oral macrolide from 20 weeks of gestation to the day before the delivery date and without any days of supply of penicillins (oral or parenteral) from pregnancy start to 30 days after delivery date. Similarly, a pregnant woman will be defined as a penicillin user if she had at least one dispensing of an oral penicillin from 20 weeks of gestation to the day before delivery date without any days of supply of macrolides (oral or parenteral) from pregnancy start to 30 days after delivery date. The index date is the earliest date of the macrolides or penicillins dispensing during the assessed window. We exclude women with injectable macrolide or penicillin use (Appendix D: Generic and Brand Names of Medical Products to Identify Injectable Macrolides and Penicillins and Appendix E: List of HCPCS codes of Injectable Macrolides and Penicillin) from pregnancy start to index date. Women with any exposure to other antibiotics from pregnancy start date to index date were also excluded (Appendix F: Generic and Brand Names of Medical Products to Identify Other Antibiotics for Exclusion and Appendix G: List of HCPCS Codes of Other Antibiotics).

4.4 Confounding Control

In pregnancy studies focusing on maternal complications, the primary confounder is gestational age at the time of treatment initiation. With a wide exposure assessment window (i.e., from 20 weeks gestation to 30 days post-partum), it is necessary to balance cohorts on this confounder specifically. For more general confounding control in observational studies, propensity score methods are typically used to condense multiple confounders into a single index value. Patients with propensity scores in non-overlapping regions are typically removed. Percentiles of the propensity score will be used to define strata for the outcomes evaluation (see Section 5.5).

We will try two approaches to balance gestational age at treatment initiation. One approach stratifies by both a) gestational age at treatment initiation and b) a propensity score that does not include this variable (two-variable stratification approach). The second approach includes gestational age at treatment initiation in the propensity score and then stratifies on the propensity score alone. If balance is not achieved, then further stratification on the gestational age at treatment initiation will be performed (two-step stratification approach). More detail is described in Section 5.6. Stratification analysis.

In signal identification, we are evaluating thousands of outcomes which makes it difficult to identify confounders in the traditional sense of being required to be associated independently with both the outcome and exposure. Therefore, the propensity score model includes covariates that have been commonly found to be associated with a variety of adverse pregnancy outcomes.

A previous Sentinel project developed a general set of covariates for a propensity score specifically for the pregnant population, including maternal demographics, pre-existing conditions, prenatal screening, and measures of previous healthcare utilization (Appendix H:

List of ICD-10-CM to identify preexisting conditions and Appendix I: List of ICD-10-CM, HCPCS, and CPT-4 Diagnosis and Procedure Codes to Identify Screening), which we will also employ for this study.[18]

Additionally, we will include the specific indications for macrolides and penicillins to control for confounding by indication. The code list was selected based on a previous study evaluating the appropriateness of antibiotic use and is based on diseases categories from the Clinical Classification Software developed by the Agency for Healthcare Research and Quality (Appendix J: List of ICD-10-CM to identify Indications).[36,37]

5 TreeScan and Analysis Methods

5.1 Maternal Outcome Selection

In this study, we focus on pregnancy-related maternal health complications contained in ICD-10-CM Chapter 15 (Pregnancy, childbirth and the puerperium, O00-O9A).[38] This chapter contains important outcomes of interest such as pre-term labor, gestational diabetes, oligo- and polyhydramnios, pre/eclampsia and chorioamnionitis. We curated this chapter into the outcome tree used by TreeScan by excluding codes that could not have been reasonably associated with exposure (e.g., multiple gestation) or that are impossible given our cohort definition of livebirth deliveries (e.g., stillbirth). The final code list in the maternal outcome tree is in Appendix K.

5.2 Defining the Hierarchical Tree Structure for Maternal Outcomes

We use the hierarchical structure inherent in the ICD-10-CM coding system to define the maternal outcomes tree with six levels based on the curated outcome list as described above. An example of the maternal outcome tree is shown in Figure 2.

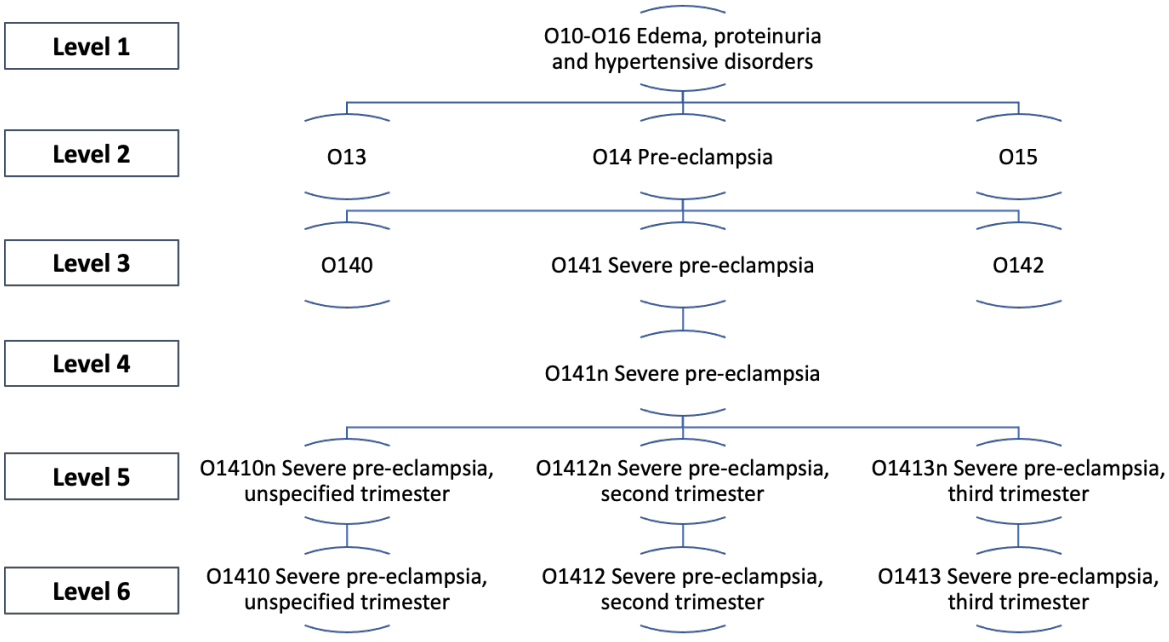

Figure 2. An example of the ICD-10-CM hierarchical tree for maternal outcomes. Each circle represents either a specific or composite clinical outcome on the tree, and is the unit of analysis in the scanning procedure for TreeScan.

5.3 Defining Incident Outcomes

In TreeScan, an incident outcome criterion is needed to identify new-onset conditions emerging after drug exposure and to prevent repeat counting the same condition that is being followed by a clinician. Incident outcomes will be identified from one day after the index date to 30 days after delivery. Incident outcomes will be defined as the first code under that level 3 clinical grouping in any care setting without any codes in the same grouping from 90 days before pregnancy start date to the outcome date. See Figure 2 to show an example of a level 3 clinical grouping (e.g., severe pre-eclampsia). A pregnant woman is allowed to have multiple incident outcomes as long as these outcomes meet the incident outcome criteria.

5.4 Calculating Expected Outcome Counts in TreeScan

For each outcome in the tree, TreeScan compares the observed and expected number of events among the exposed group to evaluate for a potential elevated frequency. TreeScan supports two probability models to calculate the expected number of events: Bernoulli and Poisson. The Bernoulli model maximizes bias control at the cost of precision through a fixed ratio matching technique.[15,18] Matched sets are required to be of uniform size throughout the analysis. This requirement is hard to maintain with varying gestational length. Thus, a proper Bernoulli analysis would need to match on both gestational age at treatment initiation and gestational age at delivery in addition to other propensity score criteria. Sample size loss in these circumstances is expected to be considerable.

For the Poisson model, background rates for each outcome are based on a control group. These rates are used to calculate the expected event count in the exposed group using indirect standardization after stratification on a propensity score.[14,18] We chose the Poisson model rather than the Bernoulli model to evaluate maternal complications because it maximizes sample size while maintaining good confounding control. The Poisson model does not require matching or the same follow-up time.

In the Poisson model, the null hypothesis is that maternal complications are expected to occur in proportion to the expected count.[14] To meet the assumptions of the Poisson distribution, the background rate should be constant over the unit time being measured. However, maternal outcome manifestation often depends on gestational age. In brief, the risk of specific maternal complications tends to have a skewed distribution (e.g., Figure 3) or is confined to only a specific time period such as at delivery or post-partum. Therefore, the required assumption of having a constant risk over follow-up time is unlikely to be satisfied for many maternal outcomes if we look at outcome rates in units of weeks or months of pregnancy.

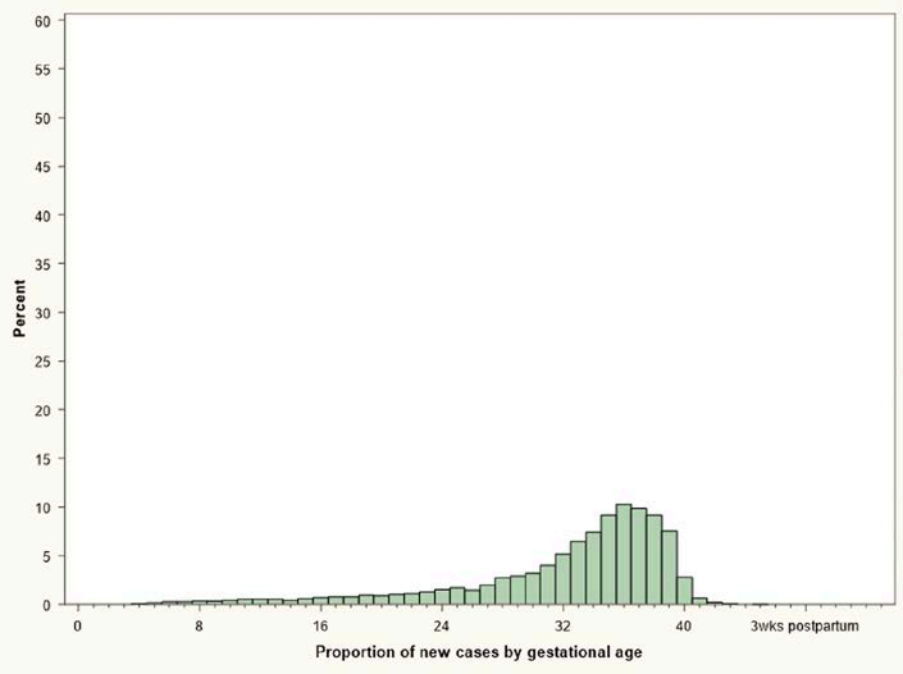

Figure 3. Proportion of women with newly diagnosed pre-eclampsia by gestational age

Instead of looking at rates per unit time, one can also use total pregnant women as the demoninator in the Poisson model (i.e., rate of gestational diabetes per pregnant women evaluated). Pregnancy duration is still variable in our study because we *only* balance pregnant women on gestational age at treatment initiation rather than gestatational age at treatment initiation *and* gestational age at delivery. Given the interchangeability of the two antibiotic classes and no documented effect on preterm deliveries, we do not expect different distributions of gestational age at delivery, therefore using total persons as the denominator in the Poisson model will be most appropriate for calculating the observed and expected outcome counts.

5.5 The Poisson TreeScan Statistic

Tree-based scan statistics can be unconditional or conditional on the total number of observed outcomes in the dataset.[14] The conditional statistic controls for increases in general healthcare utilization believed to be unrelated to the exposure of interest.[14] Because there is increased healthcare utilization around delivery that is unlikely to be related to antibiotic exposure, the conditional Poisson version is more appropriate. The conditional log likelihood ratio (LLR) based test statistic T can be calculated for the Poisson model as follows:

$$LLR(G) = \left[c_G ln\left(\frac{c_G}{n_G}\right) + (C - c_G) ln\left(\frac{C - c_G}{N - n_G}\right)\right] I\left(\frac{c_G}{n_G} > \frac{C - c_G}{N - n_G}\right)$$

$$T = \max_G LLR(G)$$

Where: T = conditional Poisson tree scan statistic
c_G = observed cases in the treatment group for a given maternal outcome
n_G = expected cases in the treatment group for a given maternal outcome
C = total number of maternal outcomes in the risk window summed over the tree
N = total number of expected maternal outcomes summed over the tree
G = maternal outcome of interest

Random datasets are generated under the null hypothesis, and the test statistic T is calculated for each random dataset. The Monte Carlo based p-value is the rank of the test statistic in the real dataset divided by the number of replicated random datasets plus 1. If the statistical significance is set as alpha=0.05, an alert for an outcome occurs if the test statistic of that outcome in the real dataset ranks in the top 5% of all test statistics among the real and replicated datasets.

5.6 Stratification Analysis

After trimming non-overlapping regions of the propensity score, the cohort will be stratified based on propensity score quartiles and gestational age at treatment initiation by the two approaches described earlier: two-variable stratification and two-step stratification.

In the two-variable stratification approach, we divide the trimmed cohort into quartiles of the propensity score and then, within each propensity score stratum, we stratify based on 4-week gestational age groups at treatment initiation: at 20-23, 24-27, 28-31, 32-35, and ≥36 weeks. As result, we have a total of 20 strata as depicted in *Figure 4*.

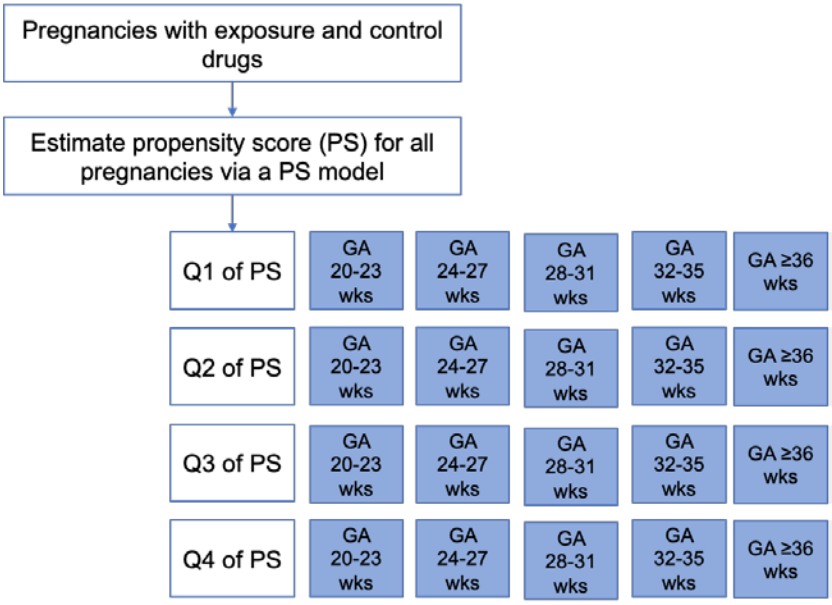

Figure 4. Two-variable stratification approach

Note: GA – gestational age; PS: propensity score; Q1-Q4 – 1st, 2nd, 3rd, and 4th quartile

In the two-step approach, we add gestational age at treatment initiation as a categorical covariate in the propensity score model (Note: gestational age at treatment initiation is not included in the propensity score model for the two-variable approach). We first stratify the trimmed cohort into quartiles of the propensity score and then examine the distribution of gestational age in each stratum. We only stratify based on gestational age at treatment initiation in a given propensity score stratum if we detect imbalanced distributions of that variable (Figure 5).

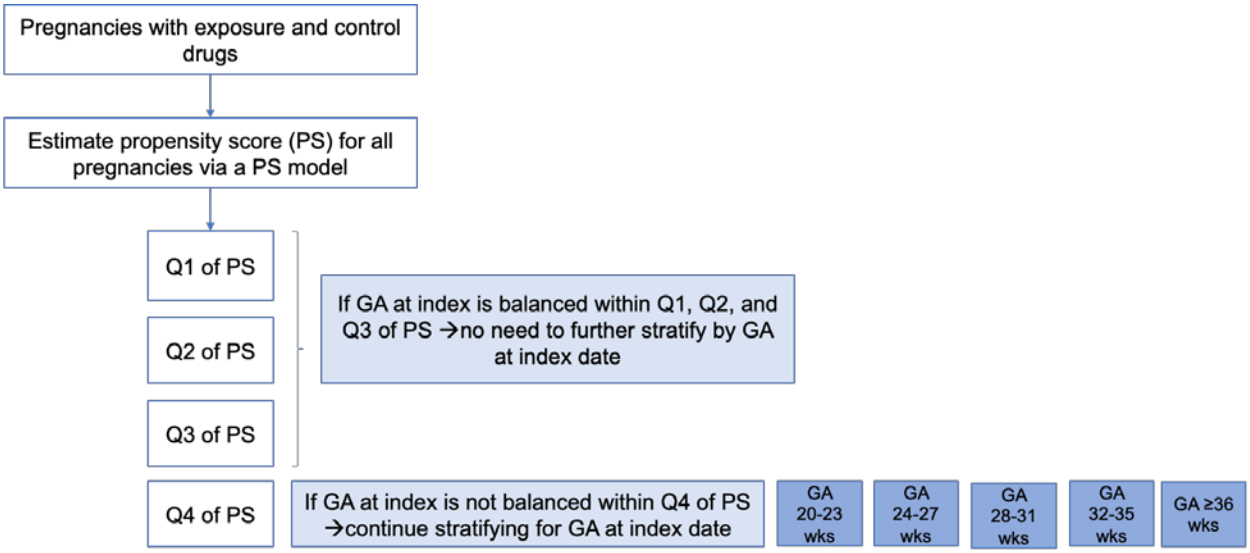

Figure 5. Two-step stratification approach

Note: GA – gestational age; PS: propensity score; Q1-Q4 – 1st, 2nd, 3rd, and 4th quartile

Outcome risk varies by gestational age so these strata create a more similar followup time within each stratum. For example, women that initiate treatment at 20-23 weeks are grouped separately than women that initiate treatment at 36 weeks. More granular strata of gestational age at treatment initiation ensures more similarity in the time available to experience certain maternal complications, and theoretically better control of confounding. However, too many strata may result in small sample sizes in each stratum and a zero count of maternal outcomes in a non-informative stratum.

The two-variable approach increases accuracy in estimating the expected count, but naturally creates more strata which has the disadvantage of smaller sample sizes per stratum. The two-step approach, in turn, aims to prevent the issue of a low count of outcomes while still addressing confounding by gestational age at treatment initiation as needed, and might offer some efficiencies. Both stratification approaches have advantages and disadvantages, so we aim to explore both methods.

5.7 Identifying Alerts Using TreeScan

In the main analysis, we will conduct the hypothesis testing at levels 3 to 5. The threshold for an alert will be $p \leq 0.05$. Because the hypothesis testing in TreeScan is one-sided, we will conduct two analyses to fully capture potential alerts: one compares macrolides vs penicillins and one compares penicillins vs macrolides. Because no previous studies have raised concerns about maternal outcomes associated with macrolide or penicillin use, alerts will be triaged to be "expected" (i.e., alerts representing labelled conditions or those related to the underlying drug indications) or unclassified. Appendix L shows a list of ICD-10-CM codes of maternal infection outcomes that are related to antibiotic indications.

5.8 Sensitivity Analyses

We will conduct several sensitivity analyses to evaluate TreeScan performance.

First, we will conduct hypothesis testing at level 2 nodes, which have a broader outcome definition, and the incident outcome criterion is redefined at level 2.

Second, an increased specificity in outcome measurement will help to decrease outcome misclassification, resulting in enhancing risk estimate accuracy. We will therefore restrict outcomes to those captured from inpatient or emergency department visits while still requiring no related outcomes in any setting prior to outcome occurrence.

Third, we will vary the number of strata of gestational age at treatment initiation and for the propensity score for both stratification approaches. We expect a larger number of strata will increase the ability to control for bias, but will also result in a loss of precision.

Table 2 summarizes all analyses in this protocol.

Table 2. Summary analysis scenarios

#	Analysis scenarios	Stratification approach	Cut-off of gestation age at index	Cut-off of propensity score	Incident outcome	Incident outcome criteria
Main analyses						
1	Vary stratification approach	Two-variable approach	Every four weeks	Quartiles	Levels 3 to 5 at any care setting	At level 3 in any care setting
2		Two-step approach	Every four weeks	Quartiles	Levels 3 to 5 at any care setting	At level 3 in any care setting
Sensitivity analyses						
3	Add level 2 for hypothesis testing	Two-variable approach	Every four weeks	Quartiles	Levels 2 to 5 at any care setting	At level 2 in any care setting
4		Two-step approach	Every four weeks	Quartiles	Levels 2 to 5 at any care setting	At level 2 in any care setting
5	Restrict to inpatient or emergency department visit	Two-variable approach	Every four weeks	Quartiles	Levels 3 to 5 at inpatient or emergency department visits	At level 3 in any care setting
6		Two-step approach	Every four weeks	Quartiles	Levels 3 to 5 at inpatient or emergency department visits	At level 3 in any care setting
7	Vary cut-off of gestational age at treatment initiation	Two-variable approach	Every two weeks	Quartiles	Levels 3 to 5 at any care setting	At level 3 in any care setting
8		Two-step approach	Every two weeks	Quartiles	Levels 3 to 5 at any care setting	At level 3 in any care setting

#	Analysis scenarios	Stratification approach	Cut-off of gestation age at index	Cut-off of propensity score	Incident outcome	Incident outcome criteria
9		Two-variable approach	Every six weeks	Quartiles	Levels 3 to 5 at any care setting	At level 3 in any care setting
10		Two-step approach	Every six weeks	Quartiles	Level 3 to 5 at any care setting	At level 3 in any care setting
11	Vary cut-off of GA propensity score	Two-variable approach	Every six weeks	Deciles	Levels 3 to 5 at any care setting	At level 3 in any care setting
12		Two-step approach	Every six weeks	Deciles	Levels 3 to 5 at any care setting	At level 3 in any care setting

6 Future Considerations

There are several limitations which are considerations for future studies. First, the follow-up time to evaluate maternal outcomes depends on gestational age at delivery. If the exposure drug increases the risk for preterm births, the risk period of having maternal outcomes is no longer the same between the exposure and control drugs and the analysis becomes a competing risk problem. An appropriate method using survival analysis and accounting for this competing risk issue will need to be investigated in the future so that TreeScan can be used for drugs with or without impact on preterm risk. Second, the current project only focuses on maternal outcomes among mothers with livebirths because pregnancy identification algorithms for non-livebirth pregnancies have not been fully developed. Evaluating maternal outcomes among mothers with livebirth and non-livebirth outcomes will provide a more complete picture of drug adverse effects during pregnancy. Third, in claims data, we use delivery date to determine pregnancy start date, which is common practice for pregnancy and birth outcomes studies. Thus, women whose pregnancies do not result in livebirth (e.g., non-livebirth outcomes or maternal demise) are excluded. If a drug increases risk of maternal death, this exclusion may create a bias that attenuates the drug's adverse effect. Using other data sources or alternative algorithms to identify start of pregnancy can address this limitation.

7 References

1. Mitchell AA, Gilboa SM, Werler MM, et al. Medication use during pregnancy, with particular focus on prescription drugs: 1976-2008. *Am J Obstet Gynecol.* Jul 2011;205(1):51.e51-58.

2. Adam MP, Polifka JE, Friedman JM. Evolving knowledge of the teratogenicity of medications in human pregnancy. *Am J Med Genet C Semin Med Genet.* Aug 2011;157C(3):175-182.

3. Lynch MM, Squiers LB, Kosa KM, et al. Making Decisions About Medication Use During Pregnancy: Implications for Communication Strategies. *Matern Child Health J.* 01 2018;22(1):92-100.

4. Food and Drug Administration. Postapproval Pregnancy Safety Studies Guidance for Industry. 2019.
5. Gliklich RE, Dreyer NA, Leavy MB. Registries for Evaluating Patient Outcomes: A User's Guide. 2014.
6. FDA Public Meeting: Study Approaches and Methods to Evaluate the Safety of Drugs and Biological Products during Pregnancy in the Post-Approval Setting.
7. Li Q, Andrade SE, Cooper WO, et al. Validation of an algorithm to estimate gestational age in electronic health plan databases. *Pharmacoepidemiol Drug Saf.* May 2013;22(5):524-532.
8. Andrade SE, Davis RL, Cheetham TC, et al. Medication Exposure in Pregnancy Risk Evaluation Program. *Matern Child Health J.* Oct 2012;16(7):1349-1354.
9. Taylor LB, Steven T, Stojanovic D, Toh S, Fazio-Eynullayeva E, Petrone A, Lyons JG. Utility of prenatal tests to estimate pregnancy start in Sentinel. *Abstracts of the 35th International Conference on Pharmacoepidemiology & Therapeutic Risk Management, Pennsylvania Convention Center, Philadelphia, PA, USA, August 24-28, 2019.* Vol 28 Suppl 2: Pharmacoepidemiol Drug Saf; 2019:5-586
10. Matcho A, Ryan P, Fife D, Gifkins D, Knoll C, Friedman A. Inferring pregnancy episodes and outcomes within a network of observational databases. *PLoS One.* 2018;13(2):e0192033.
11. Kulldorff M, Fang Z, Walsh SJ. A tree-based scan statistic for database disease surveillance. *Biometrics.* Jun 2003;59(2):323-331.
12. Yih WK, Maro JC, Nguyen M, et al. Assessment of Quadrivalent Human Papillomavirus Vaccine Safety Using the Self-Controlled Tree-Temporal Scan Statistic Signal-Detection Method in the Sentinel System. *Am J Epidemiol.* 06 2018;187(6):1269-1276.
13. Yih WK, Kulldorff M, Dashevsky I, Maro JC. Using the Self-Controlled Tree-Temporal Scan Statistic to Assess the Safety of Live Attenuated Herpes Zoster Vaccine. *Am J Epidemiol.* 07 2019;188(7):1383-1388.
14. Maro JC, Nguyen MD, Dashevsky I, Baker MA, Kulldorff M. Statistical Power for Postlicensure Medical Product Safety Data Mining. *EGEMS (Wash DC).* Jun 2017;5(1):6.
15. Wang SV, Maro JC, Baro E, et al. Data Mining for Adverse Drug Events With a Propensity Score-matched Tree-based Scan Statistic. *Epidemiology*. 11 2018;29(6):895-903.
16. Wang SV, Maro JC, Gagne JJ, et al. A General Propensity Score for Signal Identification using Tree-Based Scan Statistics. *Am J Epidemiol.* Feb 2021.
17. Kulldorff M, Dashevsky I, Avery TR, et al. Drug safety data mining with a tree-based scan statistic. *Pharmacoepidemiol Drug Saf.* May 2013;22(5):517-523.
18. Suarez EA, Nguyen M, Zhang D, et al. Use of the Tree-Based Scan Statistic for Surveillance of Infant Outcomes Following Maternal Perinatal Medication Use 2021.
19. Huybrechts KF, Kulldorff M, Hernández-Díaz S, et al. Active Surveillance of the Safety of Medications Used During Pregnancy. *Am J Epidemiol.* Jun 2021;190(6):1159-1168.
20. Maro JC, Inna D, Kulldorff M. Postlicensure Medical Product Safety Data-mining: Power Calculation for Bernoulli Data. 2017.
21. Palmsten K, Hernández-Díaz S, Chambers CD, et al. The Most Commonly Dispensed Prescription Medications Among Pregnant Women Enrolled in the U.S. Medicaid Program. *Obstet Gynecol.* Sep 2015;126(3):465-473.
22. Fan H, Li L, Wijlaars L, Gilbert RE. Associations between use of macrolide antibiotics during pregnancy and adverse child outcomes: A systematic review and meta-analysis. *PLoS One.* 2019;14(2):e0212212.
23. Mallah N, Tohidinik HR, Etminan M, Figueiras A, Takkouche B. Prenatal Exposure to Macrolides and Risk of Congenital Malformations: A Meta-Analysis. *Drug Saf.* 03 2020;43(3):211-221.

24. Leke AZ, Dolk H, Loane M, et al. Macrolide and lincosamide antibiotic exposure in the first trimester of pregnancy and risk of congenital anomaly: A European case-control study. *Reprod Toxicol.* 03 2021;100:101-108.
25. Muanda FT, Sheehy O, Bérard A. Use of antibiotics during pregnancy and the risk of major congenital malformations: a population based cohort study. *Br J Clin Pharmacol.* Nov 2017;83(11):2557-2571.
26. Daniel S, Doron M, Fishman B, Koren G, Lunenfeld E, Levy A. The safety of amoxicillin and clavulanic acid use during the first trimester of pregnancy. *Br J Clin Pharmacol.* 12 2019;85(12):2856-2863.
27. Jepsen P, Skriver MV, Floyd A, Lipworth L, Schønheyder HC, Sørensen HT. A population-based study of maternal use of amoxicillin and pregnancy outcome in Denmark. *Br J Clin Pharmacol.* Feb 2003;55(2):216-221.
28. Czeizel AE, Rockenbauer M, Sørensen HT, Olsen J. Augmentin treatment during pregnancy and the prevalence of congenital abnormalities: a population-based case-control teratologic study. *Eur J Obstet Gynecol Reprod Biol.* Aug 2001;97(2):188-192.
29. Cooper WO, Hernandez-Diaz S, Arbogast PG, et al. Antibiotics potentially used in response to bioterrorism and the risk of major congenital malformations. *Paediatr Perinat Epidemiol.* Jan 2009;23(1):18-28.
30. Damkier P, Brønniche LMS, Korch-Frandsen JFB, Broe A. In utero exposure to antibiotics and risk of congenital malformations: a population-based study. *Am J Obstet Gynecol.* 12 2019;221(6):648.e641-648.e615.
31. Puhó EH, Szunyogh M, Métneki J, Czeizel AE. Drug treatment during pregnancy and isolated orofacial clefts in hungary. *Cleft Palate Craniofac J.* Mar 2007;44(2):194-202.
32. Lin KJ, Mitchell AA, Yau WP, Louik C, Hernández-Díaz S. Maternal exposure to amoxicillin and the risk of oral clefts. *Epidemiology.* Sep 2012;23(5):699-705.
33. Kenyon SL, Taylor DJ, Tarnow-Mordi W, Group OC. Broad-spectrum antibiotics for preterm, prelabour rupture of fetal membranes: the ORACLE I randomised trial. ORACLE Collaborative Group. *Lancet.* Mar 2001;357(9261):979-988.
34. Kenyon S, Taylor DJ, Tarnow-Mordi WO, Group OC. ORACLE--antibiotics for preterm prelabour rupture of the membranes: short-term and long-term outcomes. *Acta Paediatr Suppl.* 2002;91(437):12-15.
35. Sentinel. Cohort Identification and Descriptive Analysis (CIDA) Module - Pregnancy Episodes Cohort Identification Strategy.
36. Chua KP, Fischer MA, Linder JA. Appropriateness of outpatient antibiotic prescribing among privately insured US patients: ICD-10-CM based cross sectional study. *BMJ.* 01 2019;364:k5092.
37. Agency for Healthcare Research and Quality. Tools Archive for the Chronic Condition Indicators for ICD-10-CM.
38. National Center for Health Statistics. International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM). 2021.