Emulating a Target Trial of Interventions Initiated During Pregnancy with Healthcare Databases: The Example of COVID-19 Vaccination

Sonia Hernández-Díaz,[a,b] Krista F. Huybrechts,[c] Yu-Han Chiu,[a] Jennifer J. Yland,[b,d] Brian T. Bateman,[c] and Miguel A. Hernán[a,b,e]

Background: Observational studies are often the only option to estimate effects of interventions during pregnancy. Causal inference from observational data can be conceptualized as an attempt to emulate a hypothetical pragmatic randomized trial: the target trial.

Objective: To provide a step-by-step description of how to use healthcare databases to estimate the effects of interventions initiated during pregnancy. As an example, we describe how to specify and emulate a target trial of COVID-19 vaccination during pregnancy, but the framework can be generally applied to point and sustained strategies involving both pharmacologic and non-pharmacologic interventions.

Methods: First, we specify the protocol of a target trial to evaluate the safety and effectiveness of vaccination during pregnancy. Second, we describe how to use observational data to emulate each component of the protocol of the target trial. We propose different target trials for different gestational periods because the outcomes of interest vary by gestational age at exposure. We identify challenges that affect (i) the target trial and thus its observational emulation (censoring and competing events), and (ii) mostly the observational emulation (confounding, immortal time, and measurement biases).

Conclusion: Some biases may be unavoidable in observational emulations, but others are avoidable. For instance, immortal time bias can be avoided by aligning the start of follow-up with the gestational age at the time of the intervention, as we would do in the target trial. Explicitly emulating target trials at different gestational ages can help reduce bias and improve the interpretability of effect estimates for interventions during pregnancy.

Keywords: Target trial; pregnancy; Healthcare databases: Immortal time bias; Methods; COVID-19; Vaccines; Drugs; safety; Effectiveness

Most randomized trials, like the preapproval trials of COVID-19 vaccines,[1–4] do not include pregnant women. Even when a few pregnant women are intentionally or unintentionally included,[5] the number is insufficient for precisely estimating effects at all gestational ages in both mothers and their offspring.[6,7]

Consequently, observational healthcare databases—electronic medical records and insurance claims—are increasingly used to evaluate interventions during pregnancy.[8] Causal inference from observational databases has two steps. First, the causal question of interest needs to be unambiguously formulated, which can be achieved by expressing it in terms of a hypothetical pragmatic randomized trial—the target trial. Second, the observational data can be used to explicitly emulate each component of the target trial protocol as closely as possible.[9–11]

Here we provide a step-by-step description of the use of healthcare databases to estimate the effects of interventions initiated during pregnancy, challenges that frequently arise, and ways to address them. The framework can be generally applied to point and sustained strategies, as well as to pharmacologic and non-pharmacologic interventions. As an

example, we specify and emulate a target trial of COVID-19 vaccination.

DESIGN OF RANDOMIZED TRIALS DURING PREGNANCY

The design of randomized trials in pregnancy—and therefore its observational emulation— requires explicit considerations about gestational age and pregnancy losses. Analogous considerations apply to trials in non-pregnant populations in which age plays the role of gestational age and death plays the role of pregnancy loss or the end of pregnancy. However, the condensed biologic timeframe makes these issues more prominent for trials in pregnancy.

Gestational Age at Enrollment

Gestational age is typically measured in weeks since the last menstrual period (LMP). The range of gestational age at enrollment among the trial's participants must be relevant to the outcomes of interest. When the outcome is a congenital malformation, the trial must start before week 12 of pregnancy, when most organogenesis ends. When the outcome is spontaneous abortion, the trial must start before week 20 because, by most definitions, no spontaneous abortions occur after week 20. When the outcome is preterm birth, the trial must start before week 37 because, by definition, no preterm births occur after week 37. When the outcome is a maternal outcome that may happen at any time (e.g., COVID-19 infection), the trial can enroll participants at any gestational age, but the effect of interest may need to be evaluated separately in groups defined by gestational age (if, for example, vaccine immunogenicity depends on gestational age at vaccination).

Lack of attention to gestational age at enrollment is a common source of confusion in pregnancy studies. When evaluating outcomes that accumulate during specific windows (e.g., spontaneous abortions, preterm delivery), the risk depends on the distribution of gestational age at enrollment because the time left for outcome accumulation is shorter for later gestational ages. Consider two randomized trials: in the first one all participants enroll at 36 weeks of pregnancy, while in the second one all participants enroll at 25 weeks. The risk of preterm birth will be greater in the second trial because there is a longer period to diagnose preterm births. In the extreme, a study of pregnancies after 37 weeks will have no preterm births and a study of pregnancies after 20 weeks will have no spontaneous abortions. Comparing average rates or hazards does not solve this problem because rates change substantially over gestational age. A common mistake is comparing results between studies with different distributions of gestational age at enrollment or, worse, with that in a population that followed pregnancies from LMP.

Types of Pregnancy Trials According to Gestational Age at Enrollment

We can classify target trials in pregnancy into four classes depending on the range of gestational age implied by the outcome of interest (Table 1)[12–14]: trials that start early in the first trimester for malformations (periconceptional trials), trials that start before 20 weeks for spontaneous abortions (early pregnancy trials), trials that start after 20 weeks for other outcomes (late pregnancy trials), and trials that start at any gestational age for non-pregnancy-specific maternal outcomes such as COVID-19 infection (any-trimester pregnancy trials). The early pregnancy trial could also evaluate the effects on the later outcomes of early pregnancy exposure (e.g., through a potential effect on placentation). Some outcomes (e.g., maternal severe COVID-19) can be studied in all four trials, but the most flexible approach would be to enroll at any gestational age. The implementation of each type of trial varies because the procedures necessary to identify participants in the first trimester of pregnancy are different from those used to identify participants after 20 weeks of gestation.

Gestational Age at Administrative End of Follow-Up

The period between enrolment and administrative end of follow-up needs to be long enough to cover a term pregnancy (i.e., a gestational age over 39 weeks). Otherwise, if only completed pregnancies are included in the analyses, short pregnancies (spontaneous abortions, preterm births) would be overrepresented towards the end of the study period.[15]

To avoid this problem, we can require that participants are enrolled only if their LMP is at least 12 months before the administrative end of follow-up. As illustrated in Figure 1, for a trial expected to end in December 2022, recruitment would end in December 2021. The 12-month period is the sum of the approximately 9 months required for a term pregnancy plus 3 months after birth that is often necessary for ascertainment of many outcomes (e.g., some malformations). This period would need to be extended for other outcomes (e.g., neurodevelopmental outcomes) that require longer periods for diagnosis. In some trials, the length of the period is chosen for practical reasons. For example, in vaccine effectiveness trials, 3 months may be chosen as a compromise between the recording of events affected by vaccination and timely completion of the trial.[8]

Pregnancy Losses

Pregnancy losses are competing events for later outcomes in pregnancy trials.[16,17] Until a consensus is reached on how to handle competing events in pregnancy trials, we propose reporting the risk of livebirth together with the joint and conditional probabilities of neonatal outcomes, e.g., the joint risk of "having a livebirth and admission to a neonatal intensive care unit (NICU)" and the conditional risk of NICU admission among livebirths.[17,18] Continuous outcomes (e.g., birth weight) are undefined following pregnancy losses.

Besides being competing events for later outcomes (e.g., NICU admission), pregnancy losses may be caused by earlier outcomes (e.g., malformations), which complicates the causal interpretation of the estimates in studies restricted

TABLE 1. Four Types of Target Trials Initiated During Pregnancy Depending on the Range of Gestational Age at Enrolment Implied by the Outcome of Interest

Outcome Examples	Main Gestational Exposure Window	Type of Pregnancy Trial
Major congenital malformations	Before 12 weeks	Periconceptional
Spontaneous abortion	Before 20 weeks	Early pregnancy
Elective Termination		
Preterm Delivery	Before 37 weeks	Early pregnancy or late pregnancy
Stillbirth	During pregnancy	Early pregnancy or late pregnancy
Low birth weight (birth weight)		
Small for Gestational Age		
Microcephaly (head circumference)		
Maternal obstetric complications (gestational diabetes, preeclampsia, postpartum hemorrhage, labor induction, Cesarean section, maternal death)		
Mode of delivery	After 20 weeks	Late pregnancy
Neonatal complications (neonatal admission to intensive care unit, neonatal death)		
Infant neurodevelopment		
COVID-19	At any time	Any-trimester
Adverse events (e.g., thrombotic)		
Tolerability		
Immunogenicity		
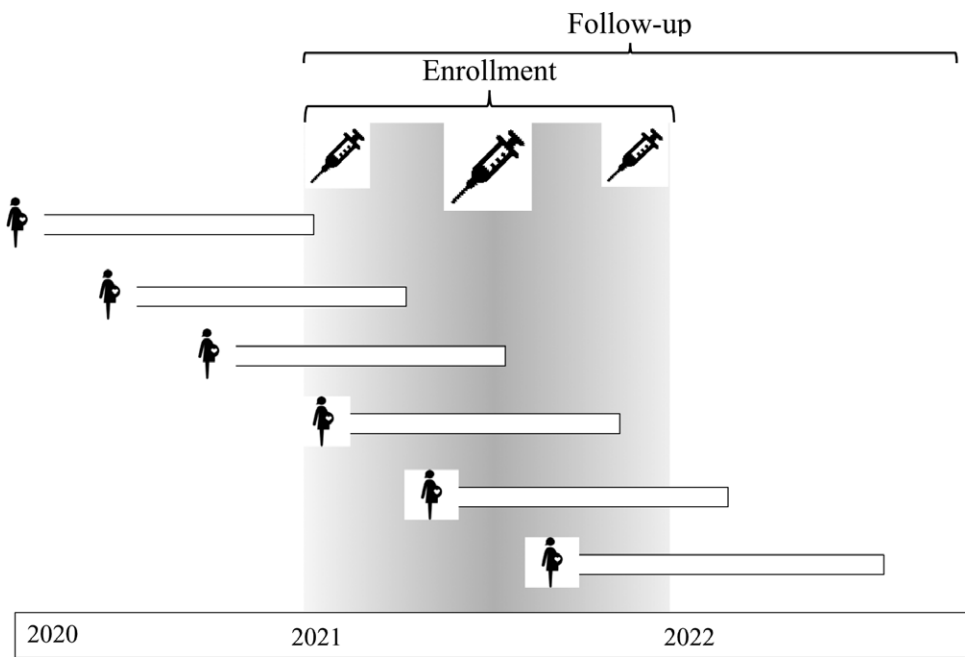
FIGURE 1. Target trial of vaccination in pregnancy during 2021 and follow-up of pregnancies throughout 2022. If we had required a "completed pregnancy" in 2021, the study would capture a disproportionally large number of short pregnancies (spontaneous abortions, preterm births) towards the end.

to livebirths. As an extreme scenario, pregnancy losses may arise in periconceptional trials when an intervention causes immediately lethal malformations, a phenomenon known as terathanasia.[19] Then, because exposed embryos with the malformation die before the malformation is diagnosed, the target trial may find a lower risk of the malformation in exposed livebirths than in unexposed livebirths. The inference would then

be that the intervention reduces the risk of the malformation when it increases it.

Even if the study is not restricted to livebirths, pregnancy losses may prevent the ascertainment of malformations because a pathology study is rarely available for spontaneous abortions and may not be recorded for terminations. Therefore, the prevalence of malformations identified at birth may underestimate the true

risk (incidence) of malformations, particularly for those that may themselves result in spontaneous abortions or for which termination is often chosen after prenatal diagnosis (*e.g.*, neural tube defects).[20,21] (eAppendix 1, http://links.lww.com/EDE/B980). The impact of pregnancy losses may be explored by estimating risk bounds[22] or conducting probabilistic bias analyses.[23,24]

Periconceptional Trials may be Specified but Cannot be easily Carried Out

Ideally, periconceptional target trials would assign the interventions at conception. This design would be straightforward in experiments with animals whose time of impregnation is known, but not in human trials because the time of conception is not under the investigators' control or is not even known. One approach is the recruitment of women who are planning to conceive,[25] but there is no guarantee that they will do so. The next best design for periconceptional trials would be to assign the interventions 2 weeks after conception (an average of 4 weeks after LMP) when the embryo is susceptible to teratogens and women may already suspect pregnancy. However, this design is also impractical for human trials because most pregnancies remain undiagnosed during that period and thus participants would not be enrolled in time.

Even when a periconceptional trial cannot be realistically carried out, we may still be able to emulate it using observational healthcare databases. If the database records accurate treatment and LMP dates, we can identify exposures initiated in the weeks following conception even when the pregnancy had not yet been diagnosed.

An increased risk of malformations would be translated into a recommendation to avoid exposure. But, because about half of pregnancies are unplanned, the recommendation would not be to avoid exposure in the first weeks after conception but rather to avoid exposure at any time in which pregnancy is a possibility. Like periconceptional trials, periconceptional interventions may be specified but not easily implemented.

SPECIFICATION OF A TARGET TRIAL DURING PREGNANCY

With the above considerations in mind, we can proceed to specify the protocol of the target trial(s) that we would like to emulate using the observational database. As an illustration, we outline the protocol of four target trials to evaluate the effectiveness and safety of a third dose (booster) of COVID-19 vaccination. All four trials are identical except for the gestational age at enrollment and the outcomes of interest. We consider target trials that rely exclusively on the healthcare database to ascertain potential eligibility (later confirmed through an interview) and to ascertain outcomes. A brief description of each component of the protocol of the target trials follows (Table 2).

Eligibility Criteria

Pregnant women (defined as human females) aged 18–50 years in 2021 with primary vaccination (2 doses) completed at least 6 months ago, no previous booster dose, no positive SARS-CoV-2 test and enrolled in the healthcare system for at least 12 months (to ensure records exist). Pregnant Women are eligible until 12 weeks after LMP for the periconceptional trial, until 20 weeks for the early pregnancy trial, after 20 weeks for the late pregnancy trial, and throughout pregnancy for the any-trimester pregnancy trial.

Treatment (Vaccination) Strategies

(1) An mRNA vaccine booster dose at enrollment, and (2) No vaccine doses during pregnancy.

Assignment Procedures

Individuals are randomly assigned to one strategy and are aware of their assignment. To guarantee perfect balance in the distribution of gestational age, calendar time, and region, we could select pairs of women with the same values of these factors and randomly assign one member of the pair to vaccination and the other one to no vaccination.

Outcomes

The effectiveness outcomes are laboratory-confirmed maternal or infant COVID-19 diagnosis and severe COVID-19 requiring hospitalization, intensive care unit (ICU) admission, or death.[26] The safety outcomes[12–14] include a major congenital malformation, spontaneous abortion, and other maternal or infant complications (Table 1).

Follow-up Period

Follow-up starts at assignment and ends at either the occurrence of an outcome of interest, 140 days after LMP (for spontaneous abortions) or 90 days after birth (for other outcomes), pregnancy loss (for safety outcomes), death, or loss to follow-up, whichever occurs earliest.

Causal Contrast

Intention-to-treat effect and per-protocol effect.[27]

Data Analysis

In the intention-to-treat analysis, for each outcome, we compare the risks (cumulative incidences) in each group defined by assignment through differences and ratios (vaccine effectiveness is traditionally defined as one minus the risk ratio or hazard ratio). We can estimate cumulative incidence curves from assignment via the Kaplan–Meier estimator or a pooled logistic model and compare means for continuous outcomes via linear regression models. We can adjust for selection bias due to loss of follow-up under the assumption that the measured variables (in pregnancy trials often only baseline variables measured at time zero) include approximately all risk factors that predict loss to follow-up.

The per-protocol analysis is the same as the intention-to-treat analysis except that individuals are censored if they deviate from the protocol, e.g., by declining the booster if assigned to booster or obtaining it outside of the trial if assigned to no booster. We can adjust for selection bias due

TABLE 2. Specification and Emulation of a Target Trial During Pregnancy Using a Healthcare Administrative Database. As an example, the table describes 4 trials of COVID-19 booster vaccination with different outcomes: periconceptional trial, early pregnancy trial, late pregnancy trial and any-trimester trial.

Protocol Component	Target Trial	Emulation
Eligibility Criteria	• Enrollment period: January to December 2021 • Pregnant: Gestation under 12 weeks for periconceptional trial, under 20 for early trial, over 20 for late trial, and unrestricted for any-trimester trial. • Aged 18–50 years • No active SARS-CoV-2 infection (past infection allowed) • Primary vaccination schedule completed at least 6 months ago • Enrolled in the healthcare system of interest (e.g., insurance with prescription benefits or electronic health records) at least 12 months	Same. Eligibility criteria are identified via codes in the database
Treatment Strategies	1) Vaccine booster at enrolment 2) No booster during pregnancy	Same. Vaccination, including brand and date, is identified based on pharmacy dispensations and procedure codes
Assignment Procedures	Individuals are randomly assigned to one of the two vaccination strategies and are aware of the strategy to which they have been assigned.	Individuals assigned to each vaccination strategy are assumed to be comparable conditional on baseline covariates: gestational week, calendar month, age, month, region, chronic conditions, health care utilization, prior COVID-19, *etc.*
Follow-up Period	• Starts at vaccine assignment • Ends at the occurrence of an outcome of interest, 140 days after LMP (for spontaneous abortion) or 90 days after birth (for other outcomes), pregnancy loss, death, or loss to follow-up (disenrollment from insurance), whichever occurs earliest	• Starts at vaccine administration • Same except for loss to follow-up. Because pregnancy status is often ascertained by the end-of-pregnancy outcome, which forces a "complete case" approach
Outcome	Safety: • Periconceptional trial: a major congenital malformation • Early pregnancy trial: Spontaneous abortion • Late pregnancy trial: Stillbirth, preterm birth, microcephaly, gestational diabetes, preeclampsia, preterm delivery, labor induction, Cesarean section, postpartum hemorrhage, maternal death, small for gestational age, need for NICU admission, and neonatal death. Effectiveness: • Any-trimester trial: Maternal or infant COVID-19 with onset postvaccination	Same. Diagnoses are identified with algorithms based on combinations of codes in the database (this may be the same approach used in the target trial, which is pragmatic by definition)
Causal Contrasts of Interest	Intention-to-treat effect Per-protocol effect	Observational analog of per-protocol effect
Analyses	Intention-to-treat analysis: estimate outcome risks in each group and compare them through risk differences and risk ratios with adjustment for loss to follow-up. Per-protocol analysis: same as intention-to-treat analysis with censoring at non-adherence and further adjustment for predictors of adherence and the outcome.	Same per-protocol analysis, except for restriction to pregnancies without loss to follow-up

to protocol deviation[28] under the assumption that the measured variables include approximately all risk factors that predict adherence.[29,30] To adjust for selection bias due to loss to follow-up or protocol deviation, we can use inverse probability weighting, standardization, or, when only baseline variables are measured, methods like matching and outcome regression.

We can carry out subgroup analyses by gestational age at enrollment and by other characteristics of interest. 95% confidence intervals may be estimated via bootstrapping.

EMULATION OF A TARGET TRIAL DURING PREGNANCY

Emulating target trials in pregnancy requires two distinctive elements. First, the evaluation of neonatal outcomes requires the linkage of mother and infant in the database,[31] using a unique family number or other linkage algorithms based on the date of delivery and other identifiers. Second, the emulation requires the determination of gestational age. LMP is typically available in databases of electronic health records but not always in databases of administrative claims, in which it is estimated using algorithms[31] that combine codes for procedures

and diagnoses of pregnancy outcomes such as preterm or term livebirth or stillbirth deliveries, spontaneous abortions, and terminations.[15] With the introduction of Z3A codes based on ICD-10, claims databases may start reliably recording the number of weeks since LMP at a given pregnancy visit.

We now describe the emulation of the components of the target trial protocol and key challenges that arise during the emulation process.

Eligibility criteria

We identify the eligibility criteria of the target trial based on diagnosis and procedure codes and, when available, free text. We use data from 2020 and 2021 to assess eligibility within the 12 months before the baseline in 2021. A challenge to correctly identify eligibility criteria is the lack of direct contact with the individuals in the database. We must assume that the absence of codes for, say, a booster dose implies no booster was administered. To address this challenge, the analysis is restricted to individuals who are expected to receive all their care from the same health system and who have been enrolled for some minimum time.

Treatment Strategies

The vaccination strategies are the same as in the target trial. We identify pharmacy dispensations for vaccination and procedure codes for vaccine administration. The first challenge is the correct identification of unvaccinated individuals. We must assume that individuals without vaccine codes did not receive the vaccine. If some individuals received vaccination at public health clinics or their workplace, and the administration was not later recorded by the health care provider (for electronic health records) or not charged to their insurance (for claims), the vaccination would not appear in the database. This may be a lesser problem when studying vaccine boosters because the eligibility criteria require evidence of prior vaccination. To explore the extent of this problem for safety outcomes, we could conduct sensitivity analyses under alternative assumptions:

1. Use as control group women who did not receive a booster during pregnancy and who were not eligible for the emulation because they had a record of having received a booster before pregnancy. We expect that these women would have had a vaccination record during pregnancy if they had received another booster during pregnancy. This analysis assumes that prepregnancy boosters do not affect the pregnancy outcome of interest.
2. Use as control group pregnancies in the same month the previous year (when vaccines were not available). This analysis assumes no temporal trends in pregnancy outcomes.
3. If interested in the direction of the effect (not its magnitude), use vaccinated women as their controls via a case-crossover or case–time–control design under the assumptions described in eAppendix 2, http://links.lww.com/EDE/B980.[32–34] This analysis is restricted to events with clear onset (*e.g.*, spontaneous abortions, preterm delivery).

Finding qualitatively similar estimates in the sensitivity analyses and the main analysis would increase confidence in the main result. A second challenge is the correct identification of the relative timing of vaccination and LMP in claims databases. While LMP estimation is quite accurate when the result of pregnancy is a livebirth,[31] a reliable algorithm for estimating LMP for pregnancy losses has yet to be developed. In the absence of Z3A codes, studies have often assigned a fixed gestational age of 10 weeks to all spontaneous abortions and 28 weeks to all stillbirths.[15] As a result, vaccinations in the first weeks of pregnancies that ended in spontaneous abortion after 10 weeks may be misclassified as preconceptional exposures and preconceptional exposures may be counted as pregnancy exposures for earlier spontaneous abortions. Again, case-crossover designs may increase confidence about the direction of the effect and do not require knowledge of LMP.[35] (eAppendix 3). The situation is further complicated if either the frequency of vaccination or its effects change over gestational age, e.g., if vaccination is either avoided or recommended after pregnancy recognition, or if vaccination is only abortifacient at specific weeks.

Treatment Assignment

We assign each eligible woman to the treatment strategy (vaccination or no vaccination) compatible with their data under the assumption that the assignment is random conditional on the baseline variables gestational age, calendar month, geographic region, maternal age at LMP, obstetric characteristics (e.g., multiples, parity); and prior SARS-CoV-2 infection, coexisting conditions (e.g., obesity, smoking, pregestational diabetes, hypertension, other cardiovascular conditions, asthma, and their treatments), and proxies for healthcare utilization (e.g., number of hospitalizations and outpatient visits, flu vaccination) in the previous 6 months.

A key challenge is whether the assumption of random assignment is approximately correct. Vaccinated and unvaccinated individuals may differ on risk factors not recorded in the database. Effectiveness will be underestimated if the vaccinated are at higher risk of infection (e.g., health care workers) or have a higher prevalence of risk factors for severe COVID-19 (e.g., diabetes), and will be overestimated if the vaccinated are more likely to have health-seeking behaviors (e.g., mask use) or have better access to care. Databases contain information on medical conditions and healthcare utilization, but they typically lack information on behavioral and lifestyle factors.[36] Unmeasured confounding may be a lesser problem when comparing different vaccine brands than when comparing vaccination vs. no vaccination.

The following sensitivity analyses may be implemented to check the robustness of the effect estimates to lack of randomization:[37]

1. Confirm that there is no association when using negative controls for which the confounding structure is expected to be similar to that for the main analysis. For example, use a

negative control period outside the etiologically relevant risk window (e.g., vaccination before LMP for congenital malformations, assuming no carryover effects); a negative control outcome that is expected to be unaffected by the vaccine (e.g., urinary tract infections for effectiveness); or a negative control population (e.g., paternal vaccination for safety outcomes).

2. For events with a clear onset, confirm that the association is in the same direction using a case-crossover or case–time–control design (eAppendix 2, http://links.lww.com/EDE/B980) that attempts to eliminate confounding by unmeasured time-fixed factors.[35]

Follow-up Period

The follow-up is the same as in the target trial. A challenge when using claims databases is that we can only include women with continuous enrollment in the database during pregnancy because we identify pregnancies, and link to an infant, based on claims associated with the end of pregnancy. Conditioning on continuous enrollment (a postassignment factor) may introduce selection bias, even under the null, if vaccination is associated with continuous coverage or a successful linkage, and continuous coverage or a successful linkage are associated with the outcomes. We hope this bias to be small after adjustment for baseline variables. If future validation studies confirm the accuracy of Z3A codes for timing gestation at each visit, we will be able to identify and time pregnancies at the first prenatal visit.

Outcomes

In claims databases, outcomes are identified based on algorithms, most of which have been validated and shown to have high positive predictive values.[38] In electronic health records, outcome ascertainment may also include the inspection of free text. A challenge is potential outcome misclassification. Therefore, as in many pragmatic trials, we restrict our attention to diagnoses that are likely to be coded by the health care provider (e.g., severe COVID-19, spontaneous abortions after 6 weeks) or that can be successfully captured from free text. Even for outcomes with little misclassification, the time of onset may be unknown. For example, women may have COVID-19 symptoms or prodromal symptoms for pregnancy events before these outcomes are recorded. If these symptoms make an individual less likely to get vaccinated, we could observe a lower frequency of vaccination before the recorded outcome (e.g., COVID-19 hospitalization). One way to explore the impact of this bias is a sensitivity analysis in which the outcome onset is set, say, 7 days before the recorded date. Also, the impact of misclassification of diagnosis or time of onset can be explored using probabilistic bias analysis.[39]

Causal Contrast of Interest

Observational analog of the per-protocol effect (i.e., the effect of receiving the vaccine booster versus receiving no booster during pregnancy).

Data Analysis

The per-protocol analysis of the target trial and of the observational analysis that emulates it is the same except for one thing: In the observational data (unlike in a true target trial), there is no date of assignment to booster or no booster. Therefore, to prevent immortal time bias, we need to choose the start of the follow-up (time zero) of each pregnancy in such a way that the distribution of gestational age at time zero is the same in both groups.[11] One approach is the emulation of sequential target trials with weekly recruitment: each gestational week, we identify eligible women who received a booster in that week and match each of them with an eligible woman who does not receive a booster in that week (a control). To adjust for confounders, we could also match women on risk factors measured in that week, such as diabetes or calendar month. Unvaccinated women may be eligible as controls for multiple weeks up until they received a vaccine (when they become eligible for inclusion in the booster group). We follow women in the booster and control groups until the outcome or the end of follow-up. We censor both members of the matched pair if/when the control receives a booster. We then pool the data across all sequential trials and estimate the outcome risk in the booster and control groups. We proposed weekly trials considering that wider intervals are more at risk of immortal time bias, but that too fine intervals may be unnecessary or unrealistic. 95% confidence intervals can be computed via bootstrapping.

DISCUSSION

We have reviewed the specification of target trials during pregnancy and their emulation using observational healthcare databases. Using vaccine boosters as an example, we described challenges and proposed sensitivity analyses for each step of the emulation process. While some limitations may be unavoidable (e.g., bias due to missing data and residual confounding), the target trial framework helps clarify that some common biases are preventable in observational analyses (e.g., including only vaccinations before week 20 to evaluate the risk of abortion, avoiding immortal time bias by aligning vaccination administration and start of follow-up).

Randomized trials cannot typically be relied upon to inform decisions during pregnancy. For example, COVID-19 vaccine trials in pregnancy[5] studied vaccinations administered only in weeks 24 to 34 of gestation (which is not the etiologically relevant window for implantation, placentation, and organogenesis), are too small to study severe COVID-19 outcomes or key obstetric and fetal outcomes (e.g., specific malformations), and too short to study long-term effects (e.g., vaccine effectiveness throughout the year postvaccination). We need analyses of large healthcare databases for a precise estimation of effects of interventions at all gestational ages in both mothers and their offspring.[5]

For the observational emulation of target trials during pregnancy, healthcare databases have advantages and

disadvantages compared with traditional observational studies with primary data collection. The rich prospectively recorded information in healthcare databases may result in less confounding, no recall bias, and low risk of some selection biases (e.g., outcome does not affect participation). Also, analyses of healthcare databases readily avoid immortal time bias by starting the follow-up at the gestational week of vaccination. In contrast, in traditional prospective studies that enroll women while pregnant, avoiding immortal time bias requires starting the follow-up at the time of enrolment, which may be several weeks after vaccination. Because pregnancy losses that occur between vaccination and enrolment are never included, the estimates may be affected by selection bias and will miss acute transient effects of the pre-enrolment vaccinations on the risk of pregnancy losses. Retrospective enrollment is avoided because it may select pregnancy losses following a vaccination. On the other hand, the use of healthcare databases poses several methodologic challenges,[37,40] such as the ascertainment of gestational age.

Our example focused on a simple iscenario: a point intervention (receiving an additional vaccine dose at time zero) and a sustained strategy (not receiving any vaccine doses during the follow-up). The framework we presented, however, can be applied to the emulation of target trials of more complex treatment strategies that are sustained over time, i.e., dynamic strategies under which the intervention received at each time during pregnancy depends on the evolving characteristics of the women under study. In our example, we made the simplifying assumption that confounders measured at baseline were sufficient to estimate the comparative effect of interest. Yet emulating target trials of sustained treatment strategies generally requires detailed longitudinal information on time-varying treatments and confounders, as well as the use of g-methods in the presence of treatment-confounder feedback.[41] Our approach can also be applied to non-pharmacologic interventions as long as the required data are available in the healthcare database.

In conclusion, explicitly emulating target trials helps reduce bias, improves the interpretability of effect estimates, and clarifies the nature of the remaining challenges to estimate the effects of interventions initiated during pregnancy.

ACKNOWLEDGMENTS

REFERENCES

1. Polack FP, Thomas SJ, Kitchin N, et al; C4591001 Clinical Trial Group. Safety and efficacy of the BNT162b2 mRNA Covid-19 vaccine. *N Engl J Med*. 2020;383:2603–2615.
2. Baden LR, El Sahly HM, Essink B, et al; COVE Study Group. Efficacy and safety of the mRNA-1273 SARS-CoV-2 vaccine. *N Engl J Med*. 2021;384:403–416.
3. Voysey M, Clemens SAC, Madhi SA, et al. Safety and efficacy of the ChAdOx1 nCoV-19 vaccine (AZD1222) against SARS-CoV-2: an interim analysis of four randomised controlled trials in Brazil, South Africa, and the UK. *Lancet*. 2021;397:99–111.
4. Sadoff J, Le Gars M, Shukarev G, et al. Interim results of a phase 1-2a trial of Ad26.COV2.S covid-19 vaccine. *N Engl J Med*. 2021;384:1824–1835.
5. Study to Evaluate the Safety, Tolerability, and Immunogenicity of SARS CoV-2 RNA Vaccine Candidate (BNT162b2) Against COVID-19 in Healthy Pregnant Women 18 Years of Age and Older. ClinicalTrialsgov Identifier: NCT04754594 2021.
6. Heath PT, Le Doare K, Khalil A. Inclusion of pregnant women in COVID-19 vaccine development. *Lancet Infect Dis*. 2020;20:1007–1008.
7. Guidance for Industry. Pregnant women: Scientific and ethical considerations for inclusion in clinical trials. 2018. Accessed 3/13/2021.
8. Dagan N, Barda N, Biron-Shental T, et al. Effectiveness of the BNT162b2 mRNA COVID-19 vaccine in pregnancy. *Nat Med*. 2021;27:1693–1695.
9. Hernán M, Alonso A, Logan R, et al. Observational studies analyzed like randomized experiments: an application to postmenopausal hormone therapy and coronary heart disease. *Epidemiology*. 2008;19:766–779.
10. Hernán M, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. *Am J Epidemiol*. 2016;183:758–764.
11. Hernán M, Sauer BC, Hernandez-Diaz S, Platt R, Shrier I. Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses. *J Clin Epidemiol*. 2016;79:70–75.
12. Bonhoeffer J, Kochhar S, Hirschfeld S, et al; GAIA project participants. Global alignment of immunization safety assessment in pregnancy - The GAIA project. *Vaccine*. 2016;34:5993–5997.
13. Jones CE, Munoz FM, Kochhar S, et al. Guidance for the collection of case report form variables to assess safety in clinical trials of vaccines in pregnancy. *Vaccine*. 2016;34:6007–6014.
14. Rasmussen SA, Olney RS, Holmes LB, Lin AE, Keppler-Noreuil KM, Moore CA; National Birth Defects Prevention Study. Guidelines for case classification for the National Birth Defects Prevention Study. *Birth Defects Res A Clin Mol Teratol*. 2003;67:193–201.
15. MacDonald SC, Cohen JM, Panchaud A, McElrath TF, Huybrechts KF, Hernandez-Diaz S. Identifying pregnancies in insurance claims data: methods and application to retinoid teratogenic surveillance. *Pharmacoepidemiol Drug Saf*. 2019;28:1211–1221.
16. Young JG, Stensrud MJ, Tchetgen Tchetgen EJ, Hernán MA. A causal framework for classical statistical estimands in failure-time settings with competing events. *Stat Med*. 2020;39:1199–1236.
17. Chiu YH, Stensrud MJ, Dahabreh IJ, et al. The effect of prenatal treatments on offspring events in the presence of competing events: an application to a randomized trial of fertility therapies. *Epidemiology*. 2020;31:636–643.
18. Yland JJ, Chiu YH, Rinaudo P, Hsu J, Hernán MA, Hernandez-Diaz S. Emulating a target trial of the comparative effectiveness of clomiphene citrate and letrozole for ovulation induction. *Hum Reprod*. 2022;37:793–805.
19. Hook EB, Czeizel AE. Can terathanasia explain the protective effect of folic-acid supplementation on birth defects? *Lancet*. 1997;350:513–515.
20. Svensson E, Ehrenstein V, Norgaard M, et al. Estimating the proportion of all observed birth defects occurring in pregnancies terminated by a second-trimester abortion. *Epidemiology*. 2014;25:866–871.
21. Friedberg MK, Silverman NH, Moon-Grady AJ, et al. Prenatal detection of congenital heart disease. *The Journal of pediatrics*. 2009;155:26–31, e1.
22. Heinke D, Rich-Edwards JW, Williams PL, et al; National Birth Defects Prevention Study. Quantification of selection bias in studies of risk factors for birth defects among livebirths. *Paediatr Perinat Epidemiol*. 2020;34:655–664.
23. Greenland S. Basic methods for sensitivity analysis of biases. *Int J Epidemiol*. 1996;25:1107–1116.
24. Khoury MJ, Flanders WD, James LM, Erickson JD. Human teratogens, prenatal mortality, and selection bias. *Am J Epidemiol*. 1989;130:361–370.
25. Laurence K, James N, Miller MH, Tennant G, Campbell H. Double-blind randomised controlled trial of folate treatment before conception to prevent recurrence of neural-tube defects. *British medical journal*. 1981;282:1509–1511.
26. Coronavirus disease 2019 (COVID-19) treatment guidelines. 2021. Accessed 3/17/2021.
27. Hernán M, Robins JM. Per-Protocol Analyses of Pragmatic Trials. *N Engl J Med*. 2017;377:1391–1398.
28. Hernán M, Hernandez-Diaz S, Robins JM. A structural approach to selection bias. *Epidemiology*. 2004;15:615–625.

29. Robins JM, Hernán MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology*. 2000;11:550–560.

30. Cole SR, Hernán MA. Adjusted survival curves with inverse probability weights. *Comput Methods Programs Biomed*. 2004;75:45–49.

31. Margulis AV, Palmsten K, Andrade SE, et al. Beginning and duration of pregnancy in automated health care databases: review of estimation methods and validation results. *Pharmacoepidemiol Drug Saf*. 2015;24:335–342.

32. Suissa S. The case–time–control design: Further assumptions and conditions. *Epidemiology*. 1998;9:441–445.

33. Hernandez-Diaz S, Hernán M, Meyer K, Werler MM, Mitchell AA. Case–crossover and case–time–control designs in birth defects epidemiology. *Am J Epidemiol*. 2003;158:385–391.

34. Shahn Z, Hernán MA, Robins JM. A formal causal interpretation of the case–crossover design. *Biometrics*. 2022. doi: 10.1111/biom.13749.

35. Greenland S. Confounding and exposure trends in case–crossover and case–time–control designs. *Epidemiology*. 1996;7:231–239.

36. Skjefte M, Ngirbabul M, Akeju O, et al. COVID-19 vaccine acceptance among pregnant women and mothers of young children: results of a survey in 16 countries. *Eur J Epidemiol*. 2021;36:197–211.

37. Huybrechts KF, Bateman BT, Hernandez-Diaz S. Use of real-world evidence from healthcare utilization data to evaluate drug safety during pregnancy. *Pharmacoepidemiol Drug Saf*. 2019;28:906–922.

38. Palmsten K, Huybrechts KF, Kowal MK, Mogun H, Hernandez-Diaz S. Validity of maternal and infant outcomes within nationwide Medicaid data. *Pharmacoepidemiol Drug Saf*. 2014;23:646–655.

39. Fox MP, Lash TL, Greenland S. A method to automate probabilistic sensitivity analyses of misclassified binary variables. *Int J Epidemiol*. 2005;34:1370–1376.

40. Savitz DA, Fell DB, Ortiz JR, Bhat N. Does influenza vaccination improve pregnancy outcome? Methodological issues and research needs. *Vaccine*. 2015;33:6430–6435.

41. Hernán MA, Robins JM. *Causal Inference: What If*. Boca Raton: Chapman & Hall/CRC; 2020.