American Journal of Epidemiology **Submitted Manuscript**

Title: International Registry of Coronavirus Exposure in Pregnancy (IRCEP) – Cohort Description and Methodological Considerations

Authors: Sonia Hernandez-Diaz, Louisa H. Smith, Camille Dollinger, Sonja A. Rasmussen, Enrique F. Schisterman, Rino Bellocco, Diego F. Wyszynski

Correspondence Address: Dr. Sonia Hernandez-Diaz, Department of Epidemiology, Harvard T.H. Chan School of Public Health, 677 Huntington Avenue. Boston, MA 02115 (e-mail: shernan@hsph.harvard.edu)

Affiliations: Harvard T.H. Chan School of Public Health, Boston, Massachusetts, United States (Sonia Hernandez-Diaz, Louisa H. Smith, and Camille Dollinger); Department of Pediatrics, Obstetrics and Gynecology, & Epidemiology, University of Florida, United States (Sonja A. Rasmussen); Perelman School of Medicine, University of Pennsylvania, United States (Enrique F. Schisterman); Karolinska Institutet, Stockholm, Sweden and University of Milano-Bicocca, Milan, Italy (Rino Bellocco); and Pregistry, Los Angeles, California, United States (Diego F. Wyszynski).

Funding: The International Registry of Coronavirus Exposure in Pregnancy receives financial support from Regeneron Pharmaceuticals, Inc. and Johnson & Johnson Services, Inc. The authors did not receive funding for this work.

Data Availability Statement: The data underlying this article are available through Pregistry conditional on data use agreements and fees

Thanks: We thank all the study participants for their trust, time and efforts contributing to the study. We thank Harpreet "Harry" and Gurpreet Singh for the many hours spent developing the survey platform. We also appreciate all the valuable discussions with our Scientific Advisory Committee: Drs Kjersti Aagaard, Thomas F. McElrath, Susan Nicholson, Claudia Perandones, Sonja A. Rasmussen, and Jennita Reefhuis.

Conference presentation: NA

Disclaimer: The funders had no role in the data collection, analysis, interpretation, or writing of this manuscript**.**

Conflict of Interest: None

Running Head: International cohort of COVID-19 in Pregnancy

Key words: COVID-19, pregnancy, cohort, registry, methods

ABSTRACT

Limited data are available about the potential health effects of infection with the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) on pregnant women and their developing offspring. We developed the International Registry of Coronavirus Exposure in Pregnancy (IRCEP) to provide data on the risk of major adverse obstetric and neonatal outcomes among women with varying degrees of severity and timing of COVID-19 exposure during pregnancy. We describe here the cohort and share the lessons learned. The IRCEP enrolls women tested for SARS-CoV-2 or with a clinical diagnosis of COVID-19 during pregnancy and obtains information using an online data collection system. By March 2021, 17,532 participants from 77 countries had enrolled; 54% enrolled during pregnancy and 46% afterwards. Among women with symptomatic COVID-19 with a positive SARS-CoV-2 test (N=4,934), symptoms were mild in 41%, moderate in 52% and severe in 7%; 7.7% were hospitalized for COVID-19 and 1.7%

were admitted to an intensive care unit. The biggest challenges were retention of participants enrolled during pregnancy, and the potential bias introduced when participants enroll after pregnancy outcomes are known. Multiple biases need to be considered and addressed when estimating and interpreting the effects of COVID-19 in pregnancy in these types of cohorts.

BACKGROUND

Over one hundred million women have given birth worldwide since the Coronavirus Disease 2019 (COVID-19) pandemic started in China in late 2019[1,2]. To date there is no consistent evidence that pregnant women are more susceptible to severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection[3-6]; and conflicting findings have been published on whether they have more severe COVID-19, if infected, than non-pregnant women of similar age[3,5,6]. In early reports, based on small samples, the risk of both symptomatic infections and death in pregnant women with COVID-19 was similar to that in non-pregnant women[7,8]. Subsequently, larger studies suggested that hospitalization rates, intensive care unit (ICU) admissions and mortality in pregnant women diagnosed with COVID-19 are higher than in non-pregnant women of similar age[8-12]. More recent studies concluded that pregnant women are not more likely to get seriously ill from SARS-CoV-2 infection.[13] These discrepancies may be at least partially explained by international and temporal differences in SARS-CoV-2 virulence, population vulnerability to its effects, or COVID-19 treatments.

Regardless, even if infection and severity risks were not above general population levels, consequences of severe COVID-19 are likely exacerbated when the patient is pregnant and has to deliver and take care of a newborn[14]. Moreover, severe COVID-19 may not only cause maternal morbidity and mortality, but also lead to iatrogenic preterm delivery due to concerns about COVID-19 transmission and progression,[3,6,15-18] and potential coagulation disorders[19] that may cause placental-related complications such as miscarriages.[20] Vertical SARS-CoV-2 transmission is rare but can occur.[3,5,6,21-23]

A dearth of information, especially during the first months of the COVID-19 pandemic, led to increased levels of anxiety among pregnant women[24], as well as unnecessary cesarean deliveries[3] and elective terminations.[25] Therefore, there was, and still is, an urgent need to gather and

communicate reliable information. The direct-to-participant International Registry of Coronavirus Exposure in Pregnancy (IRCEP) was established to describe COVID-19 in pregnant women worldwide and to assess the relative risk of major adverse obstetric and neonatal outcomes in pregnancies exposed to varying degrees of severity and timing of SARS-CoV-2 infection. We share here the lessons we learned from designing and conducting IRCEP.

METHODS

Study Design

The IRCEP is an ongoing observational cohort that began enrollment in June 2020. Data on baseline characteristics, reproductive history, chronic medical conditions and medication use, exposure to SARS-CoV-2, COVID-19 manifestations, prenatal care, pregnancy outcomes, and neonatal outcomes are collected during pregnancy and postpartum. The IRCEP allows enrollment throughout pregnancy and during the first 180 days after the end of pregnancy.

Study Population

The study population includes women 18 years of age or older from around the world with a current or recent pregnancy and tested for SARS-CoV-2 (regardless of the result) or with a clinical diagnosis of COVID-19 between the first day of the last menstrual period (LMP) and end of pregnancy. Inclusion criteria also require willingness to provide responses to a minimum set of demographic questions. The rationale for including cases without a positive SARS-CoV-2 test is that, in many countries, early in the pandemic, the reverse transcriptase-polymerase chain reaction (RT-PCR) test was inaccessible to a large proportion of the population. Therefore, the presence of clinical symptoms (e.g., typical pulmonary lesions on chest CT, loss of sense of smell or taste), confirmed by a health care professional, was sufficient to be considered a COVID-19 clinical diagnosis.

Enrollment

Information about the IRCEP is available on a dedicated website. Venues for increasing awareness include social media channels frequently visited by pregnant women (e.g., Facebook,

Instagram) and online parenting forums. The advertisements were translated to the most frequently spoken local languages and the awareness campaigns targeted websites frequently visited by pregnant women within each country. Women self-enroll using the IRCEP website without the participation of physicians involved in the testing or treatment of COVID-19, obstetric or neonatal care. The enrollment process validates the contact information by sending a random code to the submitted mobile phone number.

Follow Up

Follow-up for analyses begins on the date of the first SARS-CoV-2 test or first COVID-19 symptoms during pregnancy, whichever occurs earlier, and continues until a pregnancy loss, loss to follow-up, or 90 days after delivery. Information on pregnancy outcome (i.e., SAB, termination, stillbirth, live birth) is collected and livebirths are followed until three months after birth. Optional email and short message service (SMS) phone reminders are available throughout the study to facilitate retention. The web and mobile app also generate automated reminders to provide medical records. In addition, contact information is collected for a close friend or family member who will serve as an alternate contact that can provide minimal information (i.e., is participant healthy?) if we lose contact with the participant for unknown reasons before COVID-19 clinical resolution.

Data Collection and Timing of Assessments

Data are collected via self-administered online, easy-to-complete modules. Each module usually takes 5 to 15 minutes to complete. Given the international nature of the IRCEP, the questionnaires are available in 10 languages (English, Spanish, French, Italian, Portuguese, Russian, Urdu, German, Marathi, and Hindi). The translations were done with professional software followed by review and correction by interpreters. At enrollment, women complete 2 modules (Web Figure 1). These modules collect data on baseline characteristics (e.g., demographics, illnesses, reproductive history, prenatal screening and its results); and on timing of SARS-CoV-2 testing and COVID-19 clinical signs, their duration, severity, and treatment. For women enrolling during pregnancy, monthly follow-up questionnaires are collected until delivery, when obstetric and neonatal outcomes are collected. The follow-up modules include COVID-19-related questions to assess active COVID-19 cases as well as to detect potential new

positive SARS-CoV-2 tests in women that initially tested negative. In case of an early pregnancy loss (SAB or termination), the participant is directed to an End-of-Pregnancy brief questionnaire and is not asked to complete any more modules. Finally, around 90 days after delivery, a module collects information on postpartum and neonatal outcomes, with questions targeted to the most common specific obstetric and pediatric conditions and events. Women enrolling retrospectively within 90 days post-delivery complete the Baseline, COVID-19, and Delivery modules upon enrollment, and the postpartum outcomes at 90 days after end of pregnancy. Women who enroll between 90 and 180 days after the end of pregnancy can complete all modules, including information on pregnancy and the first 90 days post-partum, at enrollment.

In addition to self-reported information, participants are asked to photograph and upload all available SAS-CoV-2 test results, delivery and neonatal medical records, and any other health care records they consider relevant, after redacting personal identifiers. These records could be used to validate maternally reported diagnoses and to allow for potential adjudication of outcomes by experts blinded to the maternal COVID-19 status for specific studies. Uploaded records are checked for redaction prior to storage and will be translated with professional software followed by human review as they became relevant for analyses (e.g., when evaluating malformations).

The identifying and anonymized data of individual study participants are linked via a unique subject identifier. The file containing the personal identifiers used during the informed consent process is securely stored in a separate server. Files for analysis do not include any personal identifiers. Once the participant completes the study, the key used to link analytic files with identifiers is deleted.

Exposure definitions

A registry participant is considered *infected* if she had a positive RT-PCR or serologic test for SARS-CoV-2 between the first day of the LMP and end of pregnancy. In sensitivity analysis, those with clinical diagnoses of COVID-19 are also considered to be infected, regardless of SARS-CoV-2 testing. The primary reference group consists of women with a negative SARS-CoV-2 test and no positive test nor clinical diagnosis of COVID-19 during pregnancy. The goal of the reference group is to provide an estimate of the expected incidence or prevalence of

obstetric and neonatal events in women from the same source population as the exposed, upon adjustment for potential confounders associated with either infection or testing. Women in the reference group that test positive later in pregnancy would be considered exposed at that point.

Sub-cohorts defined by COVID-19 severity are identified. The classification of severity is summarized in Web table 1**.** To evaluate the effect of disease severity on adverse pregnancy outcomes among women exposed to the virus, severe cases will be compared to those with mild COVID-19 presentations. The IRCEP will also be able to assess the potential effects of specific COVID-19 characteristics or treatments, predictors of COVID-19 severity, and frequency of infections in newborns.

Timing of Pregnancy

The first day of the LMP is used to define the timing of pregnancy and the length of gestation. Gestational age is determined by an algorithm using the best available information, including reported LMP, due date based on LMP, and due date based on ultrasound. Gestational timing is needed to define exposure since the etiologically relevant period for each outcome of interest varies. Web Table 2.

Outcomes

The main outcomes of interest include SAB (spontaneous pregnancy loss prior to 20 weeks gestation), stillbirth (fetal death after 20 weeks gestation), major congenital anomalies (specific types will be explored where sample size permits), preterm delivery, maternal obstetric complications (e.g., preeclampsia, gestational hypertension, cesarean delivery, postpartum hemorrhage, postpartum depression), small for gestational age (≤10th centile on birth weight for the infant's sex and gestational age), head circumference at birth, admission to the neonatal intensive care unit, vertical transmission of SARS-CoV-2, and neonatal death. Questions on access to care during the COVID-19 pandemic, maternal mental health, and breastfeeding will allow evaluation of other aspects of wellbeing.

Primary analyses will be restricted to participants that enrolled before specific outcomes occur or can be known, to avoid both selection and recall biases[26]. For example, in the case of congenital anomalies, before any informative prenatal test. The exposure windows of interest will also

depend on the outcome; for example, to evaluate whether SARS-CoV-2 infection affects the risk for preterm delivery, those exposed in the month before a preterm delivery may be compared to the unexposed in that risk set defined by gestational age.

Covariates

The IRCEP collects data on a wide range of covariates including maternal demographic characteristics, comorbid medical conditions, habits, reproductive history, obstetric characteristics, use of medications, and measures of healthcare utilization. Women may enroll after the onset of COVID-19 infection but baseline questions collect information on characteristics existing *before* the infection to allow proper adjusted in future analysis[27].

Study termination and loss to follow-up

Participating women may withdraw from the IRCEP at any time at their own request, at which time all their information will be deleted from the database (except for select de-identified sociodemographic characteristics, which are retained to assess data representativeness in aggregate analyses relative to the initial population and correct for informative censoring if needed). The website inquires about the reason for withdrawal. When a participant does not respond to repeated online prompts requesting further information on the pregnancy, they will be considered lost to follow-up. Participants that do not complete all the modules by 180 days after the end of pregnancy will be censored from the cohort at the time of last contact.

Human Subjects

The IRCEP protocol was approved by the Harvard Longwood Campus Institutional Review Board (IRB20-0622). Eligible participants provide consent electronically with an e-Form on the IRCEP website.

RESULTS

Study Population

As of end of March 2021, 17,532 participants from 77 countries had enrolled in IRCEP. Of those, 54% enrolled during pregnancy and 46% after the end of pregnancy, i.e., postpartum or

after a pregnancy loss (Table 1). The frequency of participants by country of residence resembles the distribution of COVID-19 worldwide at that time, consistent with the IRCEP awareness campaigns online. Participants in the IRCEP are racially and socioeconomically diverse (Table 2). The frequency of both negative and positive tests in asymptomatic women reflect screening intensity in those groups (e.g., more in those living in North America or Europe, white race, having asthma, or smoking tobacco). The median number of weeks post-LMP at enrollment was 26 (interquartile range [IQR] 17, 34) for those enrolled during pregnancy and 49 (IQR 44, 54) for those enrolled after the end of pregnancy (Figure 1).

COVID-19 during pregnancy

At enrollment, 5,858 (34%) participants reported a positive test for SARS-CoV-2 (84% were symptomatic), 2.5% had a clinical diagnosis of COVID-19 without a positive test, and 10,215 (58%) had a negative test and no clinical diagnosis of COVID-19 (Table 3). A small number of participants reported neither clinical diagnoses nor a SARS-CoV-2 test (n=20) and were excluded.

The timing of COVID-19 diagnosis was equally distributed throughout pregnancy (Web Figure 2). However, testing tended to cluster around the time of delivery, since it has become standard practice in many countries to screen for SARS-CoV-2 at the time of hospital admission for delivery. Increased screening of asymptomatic women pre-delivery resulted in more negative tests, and positive tests in women without symptoms, towards the end of pregnancy.

Among those with symptomatic COVID-19 confirmed with a positive test, symptoms were mild in 41%, moderate in 52% and severe in 7%. Among 1,235 women with a clinical diagnosis of COVID-19 without a positive test confirmation, 38% had mild, 58% had moderate, and 3% had severe symptoms. Overall, the most common symptoms were upper respiratory manifestations (e.g., cough), fatigue, loss of taste or smell, and gastrointestinal symptoms (e.g., nausea, diarrhea). Moderate and severe presentations frequently included shortness of breath, fever, and muscle aches. Overall, among symptomatic women with a confirmed positive test, 7.7% were hospitalized for COVID-19, 2.2% required oxygen, ventilator assistance or extracorporeal membrane oxygenation (ECMO), and 1.7% were admitted to an intensive care unit (ICU). Other

than analgesics, the most common pharmacotherapies used to treat COVID-19 were azithromycin, oseltamivir, corticosteroids, and hydroxychloroquine/chloroquine.

Outcomes

Of the 9,471 participants enrolled during pregnancy, 5% have completed participation in the study and 6% are still pregnant and completing modules, as of March 31, 2021. Of the 8,061 participants enrolled postpartum or after fetal loss, 74% have completed participation (Figure 2). Among those participants with complete follow-up, the frequencies of the most common obstetric outcomes were similar to what would be expected in the general population (i.e., 0.8% stillbirths, 6% preeclampsia, 1.9% twins or higher-level multiples, and 2.6% major congenital malformations) (Table 4.) The frequency of pregnancy losses is lower than the usual cumulative risk throughout pregnancy in the population since women that enroll late in pregnancy represent a survivor cohort and those that enroll retrospectively had shorter opportunities to have a test or an infection.

DISCUSSION

We have enrolled a large international cohort of pregnant women with and without COVID-19 in the IRCEP. Future studies will be able to assess a variety of specific research questions related to COVID-19 and pregnancy, such as the effect of COVID-19 severity on multiple outcomes. We learned many lessons designing and conducting the IRCEP, including the following:

- **Enrollment and retention of participants in multinational pregnancy registries**

The increasing number of pregnancies affected by COVID-19 and the wider availability of tests for SARS-CoV-2 facilitated enrollment over the study period. Nonetheless, we learned that social media awareness campaigns are key to enrolling participants in internet-based studies. Enrollment in the IRCEP dropped substantially between campaigns, despite a large amount of pregnancy and COVID-19-related resources available on the Pregistry website. Web Figure 3.

While completion of modules among retrospective enrollees has been over 70%, retention of participants registered during pregnancy is below 10%. The IRCEP tries to increase retention and data completeness by fostering an online community where participants not only share data but also receive relevant information. Optional automatic reminders are also sent via SMS to remind

participants to upload information. We did not offer economic incentives to participate in the study to minimize the risk of ineligible subjects enrolling for money. The study appeals to the altruism and solidarity of volunteers by conveying the importance of their loyalty to the study for the generation of evidence that will help other pregnant women. Unfortunately, these approaches were insufficient. Similarly, a small proportion of participants has submitted redacted photos of their medical records despite an easy system to upload documents. Consequently, our intention to validate and adjudicate self-reported outcomes may not be feasible.

Despite the attrition, this remains one of the largest cohorts of pregnant women with COVID-19. However, the substantial losses to follow up might select a biased sample. Characteristics of participants lost to follow-up will be compared to the observed cohort and weights may be applied if censoring is not random. For example, among prospective enrollees, retention was higher among those negative for COVID-19, from North America or Europe, and of White race, higher education, and higher income (Web Table 3). Future studies should strive to attain higher retentions. For example, by allocating sufficient budget to a more personalized and proactive contact with enrollees or demanding stronger commitment to research from the beginning (e.g., making uploading test results an inclusion criteria).

- **Generalizability of results**

While the multinational design is meant to facilitate generalizability of results across the globe, it also represents a challenge a) logistically, because of the need to translate the materials to many languages and to make them culturally appropriate across countries (e.g., race categories, health coverage modalities); and b) methodologically, because country of residence is a strong determinant of both COVID-19 infection and of testing, and of incidence and diagnosis of the outcomes. Therefore, the statistical analyses will need to take country of residence into consideration, for example with stratification or introducing a random effect component. In retrospect, a multinational study in fewer selected countries might have been preferable from a research perspective to guarantee sufficient numbers within strata while still providing a global perspective.

In addition, enrollment of women in pregnancy registries is voluntary and, therefore, participants are a non-random sample of all women with COVID-19. Consequently, the characteristics and experience of women who participate in a registry may differ from those of non-participants, and

these characteristics may modify the observed effects of SARS-CoV-2. Our primary awareness campaigns were designed to appeal to people of a wide variety of backgrounds, but our enrolled sample turned out to be more educated than the general population (65% with at least a college education). Although biological effects of viruses tend to be universal, the health consequences of COVID-19 among those volunteering to participate in the registry, who tend to be more educated and health conscious, is likely to underestimate the absolute impact on more vulnerable populations.

- **Estimation of SARS-CoV-2 prevalence and COVID-19 incidence in pregnancy**

Non-random samples cannot provide an estimate of the distribution of SARS-CoV-2 seroprevalence or COVID-19 severity in the source population. The proportion of participants with SARS-CoV-2, or with COVID-19, should not be interpreted as a "risk" and should not be compared with the risk in non-pregnant populations. Asymptomatic pregnant women are tested more often than asymptomatic non-pregnant young women. This targeted screening results in higher detection of SARS-CoV-2 in pregnant women with asymptomatic infection (Web Figures 4 and 5). During pregnancy, more participants reported asymptomatic infection only detected by a positive test around delivery, when more screening is done. Similarly, because of preferential screening, pregnant women receive more negative test results, particularly around delivery. Consequently, the participants who joined after delivery included more negative tests, because the prospective participants enrolled based on tests conducted before screening at delivery. Therefore, study designs like ours are not appropriate to estimate the incidence of SARS-CoV-2 or COVID-19.

- **Using individuals with negative tests as the reference group**

Non-random samples selected based on testing can provide valid estimates of COVID-19 effects on pregnancy outcomes if risk factors associated not only with the infection but also with testing are controlled. Asymptomatic participants with test results for SARS-CoV-2 (either positive or negative) represent populations with increased access to screening (e.g., more affluent) or high-risk groups (e.g., women with asthma). Therefore, the unbalanced characteristics observed between participants with COVID-19 and the reference group with negative tests may be risk factors for infection, or risk factors for testing. This selection introduced by the inclusion criteria (i.e., requiring a test) can be conceptualized as conditioning on a collider.[28] Although, this

potential selection bias is less likely to affect the assessment of COVID-19 severity within symptomatic cases, if milder cases with risk factors were still preferentially tested, we would underestimate the effect of severity. While we will adjust for factors associated with testing, residual confounding remains a concern. Web Figure 6.

- **Pregnant women as a vulnerable population**

Some studies that evaluated the effect of pregnancy on people with SARS-CoV-2 infection suggested a higher frequency of ICU admissions and hospitalizations among pregnant women with COVID-19, even after accounting for age and healthier overall status of pregnant women.[8,12] However, results might be explained by preferential hospitalization and ICU admissions of pregnant women given the same disease severity and by inclusion of hospitalizations for pregnancy-related reasons in the outcome (Web figure 5). That is, even without COVID-19, pregnancy increases the likelihood of hospitalizations relative to non-pregnant women of similar age, even if only for obstetric reasons (e.g., delivery, preeclampsia). We did not include non-pregnant women in our study and, therefore, the IRCEP cannot answer the question of whether pregnant women constitute a "vulnerable population" with respect to COVID-19. We designed this study not to answer the question *what would have been the outcome had the person not been pregnant*, but *what would have been the outcome had the pregnant person not become infected, or not had severe COVID-19.*

- **Confounders, mediators, and colliders**

When assessing the effect of SARS-CoV-2 infection on pregnancy outcomes in observational studies, an association could be explained by a direct effect of the virus on the outcomes, an effect mediated through maternal symptoms (e.g., pneumonia or fever), or by confounding (e.g., women more likely to be infected might also be at higher risk of adverse pregnancy outcomes). It is also important to note that a factor may cause severe COVID-19 by increasing exposure to SARS-CoV-2, increasing susceptibility to being infected if exposed, or increasing the likelihood of progression to severe COVID-19 if infected. Studies that condition on the steps in the causal pathway (e.g., studying hospitalized participants) may be conditioning on colliders.[29] For example, when evaluating the effect of COVID-19 on preterm delivery and using a reference group recruited in the same center, within pregnancies hospitalized, those not admitted for COVID-19 would have other reasons for admissions (e.g., preeclampsia) that may be risk factors

for preterm birth. Web Figure 6. Similarly, when testing is required to classify whether the person has COVID-19, if the outcome triggers testing, there would be a higher proportion of confirmed COVID-19 (and negative tests) among those with the outcome.[28] Future statistical analyses using the IRCEP data will consider these explanations for observed associations and carefully classify confounders, mediators, and colliders depending on the research question.

- **Enrollment post-infection**

Of concern in cohort studies with primary data collection is the selection of non-lethal COVID-19 since most women will enroll after COVID-19 resolves. Although maternal mortality is expected to be low (<1%), this selection will result in optimistic descriptions of the nature of COVID-19 during pregnancy. Studies with population-based samples enrolled before infection (e.g., healthcare databases) will be able to provide the full picture.

- **Enrollment post-outcome**

Enrollment after pregnancy outcomes are known (during or after the end of pregnancy) may self-select a group with adverse outcomes and more eager to share their experience in a study, thus overestimating risks; or it might underestimate the risk if the distressing event reduces the likelihood of participation.[30] Overall, when we compared participants that enrolled after vs before the end of pregnancy, the frequency of pregnancy outcomes was slightly higher for some (e.g., preterm delivery 10% vs 9%) and slightly lower for others (e.g., major malformations 2.6% vs 2.7%). Table 4. If participation after an adverse outcome is diagnosed is more (or less) likely for patients with COVID-19, retrospective participation may lead to spurious associations. Therefore, primary analyses will be restricted to women enrolled before the pregnancy outcome of interest is known. For example, analyses of malformations will be restricted to women that enroll before the results from informative prenatal screening tests are known.

However, events that occur right after COVID-19 may only enroll retrospectively. For example, although inclusion of known SABs might introduce selection bias if participants were more likely to enroll after suffering an SAB, particularly if they attribute it to COVID-19, enrollment of retrospective SABs is necessary to capture SABs that could potentially occur immediately after an infection (i.e., pregnant women that become eligible at the time of infection might not have time to enroll prospectively before the outcome). Similarly, there are few participants who had severe COVID-19 in the third trimester and enrolled before delivery – likely because they

remained hospitalized through the end of pregnancy or had a natural or induced delivery and could only enroll after delivery. The above two scenarios challenged our plans to focus on prospective participants when evaluating the effects of infections near end of pregnancy. Studies that enroll the population before the infection (e.g., healthcare databases) will avoid the potential bias introduced by retrospective enrollment.

- **Misclassification**

The IRCEP collects information directly from participants. Women often know more about their habits, occupations, and compliance with medication use than their health care providers; however, clinicians might provide more complete and accurate information regarding diagnoses.[31] The accuracy of recall in the IRCEP is facilitated by using structured questionnaires, detailed questions that allow only plausible responses, and calendars to help establish gestational timing and enhance recall of dates.[32] To reduce misclassification of infection, primary analyses can be restricted to COVID-19 confirmed with laboratory testing, which became more available over time. Web Figure 7. Misclassification of COVID-19 severity (e.g., need for respiratory assistance or ICU) is unlikely. Outcome misclassification could be non-differential or differential between COVID-19 cases and the reference group. Concern that COVID-19 might pose a risk could lead to more prenatal diagnostic measures such as ultrasound and to more careful examination of infants for defects postnatally, potentially leading to differential accuracy in detection and classification of defects among exposed and unexposed. This potential surveillance bias can be minimized by focusing on major outcomes that are less vulnerable to differential misclassification (e.g., prematurity, SABs, malformations).

CONCLUSIONS

Many publications on COVID-19 during pregnancy that shared our limitations failed to address important sources of biases. We summarized the lessons we learned so that future studies can do better. Figure 3. In our experience with an online international pregnancy cohort, the biggest challenges were retention of participants during follow-up and the potential bias introduced when participants are enrolled retrospectively. Given the large sample size, restricting our analyses to prospectively enrolled participants and standardizing risk factors associated with censoring may preserve the utility of the current study for at least some of the questions of interest.

REFERENCES

1. UNICEF. Pregnant mothers and babies born during COVID-19 pandemic threatened by strained health systems and disruptions in services 2020 2020; https://www.unicef.org/press-releases/pregnant-mothers-and-babies-born-during-covid-19-pandemic-threatened-strained-health. Accessed August, 2020.
2. WHO. Pneumonia of unknown cause – China 2020. 2020; https://www.who.int/csr/don/05-january-2020-pneumonia-of-unkown-cause-china/en/. Accessed August, 2020.
3. Chen H, Guo J, Wang C, et al. Clinical characteristics and intrauterine vertical transmission potential of COVID-19 infection in nine pregnant women: a retrospective review of medical records. *Lancet.* 2020;395(10226):809-815.
4. Liu D, Li L, Wu X, et al. Pregnancy and Perinatal Outcomes of Women With Coronavirus Disease (COVID-19) Pneumonia: A Preliminary Analysis. *AJR Am J Roentgenol.* 2020:1-6.
5. Breslin N, Baptiste C, Gyamfi-Bannerman C, et al. COVID-19 infection among asymptomatic and symptomatic pregnant women: Two weeks of confirmed presentations to an affiliated pair of New York City hospitals. *Am J Obstet Gynecol MFM.* 2020:100-118.
6. Chen L, Li Q, Zheng D, et al. Clinical Characteristics of Pregnant Women with Covid-19 in Wuhan, China. *The New England journal of medicine.* 2020;382(25):e100.
7. Karami P, Naghavi M, Feyzi A, et al. Mortality of a pregnant patient diagnosed with COVID-19: A case report with clinical, radiological, and histopathological findings. *Travel Med Infect Dis.* 2020:101665.
8. Ellington S, Strid P, Tong VT, et al. Characteristics of Women of Reproductive Age with Laboratory-Confirmed SARS-CoV-2 Infection by Pregnancy Status - United States, January 22-June 7, 2020. *MMWR Morbidity and mortality weekly report.* 2020;69(25):769-775.
9. Collin J, Bystrom E, Carnahan A, Ahrne M. Public Health Agency of Sweden's Brief Report: Pregnant and postpartum women with severe acute respiratory syndrome coronavirus 2 infection in intensive care in Sweden. *Acta obstetricia et gynecologica Scandinavica.* 2020;99(7):819-822.
10. Delahoy MJ, Whitaker M, O'Halloran A, et al. Characteristics and Maternal and Birth Outcomes of Hospitalized Pregnant Women with Laboratory-Confirmed COVID-19 - COVID-NET, 13 States, March 1-August 22, 2020. *MMWR Morbidity and mortality weekly report.* 2020;69(38):1347-1354.
11. Panagiotakopoulos L, Myers TR, Gee J, et al. SARS-CoV-2 Infection Among Hospitalized Pregnant Women: Reasons for Admission and Pregnancy Characteristics - Eight U.S. Health Care Centers, March 1-May 30, 2020. *MMWR Morbidity and mortality weekly report.* 2020;69(38):1355-1359.
12. Zambrano LD, Ellington S, Strid P, et al. Update: Characteristics of Symptomatic Women of Reproductive Age with Laboratory-Confirmed SARS-CoV-2 Infection by Pregnancy Status - United States, January 22-October 3, 2020. *MMWR Morbidity and mortality weekly report.* 2020;69(44):1641-1647.
13. Knight M, Bunch K, Vousden N, et al. Characteristics and outcomes of pregnant women admitted to hospital with confirmed SARS-CoV-2 infection in UK: national population based cohort study. *BMJ.* 2020;369:m2107.

14. Rasmussen SA, Kissin DM, Yeung LF, et al. Preparing for influenza after 2009 H1N1: special considerations for pregnant women and newborns. *American journal of obstetrics and gynecology.* 2011;204(6 Suppl 1):S13-20.
15. Qiao J. What are the risks of COVID-19 infection in pregnant women? *Lancet.* 2020;395(10226):760-762.
16. Liu Y, Chen H, Tang K, Guo Y. Clinical manifestations and outcome of SARS-CoV-2 infection during pregnancy. *J Infect.* 2020.
17. Zhu H, Wang L, Fang C, et al. Clinical analysis of 10 neonates born to mothers with 2019-nCoV pneumonia. *Transl Pediatr.* 2020;9(1):51-60.
18. Cosma S, Carosso AR, Cusato J, et al. Coronavirus disease 2019 and first-trimester spontaneous abortion: a case-control study of 225 pregnant patients. *American journal of obstetrics and gynecology.* 2020.
19. Levi M, Thachil J, Iba T, Levy JH. Coagulation abnormalities and thrombosis in patients with COVID-19. *Lancet Haematol.* 2020;7(6):e438-e440.
20. Vlachodimitropoulou Koumoutsea E, Vivanti AJ, Shehata N, et al. COVID-19 and acute coagulopathy in pregnancy. *J Thromb Haemost.* 2020;18(7):1648-1652.
21. Fan C, Lei D, Fang C, et al. Perinatal Transmission of 2019 Coronavirus Disease-Associated Severe Acute Respiratory Syndrome Coronavirus 2: Should We Worry? *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America.* 2021;72(5):862-864.
22. Dong L, Tian J, He S, et al. Possible Vertical Transmission of SARS-CoV-2 From an Infected Mother to Her Newborn. *JAMA : the journal of the American Medical Association.* 2020;323(18):1846-1848.
23. Zeng H, Xu C, Fan J, et al. Antibodies in Infants Born to Mothers With COVID-19 Pneumonia. *JAMA : the journal of the American Medical Association.* 2020;323(18):1848-1849.
24. Sinaci S, Ozden Tokalioglu E, Ocal D, et al. Does having a high-risk pregnancy influence anxiety level during the COVID-19 pandemic? *European journal of obstetrics, gynecology, and reproductive biology.* 2020;255:190-196.
25. Wu YT, Li C, Zhang CJ, Huang HF, Frcog. Is termination of early pregnancy indicated in women with COVID-19? *European journal of obstetrics, gynecology, and reproductive biology.* 2020;251:271-272.
26. MacDonald SC, Hernan MA, McElrath TF, Hernandez-Diaz S. Assessment of recording bias in pregnancy studies using health care databases: An application to neurologic conditions. *Paediatric and perinatal epidemiology.* 2018;32(3):281-286.
27. Hernan MA, Hernandez-Diaz S, Werler MM, Mitchell AA. Causal knowledge as a prerequisite for confounding evaluation: an application to birth defects epidemiology. *Am J Epidemiol.* 2002;155(2):176-184.
28. Messinger CJ, Lipsitch M, Bateman BT, et al. Association Between Congenital Cytomegalovirus and the Prevalence at Birth of Microcephaly in the United States. *JAMA pediatrics.* 2020.
29. Hernandez-Diaz S, Schisterman EF, Hernan MA. The birth weight "paradox" uncovered? *Am J Epidemiol.* 2006;164(11):1115-1120.
30. Margulis AV, Mittleman MA, Glynn RJ, Holmes LB, Hernandez-Diaz S. Effects of gestational age at enrollment in pregnancy exposure registries. *Pharmacoepidemiology and drug safety.* 2015;24(4):343-352.

31. Reiff_Eldridge R, Heffner CR, Ephross SA, Tennis PS, White A, Andrews E. Monitoring pregnancy outcomes after prenatal drug exposure through prospective pregnancy registries: A pharmaceutical company commitment. *American journal of obstetrics and gynecology*. 2000;182:159-163.
32. Mitchell AA, Cottler LB, Shapiro S. Effect of questionnaire design on recall of drug exposure in pregnancy. *American Journal of Epidemiology*. 1986;123(4):670-676.

Table 1. Number of women enrolled during pregnancy or within 180 days after end of pregnancy by country of residence by March 2021

Country	Pregnant at Enrollment (N = 9471)		Enrollment After End of Pregnancy (N = 8061)		Overall (N = 17,532)		WHO Prevalence ranking [a]
	No	%	No	%	No	%	
United States	2,013	(21.0)	2,814	(35.0)	4,827	(28.0)	1
Brazil	1,664	(18.0)	467	(5.8)	2,131	(12.0)	3
United Kingdom	484	(5.1)	810	(10.0)	1,294	(7.4)	5
Russia	690	(7.3)	526	(6.5)	1,216	(6.9)	4
Mexico	641	(6.8)	317	(3.9)	958	(5.5)	13
Spain	370	(3.9)	540	(6.7)	910	(5.2)	7
India	435	(4.6)	359	(4.5)	794	(4.5)	2
France	559	(5.9)	182	(2.3)	741	(4.2)	6
Chile	275	(2.9)	452	(5.6)	727	(4.1)	24
Italy	297	(3.1)	412	(5.1)	709	(4.0)	8
South Africa	320	(3.4)	360	(4.5)	680	(3.9)	15
Peru	338	(3.6)	244	(3.0)	582	(3.3)	18
Colombia	427	(4.5)	89	(1.1)	516	(2.9)	11
Philippines	250	(2.6)	169	(2.1)	419	(2.4)	32
Argentina	331	(3.5)	84	(1.0)	415	(2.4)	12
Germany	240	(2.5)	150	(1.9)	390	(2.2)	10
Others [b]	137	(1.4)	86	(1.1)	223	(1.3)	

[a] World Health Organization (WHO) ranking of top countries in number of cases reported as a reference.

[b] Other countries: Algeria, Andorra, Angola, Armenia, Australia, Austria, Bahamas, Bahrain, Belgium, British Indian Ocean Territory, British Virgin Islands, Burkina Faso, Cameroon, Canada, China, Comoros, Costa Rica, Democratic Republic of the Congo, Denmark, Dominican Republic, Ecuador, Ethiopia, Georgia, Ghana, Gibraltar, Greece, Guam, Guatemala, Hungary,

Iran, Ireland, Israel, Jordan, Kenya, Libya, Lithuania, Malawi, Mayotte, Montenegro, Namibia, Netherlands, Nigeria, Pakistan, Portugal, Puerto Rico, Qatar, Romania, Saudi Arabia, Senegal, Singapore, Slovakia, South Korea, Sweden, Switzerland, U.S. Virgin Islands, Ukraine, United Arab Emirates, Uruguay, Zambia, Zimbabwe

Table 2. Distribution of baseline characteristics by SARS-CoV-2 test results and presence of COVID-19 symptoms among those with positive test results.

Characteristics [a]	SARS-CoV-2 Test				Symptoms among SARS-CoV-2 Positive			
	Positive N = 7,148		Negative N = 10,139		Asymptomatic N = 1,437		Symptomatic N = 5,654	
	No.	%	No.	%	No.	%	No.	%
Age [b]	30.0 (5.0)		30.5 (5.0)		29.5 (5.4)		30.1 (4.9)	
Continent								
Africa	293	4.1	440	4.3	86	6.0	205	3.6
Asia	644	9.0	618	6.1	271	19	366	6.5
Europe	1,656	23	3,547	35	335	23	1,298	23
North America	1,841	26	3,914	39	246	17	1,580	28
Oceania	3	<0.1	3	<0.1	0	0	3	<0.1
South America	2,711	38	1,614	16	499	35	2,202	39
Education								
Less than high school	334	6.4	356	4.3	99	10.0	236	5.6
High school	1,665	32	2,361	29	311	31	1,344	32
College	1,754	33	3,039	37	302	30	1,433	34
Graduate education	1,488	28	2,470	30	282	28	1,183	28
Race/ethnicity								
American Indian/Alaska Native	21	0.4	27	0.3	6 0.6	0.6	15	0.4
Asian	281	5.3	339	4.1	86	8.6	193	4.6
Black	343	6.5	334	4.0	71	7.1	271	6.4
Latina	1,215	23	1,201	15	242	24	964	23
Middle Eastern	32	0.6	18	0.2	11	1.1	21	0.5

Mixed	551	10	635	7.7	102	10	445	11
Native Hawaiian/Pacific Islander	4	<0.1	6	<0.1	1	<0.1	3	<0.1
South Asian	65	1.2	110	1.3	20	2.0	43	1.0
White	2,760	52	5,591	68	463	46	2,263	54
Economic status								
Poor	605	12	925	11	143	15	459	11
Lower-middle class	1,477	29	1,962	24	267	27	1,198	29
Middle class	2,445	47	3,808	47	460	47	1,955	47
Wealthy	643	12	1,371	17	101	10	535	13
Health insurance	4,385	84	7,238	88	776	78	3,561	85
Current employment								
Not working	1,746	33	2,902	36	390	39	1,343	32
Working in an office	507	9.7	619	7.6	82	8.3	420	10
Working in food services	114	2.2	177	2.2	18	1.8	95	2.3
Working in health care	922	18	1,156	14	133	13	784	19
Working from home	1,447	28	2,551	31	283	29	1,144	27
Other work	497	9.5	746	9.2	83	8.4	407	9.7
Smoking								
No	3,558	75	5,250	68	627	74	2,895	76
Before pregnancy	969	21	1,812	23	186	22	771	20
During	188	4.0	654	8.5	39	4.6	146	3.8

pregnancy								
Vaping								
No	4,183	89	6,465	84	755	89	3,385	89
Before pregnancy	462	9.8	1,047	14	88	10	366	9.6
During pregnancy	62	1.3	203	2.6	7	0.8	55	1.4
Recreational drugs								
No	4,072	87	6,249	81	737	87	3,295	87
Before pregnancy	573	12	1,260	16	103	12	460	12
During pregnancy	62	1.3	203	2.6	9	1.1	52	1.4
Alcohol in first trimester								
Never	3,573	76	6,045	78	627	74	2,910	76
Once per month	772	16	1,152	15	143	17	618	16
Weekly or more	358	7.6	514	6.7	78	9.2	276	7.3
Vitamin in first trimester								
Never	541	12	684	8.9	124	15	411	11
Some days	892	19	1,536	20	140	17	742	20
Every day	3,265	69	5,483	71	584	69	2,646	70
Pre-pregnancy BMI[c]								
< 18.5	137	3.2	236	3.3	26	3.4	106	3.1
18.5-24.9	1,938	45	3,075	44	366	47	1,549	45
25.0-29.9	1,177	27	1,848	26	207	27	955	28
≥30	1,032	24	1,905	27	172	22	851	25

Primiparous	2,042	43	3,549	46	429	50	1,590	41
Multiple gestations	83	1.8	147	1.9	16	1.9	66	1.8
Fertility treatment used	199	8.0	424	8.7	45	10	149	7.4
Pregnancy planning								
Trying for 6-12 months	428	9.2	698	9.2	64	7.5	360	9.7
Trying for at least 12 months	602	13	1,067	14	127	15	468	13
Unplanned	2,142	46	2,721	36	410	48	1,706	46
Trying for < 6 months	1,465	32	3,106	41	256	30	1,193	32
Thyroid disease	284	6.0	529	6.8	57	6.7	223	5.8
Pre-pregnancy hypertension	179	3.8	310	4.0	36	4.2	141	3.7
Diabetes (I or II, pre-pregnancy)	77	1.6	143	1.8	18	2.1	57	1.5
Any cardiovascular condition	94	2.0	148	1.9	10	1.2	82	2.1
Asthma	345	7.3	820	11	58	6.8	283	7.4
Autoimmune disease	92	1.9	157	2.0	18	2.1	73	1.9

[a] The proportions are estimated among those that responded to that specific module or question.
[b] Values are expressed as mean (standard deviation)
[c] Weight (kg)/height $(m)^2$.

Table 3. Distribution of participants according to COVID-19 clinical diagnosis and SARS-CoV-2 test results during pregnancy as reported at enrollment.

COVID-19	SARS-CoV-2 Test at Enrollment									
	Positive		Negative		Inconclusive		Test not done		Total	
	No.	%	No.	%	No.	%	No.	%	No.	%
Clinical diagnosis	3,756	64	704	6.4	87	30	441	96	4,988	28
Self-reported symptoms only	1,584	27	1,625	15	56	19	8	1.7	3,273	19
No	473	8.1	8,465	78	80	27	0	0	9,018	51
Uncertain	45	0.8	125	1.1	71	24	12	2.6	253	1.4
Total	5,858	33	10,919	62	294	2	461	3	17,532	100

Table 4. Obstetric and neonatal outcomes among women with completed pregnancies by March 2021.

Outcomes	Pregnant at Enrollment N = 438		Enrollment After End of Pregnancy N = 5978		Overall N = 6416	
	No.	%	No.	%	No.	%
Pregnancy outcomes						
Live birth	416	95	5,844	98	6,260	98
Spontaneous abortion	13	3.0	56	0.9	69	1.1
Stillbirth	7	1.6	43	0.7	50	0.8
Termination (fetal problem)	0	0	20	0.3	20	0.3
Termination (other)	2	0.5	15	0.3	17	0.3
Gestational diabetes	37	8.9	613	10	650	10
(Pre-) Eclampsia	25	6.0	363	363	388	6.2
Caesarean section	168	41	2,466	43	2,634	43
Postpartum hemorrhage	22	5.3	239	4.1	261	4.2
Breastfed in hospital	350	89	4,637	85	4,987	85
Multiples	7	1.6	115	2.0	122	1.9
Neonatal outcomes[a]						
Birthweight in grams[b]	3,336 (535)		3,249 (614)		3,257 (608)	
Gestational weeks at delivery[b]	38.8 (2)		38.8 (2)		38.8 (2)	
Gestational age at delivery						

≥42.0 weeks	10	2.4	95	1.6	105	1.6
39.0-41.9 weeks	275	65	3,504	59	3,779	59
37.0-38.9 weeks	102	24	1,774	30	1,876	29
32.0-36.9 weeks	34	8.0	527	8.8	561	8.8
28.0-31.9 weeks	3	0.7	57	1.0	60	0.9
<28.0 weeks	0	0	27	0.5	27	0.4
Small for gestational age	33	7.8	482	8.1	515	8.0
Large for gestational age	49	12	726	12	775	12
NICU admission	52	13	796	14	848	14
Major congenital malformation	11	2.7	147	2.6	158	2.6
Roomed-in with mother						
Every day	337	84	4,333	77	4,670	78
Some days	19	4.7	326	5.8	345	5.8
No	42	10	829	15	871	15
I don't know	4	1.0	109	1.9	113	1.9
Tested for COVID-19	38	9.5	684	12	722	12
Positive for COVID-19	2	0.5	48	0.9	50	0.8
Neonatal death	1	0.2	12	0.2	13	0.2

[a] n = 6410 infants.

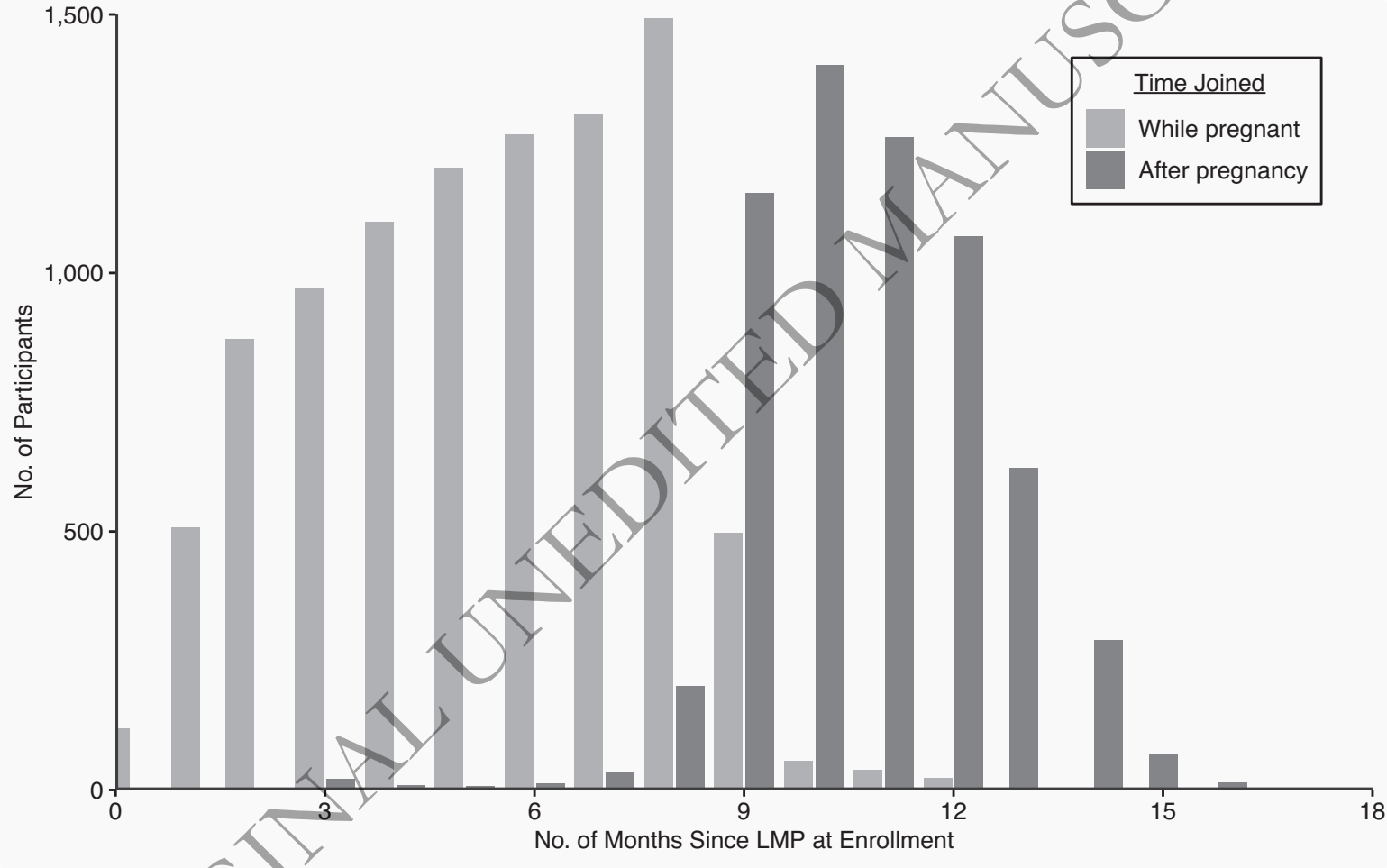
[b] Values are expressed as mean (standard deviation)

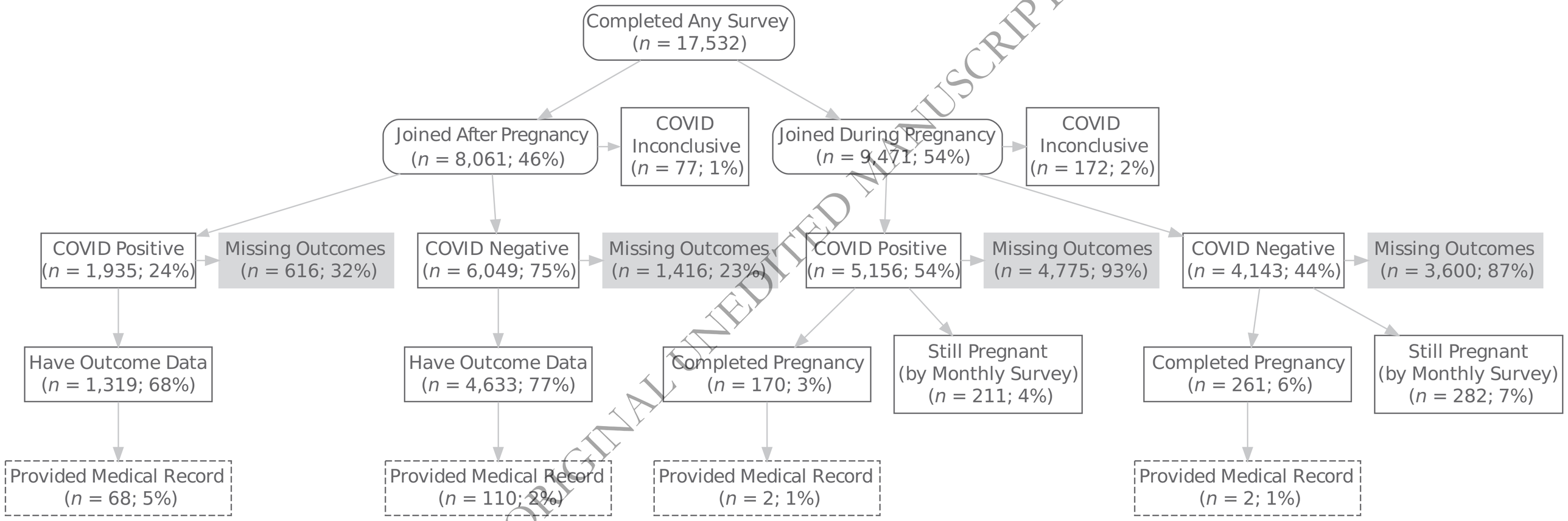
Figure 1. Distribution of months from the last menstrual period (LMP) at enrollment for participants that enrolled during or after end of pregnancy by March 2021.

Figure 2. Study population flow chart from consent to study completion by March 2021. Eligible for these analyses was everyone who consented to participate, filled out the initial module on COVID-19 tests and clinical diagnosis, and met the inclusion criteria.

Figure 3: Lessons learned from the International Registry of Coronavirus Exposure in

Pregnancy.





Lessons Learned

- A key implementation challenge for online international pregnancy cohorts is the retention of prospectively enrolled participants during follow-up.
- A key methodological challenge for pregnancy cohorts is the potential bias introduced when participants are enrolled retrospectively.
- Multiple biases need to be considered and addressed when estimating and interpreting the effects of COVID-19 in pregnancy in these types of cohorts.