The effect of maternal SARS-CoV-2 infection timing on birth outcomes: a retrospective multicentre cohort study

Samantha N Piekos, Ryan T Roper, Yeon Mi Hwang, Tanya Sorensen, Nathan D Price, Leroy Hood, Jennifer J Hadlock

Summary

Background The impact of maternal SARS-CoV-2 infection remains unclear. In this study, we evaluated the risk of maternal SARS-CoV-2 infection on birth outcomes and how this is modulated by the pregnancy trimester in which the infection occurs. We also developed models to predict gestational age at delivery for people following a SARS-CoV-2 infection during pregnancy.

Methods We did a retrospective cohort study of the impact of maternal SARS-CoV-2 infection on birth outcomes. We used clinical data from Providence St Joseph Health electronic health records for pregnant people who delivered in the USA at the Providence, Swedish, or Kadlec sites in Alaska, California, Montana, Oregon, or Washington. The SARS-CoV-2 positive cohort included people who had a positive SARS-CoV-2 PCR-based test during pregnancy, subdivided by trimester of infection. No one in this cohort had been vaccinated for COVID-19 at time of infection. The SARS-CoV-2 negative cohort were people with at least one negative SARS-CoV-2 PCR-based test and no positive tests during pregnancy. Cohorts were matched on common covariates impacting birth outcomes, and univariate and multivariate analysis were done to investigate risk factors and predict outcomes. The primary outcome was gestational age at delivery with annotation of preterm birth classification. We trained multiple supervised learning models on 24 features of the SARS-CoV-2 positive cohort to evaluate performance and feature importance for each model and discuss the impact of SARS-CoV-2 infection on gestational age at delivery.

Findings Between March 5, 2020, and July 4, 2021, 73 666 pregnant people delivered, 18 335 of whom had at least one SARS-CoV-2 test during pregnancy before Feb 14, 2021. We observed 882 people infected with SARS-CoV-2 during their pregnancy (first trimester n=85; second trimester n=226; and third trimester n=571) and 19769 people who have never tested positive for SARS-CoV-2 and received at least one negative SARS-CoV-2 test during their pregnancy. SARS-CoV-2 infection indicated an increased risk of preterm delivery (p=0·05) and stillbirth (p=0·05), accounted for primarily by first and second trimester SARS-CoV-2 infections. Gestational age at SARS-CoV-2 infection was correlated with gestational age at delivery (p=0·01) and had the greatest impact on predicting gestational age at delivery. The people in this study had mild or moderate SARS-CoV-2 infections and acute COVID-19 severity was not correlated with gestational age at delivery (p=0·31).

Interpretation These results suggest that pregnant people would benefit from increased monitoring and enhanced prenatal care after first or second trimester SARS-CoV-2 infection, regardless of acute COVID-19 severity.

Funding US National Institutes of Health.

Copyright © 2022 The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY 4.0 license.
Research in context

Evidence before this study
We searched PubMed for the keywords “SARS-CoV-2” or “COVID-19” combined with “preterm birth” or “trimester,” including all articles published before March 21, 2021. No language restrictions were applied. We found several studies that examined the impact of maternal SARS-CoV-2 infection on pregnancy outcomes using electronic health records (EHRs) or insurance claims. However, these studies were limited by small sample size or narrow scope of the data. Also, most studies did not differentiate between the timing of preterm birth, which has clinically significant implications in terms of neonatal survival, care, and potential lifelong health problems. To our knowledge, as of November, 2021, there has not yet been any primary study that accounts for the trimester of maternal SARS-CoV-2 infection, despite the awareness of the crucial periods during pregnancy when there exists a greater potential for in utero shocks on fetal development.

Added value of this study
We did a retrospective cohort study using EHRs from hospitals and clinics across five states in the USA. We compared pregnancy outcomes of unvaccinated people with a positive SARS-CoV-2 test during pregnancy to a matched control cohort with negative SARS-CoV-2 test results. The positive cohort was stratified by trimester of infection. To our knowledge, our study is the first to modulate birth outcomes by the trimester of maternal infection. Additionally, using propensity score matching, we controlled for several confounding variables known to be associated with negative pregnancy outcomes, which were enriched in the SARS-CoV-2 positive cohort. We evaluated outcomes using delivery events, including preterm birth classifications. There was an increase in preterm birth and stillbirth following SARS-CoV-2 infection, primarily driven by increases following first or second trimester maternal infection. We built supervised learning models that predict gestational age at delivery using demographics, maternal comorbidities, information related to SARS-CoV-2 infection, and fetal characteristics. The single greatest predictor of gestational age at delivery is gestational age at infection, with earlier age at infection associated with earlier age at delivery. There was no correlation between severity of COVID-19 and gestational age at delivery.

Implications of all the available evidence
These results suggest that pregnant people are at increased risk of preterm birth following SARS-CoV-2 infection regardless of COVID-19 severity. We recommend enhanced prenatal care and increased monitoring for pregnant people following a SARS-CoV-2 infection. Due to increased risk of maternal–fetal health of SARS-CoV-2 infection, we propose prioritisation of vaccination of pregnant people in areas where vaccine distribution is scarce.

Methods
Study setting and participants
In this retrospective cohort study, we used clinical data from Providence St Joseph Health (PSJH) electronic health records for pregnant people (aged 18 years and older and younger than 45 years) who delivered in the USA at the Providence, Swedish, or Kadlec sites in Alaska, California, Montana, Oregon, and Washington between March 5, 2020, and July 4, 2021 (appendix pp 1–2). PSJH is an integrated health-care system that has strong continuity of care across inpatient and outpatient settings; the US organisation’s 51 hospitals, 1085 clinics, and 120 000 caregivers collaborate to provide health and social services across seven states. We provide the COVID-19 seroprevalence rates for each of the five states in the study and the USA at the national level during the Centre for Disease Control and Prevention’s monitoring period containing the date Feb 14, 2021 (appendix p 12).9

The SARS-CoV-2 positive cohort included people who had a positive SARS-CoV-2 PCR-based test during pregnancy, subdivided by trimester of infection (appendix pp 1–2). No patients in the SARS-CoV-2 positive cohort were vaccinated for COVID-19 at the time of SARS-CoV-2 infection. The SARS-CoV-2 negative cohort were people with at least one negative PCR-based SARS-CoV-2 test and no positive tests during pregnancy. The cohort was limited to women with singleton pregnancies who delivered after 140 days’ gestational age (20 weeks), who had either commercial or state-provided Medicaid insurance. At the time of this study, PSJH did not routinely scan pregnant people for active SARS-CoV-2 infection or antibody presence during prenatal care or at
delivery (appendix p 3). Due to scarce testing availability at the beginning of the pandemic, testing was restricted to the sickest patients. As testing became more accessible, PSJH began testing anyone with COVID-19 symptoms and screening patients 2 days before admission for a planned procedure, including scheduled caesarean delivery or labour induction. The date after which broader testing began was site specific.

To account for covariates associated with preterm birth, propensity score matching was done to generate a SARS-CoV-2 negative matched control cohort. The unsupervised learning model nearest neighbours with replacement (k=1) was done to match across ten common covariates using the Python library sklearn (version 0.22.1; appendix pp 13–14). For each SARS-CoV-2 positive pregnant patient, this identifies the single most similar SARS-CoV-2 negative pregnant person across these ten variables. For people who were missing information in the pregravid body-mass index (BMI) field, a value was imputed using the median of the pregravid BMI reported for other people in their cohort (either the cohort of 882 people who had positive SARS-CoV-2 test results or the cohort of 17453 people in the SARS-CoV-2 negative cohort before matching). Absence of data for race, ethnicity, fetal sex, parity, and delivery method was encoded as −1. This resulted in a control cohort of 889 people that had improved representation of the matched covariates to the SARS-CoV-2 positive cohort.

All procedures were reviewed and approved by the Institutional Review Board at the PSJH through expedited review (study number STUDY2020000196). Consent was waived because disclosure of protected health information for the study was determined to involve no more than a minimal risk to the privacy of individuals.

Outcomes
The primary outcome evaluated in this study was gestational age at delivery with annotation of preterm birth classification. Secondary outcomes assessed included stillbirth, birthweight, fetal growth percentile, SGA, and rates of common pregnancy-related conditions.

Statistical analysis
The composition of each cohort’s demographics, comorbidities, and birth characteristics was represented as proportions or median along with the IQR (table 1). A Fisher’s exact test (for categorical data) or a Mann-Whitney U test (for continuous data) were used to evaluate differences between cohorts, using the R stats (version 3.6.3) and Python scipy (version 1.4.1) packages, respectively (appendix p 13). Comorbidities were identified by patient diagnosis codes, using SNOMED-CT (version 20200901; appendix p 15). Stillbirth is defined as pregnancy loss at 20 or more weeks of gestation. Term birth is defined as birth from 37 or more weeks of gestation; late preterm birth as less than 37 weeks and 34 or more weeks of gestation; moderate preterm birth as less than 34 weeks and 32 or more weeks of gestation; very preterm birth as less than 32 weeks and 28 or more weeks of gestation; and extremely preterm birth as less than 28 weeks of gestation.

Fetal growth percentile was calculated using the WHO Fetal Growth Charts based on gestational age and weight. A moving points average, which creates a series of averages across different subsets of the full dataset, was used to calculate weight by percentile at gestational age. The fetal growth percentile was then calculated by

|                         | SARS-CoV-2 positive (n=882) | SARS-CoV-2 negative (n=17 453) | SARS-CoV-2 negative matched (n=889) |
|-------------------------|-----------------------------|--------------------------------|----------------------------------|
| Demographics            |                             |                                |                                  |
| Maternal age at birth, years | 27.6 (9.4)                 | 31.5 (8.2)                    | 27.9 (9.0)                      |
| Pregravid body-mass index, kg/m² | 28.2 (8.6)                 | 25.8 (9.0)                    | 28.4 (2.1)                      |
| Race                    |                             |                                |                                  |
| Total number known      | 849                         | 16 920                        | 869                              |
| American Indian or Alaska Native | 20 (2.4%)                | 232 (1.4%)                    | 16 (1.8%)                       |
| Asian                   | 29 (3.4%)                   | 1571 (9.3%)                   | 38 (4.4%)                       |
| Black                   | 46 (5.4%)                   | 708 (4.2%)                    | 45 (5.2%)                       |
| Multiracial             | 9 (1.1%)                    | 185 (1.1%)                    | 9 (1.0%)                        |
| Native Hawaiian or Pacific Islander | 19 (2.2%)               | 215 (1.3%)                    | 17 (2.0%)                       |
| Other*                  | 306 (36.0%)                 | 2851 (16.8%)                  | 295 (33.9%)                     |
| White                   | 420 (49.5%)                 | 11 158 (65.9%)                | 449 (51.7%)                     |
| Ethnicity               |                             |                                |                                  |
| Total number known      | 863                         | 16 898                        | 873                              |
| Hispanic or Latino      | 453 (52.5%)                 | 3607 (21.3%)                  | 449 (51.4%)                     |
| Not Hispanic or Latino  | 410 (47.5%)                 | 13 291 (78.7%)                | 424 (48.6%)                     |
| Insurance               |                             |                                |                                  |
| Total number known      | 882                         | 17 453                        | 889                              |
| Medicaid                | 661 (74.9%)                 | 8 191 (46.9%)                 | 668 (75.1%)                     |
| Private                 | 221 (25.1%)                 | 9 262 (53.1%)                 | 221 (24.9%)                     |
| Smoker                  | 79 (9.0%)                   | 1 903 (10.9%)                 | 86 (9.7%)                       |
| Illicit drug user       | 93 (10.5%)                  | 2 146 (12.3%)                 | 93 (10.5%)                      |
| Preterm history         | 44 (5.0%)                   | 695 (4.0%)                    | 43 (4.8%)                       |
| Parity                  |                             |                                |                                  |
| Total number known      | 874                         | 17 156                        | 882                              |
| Nulliparity             | 426 (46.7%)                 | 10 288 (60.0%)                | 439 (49.8%)                     |
| Low multiparity         | 413 (47.3%)                 | 6 522 (38.0%)                 | 410 (46.5%)                     |
| Grand multipara         | 35 (4.0%)                   | 346 (2.0%)                    | 33 (3.7%)                       |
| Gravity                 |                             |                                |                                  |
| Total number known      | 874                         | 17 156                        | 882                              |
| Nulligravidity          | 239 (27.3%)                 | 5 757 (33.6%)                 | 239 (27.1%)                     |
| Low multigravidity      | 568 (65.0%)                 | 10 451 (60.9%)                | 583 (66.1%)                     |
| Grand multigravidity    | 67 (7.7%)                   | 948 (5.5%)                    | 60 (6.8%)                       |
| Educational attainment  |                             |                                |                                  |
| Total number known      | 40                          | 567                            | 45                               |
| Less than high school   | 8 (20.0%)                   | 53 (9.3%)                     | 7 (15.6%)                       |
| High school             | 26 (65.0%)                  | 229 (40.4%)                   | 23 (51.1%)                      |
| Undergraduate degree    | 5 (12.5%)                   | 203 (35.8%)                   | 14 (31.1%)                      |
| Graduate degree         | 1 (2.5%)                    | 82 (14.5%)                    | 1 (2.2%)                        |

(Table 1 continues on next page)
interpolation. SGA is defined as the babies in the bottom 10th percentile of fetal growth at birth.

**Predictive models**

To evaluate which variables are most predictive of gestational age at delivery, we trained multiple supervised learning models on 24 features of the SARS-CoV-2 positive cohort (n=882; appendix p 16). SARS-CoV-2 encounters were only observed for people who engaged with the health-care system. These features included patient demographics, fetal characteristics, comorbidities, and treatment following SARS-CoV-2 infection before delivery (including encounters, diagnoses, medications, and COVID-19 severity score). COVID-19 severity was defined as the patient’s maximum score on the WHO Ordinal Scale.19 Models were generated using the Python package sklearn with default settings for linear regression, ridge regression, gradient boosting regression, and random forest.19 The models were trained on 80% of the data, with 20% of the data held out for performance testing of the final model. Performance was evaluated on the test set for accuracy within 1 week of the actual gestational age at delivery, the coefficient of determinant (R² value), and root mean square deviation. Models were also evaluated by plotting the predicted versus the observed gestational age at delivery and calculating the trendline and R² value. Gini importance and Shapley additive explanations (SHAP) were applied to understand each feature’s marginal contribution and influence on model prediction using the Python libraries sklearn and SHAP (version 0.37.0), respectively.19 Both these approaches evaluate the degree of influence of the feature on the sample’s outcome, which provides interpretation of the machine learning models. Additionally, the variance captured by each non-binary feature was individually evaluated via R² value, and the correlation of each feature with gestational age at delivery was evaluated via Pearson’s correlation using the Python scipy package. Predictive models were reported following the TRIPOD guidelines15 (appendix pp 18–19).

**Role of the funding source**

The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

**Results**

Between March 5, 2020, and July 4, 2021, 73666 people delivered, 18335 (24.9%) of whom had at least one SARS-CoV-2 test during pregnancy (appendix pp 1–2). We observed 882 people infected with SARS-CoV-2 during their pregnancy (first trimester: n=85; second trimester: n=226; and third trimester: n=571) and 19769 people who have never tested positive for

**Table 1: Demographic and birth characteristics of people with SARS-CoV-2 test results during pregnancy**

| Comorbidities                  | SARS-CoV-2 positive (n=882) | SARS-CoV-2 negative (n=17453) | SARS-CoV-2 negative matched (n=889) |
|-------------------------------|------------------------------|-------------------------------|-----------------------------------|
| Chronic diabetes              | 87 (9.9%)                    | 1542 (8.8%)                   | 101 (11.3%)                       |
| Chronic hypertension          | 13 (1.5%)                    | 448 (2.6%)                    | 23 (2.6%)                         |
| Gestational diabetes          | 79 (9.0%)                    | 1404 (8.0%)                   | 89 (10.0%)                        |
| Gestational hypertension      | 52 (5.9%)                    | 1078 (6.2%)                   | 63 (7.1%)                         |
| Preeclampsia                  | 41 (4.6%)                    | 41 (0.2%)                     | 41 (4.6%)                         |
| Severe pre-eclampsia          | 3 (0.3%)                     | 23 (0.1%)                     | 0 (0.0%)                          |

**Birth characteristics**

|                         | SARS-CoV-2 positive (n=882) | SARS-CoV-2 negative (n=17453) | SARS-CoV-2 negative matched (n=889) |
|-------------------------|------------------------------|-------------------------------|-----------------------------------|
| Fetal sex               |                               |                               |                                   |
| Total number known      | 828                           | 16 840                        | 841                               |
| Female                  | 396 (47.8%)                   | 8246 (49.0%)                  | 409 (48.6%)                       |
| Male                    | 432 (52.2%)                   | 8594 (51.0%)                  | 432 (51.4%)                       |
| Mode of delivery        |                               |                               |                                   |
| Total number known      | 834                           | 16 869                        | 842                               |
| Caesarean section       | 265 (31.8%)                   | 6235 (37.0%)                  | 270 (32.1%)                       |
| Vaginal                 | 569 (68.2%)                   | 10 634 (63.0%)                | 572 (67.9%)                       |

Median (IQR) of the maternal age at birth (years) and pregravid body-mass index for people with maternal SARS-CoV-2 infection during pregnancy (n=882), with no SARS-CoV-2 infection and at least one negative SARS-CoV-2 test during pregnancy (n=17 453), and propensity score matched negative control (n=889). Distribution of race, ethnicity, insurance type, parity, gravidity, educational attainment, fetal sex, and mode of delivery. Number of people who smoke, use illicit drugs, have previously delivered prematurely, or have common pregnancy-related comorbidities.

**Table 2: Birth outcomes of people with SARS-CoV-2 infection during their first, second, or third trimester**

|                       | First trimester (n=85) | Second trimester (n=226) | Third trimester (n=571) | Total (n=882) | Matched negative control (n=889) |
|-----------------------|------------------------|--------------------------|-------------------------|---------------|-------------------------------|
|                       | N or median            | p value                  | N or median             | p value       | N or median                  | p value                      |
| Stillbirth*           | 1 (1.2%)               | 0.05                     | 5 (2.2%)                | 0.06          | 1 (0.2%)                     | 0.36                        |
| Small for gestational age† | 9 (10.6%)            | 0.79                     | 29 (12.8%)              | 0.37          | 78 (13.7%)                   | 0.08                        |
| Fetal growth percent1 | 49 (49.2%)             | 0.29                     | 47 (49.4%)              | 0.25          | 45 (48.6%)                   | <0.05                       |
| Gestational days1     | 273 (13.0%)            | <0.05                    | 273 (16.0%)             | <0.05         | 275 (13.0%)                  | 0.38                        |
| Birthweight, oz       | 118 (23.4%)            | <0.05                    | 117 (23.9)              | 0.07          | 118 (23.5)                   | <0.05                       |

Data are n (%) or median (IQR), unless stated otherwise. Stillbirth and small for gestational age rates for people with maternal SARS-CoV-2 infections during their first, second, or third trimester compared with the matched negative control group. Median (IQR) of the fetal growth percentile at delivery, gestational age at delivery, and weight at delivery. *Categorical variables: Fisher’s exact test. †Percentages reported are calculated from the total of infants whose birthweight at delivery was recorded. ‡Continuous variables: Mann-Whitney U test.
SARS-CoV-2 and received at least one negative SARS-CoV-2 test during their pregnancy (table 1). Compared with the SARS-CoV-2 negative cohort, patients in the SARS-CoV-2 positive cohort were more likely to have the following characteristics: Hispanic ethnicity (p<0·0001), race other than Asian or White (p<0·001), lower age (p<0·0001), higher BMI (p<0·0001), higher parity (p<0·0001), higher gravidity (p<0·01), Medicaid insurance (p<0·001), and lower educational attainment (85% attained a high school degree or less, p<0·01; table 1; appendix p 13). To account for these covariates, which are known to be associated with negative birth outcomes, propensity score matching was used to create a control SARS-CoV-2 negative cohort of 889 people (table 1; appendix pp 13–14).

Patients with SARS-CoV-2 infection during the first and second trimester of pregnancy had decreased gestational age at delivery (first trimester p<0·05, second trimester p<0·05; table 2). There was an increase in preterm delivery across all trimesters of pregnancy infections (all pregnancies p<0·05) with the most pronounced increase occurring following first trimester infection (figure 1; first trimester p<0·22, second trimester, p<0·22, and third trimester p<0·06). Pregnant patients with SARS-CoV-2 infection gave birth to infants with lower birthweight (first trimester p<0·05, second trimester p<0·07, third trimester p<0·05, all pregnancies p<0·01); however, lower fetal growth percentile at delivery was observed only following third trimester exposure (first trimester p<0·29, second trimester p<0·25, third trimester p<0·05, all pregnancies p<0·05; table 2; appendix p 4). There were higher rates of SGA among infants born to mothers who had a SARS-CoV-2 positive test during pregnancy (p<0·05), with the most pronounced increase following third trimester infection (table 2; appendix p 4). Additionally, people with SARS-CoV-2 infection during pregnancy had no increased risk of common pregnancy disorders including gestational diabetes, gestational hypertension, and pre-eclampsia (table 2; appendix p 17).

Multivariable predictive modeling was deployed to evaluate the ability of demographic data, comorbidities, health following SARS-CoV-2 infection, and fetal characteristics features to predict the outcome gestational age at delivery. Supervised machine learning models predicting gestational age at delivery were trained on 24 features with model performance evaluated on the held-out test dataset (appendix pp 16–17). The highest performing model was random forest, which accurately predicted gestational age at delivery within a week 58% of the time, and the model accounted for 28% of the data variance (appendix pp 5, 17). The most important feature of the random forest model was gestational age at infection (figure 2; appendix pp 6–7). Earlier gestational age at infection is associated with earlier gestational age at delivery (figure 3). The second most important feature in the model is the number of unique active ingredients in medications prescribed following SARS-CoV-2 infection. Earlier gestational age at delivery was
associated with higher numbers of unique active ingredients in medications prescribed (figure 3). The second highest performing model was the gradient boosting regression model, which accurately predicted gestational age at delivery within a week 56% of the time and accounted for 28% of the data variance (appendix pp 8, 17). The most important feature in this model was also gestational age at infection (appendix pp 9–11).

Gestational age at infection and the count of unique active ingredients in medications accounted for 1·2% and 5·6%, respectively, of the variance in gestational days at delivery alone (figure 3). People who delivered prematurely tended to be infected with SARS-CoV-2 earlier in pregnancy, but there are people who had early infections who went on to deliver at term. Similarly, patients with higher numbers of unique active ingredients in medications prescribed following SARS-CoV-2 infection tended to deliver prematurely. However, some people with no medications delivered prematurely and some with dozens of medications prescribed delivered at term. There was a slight uptick in the number of SARS-CoV-2 encounters among people who had extremely and very preterm birth, but the people with the highest numbers of SARS-CoV-2 encounters had late preterm or term deliveries.

Gestational age at infection (r=0·11, p<0·01), preterm history (r=0·17, p<0·001), commercial insurance (r=0·09, p<0·01), and number of cord vessels (r=0·15, p<0·001) were associated with higher gestational age at delivery (table 3). Count of unique medication active ingredients (r=–0·21, p<0·0001), gestational hypertension status (r=–0·24, p<0·0001), gestational diabetes status (r=–0·07, p<0·05), chronic hypertension status (r=–0·11, p<0·01), chronic diabetes status (r=–0·09, p<0·05), and unique diagnoses count (r=–0·08, p<0·05) were associated with lower gestational age at delivery (table 3).

Finally, the COVID-19 severity score was not correlated with gestational age at delivery (p=0·31; table 3; appendix p 12), although this cohort only had patients who had mild or moderate acute COVID-19, and no patients who required mechanical ventilation. There was no correlation between gestational age at delivery and the number of SARS-CoV-2 encounters (p=0·98), need for supplemental oxygen (p=0·18; WHO COVID-19 severity score 4–5), or maximum patient class (p=0·13; outpatient, emergency or urgent care, or inpatient; table 3; appendix p 12). Also, there was no correlation between COVID-19 severity and the number of days before delivery SARS-CoV-2 infection occurs (p=0·25; appendix p 3). Altogether, people who delivered the
Discussion

In this study, we examined the impact of maternal SARS-CoV-2 infection on birth outcomes modulated by gestational age, which resulted in the following key findings. First, more negative birth outcomes were observed when infections occurred earlier in gestation, including increased risk for preterm birth and stillbirth. Second, there were increased rates of SGA infants born to people who had a positive SARS-CoV-2 test result during pregnancy, suggesting that preterm delivery is induced via a mechanism that could impact fetal growth. Third, there appears to be two distinct populations of pregnant people: a subset with a negative correlation between the time of maternal SARS-CoV-2 infection and gestational age at delivery and a second population appearing to be unaffected. Fourth, the difference between mild and moderate severity of the SARS-CoV-2 infection does not appear to play a part in whether a pregnancy is likely to be negatively affected. Taken together, these findings suggest that SARS-CoV-2 infection early in pregnancy is an important risk factor that should be monitored in health systems.

The biggest predictor of gestational age at delivery is gestational age at maternal SARS-CoV-2 infection followed by the number of unique medication active ingredients prescribed following a SARS-CoV-2 infection. Notably, there appeared to be two subpopulations when comparing the gestational age at delivery versus gestational age at SARS-CoV-2 infection. One subpopulation shows a positive correlation, but for most individuals, the timing of infection did not show a direct association with gestational age at delivery (figure 3). This finding suggests there could be some event, or class of events, that happen during some pregnancies and not others, triggering the chance of negative effects on outcomes. Surprisingly, there was no correlation between gestational age at delivery and COVID-19 severity or the number of SARS-CoV-2 encounters, suggesting that the mechanism behind the effect of SARS-CoV-2 on preterm birth might differ from those that drive severity. Pregnant people with preterm birth tended to have mild SARS-CoV-2 infections and limited numbers of encounters during an active SARS-CoV-2 infection. Further studies are needed to investigate the mechanisms of how SARS-CoV-2 infection affects preterm birth and stillbirth.

Previous studies reported more severe birth outcomes among symptomatic than asymptomatic people. However, these studies did not account for trimester of infection. Cohorts might have been skewed towards people infected during their third trimester due to the state of testing early in the pandemic and studies' earlier reporting, which would mean a low number of observations of first or second trimester SARS-CoV-2 infection. We showed that gestational timing of infection is the most important factor of birth outcome. Further, COVID-19 disproportionately impacts low-income and minority communities, resulting in many covariates known to increase the risk of preterm birth and other negative birth outcomes. There were significant differences in the demographics, comorbidities, and birth characteristics of people in the SARS-CoV-2 positive and SARS-CoV-2 negative cohorts, including ethnicity, race, and Medicaid insurance status.

Table 3: Contribution of features towards the predicted gestational age

| Pearson's correlation | p value |
|-----------------------|---------|
| Demographic           |         |
| Maternal age, years   | -0.1    | <0.01  |
| Pre gravid body-mass index, kg/m² | -0.05 | 0.17 |
| Race                  | -0.05   | 0.15   |
| Ethnicity             | -0.03   | 0.42   |
| Insurance             | 0.09    | <0.01  |
| Parity                | -0.03   | 0.35   |
| Gravidity             | -0.03   | 0.35   |
| Preterm history       | 0.17    | <0.0001|
| Smoker                | 0.01    | 0.85   |
| Illicit drug user     | -0.01   | 0.83   |
| Comorbidities         |         |
| Chronic diabetes status | -0.09 | <0.05 |
| Chronic hypertension status | -0.11 | <0.01 |
| Gestational diabetes status | -0.07 | <0.05 |
| Gestational hypertension status | -0.24 | <0.0001|
| Health following SARS-CoV-2 infection |         |
| Max COVID-19 severity index | 0.03 | 0.31 |
| Gestational age at maternal SARS-CoV-2 infection | 0.11 | <0.01 |
| Max patient class     | 0.05    | 0.13   |
| Number of SARS-CoV-2 encounters | 7.5 x 10⁴ | 0.98 |
| Max oxygen device     | -0.04   | 0.18   |
| Unique medication active ingredients | -0.21 | 0.0001|
| Total number of diagnoses | -0.07 | 0.05 |
| Unique number of diagnoses | -0.08 | 0.05 |
| Fetal characteristics  |         |
| Fetal sex             | 0.05    | 0.10   |
| Delivery cord vessels | 0.15    | <0.0001|
Because these variables have previously been associated with differences in birth outcomes in the USA, we control for these cofactors using propensity score matching.

A potential mechanism that could account for worse outcomes with SARS-CoV-2 infections earlier in gestation is increased levels of angiotensin converting enzyme 2 (ACE2) in the placenta earlier in gestation.\(^9\) The SARS-CoV-2 spike protein interacts with ACE2 for entry into human cells,\(^18\) and ACE2 placental levels are dependent on gestational age, with the highest levels observed early in gestation and near undetectable levels observed near term.\(^17,18\) This difference implicates an increased risk of SARS-CoV-2 infecting the placenta via ACE2 binding. Vertical SARS-CoV-2 transmission has been reported, but remains a rare mode of infection.\(^21,22\) However, an intrauterine infection can lead to ACE2-expressing neutrophils and monocytes (macrophage) to invade the placenta, which could result in increased risk for fetal distress.\(^17\) This is a potential mechanism by which increased stress occurs on the placenta during SARS-CoV-2 infection in a gestational age dependent manner. The histological features of SARS-CoV-2 infected placenta remain ill-defined due to scarce placenta histopathological reports. However, given that patients with COVID-19 have an increased risk of thromboembolic events, additional investigations into placental thrombosis, thromboembolic events, and anticoagulation in pregnant patients following maternal SARS-CoV-2 infection are justified.\(^21\) Only 35 people in the SARS-CoV-2 positive cohort received anticoagulants as a treatment for COVID-19 in this study. Because few people in this study received anticoagulation during acute COVID-19 infection, it is not possible to do a rigorous analysis on the association between this treatment and birth outcomes. This remains an important area for future investigation. Future studies examining the impact of SARS-CoV-2 infection on placenta health are needed to provide insight into the mechanism by which maternal SARS-CoV-2 infection might promote negative pregnancy outcomes.

The trend of increased rates of preterm birth and stillbirth among women who had a first or second trimester SARS-CoV-2 infection is concerning. Extremely preterm birth is associated with lower fetal survival and lifelong health problems, including neuropsychiatric impairment and elevated levels of major comorbidities.\(^21,24\) Furthermore, extremely preterm infants require substantial resources for care, with costs nearly 100 times that of care for full-term infant in the first 6 months of life.\(^26\) In addition to increased risk of negative pregnancy outcomes following SARS-CoV-2 infection, pregnant people have higher infection rates and are more likely to develop severe COVID-19.\(^1,2,25\) Taken together, both maternal and fetal health are at increased risk following maternal SARS-CoV-2 infection. It might, therefore, be prudent to consider pregnant people as a prioritised population for SARS-CoV-2 vaccination in areas where vaccine dissemination is scarce. There is also evidence that vaccine antibodies can pass through the umbilical cord to the fetus in utero, meaning that maternal vaccination can have the additional benefit of protecting the subsequent neonate.\(^20,21\)

The results suggest that additional monitoring of birth outcomes following first or second trimester maternal SARS-CoV-2 infection is warranted. The pregnant people who tested positive for SARS-CoV-2 in this study were unvaccinated against COVID-19; therefore, it would be interesting to examine in future studies whether vaccination helps to prevent negative birth outcomes in breakthrough cases. Also, modifiable features that are predictive of premature delivery can suggest hypotheses for clinical interventions, and additional features can be added to improve model accuracy and explore additional potential risk factors. The impact of in utero SARS-CoV-2 exposure on neonatal outcomes, including meeting developmental milestones, also needs further investigation. Finally, the impact of COVID-19 on subsequent pregnancy outcomes should be studied by examining the birth outcomes of non-pregnant women who had COVID-19 and later went on to become pregnant. Such studies are needed to assess the long-term impact of COVID-19 and the effect of post-acute sequelae of SARS-CoV-2 on pregnancy.

Here we examine outcomes of pregnancies that reached at least 20 weeks of gestation. It is possible that first trimester SARS-CoV-2 infection increases the risk of miscarriage, an outcome that we do not evaluate in this study. People in our study cohort had mild or moderate SARS-CoV-2 infections, so our findings might not extrapolate to a pregnant person with severe COVID-19, which remains rare in pregnant people. We have identified features that are predictive of gestational age at delivery, but they are not necessarily risk factors for premature delivery. Furthermore, this study was done at PSJH without validation at an independent health-care system. However, concerns regarding generalisability of this study are mitigated by the size and diversity of PSJH, which serves patients at 51 hospitals and 1085 clinics across five western US states. Additionally, this study focused on birth outcomes and does not account for neonatal health. Thus far, there have been no serious outcomes observed among infants exposed to SARS-CoV-2 as neonates.\(^22\) However, it is too early to observe the impact of in utero SARS-CoV-2 exposure on subsequent neonate and early development. Previously, in utero exposure to viruses have been shown to increase later risk of developing autism or other neuropsychiatric diseases.\(^25\) Furthermore, in utero exposure to infection is associated with the development of subseqent autoimmune diseases, including asthma and type 1 diabetes.\(^27,28\) Therefore, it is important to follow
This cohort study found that first and second trimester maternal SARS-CoV-2 infection was a risk factor for preterm birth and stillbirth. The greatest predictor of gestational age at delivery following a maternal SARS-CoV-2 infection was the gestational age at infection. In this cohort of patients with mild or moderate COVID-19, there was no correlation between COVID-19 severity and gestational age at delivery. These findings suggest that increased monitoring and enhanced prenatal care could be appropriate for pregnant people who have had a SARS-CoV-2 infection during the first or second trimester of pregnancy, regardless of infection severity.

Contributors
SNP, NPD, LH, and JJH conceptualised the study. LH and JJH supervised the study implementation. Funding was provided by NPD and JJH. Administrative and material support was provided by LH. SNP and RTR did the data cleaning and transformation. SNP did the data analyses, including statistical analysis and machine learning. SNP, YMH, and TS were involved in data interpretation. SNP and RTR, YMH, TS, and JJH had full access to the data in the study. All authors reviewed and approved the final version of the manuscript. SNP, RTR, YMH, TS, and JJH had full access to all the data in the study. All authors had full responsibility for the decision to submit for publication.

Declarations of interest
LH and NPD are scientific advisors for Sera Prognostics, a pregnancy diagnostics company, and hold stock options. The company is not associated with this study or any of the findings. All other authors declare no competing interests.

Data sharing
All clinical logic has been shared. Results have been aggregated and reported within this paper to the extent possible while maintaining privacy from personal health information as required by law. All data are archived within Providence St Joseph Health systems in a HIPAA-secure audited compute environment to facilitate verification of study conclusions.

Acknowledgements
This work was funded by a grant from the National Institutes of Health (HD091527), supplied by NPD. JHH has been funded in part with federal funds from the Department of Health and Human Services, Office of the Assistant Secretary for Preparedness and Response, Biomedical Advanced Research and Development Authority (contract number HHSO100201600031C). We thank Chengzhen Dai for code extracting patients’ insurance status and Stephen Lye for his insight into angiotensin converting enzyme 2 placental levels throughout pregnancy. We are grateful to Providence St Joseph Health for sharing their data engineering expertise and computational resources. We would also like to acknowledge SNOMED International for developing and maintaining SNOMED-CT.

References
1 Delahunty 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 Morb Mortal Wkly Rep 2020; 69: 147–54.
2 Jering KS, Cleggert BL, Cunningham JW, et al. Clinical characteristics and outcomes of hospitalized women giving birth with and without COVID-19. JAMA Intern Med 2021; 181: 714–17.
3 Lokken EM, Huebner EM, Gray Taylor G, et al. Disease severity, pregnancy outcomes and maternal deaths among pregnant patients with SARS-CoV-2 disease in Washington state. Am J Obstet Gynecol 2021; 225: e1–77.
4 Little C, Alsen M, Barlow J, et al. The impact of socioeconomic status on the clinical outcomes of COVID-19: a retrospective cohort study. J Community Health 2021; 46: 794–802.
5 Tai DBG, Shah A, Doubeni CA, Sia IG, Wieland ML. The disproportionate impact of COVID-19 on racial and ethnic minorities in the United States. Clin Infect Dis 2021; 72: 703–06.
6 Goldenberg RL, Culhane JF, Iams JD, Romero R. Epidemiology and causes of preterm birth. Lancet 2008; 371: 75–84.
7 Class QA, Lichtenstein P, Långström N, D’Onofrio BM. Timing of prenatal maternal exposure to severe life events and adverse pregnancy outcomes: a population study of 2.6 million pregnancies. Psychosom Med 2011; 73: 234–41.
8 Doré J, The effects of in utero exposure to influenza on birth and infant outcomes in the US. Pediatr Dev Rev 2019; 45: 489–523.
9 Havers FP, Reed C, Lim T, et al. Serorecovery of antibodies to SARS-CoV-2 in 30 Sites in the United States, March 23–May 12, 2020. JAMA Intern Med 2020; 180: 1576–86.
10 Austin PC, Fine JP. Propensity-score matching with competing risks in survival analysis. Stat Med 2019; 38: 751–77.
11 Pedregrosa F, Grisel O, Blondel M, et al. Scikit-learn: machine learning in Python. J Mach Learn Res 2011; 12: 2825–30.
12 Kiserud T, Piaggio G, Carroli G, et al. The World Health Organization fetal growth charts: a multinational longitudinal study of ultrasound biometric measurements and estimated fetal weight. PLoS Med 2017; 14: e1002220.
13 WHO. COVID-19 therapeutic trial synopsis. 2020. https://www.who.int/publications/i/item/covid-19-therapeutic-trial-synopsis (accessed Feb 19, 2021).
14 Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. 31st conference on Neural Information Processing Systems (NIPS 2017); Dec 4–9, 2017.
15 Collins GS, Reitsma GB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD Statement. BMJ 2015; 350: g7594.
16 Vogel JP, Chawanaalboon S, Moller A-B, et al. The global epidemiology of preterm birth. Best Pract Res Clin Obstet Gynaecol 2018; 52: 3–12.
17 Lye P, Dunk CE, Zhang J, et al. ACE2 is expressed in immune cells that infiltrate the placenta in infection-associated preterm birth. Cells 2021; 10: 1724.
18 Hoffmann M, Kleine-Weber H, Schroeder S, et al. SARS-CoV-2 cell entry depends on ACE2 and TMEM127 and is blocked by a clinically proven protease inhibitor. Cell 2020; 181: 271–80.
19 Wang Q, Zhang Y, Wu L, et al. Structural and functional basis of SARS-CoV-2 entry by using human ACE2. Cell 2020; 181: 894–904.
20 Zhou P, Yang X-L, Wang X-G, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature 2020; 579: 270–73.
21 Bloise E, Zhang J, Nakpu J, et al. Expression of severe acute respiratory syndrome coronavirus 2 cell entry genes, angiotensin-converting enzyme 2 and transmembrane protease serine 2, in the placenta across gestation and at the maternal-fetal interface in pregnancies complicated by preterm birth or pre eclampsia. Am J Obstet Gynecol 2021; 224: 298.
22 Wastnedge EAN, Reynolds RM, van Boeckel SR, et al. Pregnancy and COVID-19. Physiol Rev 2021; 101: 303–18.
23 Crump C, Winkleby MA, Sundquist J, Sundquist K. Prevalence of survival without major comorbidities among adults born prematurely. JAMA 2019; 322: 1380–88.
24 Rogers EE, Hintze SR. Early neurodevelopmental outcomes of extremely preterm infants. Semin Perinatol 2016; 40: 497–509.
25 Zambrano LD, Ellington S, Strid P, et al. Update: characteristics of asymptomatic women of reproductive age with laboratory-confirmed SARS-CoV-2 infection by pregnancy status—United States, January 22–October 1, 2020. MMWR Morb Mortal Wkly Rep 2020; 69: 1641–47.
26  Beam AL, Fried I, Palmer N, et al. Estimates of healthcare spending for preterm and low-birthweight infants in a commercially insured population: 2008-2016. *J Perinatol* 2020; 40: 1091–99.
27  Lokken EM, Gray Taylor G, Huebner EM, et al. Higher SARS-CoV-2 infection rate in pregnant patients. *Am J Obstet Gynecol* 2021; 225: 31–75.
28  Gilbert P, Rudnick C. Newborn antibodies to SARS-CoV-2 detected in cord blood after maternal vaccination—a case report. *BMC Pediatr* 2021; 21: 118.
29  Gill L, Jones CW. Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) antibodies in neonatal cord blood after vaccination in pregnancy. *Obstet Gynecol* 2021; 137: 894–96.
30  Al-Haddad BJS, Jacobsson B, Chakra S, et al. Long-term risk of neuropsychiatric disease after exposure to infection in utero. *JAMA Psychiatry* 2019; 76: 594–602.
31  Algert CS, Bowen JR, Lain SL, Allen HD, Vivian-Taylor JM, Roberts CL. Pregnancy exposures and risk of childhood asthma admission in a population birth cohort. *Pediatr Allergy Immunol* 2011; 22: 836–42.
32  Adams Waldorf KM, McAdams RM. Influence of infection during pregnancy on fetal development. *Reproduction* 2013; 146: R151–62.