Title
Air quality and preterm birth: distance to highways, exposure to wildfires, and effect modification by COVID-19

Author
Costello, Jean Michelle

Publication Date
2022

Peer reviewed|Thesis/dissertation

Air quality and preterm birth: distance to highways, exposure to wildfires, and effect modification by COVID-19

by
Jean Costello

DISSERTATION
Submitted in partial satisfaction of the requirements for degree of
DOCTOR OF PHILOSOPHY

in

Biological and Medical Informatics

in the

GRADUATE DIVISION
of the
UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Approved:

DocuSigned by:
John Witte
5991B9E9C971446...

John Witte
Chair

DocuSigned by:
Laura Jelliffe-Pawlowski

Laura Jelliffe-Pawlowski

DocuSigned by:
Martina Steurer

Martina Steurer

DocuSigned by:
Thomas Hoffmann
1DCDB1C1731941B...

Thomas Hoffmann

Committee Members

ACKNOWLEDGEMENTS

This work would not have been possible without the support and guidance of many people. I am grateful to all those who have helped me during the course of my degree.

I would like to thank both of my advisors, Laura Jelliffe-Pawlowski and John Witte. Laura, I am so grateful for all your support throughout my PhD. I am so glad you were willing to work with a student from outside your program. I learned so much from you, from your extensive preterm birth knowledge, to how to design a project and write a paper. John, thank you for your guidance and willingness to collaborate outside of the cancer space. I am deeply appreciative of all the statistical knowledge and support you gave. I am incredibly lucky to have had you both as mentors in graduate school.

I am also grateful to Dr. Martina Steurer, who, in addition to sitting on my committee, shared her extensive medical knowledge, helped with my projects, and supported me through the process of submitting my first paper. I would also like to thank my final committee member, Thomas Hoffmann, who offered vital statistical support. Additionally, I am grateful for Iona Cheng, who provided valuable geospatial knowledge which was critical to all of my projects.

I am thankful for the support I received from the Witte lab. I am grateful to my cohort member Taylor Cavazos, who helped me through the difficult phases of graduate school and answered all my tech questions. Taylor, I wish you all the best in moving towards graduation and onto the next stage. Linda Kachuri was always willing to share her epidemiology expertise, recommend methods, and review my DAGs. Many thanks to everyone who in the broader lab group for their feedback and contributions: Jovia Nierenberg, Pooja Suresh, Yulie Jiang, Geno Guerra, Rebecca Graff, Emmalyn Chen, Travis Meyers, Chenan Zhang, and Rosalyn Sayaman.

In addition, I would like to thank everyone at the Preterm Birth Initiative. I learned so

much from everyone in the group, from the clinical researchers to the wonderful members of the Community Advisory Board. I am particularly grateful to Rebecca Baer, for her dataset expertise and unnerving efficiency, Scott Oltman, for his insight and sense of humor, and Bri Taylor, for her warmth and incredible ability to build community over videocalls. I am especially lucky that my time in graduate school overlapped with that of Safyer McKenzie-Sampson. Safyer, I wish you the very best in graduate school and in your future career.

I am deeply grateful for my parents and my sister, for their encouragement and support.

Finally, I am so thankful for the patience and support I received from my partner, Mike.

CONTRIBUTIONS

Chapter 2 of this thesis is a reprint of the material as it appears in the journal *Paediatric and Perinatal Epidemiology*.[1] The published material is substantially the product of Jean Costello's period of study of study at UCSF and was primarily conducted and written by her. The work she completed for this published manuscript is comparable to a standard dissertation chapter.

[1]Costello, Jean M., et al. "Residential particulate matter, proximity to major roads, traffic density and traffic volume as risk factors for preterm birth in California." *Paediatric and perinatal epidemiology* 36.1 (2022): 70-79.

ABSTRACT

Air quality and preterm birth: distance to highways, exposure to wildfires, and effect modification by COVID-19

Jean M. Costello

Preterm birth, defined as birth prior to 37 weeks of gestation, affects approximately 1 in 10 live births and is the leading cause of neonatal morbidity and mortality in the United States. One known risk factor associated with preterm delivery is perinatal exposure to air pollution, and more specifically, fine particulate matter (particulate matter <2.5µm in diameter ($PM_{2.5}$)).

In this dissertation, I present three retrospective cohort studies investigating the relationship between $PM_{2.5}$ and preterm birth. The full cohort consisted of all births in the state of California, from 2007 to 2020. Data included maternal sociodemographic characteristics, residential address at the time of delivery, hospital delivery records, and infant characteristics. I combined this cohort data with publicly available data sources, including census data, transportation data, and environmental data repositories.

In the first study (Chapter 2), I examined the role that exposure to major highways plays in preterm birth risk. Comparing the risk associated with residential distance to major roads to that of fine particulate matter exposure, I found that the latter was significantly associated with preterm birth, while the former was not.

In the second (Chapter 3), I considered the impacts of wildfire-driven high $PM_{2.5}$ values. This study focused specifically on the San Francisco Bay Area during the 2018 California Camp Fire. Both exposure to the wildfire smoke and the degree of wildfire smoke experienced were significantly associated with preterm birth.

Finally, in the third study (Chapter 4), I investigated whether a positive test for SARS-CoV-2 modified the association between $PM_{2.5}$ and preterm birth. I found that such an effect modification was not present.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Preterm birth

Preterm birth (PTB) is defined as live birth prior to 37 completed weeks of gestation. Because of their interrupted fetal development, preterm infants face a range of complications. Mortality among preterm infants is significantly higher than in their term counterparts.[1] Preterm birth is also associated with a range of neonatal morbidities. Major examples include respiratory distress syndrome[2] and necrotizing enterocolitis,[3] but acute complications are wide-ranging, and may affect a variety of systems.[4]

Preterm birth is also associated with long-term health consequences. Studies investigating the medical and social consequences of preterm birth have identified associations with a range of conditions, including cognitive impairment, chronic lung conditions, cerebral palsy, and severe medical disability.[5, 6]

The impacts of preterm birth are not limited to the newborn; preterm birth has been shown to negatively impact the mental health of women and birthing people. Parents of preterm infants are at higher risk of postpartum depression.[7]

Given the risk associated with preterm birth and its prevalence (approximately 1 in 10 live births in the United States[8]), much work has been done to understand the risks of delivering preterm. Some of these risk factors are related to the demographic characteristics or socioeconomic status of the mother or birthing person. These include educational attainment, race/ethnicity, and insurance status: those with fewer completed years of education are at higher risk for delivering preterm than those with more completed years, those who are Black face higher rates of preterm birth than those who are white, and those on public insurance face higher rates than those on private insurance.[9] It is important to note that these risk factors should not be taken as causes of preterm birth, and many of these variables serve as well-recorded proxies for conditions that are more abstract or difficult to measure. For instance, completing less than twelve years of education does not lead to preterm birth, rather, it points to some level of socioeconomic deprivation that may have negative consequences. Similarly, Black race is not a "risk" in and of itself. Instead, it serves as an indicator of experienced stress and racism that may negatively impact outcomes.[10]

Some clinical factors also act as risk indicators for preterm birth. Among them are preexisting and gestational hypertension, preexisting and gestational diabetes, and drug or alcohol dependence or abuse.[9] Preterm birth is more common among twins and other multiple gestations than it is among singletons;[11] however, the present work focuses solely on singleton gestations, as the drivers of preterm birth in multiple gestations have been shown to be quite different.[12]

In addition to clinical and sociodemographic characteristics, research has shown that environmental exposures also present a risk for preterm birth. Perinatal exposure to lead, for instance, has been linked with preterm birth.[13] Multiple studies have shown that perinatal exposure to air pollution is associated with preterm birth.[14–16] Here, while investigating the relationship between air pollution and preterm birth, I will focus specifically on particulate matter (PM).

1.2 Particulate matter

Particulate matter (PM) is a mixture of solid particles and liquid droplets found in the air. Natural sources of PM include wildfires, dust storms, and volcanoes.[17–19] Manmade sources of PM include power generation, traffic, and agricultural emissions.[20–22]

PM can be classified in multiple ways. While the exact composition (or speciation) of PM is useful for examining potential health effects, this is difficult to determine, and varies by location and season.[23] In health contexts, it is common to consider the effects of "fine particulate matter," defined as PM less than 2.5 microns in diameter ($PM_{2.5}$). $PM_{2.5}$ is a concern because its small size allows it to penetrate into deeper parts of the lung and even enter the bloodstream.[24] In addition, high concentrations of $PM_{2.5}$ provide a large surface area of interaction between PM and the lungs.[25] Unlike PM speciation, $PM_{2.5}$ concentrations can be easily determined using scattering methods.

1.3 Biological pathways

While the focus of this dissertation is not the pathophysiology linking air pollution and preterm birth, it is important to establish the biological plausability of such an association. The pathway connecting $PM_{2.5}$ exposure with preterm birth is not fully understood; however, multiple mechanisms have been proposed. In their 2019 paper, Li and colleagues outlined three possible mechanisms.[26] The first of these is oxidative stress: $PM_{2.5}$ has been shown to create an imbalance between oxidants and antioxidants, which in turn may be associated with preterm birth.[27] The second proposed mechanism is mitochondrial DNA alteration and DNA methylation. $PM_{2.5}$ pollution has been linked with the methylation patterns of genes in the circadian pathway;[28] disruption or alteration of circadian rhythms may affect pregnancy processes. Finally, the authors propose that persistent organic pollutants (POPs) and other components of $PM_{2.5}$ can disrupt the endocrine system,[29] possibly leading to

preterm birth.

1.4 Dissertation overview

In this dissertation, I present three projects exploring the link between air pollution and preterm birth. First, I consider the role of highways in relation to preterm birth to determine whether distance to major highway serves as a proxy for $PM_{2.5}$ in assessing preterm birth risk.[30] Second, I investigate the impact of wildfire smoke from a 2018 California wildfire on preterm birth rates in the San Francisco Bay Area. Finally, I present work examining whether testing positive for COVID-19 modifies the association between $PM_{2.5}$ and PTB.

Bibliography

[1] Richard E Behrman, Beverly L Koops, Linda J Morgan, and Frederick C Battaglia. Neonatal mortality risk in relation to birth weight and gestational age: update. *The Journal of pediatrics*, 101(6):969–977, 1982.

[2] Patricia A Robertson, Susan H Sniderman, Russell K Laros Jr, Ronald Cowan, David Heilbron, Robert L Goldenberg, Jay D Iams, and Robert K Creasy. Neonatal morbidity according to gestational age and birth weight from five tertiary care centers in the united states, 1983 through 1986. *American journal of obstetrics and gynecology*, 166 (6):1629–1645, 1992.

[3] Shoo K Lee, Douglas D McMillan, Arne Ohlsson, Margaret Pendray, Anne Synnes, Robin Whyte, Li-Yin Chien, Joanna Sale, et al. Variations in practice and outcomes in the canadian nicu network: 1996–1997. *Pediatrics*, 106(5):1070–1079, 2000.

[4] Richard E Behrman and Adrienne Stith Butler. Committee on understanding premature birth and assuring healthy outcomes board on health sciences policy: Preterm birth: causes, consequences, and prevention, 2007.

[5] MJ Platt. Outcomes in preterm infants. *Public health*, 128(5):399–403, 2014.

[6] Dag Moster, Rolv Terje Lie, and Trond Markestad. Long-term medical and social consequences of preterm birth. *New England Journal of Medicine*, 359(3):262–273, 2008.

[7] Juliana Arantes Figueiredo de Paula Eduardo, Marcos Goncalves de Rezende, Paulo Rossi Menezes, and Cristina Marta Del-Ben. Preterm birth as a risk factor for postpartum depression: A systematic review and meta-analysis. *Journal of affective disorders*, 259:392–403, 2019.

[8] J.A. Martin, B.E. Hamilton, M.J.K. Osterman, and A.K. Driscoll. Births: Final data for 2018. *National Vital Statistics Reports*, 68(13).

[9] Rebecca J Baer, Monica R McLemore, Nancy Adler, Scott P Oltman, Brittany D Chambers, Miriam Kuppermann, Matthew S Pantell, Elizabeth E Rogers, Kelli K Ryckman, Marina Sirota, et al. Pre-pregnancy or first-trimester risk scoring to identify women at high risk of preterm birth. *European Journal of Obstetrics & Gynecology and Reproductive Biology*, 231:235–240, 2018.

[10] Zinzi D Bailey, Nancy Krieger, Madina Agénor, Jasmine Graves, Natalia Linos, and Mary T Bassett. Structural racism and health inequities in the usa: evidence and interventions. *The lancet*, 389(10077):1453–1463, 2017.

[11] Anna Heino, Mika Gissler, Ashna D Hindori-Mohangoo, Béatrice Blondel, Kari Klungsøyr, Ivan Verdenik, Ewa Mierzejewska, Petr Velebil, Helga Sól Ólafsdóttir, Alison Macfarlane, et al. Variations in multiple birth rates and impact on perinatal outcomes in europe. *PloS one*, 11(3):e0149252, 2016.

[12] Sarah Stock and Jane Norman. Preterm and term labour in multiple pregnancies. In *Seminars in Fetal and Neonatal Medicine*, volume 15, pages 336–341. Elsevier, 2010.

[13] LL Jelliffe-Pawlowski, SQ Miles, JG Courtney, B Materna, and V Charlton. Effect of magnitude and timing of maternal pregnancy blood lead (pb) levels on birth outcomes. *Journal of Perinatology*, 26(3):154–162, 2006.

[14] Bruce Bekkar, Susan Pacheco, Rupa Basu, and Nathaniel DeNicola. Association of air pollution and heat exposure with preterm birth, low birth weight, and stillbirth in the us: a systematic review. *JAMA network open*, 3(6):e208243–e208243, 2020.

[15] Mary Huynh, Tracey J Woodruff, Jennifer D Parker, and Kenneth C Schoendorf. Relationships between air pollution and preterm birth in california. *Paediatric and perinatal epidemiology*, 20(6):454–461, 2006.

[16] Jun Wu, Cizao Ren, Ralph J Delfino, Judith Chung, Michelle Wilhelm, and Beate Ritz. Association between local traffic-generated air pollution and preeclampsia and preterm delivery in the south coast air basin of california. *Environmental health perspectives*, 117(11):1773–1779, 2009.

[17] Mary K Butwin, Sibylle von Löwis, Melissa A Pfeffer, and Throstur Thorsteinsson. The effects of volcanic eruptions on the frequency of particulate matter suspension events in iceland. *Journal of Aerosol Science*, 128:99–113, 2019.

[18] Throstur Thorsteinsson, Guđrún Gísladóttir, Joanna Bullard, and Grant McTainsh. Dust storm contributions to airborne particulate matter in reykjavík, iceland. *Atmospheric Environment*, 45(32):5924–5933, 2011.

[19] Shuhui Guan, David C Wong, Yang Gao, Tianqi Zhang, and George Pouliot. Impact of wildfire on particulate matter in the southeastern united states in november 2016. *Science of The Total Environment*, 724:138354, 2020.

[20] Elizabeth Pattey and Guowang Qiu. Trends in primary particulate matter emissions from canadian agriculture. *Journal of the Air & Waste Management Association*, 62 (7):737–747, 2012.

[21] Pallavi Pant and Roy M Harrison. Estimation of the contribution of road traffic emissions to particulate matter concentrations from field measurements: A review. *Atmospheric environment*, 77:78–97, 2013.

[22] Steven G Brown, Anna Frankel, Sean M Raffuse, Paul T Roberts, Hilary R Hafner, and Darcy J Anderson. Source apportionment of fine particulate matter in phoenix, az, using positive matrix factorization. *Journal of the Air & Waste Management Association*, 57 (6):741–752, 2007.

[23] S Hasheminassab, N Daher, A Saffari, D Wang, BD Ostro, and C Sioutas. Spatial and

temporal variability of sources of ambient fine particulate matter (pm 2.5) in california. *Atmospheric Chemistry and Physics*, 14(22):12085–12097, 2014.

[24] Shaolong Feng, Dan Gao, Fen Liao, Furong Zhou, and Xinming Wang. The health effects of ambient pm2. 5 and potential mechanisms. *Ecotoxicology and environmental safety*, 128:67–74, 2016.

[25] Flavia Mazzoli-Rocha, Silviane Fernandes, Marcelo Einicker-Lamas, and Walter Araújo Zin. Roles of oxidative stress in signaling and inflammation induced by particulate matter. *Cell biology and toxicology*, 26(5):481–498, 2010.

[26] Zhou Li, Yuqing Tang, Xin Song, Lissy Lazar, Zhen Li, and Jinshun Zhao. Impact of ambient pm2.5 on adverse birth outcome and potential molecular mechanism. *Ecotoxicology and environmental safety*, 169:248–254, 2019.

[27] Tiffany A Moore, Iman M Ahmad, and Matthew C Zimmerman. Oxidative stress and preterm birth: an integrative review. *Biological research for nursing*, 20(5):497–512, 2018.

[28] Tim S Nawrot, Nelly D Saenen, Julie Schenk, Bram G Janssen, Valeria Motta, Letizia Tarantini, Bianca Cox, Wouter Lefebvre, Charlotte Vanpoucke, Cristina Maggioni, et al. Placental circadian pathway methylation and in utero exposure to fine particle air pollution. *Environment international*, 114:231–241, 2018.

[29] Quan Zhang, Jinghua Wang, Jianqiang Zhu, Jing Liu, and Meirong Zhao. Potential glucocorticoid and mineralocorticoid effects of nine organophosphate flame retardants. *Environmental science & technology*, 51(10):5803–5810, 2017.

[30] Jean M Costello, Martina A Steurer, Rebecca J Baer, John S Witte, and Laura L Jelliffe-Pawlowski. Residential particulate matter, proximity to major roads, traffic density and traffic volume as risk factors for preterm birth in california. *Paediatric and perinatal epidemiology*, 36(1):70–79, 2022.

Chapter 2

Proximity to highways

2.1 Abstract

Background: While pollution from vehicle sources is an established risk factor for preterm birth, it is unclear if distance of residence to the nearest major road or related measures like major road density represent useful measures for characterizing risk.

Objective: To determine whether major road proximity measures (including distance to major road, major road density, and traffic volume) are more useful risk factors for preterm birth than other established vehicle-related measures (including particulate matter $< 2.5\mu m$ in diameter (PM2.5) and diesel particulate matter (diesel PM)).

Methods: This retrospective cohort study included 2.7 million births across the state of California from 2011 - 2017; each address at delivery was geocoded. Geocoding was used to calculate distance to the nearest major road, major road density within a 500m radius, and major road density weighted by truck volume. We measured associations with preterm birth using risk ratios adjusted for target demographic, clinical, socioeconomic, and environmental covariates (aRRs). We compared these to the associations between preterm birth and PM2.5

and diesel PM by census tract of residence.

Results: Findings showed that whereas higher mean levels of PM2.5 and diesel PM by census tract were associated with a higher risk of preterm birth, living closer to roads or living in higher traffic density areas was not associated with higher risk. Residence in a census tract with a mean PM2.5 in the top quartile compared to the lowest quartile was associated with the highest observed risk of preterm birth (aRR 1.04, 95% CI 1.04, 1.05).

Conclusions: Over a large geographical region with a diverse population, PM2.5 and diesel PM were associated with preterm birth, while measures of distance to major road were not, suggesting that these distance measures do not serve as a proxy for measures of particulate matter in the context of preterm birth.

2.2 Introduction

Preterm birth (birth before 37 weeks of gestation) affects approximately one in ten newborns in the United States.[1] Preterm birth is the leading cause of perinatal mortality and morbidity, and has been implicated in long-term health problems, including neurodevelopmental disabilities and psychological disorders.[2] Air pollution has been implicated as a risk factor for preterm birth,[3, 4] and prior studies have found that emissions from sources including coal power, oil and gas production,[5, 6] and traffic[7] are associated with a higher risk of preterm birth.

Air pollution can be difficult to evaluate and measurements may be restricted to small geographical areas near a monitor.[8] A widely applicable measure of exposure is distance to a major road, which is of particular interest given that prior research has suggested traffic-related pollutants are associated with preterm birth.[7]

Distance to major road has been examined in previous studies with mixed findings. A

study in Taiwan showed a higher risk of preterm birth in women living <500 meters from a major highway compared to women living 500 - 1,500 meters from the same highway.[9] Researchers in Japan found higher risk when comparing women who lived <200 meters from a highway to those who lived farther.[10, 11] A large cohort study in the state of North Carolina found higher risk for preterm birth based on distance of residence to road.[12] By contrast, investigators in Vancouver[13] and Rotterdam did not find an association with preterm birth.[14]

A particular challenge in previous studies lies in acquiring a sufficient sample size while simultaneously incorporating both demographic and clinical covariates. Most of these studies did not include clinical covariates associated with preterm birth, considered a small geographic area, and examined major road proximity as the only measure of air pollution.

Here, we used a large cohort of births across the state of California, in combination with state highway maps and traffic measures, to test whether living closer to major roads was associated with a higher risk of preterm birth. We incorporated both demographic and clinical covariates and assessed multiple measures of road and traffic exposure.

2.3 Methods

2.3.1 Cohort data

The sample was drawn from all California live-born infants between 2011 and 2017. Birth certificates, maintained by California Vital Statistics, were linked to a hospital discharge, emergency department, and ambulatory surgery records maintained by the California Office of Statewide Health Planning and Development. We restricted to singleton births at 20 - 44 weeks of gestational age where the mother was 12 - 55 years old. Census-level environmental data must be available (Figure 2.2). Our outcome of interest was preterm birth, defined as a live delivery prior to 37 weeks of gestation. Best obstetric estimate of gestation at birth

was obtained from birth certificate records.

2.3.2 Exposure variables

We considered six measures of exposure, detailed below. The first three (distance to major road, major road density, and major road density with truck volume) are calculated at an individual level; the remaining three are census tract-level.

Distance to major road

The address of residence was recorded at birth. Addresses were geocoded in ArcGIS[15] using files from the US Census Bureau.[16] Distance to major road was determined by computing the distance from each geocoded address to the nearest state highway (as defined by CalTrans,[17] see Figure 2.1). We considered distance both as a continuous variable and as a binary exposure (comparing those living <250m from a road to those living 250-500m from a road, and those living <250m from a road to all others).

Major road density

For each address, a density score was computed using the following model:

$$\sum_{i}^{n} \frac{1}{d_i} l_i,$$

where n is the number of roads within a 500m radius of the residence, d_i is the distance (in meters) to the nearest point on road i, and l_i is the length (in meters) of road i that falls within a 500m radius of the residence. The radius was chosen based on previously published work covering a large geography.[12] For addresses with no roads nearer than 500m, the major road density is $1/d$, where d is the distance to the nearest major road.

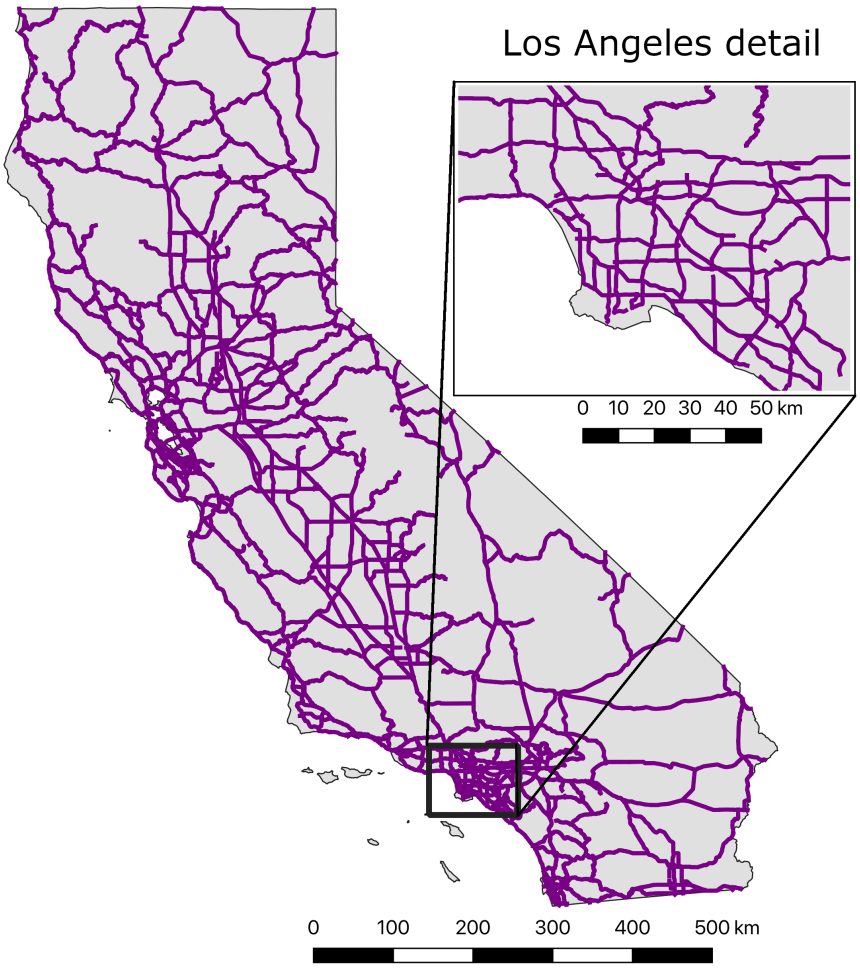

Figure 2.1: Major roads in California, as defined by Caltrans.

Major road density with truck volume

An additional major road density measure was computed to incorporate truck volume. Volume measures were obtained from CalTrans and represent annual average daily truck traffic.[18] Scores were computed in the following manner:

$$\sum_{i}^{n} \frac{1}{d_i} l_i v_i,$$

where v_i represents the truck volume of road i; all other variables are as described in the previous model. To compute the contribution of road i, we used the volume measurement on road i nearest to the residence. In a small number of cases, there were no volume measures available for a particular highway. In these cases, the nearest volume measurement was used. For addresses with no roads nearer than 500m, we instead computed v/d where d is the distance to the nearest road, and v is the truck volume of that road.

Traffic

CalEnviroScreen provides a census tract-level measure of traffic concentration. Road segments falling within a 150m census tract buffer are linked to the most recent traffic volume data (the latest year being 2013). The result is number of vehicles (adjusted by road segment lengths) per hour per kilometer of roadways within the census tract buffer. Each tract is assigned a percentile based on its rank in the statewide distribution.[19]

Particulate matter 2.5 ($PM_{2.5}$)

$PM_{2.5}$ refers to particles with a diameter of $\leq 2.5 \mu g/m^3$. The specific composition of $PM_{2.5}$ is dependent on multiple factors, including location, season, and meteorologic conditions.[20] Measures of census tract-level $PM_{2.5}$ concentration come from CalEnviroScreen.[21] CalEnviroScreen computes these measures by using ordinary kriging to estimate quarterly mean $PM_{2.5}$ concentrations at the center of the census tract from nearby air monitors, then aver-

aging quarterly means to annual means. The measure provided is the mean of three annual averages (2012-14). Each tract is assigned a percentile based on its rank in the statewide distribution.[19]

Diesel Particulate Matter (Diesel PM)

Measurements of diesel PM are provided by CalEnviroScreen on a census tract-level and include both on-road and non-road sources. CalEnviroScreen computes this measure by first taking county-wide estimates of diesel PM (from a weekday in July 2012); a measure is assigned to a census tract based on the overlap of diesel sources and the census tract population. Each tract is assigned a percentile based on its rank in the statewide distribution.[19]

2.3.3 Covariates

We included a range of covariates in our models, which we identified based on previous literature.[22] For our analyses, we grouped the covariates into four categories: demographic: maternal age (<18 years or >34 years compared to 18-34 years), maternal body mass index (BMI) (underweight (<18.5), overweight (25 - 30), or obese (>30) compared to normal (18.5 - <25)), maternal race/ethnicity (Black, Hispanic, Asian, or other compared to white non-Hispanic), maternal birthplace outside the US compared to within the US, smoking, drug abuse and alcohol abuse (defined by the presence of an ICD code from the following list on the hospital discharge record: 648.3, 305.0, 305.2, 305.3, 305.4, 305.5, 305.6, 305.7, 305.8, 305.9, 304; 303 (ICD-9); F10, F11, F12, F13, F14, F15, F16, F18, F19 (ICD-10)), male infant sex compared to female; clinical: nulliparity, interpregnancy interval (<6 months, 6-11 months, 12-17 months, 24-59 months, or >59 months compared to 18-23 months[23], mental disorder, maternal infection, preexisting diabetes, gestational diabetes, preexisting hypertension, gestational hypertension, and adequacy of prenatal care (inadequate (care begun after 4th month or <50% of recommended visits received) or intermediate (care begun by 4th month and 50%-79% of recommended visits received) compared to adequate/adequate plus (care

begun by 4th month and ≥80% of recommended visits received)[24]); socioeconomic: WIC (Women, Infants and Children - a federal nutrition assistance program) enrolment compared to none, public insurance for prenatal care compared to private insurance, and maternal education (>12 years or <12 years compared to 12 years); and environmental: season of conception (winter, spring, or autumn compared to summer), and urbanicity (Urban and Large Rural compared to Small and Isolated Small Rural). Urbanicity was assigned at the census tract-level using Rural-Urban Commuting Area Codes (RUCA).[25] We used a classification defined on the Rural Health Research Center's website to group codes into the three categories listed.[26]

2.3.4 Statistical analysis

We computed a correlation matrix to examine the relationship of exposures to one another. We calculated unadjusted risk ratios to check whether covariates were associated with preterm birth. To assess the relationship between each exposure and preterm birth, we used multivariable log-linear models, adjusting for the covariates found to be associated with preterm birth. Each exposure was considered independently. Covariates were introduced in stages. The first model included only the exposure of interest. The second included the exposure and demographic covariates. The third included the exposure and demographic and clinical covariates. The fourth included the exposure and demographic, clinical, and socioeconomic covariates. The fifth included the exposure and demographic, clinical, socioeconomic, and environmental covariates. Adjusted risk ratios (aRR) and 95% confidence intervals (95% CI) were computed in each phase. Statistical analyses were performed in R (version 4.04).[27] To understand the role of race/ethnicity in our models, we repeated the analysis steps for each category (Asian, Black, Hispanic, and White).

2.3.5 Missing data

We successfully geocoded approximately 85% of all deliveries from the time period. For those variables which have a code for "missing," we display the percent of missing data in **Table 2.1**. Approximately 9% of people in our data sample were missing at least one covariate. We performed a multiple imputation using the mice package in R[28]; we computed 50 imputations with 5 iterations each. Analysis results for both the original dataset and imputed data are in **Table 2.3**. Results presented in the paper are from the original dataset.

Table 2.1: Missing data counts.

Variable	**Missing count**	**Missing percent**
Race/ethnicity	56197	2
Insurance	2914	0.1
Education	118513	4.2
Maternal Nativity	1042	0
Infant Sex	21	0
Parity	1767	0.1
Maternal Age	26	0
BMI	115158	4.1
Prenatal care	62366	2.2

2.3.6 Sensitivity analyses

We computed E-values to assess the robustness of our associations.[29] There are two types of clustering in our data. Our dataset contains multiple singleton births to the same mother, not all of which have been properly identified. We repeated the log-linear analysis selecting only one birth from each mother and found this did not alter our results. Additionally, some of our exposure variables are calculated on a census tract-level, while other covariates are at the individual level. For the exposure $PM_{2.5}$, we compared a single-level logistic regression to mixed effects logistic regression on a subset of the data and found similar results.

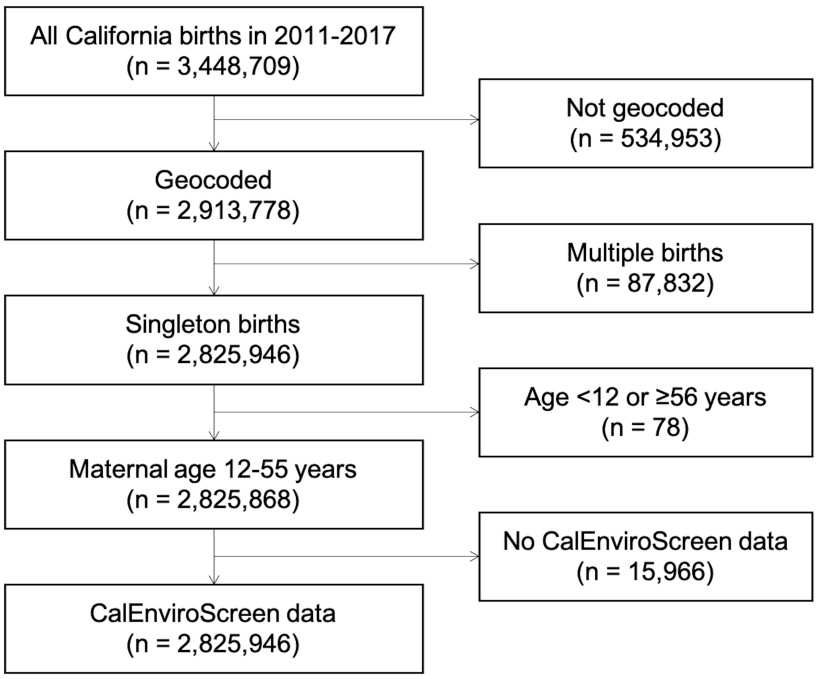

Figure 2.2: Sample selection

2.3.7 Ethics approval

Methods and protocols for the study were approved by the Committee for the Protection of Human Subjects (CPHS) which serves as the institutional review board (IRB) for the California Health and Human Services Agency (CHHSA).

2.4 Results

Among the eligible births (n = 2,809,902), the preterm birth rate was 6.6%, consistent with California's rate for singleton births. A plurality of our sample self-identified as Hispanic (49.2%). Most were between the ages of 18 and 34 at the time of delivery (77.9%) and most had more than 12 years of education (58.4%). Characteristics are displayed in Table 1.

Table 2.2: Sample overview.

Characteristic	Preterm	Term
	(n = 187,235)	n = (2,622,667)
Maternal age		
18-34	137,655 (73.5)	2,052,044 (78.2)
>34	45,976 (24.6)	527,419 (20.1)
<18	3,604 (1.9)	43,204 (1.6)
Race/Ethnicity		
White	41,455 (22.1)	709,385 (27.0)
Hispanic	96,103 (51.3)	1,285,817 (49.0)
Black	13,440 (7.2)	124,512 (4.7)
Asian	26,116 (13.9)	374,150 (14.3)
Other	10,121 (5.4)	128,803 (4.9)
Maternal BMI		
Normal	85,675 (45.8)	1,299,669 (49.6)
Underweight	7,494 (4.0)	97,133 (3.7)
Overweight	46,495 (24.8)	660,393 (25.2)
Obese	47,571 (25.4)	565,472 (21.6)
Mother born in US	116,344 (62.1)	1,604,394 (61.2)
Smoking status	9,377 (5.0)	74,670 (2.8)
Drug abuse	8,906 (4.8)	46,530 (1.8)
Alcohol abuse	753 (0.4)	4,871 (0.2)
Infant sex		
Female	84,135 (44.9)	1,294,005 (49.3)
Male	103,100 (55.1)	1,328,662 (50.7)
Nulliparous	71,703 (38.3)	1,011,031 (38.5)
Mental disorder	18,780 (10.0)	145,232 (5.5)
Any infection	31,715 (16.9)	264,457 (10.1)
Gestational diabetes	30,091 (16.1)	269,299 (10.3)
Gestational hypertension	30,434 (16.3)	153,673 (5.9)
Interpregnancy interval (IPI)		
18-23 months	7,733 (4.3)	139,936 (5.5)
<6 months	7,035 (3.9)	67,354 (2.7)
6-11 months	8,806 (4.8)	121,995 (4.8)
12-17 months	9,422 (5.2)	158,617 (6.3)
24-59 months	33,659 (19.0)	534,148 (21.4)
>59 months	31,014 (16.6)	358,706 (13.7)
Prenatal care		
Adequate	150,476 (80.4)	1,980,213 (75.5)
Intermediate	15,496 (8.3)	374,176 (14.3)
Inadequate	21,263 (11.4)	268,278 (10.2)
Preexisting diabetes	5,110 (2.7)	20,726 (0.8)
Preexisting hypertension	13,716 (7.3)	48,368 (1.8)
Maternal education		

Characteristic	Preterm	Term
12 years	48,163 (25.7)	641,221 (24.4)
>12 years	102,907 (55.0)	1,539,378 (58.7)
<12 years	36,165 (19.3)	442,068 (16.9)
Public insurance	89,112 (47.6)	1,182,639 (45.1)
Birth season		
Summer	48,765 (26.0)	696,994 (26.6)
Winter	45,593 (24.4)	631,994 (24.1)
Spring	45,888 (24.5)	636,131 (24.3)
Autumn	46,989 (25.1)	657,548 (25.1)
Urbanicity		
Small & isolated small rural	2,684 (1.4)	37,473 (1.4)
Large rural	5,617 (3.0)	79,212 (3.0)
Urban	178,934 (95.6)	2,505,982 (95.6)

Correlations between the exposures are displayed in **Figure 2.3**. Distance to major road was slightly negatively correlated with traffic and diesel PM, with coefficients of -0.35 and -0.29, respectively. Major road density and major road density with truck volume were strongly correlated with one another, but not correlated with other exposures.

In unadjusted analyses, we found that women of color faced elevated risk of preterm birth; Black women had the greatest relative risk, with an RR of 1.76 (95% CI 1.73, 1.80) compared to white women. Hispanic women, who made up the majority of our sample, had an RR of 1.26 (95% CI 1.25, 1.27). Other factors found to be associated with a higher risk of preterm birth were an age <18 or >34 years at delivery, an overweight, obese, or underweight BMI, smoking, drug/alcohol use, a short or long IPI, the presence of a mental disorder, an infection, gestational and preexisting diabetes, gestational and preexisting hypertension, <12 years of education, WIC enrolment, and public insurance. All seasons of conception showed very slightly higher risk compared to summer. Intermediate prenatal care and >12 years of education reduced the risk of preterm birth with RRs of 0.56 (95% CI 0.55, 0.57), and 0.90 (95% CI 0.89, 0.91), respectively. Unadjusted risk ratios for all covariates are shown in **Figure 2.4**. Urbanicity was not associated with preterm birth and was not included in further analyses.

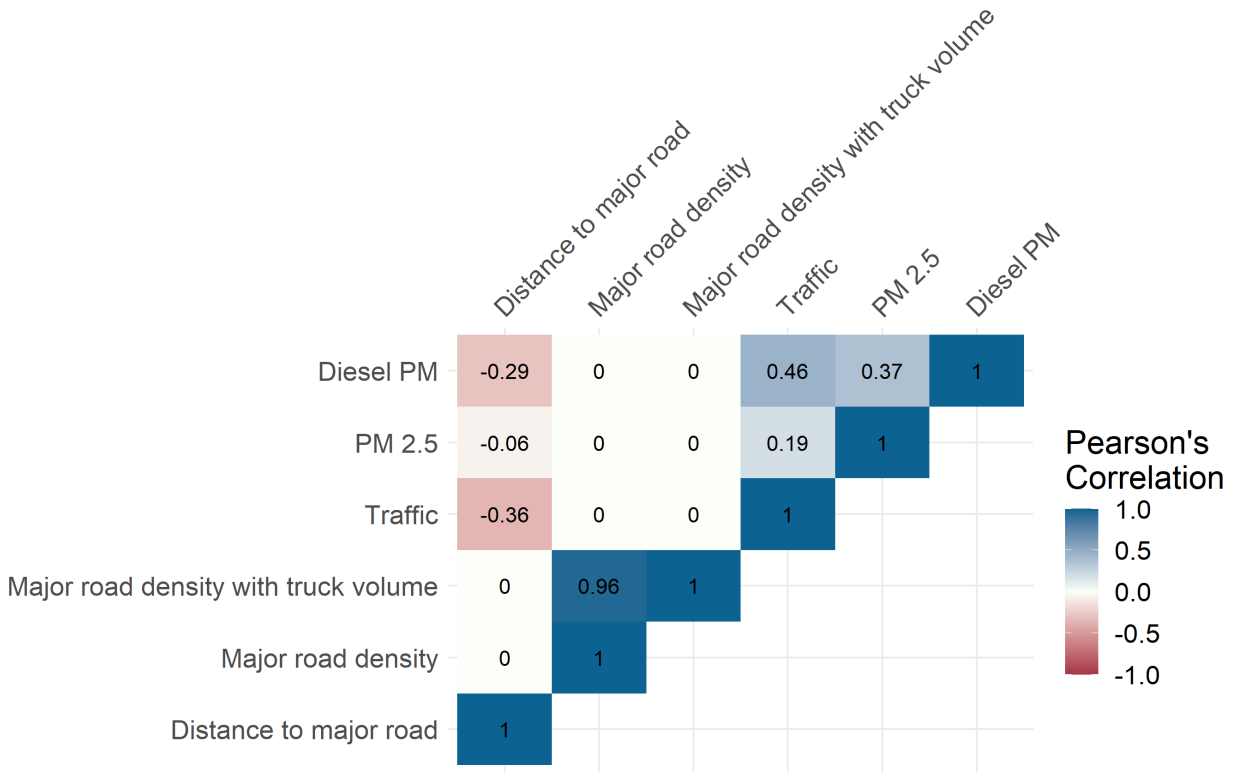

Figure 2.3: Correlation matrix of the exposures.

Unadjusted and adjusted risk ratios for each measure of exposure are shown in **Figure 2.5**. Exposures were considered separately; each row of each panel represents a single log-linear model. Panel A shows the unadjusted risk ratio for each exposure. The models shown in Panel B consist of the exposure and demographic covariates. The fully adjusted models are displayed in Panel C; each includes an exposure along with demographic, clinical, socioeconomic, and environmental covariates. Both $PM_{2.5}$ and diesel PM were associated with preterm birth in the unadjusted and in all adjusted models. The aRR (per interquartile increase) for $PM_{2.5}$ was 1.08 (95% CI 1.07, 1.08), E-value = 2.57 (E-valuelower = 1.34) and for diesel PM was 1.02 (95% CI 1.01, 1.03), E-value = 2.45 (E-valuelower = 1.11). Traffic, as measured by CalEnviroScreen, was not associated with preterm birth in the unadjusted or any adjusted model; the aRR (per interquartile increase) was .99 (95% CI .99, 1.00). Major road density and major road density with truck volume were also not associated with preterm birth; the aRRs (per interquartile increase) were .99 (95% CI .99, 1.00) and .99

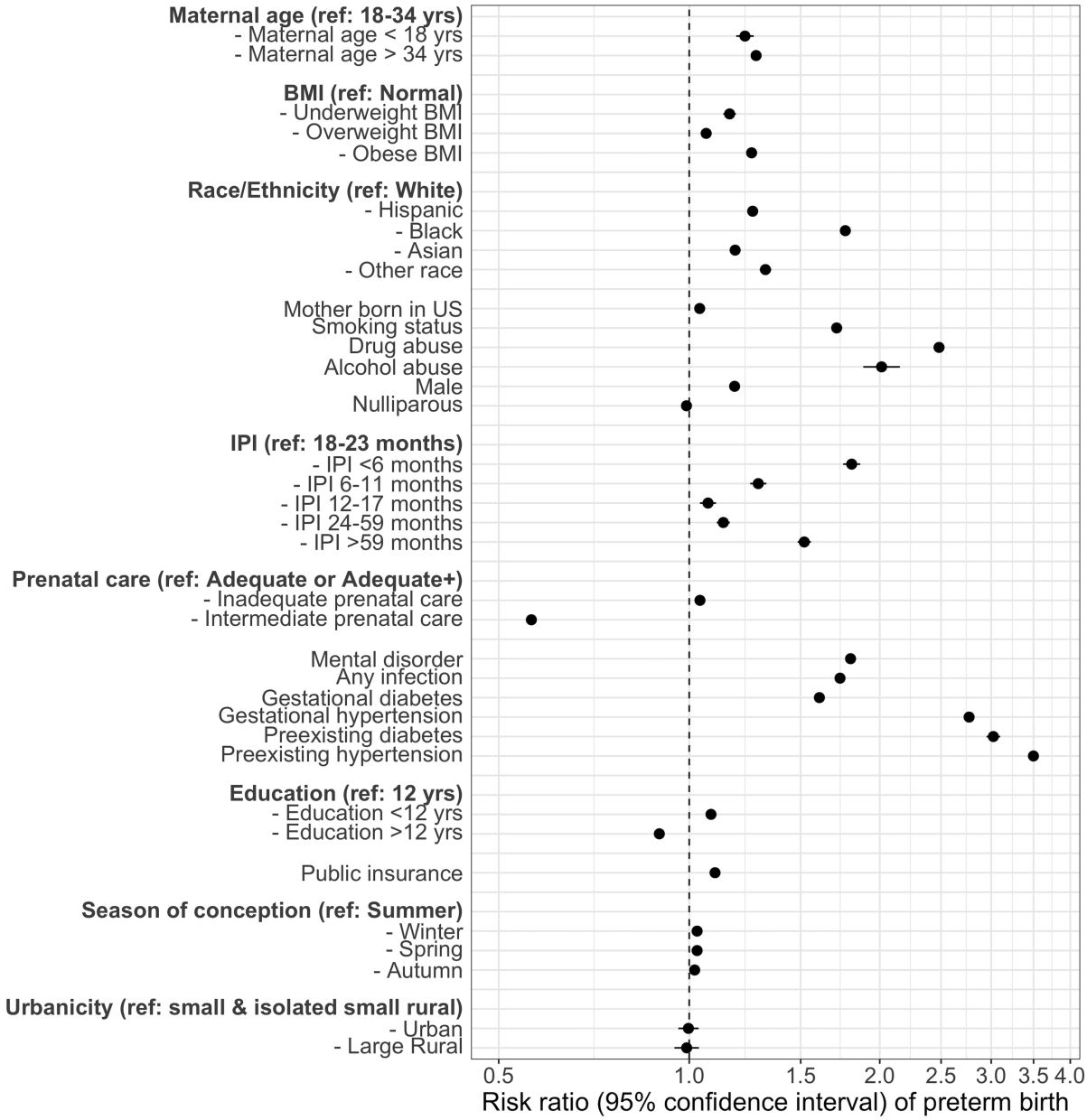

Figure 2.4: Forest plots for unadjusted associations between covariates and preterm birth. Unadjusted risk ratios and 95% confidence interval are shown on a log scale.

(95% CI .99, 1.00), respectively. Distance to major road (as a continuous variable) showed a very slightly higher risk in the adjusted model, with an aRR of 1.01 (95% CI >1.00, 1.01) per 1,000m increase in distance (**Table 2.3**).

Table 2.3: Original and imputed results.

Variable	**Original Results aRR (95% CI)**	**Imputed Results aRR (95% CI)**
Distance to major road	1.01 (1.00, 1.01)	1.01 (1.00, 1.01)
Major road density	1.00 (0.99, 1.00)	0.99 (0.98, 0.99)
Major road density with truck volume	1.00 (0.99, 1.00)	0.99 (0.98, 0.99)
Traffic	0.99 (0.99, 1.00)	0.99 (0.99, 1.00)
Diesel PM	1.02 (1.01, 1.03)	1.02 (1.01, 1.02)
PM 2.5	1.08 (1.07, 1.08)	1.07 (1.06, 1.08)

Stratified analysis by race/ethnicity revealed different strengths and directions of effects between groups. Traffic density was associated with a reduced risk of preterm birth in Asian women (aRR 0.96, 95% CI 0.94, 0.98)), and a higher risk in Hispanic women (aRR 1.01, 95% CI >1.00, 1.03).

Distance to major road as a binary exposure was not found to be associated with preterm birth in any model. Comparing births with addresses within 250m of a major road to all others, the aRR was 1.00 (95% CI .98, 1.01). Comparing births with addresses within 250m of a major road to those between 250m and 500m, the aRR was 1.00 (95% CI .99, 1.02).

2.5 Comment

2.5.1 Principal findings

Our findings show that, while census tract-level mean $PM_{2.5}$ and diesel PM are positively associated with preterm birth, this association is not seen in measures of distance to major road. This suggests that, over a large geographical area, these are not equivalent measures of pollution exposure when considering the risk of preterm birth. The association between

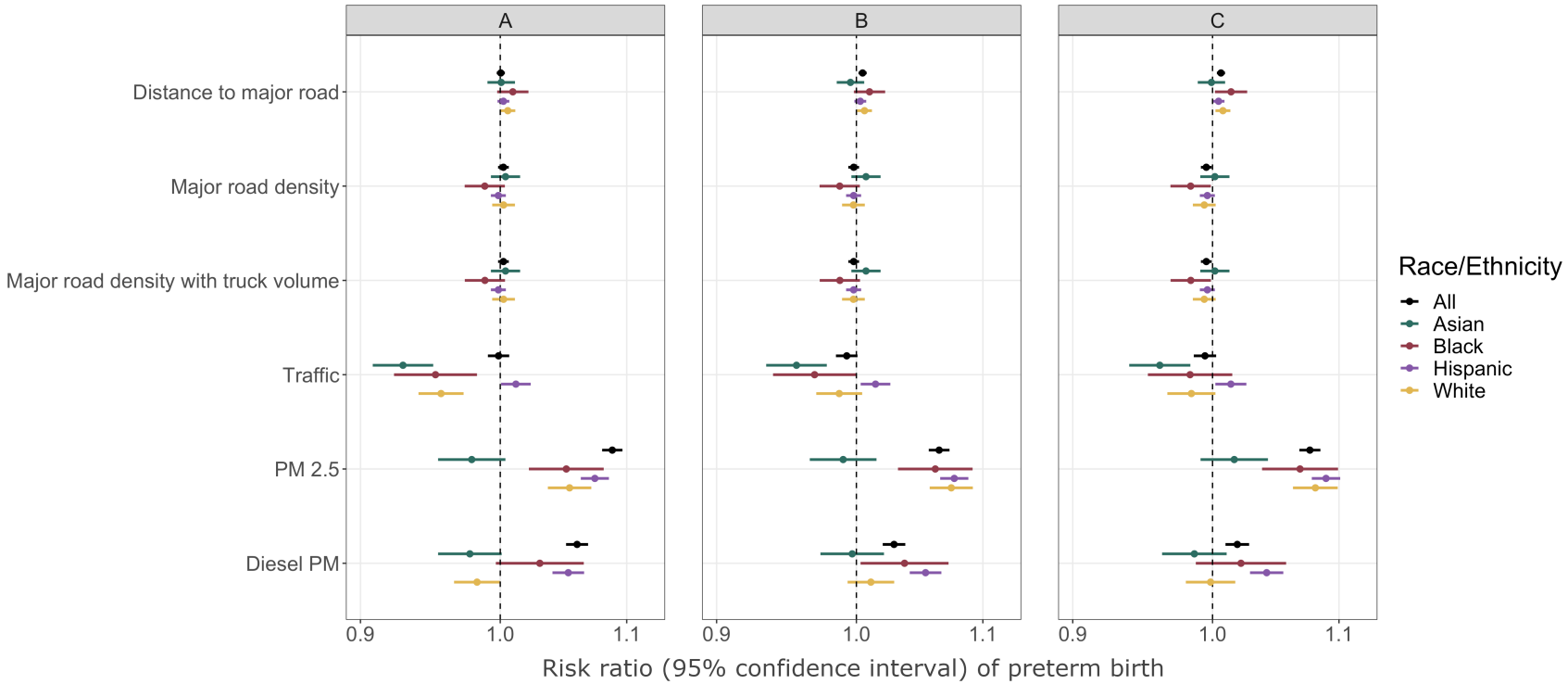

Figure 2.5: Each predictor is considered separately. For Distance to major road, the risks are shown for an increase of 1000m. For all other exposures, risks are shown for interquartile increase. A: unadjusted risk ratios. B: risk ratios adjusted for maternal age, maternal BMI, maternal nativity, smoking status, drug abuse and alcohol abuse. The 'All' category is also adjusted by race/ethnicity. C: risk ratios adjusted for all variables listed in B as well as infant sex, parity, interpregnancy interval, mental disorder, maternal infection, preexisting diabetes, gestational diabetes, preexisting hypertension, gestational hypertension, prenatal care, maternal education, insurance status, WIC enrolment, and season.

air pollution and preterm birth was not consistent across all race/ethnicity groups: neither $PM_{2.5}$ nor diesel PM showed an association with preterm birth in Asian women.

2.5.2 Strengths of the study

Major strengths of our study include access to both clinical and demographic information for the mothers and a large and diverse sample. We were able to adjust for many established risk factors of preterm birth, including preexisting conditions. We had residential addresses recorded, so we were able to calculate distance to major road, major road density, and major road density with truck volume for each individual. We were able to compare multiple measures of major road proximity and air pollution in the same large cohort of births.

2.5.3 Limitations of the data

Despite these strengths, some study limitations are important to consider. A different definition of major roads may have shown an association between proximity and preterm birth. While our large sample size was advantageous, we were not able to consider other potential sources of background pollution. Waste sites, refineries,[30] or agriculture operations[31] could alter pollution exposure independently of roadways. Meteorological patterns also contribute to variation in background pollution. Addresses used in geocoding were recorded at the time of delivery; there is no information on moves during pregnancy. However, one cohort study in Texas found that, while between one quarter and one third of mothers moved during pregnancy, the moves were short-distance and had minimal impact on pollution exposure.[32] We performed sensitivity analyses to verify that data clustering did not materially change our results, however, our use of single-level models remains a limitation.

Air pollution has been shown to have both chronic and acute effects on preterm birth.[33],[34] Further work aiming to clarify the relationship between major road proximity and preterm birth would benefit from additional large-scale studies which can use residential address

history in combination with measures of air and background pollution on a finer temporal scale.

2.5.4 Interpretation

Our finding of an association between $PM_{2.5}$, diesel PM, and preterm birth is consistent with the findings of several other investigative teams in[35, 36] and outside California.[37, 38] Our results further expand on these findings, given the size of our sample and our ability to test multiple risk factors. This suggests that these measures of air pollution are robust risk factors in a number of geographies.

The lack of association between road distance (as a binary exposure) and preterm birth echoes the results of some previous studies. In Rotterdam, researchers tested multiple distance cutoffs and did not find a significant association between road proximity and preterm birth.[14] A study in Vancouver showed no association between road proximity and preterm birth but did see an association between $PM_{2.5}$ and preterm birth.[13] The summary measure of census tract-level $PM_{2.5}$ served as a better metric than distance to major road for evaluating air pollution as a risk factor for preterm birth.

Other studies have found an association between road distance and preterm birth. Generally, these studies have considered the question of road exposure on a much smaller geographical scale. Studies focused on a single city or highway found larger associations, with aRRs of 1.30 (95% CI 1.03, 1.65)[9], 1.14 (95% CI 1.02, 1.27)[39], and 1.5 (95% CI 1.3, 1.9).[9] The largest geographical region represented in the previous work was North Carolina and this study found a smaller association than those listed above, at an aRR of 1.04 (95% CI 1.01, 1.07).[12] The present study in California covers an area roughly three times larger. As area increases, factors such as background pollution and access to care may play larger roles. Also of note is that this previous largest study included fewer covariates known be associated with preterm birth, such as maternal BMI and individual socioeconomic status. This suggests

that previous studies focusing on major road proximity and preterm birth perhaps found this association because they covered a smaller area and/or were less able to account for other known risk factors for preterm birth.

Exposures that measured traffic density did not show an association with preterm birth in the present study. This held true for the census-level exposure (traffic) and the exposures calculated at the individual-level (major road density and major road density with truck volume). The strong correlation between major road density with and without truck volume suggests that the volume measurements did not significantly alter the density score. Traffic is positively correlated with diesel PM and $PM_{2.5}$, but major road density and major road density with volume are not. The lack of association between road and traffic density measures and preterm birth echo the findings of van den Hooven, et. al.[14]

One interesting finding in the present study was that the association between distance to major roads and preterm birth shifted from no association in the unadjusted model (**Figure 2.5**, Panel A) to an association in the first adjusted model (**Figure 2.5**, Panel B), wherein further distance from major roads was found to be associated with preterm birth – a pattern that was opposite to what we had hypothesized. While this pattern seems related to race/ethnicity, it is essential that follow-up investigations more fully explore this possibility.

While it is not completely clear what may be driving the observed associations between $PM_{2.5}$, diesel PM, and preterm birth, one potential mechanism is oxidative stress,[40] which leads to an inflammatory response that can have adverse effects in multiple organ systems.[41, 42] An imbalance between oxidants and antioxidants may be associated with preterm birth.[43] Multiple studies have established a link between exposure to traffic-related particulate matter, such as black carbon and nitrous oxides, and elevated biomarkers of oxidative stress and changes in DNA methylation.[44, 45] There is also increasing concern regarding the effects of ultrafine particles, which are present in traffic-related air pollution,[46, 47] and

have been shown to enter the bloodstream[48] and induce oxidative stress.[49] However, the understanding and measurements are limited, and more research is required.

Particulate air pollution has been linked to preterm birth through other pathways as well. A meta-analysis of air pollution and diabetes found that NO_2 and $PM_{2.5}$, among other pollutants, were associated with higher risk of diabetes,[50] which has been established as a risk factor for preterm birth.[51] Multiple studies have found that traffic noise and particulate air pollution are associated with hypertension,[52, 53] another risk factor for preterm birth.[54] We still see an association between our particulate exposures and preterm birth after adjusting for these clinical factors. This may indicate residual effects not due to hypertension or diabetes.

2.6 Conclusions

Our results show that over large geographical areas, increased particulate matter concentrations are associated with preterm birth. Our results suggest that living closer to major roads or roads with heavy traffic does not serve as a proxy for particulate matter exposure as it relates to the risk of preterm birth. These findings have important implications for efforts focused on preventing preterm birth, including policies aimed at reducing air pollution.

Bibliography

[1] J.A. Martin, B.E. Hamilton, M.J.K. Osterman, and A.K. Driscoll. Births: Final data for 2018. *National Vital Statistics Reports*, 68(13).

[2] Dag Moster, Rolv Terje Lie, and Trond Markestad. Long-Term Medical and Social Consequences of Preterm Birth. *New England Journal of Medicine*, 359(3):262–273, 7 2008. ISSN 0028-4793. doi: 10.1056/NEJMoa0706475. URL `http://www.nejm.org/doi/abs/10.1056/NEJMoa0706475`.

[3] Petra Klepac, Igor Locatelli, Sara Korošec, Nino Künzli, and Andreja Kukec. Ambient air pollution and pregnancy outcomes: A comprehensive review and identification of environmental public health challenges. *Environmental research*, 167:144–159, 2018.

[4] Amy M Padula, Elizabeth M Noth, S Katharine Hammond, Fred W Lurmann, Wei Yang, Ira B Tager, and Gary M Shaw. Exposure to airborne polycyclic aromatic hydrocarbons during pregnancy and risk of preterm birth. *Environmental research*, 135: 221–226, 2014.

[5] David JX Gonzalez, Allison R Sherris, Wei Yang, David K Stevenson, Amy M Padula, Michael Baiocchi, Marshall Burke, Mark R Cullen, and Gary M Shaw. Oil and gas production and spontaneous preterm birth in the san joaquin valley, ca: a case–control study. *Environmental Epidemiology*, 4(4), 2020.

[6] Joan A Casey, Deborah Karasek, Elizabeth L Ogburn, Dana E Goin, Kristina Dang, Paula A Braveman, and Rachel Morello-Frosch. Retirements of coal and oil power plants in california: association with reduced preterm birth among populations nearby. *American journal of epidemiology*, 187(8):1586–1594, 2018.

[7] Michelle Wilhelm, Jo Kay Ghosh, Jason Su, Myles Cockburn, Michael Jerrett, and Beate Ritz. Traffic-related air toxics and preterm birth: a population-based case-control study in los angeles county, california. *Environmental Health*, 10(1):89, 2011.

[8] Kathie L Dionisio, Lisa K Baxter, Janet Burke, and Halûk Özkaynak. The importance of the exposure metric in air pollution epidemiology studies: When does it matter, and why? *Air Quality, Atmosphere & Health*, 9(5):495–502, 2016.

[9] Chun-Yuh Yang, Chih-Ching Chang, Hung-Yi Chuang, Chi-Kung Ho, Trong-Neng Wu, and Shang-Shyue Tsai. Evidence for Increased Risks of Preterm Delivery in a Population Residing near a Freeway in Taiwan. *Archives of Environmental Health*, 58(10):649–654, 2003. ISSN 0003-9896. doi: 10.3200/AEOH.58.10.649-654. URL `https://www.tandfonline.com/action/journalInformation?journalCode=vaeh20`.

[10] Takashi Yorifuji, Hiroo Naruse, Saori Kashima, Shigeru Ohki, Takeshi Murakoshi, Soshi Takao, Toshihide Tsuda, and Hiroyuki Doi. Residential proximity to major roads and preterm births. *Epidemiology*, 22(1):74–80, 1 2011. ISSN 10443983. doi: 10.1097/EDE.0b013e3181fe759f. URL `https://pubmed.ncbi.nlm.nih.gov/21052006/https://pubmed.ncbi.nlm.nih.gov/21052006/?dopt=Abstract`.

[11] Takashi Yorifuji, Hiroo Naruse, Saori Kashima, Soshi Takao, Takeshi Murakoshi, Hiroyuki Doi, and Ichiro Kawachi. Residential proximity to major roads and adverse birth outcomes: a hospital-based study. *Environmental health : a global access science source*, 12(1):34, 4 2013. ISSN 1476-069X. doi: 10.1186/1476-069X-12-34. URL `http://www.ncbi.nlm.nih.gov/pubmed/23597260http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3639132`.

[12] Marie Lynn Miranda, Sharon E. Edwards, Howard H. Chang, and Richard L. Auten. Proximity to roadways and pregnancy outcomes. *Journal of Exposure Science and Environmental Epidemiology*, 23(1):32–38, 1 2013. ISSN 15590631. doi: 10.1038/jes.2012.78. URL `https://www.nature.com/articles/jes201278`.

[13] Michael Brauer, Cornel Lencar, Lilian Tamburic, Mieke Koehoorn, Paul Demers, and Catherine Karr. A cohort study of traffic-related air pollution impacts on birth out-

comes. *Environmental Health Perspectives*, 116(5):680–686, 5 2008. ISSN 00916765. doi: 10.1289/ehp.10952.

[14] Edith H. Van Den Hooven, Vincent Wv Jaddoe, Yvonne De Kluizenaar, Albert Hofman, Johan P. MacKenbach, Eric Ap Steegers, Henk Me Miedema, and Frank H. Pierik. Residential traffic exposure and pregnancy-related outcomes: A prospective birth cohort study. *Environmental Health: A Global Access Science Source*, 8(1), 2009. ISSN 1476069X. doi: 10.1186/1476-069X-8-59.

[15] Esri Inc. Arcgis desktop 10.7.1, 2019. URL `http://www.arcgis.com`.

[16] U.S. Census Bureau. Tiger/lineshapefiles, 2019.

[17] California State Department of Transportation. Shn_lines. URL `https://gisdata.dot.ca.gov/arcgis/rest/services/Highway/SHN_Lines/MapServer`.

[18] California State Department of Transportation. Truck_volumes_aadt. URL `https://gisdata.dot.ca.gov/arcgis/rest/services/Highway/Truck_Volumes_AADT/MapServer`.

[19] J. Faust, L. August, K. Bangia, V. Galaviz, J. Leichty, S. Prasad, R. Schmitz, A. Slocombe, R. Welling, W. Wieland, and L. Zeise. Update to the california communities environmental health screening tool, January 2017. URL `https://oehha.ca.gov/media/downloads/calenviroscreen/report/ces3report.pdf`.

[20] Robert A Eldred, Thomas A Cahill, and Robert G Flocchini. Composition of pm2.5 and pm10 aerosols in the improve network. *Journal of the Air & Waste Management Association*, 47(2):194–203, 1997.

[21] OEHHA. Calenviroscreen 3.0, 6 2018. URL `https://oehha.ca.gov/calenviroscreen/report/calenviroscreen-30`.

[22] Rebecca J Baer, Monica R McLemore, Nancy Adler, Scott P Oltman, Brittany D Chambers, Miriam Kuppermann, Matthew S Pantell, Elizabeth E Rogers, Kelli K Ryckman, Marina Sirota, et al. Pre-pregnancy or first-trimester risk scoring to identify women at high risk of preterm birth. *European Journal of Obstetrics & Gynecology and Reproductive Biology*, 231:235–240, 2018.

[23] Jennifer A Hutcheon, Susan Moskosky, Cande V Ananth, Olga Basso, Peter A Briss, Cynthia D Ferré, Brittni N Frederiksen, Sam Harper, Sonia Hernández-Díaz, Ashley H Hirai, et al. Good practices for the design, analysis, and interpretation of observational studies on birth spacing and perinatal health outcomes. *Paediatric and Perinatal Epidemiology*, 33(1):O15–O24, 2019.

[24] Milton Kotelchuck. An evaluation of the kessner adequacy of prenatal care index and a proposed adequacy of prenatal care utilization index. *American journal of public health*, 84(9):1414–1420, 1994.

[25] ERS. Rural-urban commuting area codes, 2019. URL `https://www.ers.usda.gov/data-products/rural-urban-commuting-area-codes.aspx#.U9lO7GPDWHo`.

[26] WWAMI Rural Health Research Center. Ruca data: Using ruca data. URL `https://depts.washington.edu/uwruca/ruca-uses.php`.

[27] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2021. URL `https://www.R-project.org/`.

[28] Stef van Buuren and Karin Groothuis-Oudshoorn. mice: Multivariate imputation by chained equations in r. *Journal of Statistical Software*, 45(3):1–67, 2011. URL `https://www.jstatsoft.org/v45/i03/`.

[29] Tyler J VanderWeele and Peng Ding. Sensitivity analysis in observational research: introducing the e-value. *Annals of internal medicine*, 167(4):268–274, 2017.

[30] Ewa Dabek-Zlotorzynska, Tom F Dann, P Kalyani Martinelango, Valbona Celo, Jeffrey R Brook, David Mathieu, Luyi Ding, and Claire C Austin. Canadian national air pollution surveillance (naps) pm2.5 speciation program: Methodology and pm2.5 chemical composition for the years 2003–2008. *Atmospheric Environment*, 45(3):673–686, 2011.

[31] L-W Antony Chen, John G Watson, Judith C Chow, and Karen L Magliano. Quantifying pm2.5 source contributions for the san joaquin valley with multivariate receptor models. *Environmental science & technology*, 41(8):2818–2826, 2007.

[32] Philip J Lupo, Elaine Symanski, Wenyaw Chan, Laura E Mitchell, D Kim Waller, Mark A Canfield, and Peter H Langlois. Differences in exposure assignment between conception and delivery: the impact of maternal mobility. *Paediatric and perinatal epidemiology*, 24(2):200–208, 2010.

[33] Shanshan Li, Yuming Guo, and Gail Williams. Acute impact of hourly ambient air pollution on preterm birth. *Environmental health perspectives*, 124(10):1623–1629, 2016.

[34] Shiliang Liu, Daniel Krewski, Yuanli Shi, Yue Chen, and Richard T Burnett. Association between gaseous ambient air pollutants and adverse pregnancy outcomes in vancouver, canada. *Environmental health perspectives*, 111(14):1773–1778, 2003.

[35] Amy M Padula, Hongtai Huang, Rebecca J Baer, Laura M August, Marta M Jankowska, Laura L Jellife-Pawlowski, Marina Sirota, and Tracey J Woodruff. Environmental pollution and social factors as contributors to preterm birth in fresno county. *Environmental Health*, 17(1):70, 2018.

[36] Hongtai Huang, Tracey J Woodruff, Rebecca J Baer, Komal Bangia, Laura M August, Laura L Jellife-Palowski, Amy M Padula, and Marina Sirota. Investigation of association between environmental and socioeconomic factors and preterm birth in california. *Environment international*, 121:1066–1078, 2018.

[37] Marta Ottone, Serena Broccoli, Federica Parmagnani, Simone Giannini, Fabiana Scotto, Laura Bonvicini, Ferdinando Luberto, Dimitri Bacco, Arianna Trentini, Vanes Poluzzi, et al. Source-related components of fine particulate matter and risk of adverse birth outcomes in northern italy. *Environmental Research*, 186:109564, 2020.

[38] Emily DeFranco, William Moravec, Fan Xu, Eric Hall, Monir Hossain, Erin N Haynes, Louis Muglia, and Aimin Chen. Exposure to airborne particulate matter during pregnancy is associated with preterm birth: a population-based cohort study. *Environmental Health*, 15(1):1–8, 2016.

[39] M Généreux, N Auger, M Goneau, and M Daniel. Neighbourhood socioeconomic status, maternal education and adverse birth outcomes among mothers living near highways. *Journal of Epidemiology & Community Health*, 62(8):695–700, 2008. ISSN 0143-005X. doi: 10.1136/jech.2007.066167. URL `https://jech.bmj.com/content/62/8/695`.

[40] Graham J Burton and Eric Jauniaux. Oxidative stress. *Best practice & research Clinical obstetrics & gynaecology*, 25(3):287–299, 2011.

[41] Mark R Miller. Oxidative stress and the cardiovascular effects of air pollution. *Free Radical Biology and Medicine*, 2020.

[42] Paolo Giorgini, Paolo Di Giosia, Davide Grassi, Melvyn Rubenfire, Robert D Brook, and Claudio Ferri. Air pollution exposure and blood pressure: an updated review of the literature. *Current pharmaceutical design*, 22(1):28–51, 2016.

[43] Tiffany A Moore, Iman M Ahmad, and Matthew C Zimmerman. Oxidative stress and preterm birth: an integrative review. *Biological research for nursing*, 20(5):497–512, 2018.

[44] Ralph J Delfino, Norbert Staimer, Thomas Tjoa, Daniel L Gillen, Andrea Polidori, Mohammad Arhami, Micheal T Kleinman, Nosratola D Vaziri, John Longhurst, and

Constantinos Sioutas. Air pollution exposures and circulating biomarkers of effect in a susceptible population: clues to potential causal component mixtures and mechanisms. *Environmental health perspectives*, 117(8):1232–1238, 2009.

[45] Andrea Baccarelli, Robert O Wright, Valentina Bollati, Letizia Tarantini, Augusto A Litonjua, Helen H Suh, Antonella Zanobetti, David Sparrow, Pantel S Vokonas, and Joel Schwartz. Rapid dna methylation changes after exposure to traffic particles. *American journal of respiratory and critical care medicine*, 179(7):572–578, 2009.

[46] Leonidas Ntziachristos, Zhi Ning, Michael D Geller, and Constantinos Sioutas. Particle concentration and characteristics near a major freeway with heavy-duty diesel traffic. *Environmental science & technology*, 41(7):2223–2230, 2007.

[47] Barouch Giechaskiel, Leonidas Ntziachristos, Zissis Samaras, Volker Scheer, Roberto Casati, and Rainer Vogt. Formation potential of vehicle exhaust nucleation mode particles on-road and in the laboratory. *Atmospheric Environment*, 39(18):3191–3198, 2005.

[48] Mark R Miller, Jennifer B Raftis, Jeremy P Langrish, Steven G McLean, Pawitrabhorn Samutrtai, Shea P Connell, Simon Wilson, Alex T Vesey, Paul HB Fokkens, A John F Boere, et al. Inhaled nanoparticles accumulate at sites of vascular disease. *ACS nano*, 11(5):4542–4552, 2017.

[49] Ning Li, Constantinos Sioutas, Arthur Cho, Debra Schmitz, Chandan Misra, Joan Sempf, Meiying Wang, Terry Oberley, John Froines, and Andre Nel. Ultrafine particulate pollutants induce oxidative stress and mitochondrial damage. *Environmental health perspectives*, 111(4):455–460, 2003.

[50] Bo-Yi Yang, Shujun Fan, Elisabeth Thiering, Jochen Seissler, Dennis Nowak, Guang-Hui Dong, and Joachim Heinrich. Ambient air pollution and diabetes: a systematic review and meta-analysis. *Environmental Research*, 180:108817, 2020.

[51] Hayfaa A Wahabi, Samia A Esmaeil, Amel Fayed, Ghadeer Al-Shaikh, and Rasmieh A Alzeidan. Pre-existing diabetes mellitus and adverse pregnancy outcomes. *BMC Research Notes*, 5(1):496, 2012.

[52] Kateryna B Fuks, Gudrun Weinmayr, Xavier Basagaña, Olena Gruzieva, Regina Hampel, Bente Oftedal, Mette Sørensen, Kathrin Wolf, Geir Aamodt, Gunn Marit Aasvang, et al. Long-term exposure to ambient air pollution and traffic noise and incident hypertension in seven cohorts of the european study of cohorts for air pollution effects (escape). *European heart journal*, 38(13):983–990, 2017.

[53] Wolfgang Babisch, Kathrin Wolf, Markus Petz, Joachim Heinrich, Josef Cyrys, and Annette Peters. Associations between traffic noise, particulate air pollution, hypertension, and isolated systolic hypertension in adults: the kora study. *Environmental health perspectives*, 122(5):492–498, 2014.

[54] Kate Bramham, Bethany Parnell, Catherine Nelson-Piercy, Paul T Seed, Lucilla Poston, and Lucy C Chappell. Chronic hypertension and pregnancy outcomes: systematic review and meta-analysis. *Bmj*, 348:g2301, 2014.

Chapter 3

Exposure to wildfire smoke

3.1 Abstract

Background: Long term studies have shown an association between wildfire smoke exposure and preterm birth. However, there have been few studies examining the association between an isolated wildfire event and subsequent preterm birth in a small region. Further, past studies have not typically adjusted for chronic clinical conditions associated with preterm birth.

Objectives: In a retrospective cohort study, we examine the association between wildfire-driven high concentrations of fine particulate matter during the 2018 California Camp Fire and preterm birth in the San Francisco Bay Area. This fire was the first to affect the Bay Area in over a decade. We evaluate the risk of preterm birth based on smoke exposure and examine whether the level of fine particulate matter (particulate matter less than 2.5µm in diameter) was associated with preterm birth.

Methods: We obtained air quality data from Environmental Protection Agency sensors and consumer-purchased pollution sensors from PurpleAir. We interpolated daily average

fine particulate matter surfaces and assigned a mean pregnancy fine particulate matter value to each study subject using residential address. Using pregnancies during the year prior as a control, we used log-linear models to determine whether experiencing the fire was significantly associated with preterm birth. We tested whether the severity of exposure during the fire was associated with preterm birth. We adjusted for demographic and clinical conditions.

Results: Experiencing the wildfire was associated with preterm birth, with an adjusted relative risk aRR) of 1.10 (95% confidence interval 1.03, 1.17). The level of fine particulate matter exposure during the fire was associated with preterm birth: exposure in the highest tertile compared to the lowest tertile resulted in an adjusted relative risk of 1.17 (95% confidence interval 1.05, 1.29).

Conclusion: Experiencing the 2018 California Camp Fire was associated with increased mean fine particulate matter exposure during pregnancy and with preterm birth among pregnant people in the San Francisco Bay Area. Those with higher fine particulate matter exposure during the fire had higher risk, even after adjusting for chronic conditions. This provides stronger evidence for a connection between wildfire smoke exposure and preterm birth.

Keywords: EPA, fine particulate matter, particulate matter $PM_{2.5}$, PurpleAir, wildfire smoke

3.2 Introduction

Preterm birth (PTB), defined as birth prior to 37 weeks of gestation, is the leading cause of neonatal morbidity and mortality and is associated with long-term health consequences.[1, 2] There are several known risk factors for PTB, including preexisting clinical conditions, sociodemographic characteristics, and environmental exposures.[3, 4] Previous research has identified high concentrations of fine particulate matter ($PM_{2.5}$) as a risk factor for adverse pregnancy outcomes, including PTB and small for gestational age.[5] As climate change expands the duration, location, timing, and intensity of wildfires, exposure to $PM_{2.5}$ will increase.[6] Prior work examining the association between exposure to wildfire smoke during pregnancy and adverse outcomes has tended to focus on birth weight.[7] Those studies that have examined the association between either $PM_{2.5}$ or wildfire smoke and PTB have tended to investigate patterns over large geographic areas with multiple smoke events over a long time frame.[8, 9] In this study, we consider a recent smoke exposure event from one wildfire affecting an area without prior exposure to wildfire smoke and test for any association with PTB.

The Camp Fire occurred in Butte County, California in November of 2018.[10] It began November 8th, burned for 12 days through November 21st when it started raining, and was declared officially contained on the 25th of November. This wildfire is of particular interest for several reasons: it was caused by electric utility infrastructure in a remote mountainous wildland area, but because it began during a high wind event which typically only occur in the fall season, it quickly became a wildland urban interface fire and burned down most of the town of Paradise, California (nearly 19,000 structures) within several hours.[11, 12]

This investigation evaluates the impact of the smoke from the Camp Fire on the SF Bay Area, located about 200 miles southwest of Paradise, California. Severe wind pushed the smoke plume directly to the SF Bay Area within a few hours of the fire starting, inundating this population with wildfire smoke (see **Figure 3.1**) for several days of extreme pollution

for the first time in many years. The typical coastal weather patterns of the SF Bay Area generally keep the air relatively low in $PM_{2.5}$[13]. Investigation from EPA and satellite smoke data indicate that although California has experienced many large wildfires in the last two decades, the most recent fire to have inundated the SF Bay Area with smoke (prior to the Camp Fire) was in 1991. This means that for many long-term residents of the SF Bay Area, there are no lingering physiological effects from previous major smoke exposures. The Camp Fire provides an unusual opportunity to investigate how one severe wildfire and weather event may affect a normally sheltered population.

Choosing to focus on the SF Bay Area presents additional advantages. First, the SF Bay Area is a diverse and densely populated region. While areas nearer to the fire were exposed to higher concentrations of fine particulate matter, these areas are rural; using the directly affected population would mean a smaller and less diverse sample size. Second, while the SF Bay Area experienced heavy smoke, it was not directly at risk from the fire, and consequently was not under any evacuation orders. Additionally, the smoke exposure duration was so long that temporary self-evacuation became impractical, so we expect that most residents experienced high particulate matter. Third, the topography of the SF Bay Area is highly varied, so even within this small region we expect that smoke exposure differed significantly by location. We also chose the SF Bay Area because of its early adoption of the crowdsourced laser pollution sensors made by PurpleAir.[14] We used EPA and PurpleAir sensors (with careful data cleaning) to interpolate a pollution surface across the entire SF Bay Area.

There are previous studies examining the association between exposure to wildfire smoke and adverse pregnancy outcomes. While most of these have focused on birth weight, there have been few considering PTB.[15] In a recent study, Heft-Neal and colleagues found a statistically significant association between wildfire smoke and PTB, when considering California births from 2006 to 2011[8]. Similarly, a study from Abdo and colleagues found that, when examining pregnancies in Colorado from 2007 to 2015, exposure to wildfire smoke

in the second trimester of pregnancy was positively associated with PTB.[9] Both of these papers considered long time frames rather than the impact of a single acute event. One paper considering the association between a severe fire event in Victoria, Australia, and PTB considered only the stress caused by the fire and did not address pollution.[16]

Here we address two main questions in a large, diverse population, with individual clinical information available. First, we assessed whether there was any association between experiencing the 2018 Camp Fire and PTB among SF Bay Area residents. Second, we examined whether experienceing higher concentrations of $PM_{2.5}$ during the fire (compared to lower $PM_{2.5}$ concentrations during the fire) was associated with PTB among SF Bay Area residents. We hypothesized that there would be a positive association between fire exposure and PTB and we hypothesized that worse air would be associated with an increased risk of PTB.

3.3 Materials and Methods

Methods and protocols for the study were approved by the Committee for the Protection of Human Subjects (CPHS) which serves as the institutional review board (IRB) for the California Health and Human Services Agency (CHHSA).

3.3.1 Birth data

Our outcome of interest was PTB, defined as a live delivery prior to 37 weeks of gestation. Our study population consisted of singleton births in the following six counties: Alameda, Contra Costa, Marin, San Francisco, San Mateo, Santa Clara. Birth records from California Vital Statistics were linked to hospital records from the California Office of Statewide Planning and Development. These data record diagnoses and other clinical information from the birth. The records also contain residential address at the time of birth. We required that addresses were valid and could be successfully geocoded to obtain a latitude and lon-

gitude. Geocoding was performed in ArcGIS (version 10.7.1.11595).[17] All other geospatial computation was performed in R (version 4.0.4).[18–20] Additionally, we required that the gestational age (in weeks) of the infant must be known, so that we could use this information to estimate a conception date. We obtained best obstetric estimate of gestation at birth from birth certificates.

Our control group consisted of those pregnancies which overlapped with the same calendar time a year before the wildfires, that is, those who were pregnant during Nov 09, 2017 to Nov 21, 2017. We chose this group to control for seasonality. Since we are not considering the possible triggering effect smoke exposure may have on preterm delivery, we discarded those whose pregnancies only partially overlapped the fire.

3.3.2 $PM_{2.5}$ Data

In California, two natural choices for air quality data are EPA sensors, which are spatially sparse, and CalEnviroScreen measures, which do not account for acute exposures.[21, 22] Another source is modelled pollution surfaces; while these can record acute exposures, there is typically a lag period between a given year and the release of modelled surfaces. [23, 24] The modeled data for the Camp Fire does not exist at the time of publication of this research. Lastly, smoke plume coverage maps derived from satellite imagery are available, in particular from the GOES, MODIS and VIIRS satellites; however, these do not measure pollution at the ground level, which changes over the life of of the wildfire and and wind patterns. As seen from **Figure 3.1**), smoke covered the Bay Area for the life of the burn period. The objective of this paper is to use high resolution measured pollution values, which for the first time exist for a study such as this.

In just the last few years, PurpleAir, a commodity air pollution sensor company, has sold thousands of personal pollution sensors in the SF Bay Area. In November 2018, at the time of the Camp Fire, there were approximately 250 PurpleAir sensors in our study area. These can

complement the 12 EPA sensors and provide an opportunity to measure pollution exposure in greater spatial resolution than before. However, due to the nature of these crowd-sourced measurements, the quality of the measures can be variable. The EPA performed a study to establish a method for adjusting for Purple Air data when compared to the well-calibrated EPA sensors.[25]

PurpleAir sensors were identified using the PurpleAir API python package developed by ReagentX.[26] We generated a list of outdoor sensors falling within the following coordinates: north = 38.5, south = 36.55, east = -121.3, west = -122.93, covering the extent of the metropolitan SF Bay Area. Daily average measures were downloaded from each outside sensor's parent channel going back to 2017. We first applied the EPA correction. We required that each sensor have at least one measurement during the fire period or at least 50 during the study period. PurpleAir sensors sometimes reported unrealistically high $PM_{2.5}$ concentrations, even after the EPA correction. In order to remove these, we excluded PurpleAir measurements that were more than four standard deviations above the average of the SF Bay Area EPA sensors for that day. Similarly, PurpleAir sensors sometimes record extremely low $PM_{2.5}$ concentrations during severe smoke events. One possible explanation for these low measures is that during severe smoke events, people may choose to move their monitors to evaluate the quality of the air inside their homes. We applied the same process to exclude low values wherein we removed values that fell below four standard deviations of the EPA average for the day.

3.3.3 $PM_{2.5}$ Interpolation

Air pollution in the context of wildfire behaves differently than air pollution generated on the ground from local sources. This fire was driven by wind gusts exceeding 35 miles per hour at the site of its ignition.[27] Wind pushed smoke from Paradise, California to the SF Bay Area within a few hours (**Figure 3.1**). As the wind patterns changed, the smoke

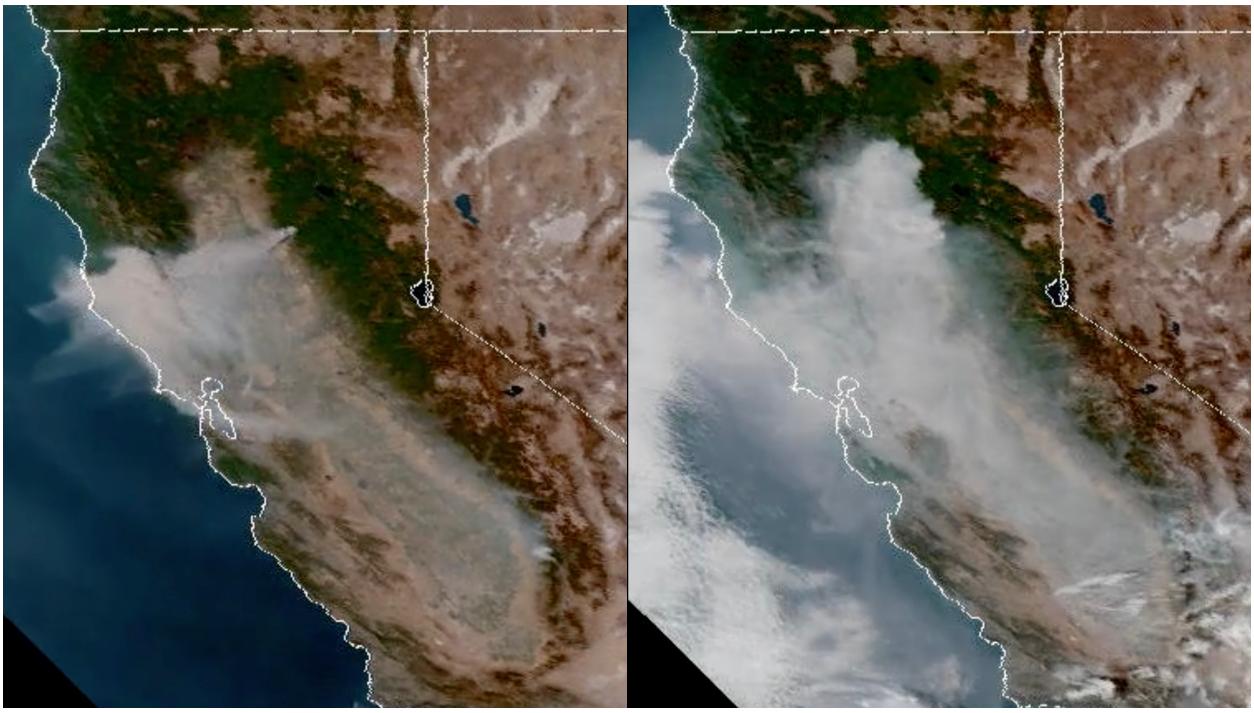

Figure 3.1: Satellite imagery of the smoke plumes on November 8, 2018 (left) and November 15, 2018 (right).

sank when wind velocities decreased. Looking at figures from the Geostationary Operational Environmental Satellite (GOES) satellite,[28] we see that much of California's land mass was covered by smoke from this fire. However, smoke plume coverage maps seen from satellite do not effectively tell us what concentrations of $PM_{2.5}$ are being measured on the ground at different stages of the fire and smoke evolution. We used kriging as a standard spatial interpolation method to calculate pollution values across the dense distribution of EPA and Purple Air sensors instead of a modeled fluid dynamics air pollution dataset. We can do this because the density of stations from Purple Air provides enough data to estimate exposures across the region.

We evaluated the following interpolation methods: splines, kriging and inverse distance weighting (IDW). We chose kriging because past experiments found this to be the most accurate method [29, 30]. It adheres to the sensor value inputs, and it can derive values greater than the inputs. It has also been used in many other pollution studies. [30–33]

We performed kriging using a first-order polynomial: $PM_{2.5} \sim$ longitude + latitude. While our focus was the SF Bay Area, we performed kriging over a larger area (north = 40.17, south = 36.03, west = -123.35, east = -118.82) by adding in EPA sensors outside of our study area. This helped us verify our air surfaces by identifying other adverse air events outside the SF Bay Area.

We tested creating these continuous surfaces by regional watersheds to take into account the topographic constraints. In doing so, we observed that the pollution stations were distributed well enough to represent topographic variability already. Knowing the topographic and climatic patterns in the SF Bay Area at the time of the wildfire has been used qualitatively to validate the resulting pollution measures due to the spatial interpolation process.[34]

3.3.4 Exposure measures

We computed multiple measures of exposure. Our first measure was a binary indicator of wildfire overlap; we assigned “exposed" to those whose pregnancies overlapped the full fire period (12 days) and lived within the study extent. The study extent was covered in wildfire smoke as seen from satellite images see **Figure 3.1**, and ground pollution sensors measured increased smoke pollution throughout. “Unexposed" was assigned to all other pregnancies, those pregnant during the same time period and study area but one year earlier, in November 2017. To assign the amount of $PM_{2.5}$ exposures to each pregnancy, we extracted the $PM_{2.5}$ concentrations for each day of the pregnancy from the kriged surfaces at the individual's home latitude and longitude. We computed the mean $PM_{2.5}$ exposure over the whole pregnancy. For those whose pregnancy overlapped the fire, we also computed the mean $PM_{2.5}$ exposure and the maximum $PM_{2.5}$ during the fire.

3.3.5 Statistical analysis

We first conducted a Welch two-sample T-test to determine whether the mean $PM_{2.5}$ exposure was the same for those who had experienced the fire and those who had not. To determine whether experiencing the fire was a risk factor for PTB, we considered multiple scenarios. We first computed a simple contingency table with the fire exposure as a binary variable. We computed separate log-linear models for our two exposures of interest: fire as a binary, and mean $PM_{2.5}$ over the course of pregnancy. We adjusted these models in stages; to see whether clinical covariates affected our results, we first added demographic covariates only, then included clinical covariates. Our demographic covariates were: maternal education (>12 years or <12 years compared to 12 years), maternal age (<18 years or >34 years compared to 18-34 years), interpregnancy interval (very short (<6 months), short (6-23 months), or long (>59 months) compared to normal (24-59 months)), insurance status (public compared to non-public), race/ethnicity (Black, Asian, Hispanic, or other compared to white non-Hispanic), and adequacy of prenatal care: inadequate (care begun after 4th month or <50% of recommended visits received) or intermediate (care begun by 4th month and 50%-79% of recommended visits received) compared to adequate/adequate plus (care begun by 4th month and ≥80% of recommended visits received).[35] Our clinical covariates were: the presence of a mental disorder, preexisting diabetes, preexisting hypertension, nulliparous, drug abuse, alcohol abuse, smoking status, and asthma. We considered timing subtypes of PTB: Very Early PTB (<28 weeks), Early PTB (28 - 32 weeks), and Late PTB (36-37 weeks) as compared to term birth.

To verify our results, we grouped pregnancies into census tracts, then divided census tracts into tertiles based on mean $PM_{2.5}$ over two non-intersecting periods. We defined baseline $PM_{2.5}$ tertiles using a one-year period starting from 2017-03-20. We calculated fire tertiles using the fire period (2018-11-09 to 2018-11-21). We assessed whether there was a difference in PTB rates between high fire tertiles and high baseline tertiles prior in the control group.

For additional testing, we performed Welch two-sample T-tests on the mean income levels of the low baseline and low fire tertiles, as well as the high baseline and high fire tertiles.

We also tested for an association between $PM_{2.5}$ exposure during the fire period and PTB. For this analysis, we restricted our population to only birthing people who had experienced the full fire period. We examined both the peak $PM_{2.5}$ measure they experienced during the fire as well as the mean $PM_{2.5}$ during the fire period using log-linear models. As before, we adjusted this model in stages, first adding in the exposure and demographic covariates, then recomputing the model with clinical covariates added.

3.4 Results

Sample selection is shown in **Figure 3.2**. Population characteristics are displayed in **Table 3.1**. Our population wass largely Asian (30.3%), White (27.5%), or Hispanic (26.5%), with Black birthing people making up less than 5% of our study population. Roughly half of the mothers and birthing people in our study were born in the US. More than two-thirds of our population had completed more than 12 years of education. While MediCal generally covers approximately half the births in California,[36] only 20% of our population was covered by MediCal; the majority of our sample used private insurance. This is likely due to the fact that the SF Bay Area is comprised of the six wealthiest counties (by per-capita income) in California. **Figure 3.3** shows the timing of pregnancies in our control and exposed population.

Table 3.1: Sample overview.

N	**68,006**
Race/ethnicity	
White	20064 (29.5)
Hispanic	17826 (26.2)
Black	3084 (4.5)
Asian	19757 (29.1)
Other race	7275 (10.7)

N	**68,006**
IPI	
Normal IPI	47658 (70.1)
Very short IPI	773 (1.1)
Short IPI	11865 (17.4)
Long IPI	7710 (11.3)
Education	
12 years	14946 (22.0)
<12 years	4949 (7.3)
>12 years	48111 (70.7)
Insurance	
Other pay	52854 (77.7)
MediCal	14001 (20.6)
Self pay	1151 (1.7)
Prenatal care	
Adequate plus	13480 (19.8)
Adequate	31780 (46.7)
Intermediate	17343 (25.5)
Inadequate	4640 (6.8)
Maternal age	
18 to 34	45978 (67.6)
>34	21724 (31.9)
<18	304 (0.4)
Mental disorder	7532 (11.1)
Preexisting diabetes	1509 (2.2)
Preexisting hypertension	1368 (2.0)
Nulliparous	31864 (46.9)
Drug abuse	0.01 (0.12)
Alcohol abuse	0.00 (0.04)
Gestational diabetes	7771 (11.4)
Gestational hypertension	6354 (9.3)
Maternal asthma	4482 (6.6)
Maternal smoking	876 (1.3)

Weather data shows that the wind pushed smoke into the SF Bay Area for the first several days of the fire. Wind patterns changed and the worst air pollution days were November 15th and 16th. Pollution spiked in all of the valleys around the SF Bay Area during the peak pollution days. Pollution is lower in the hills and the coast. See **Figure 3.4** for the constructed air surface for November 16, 2018.

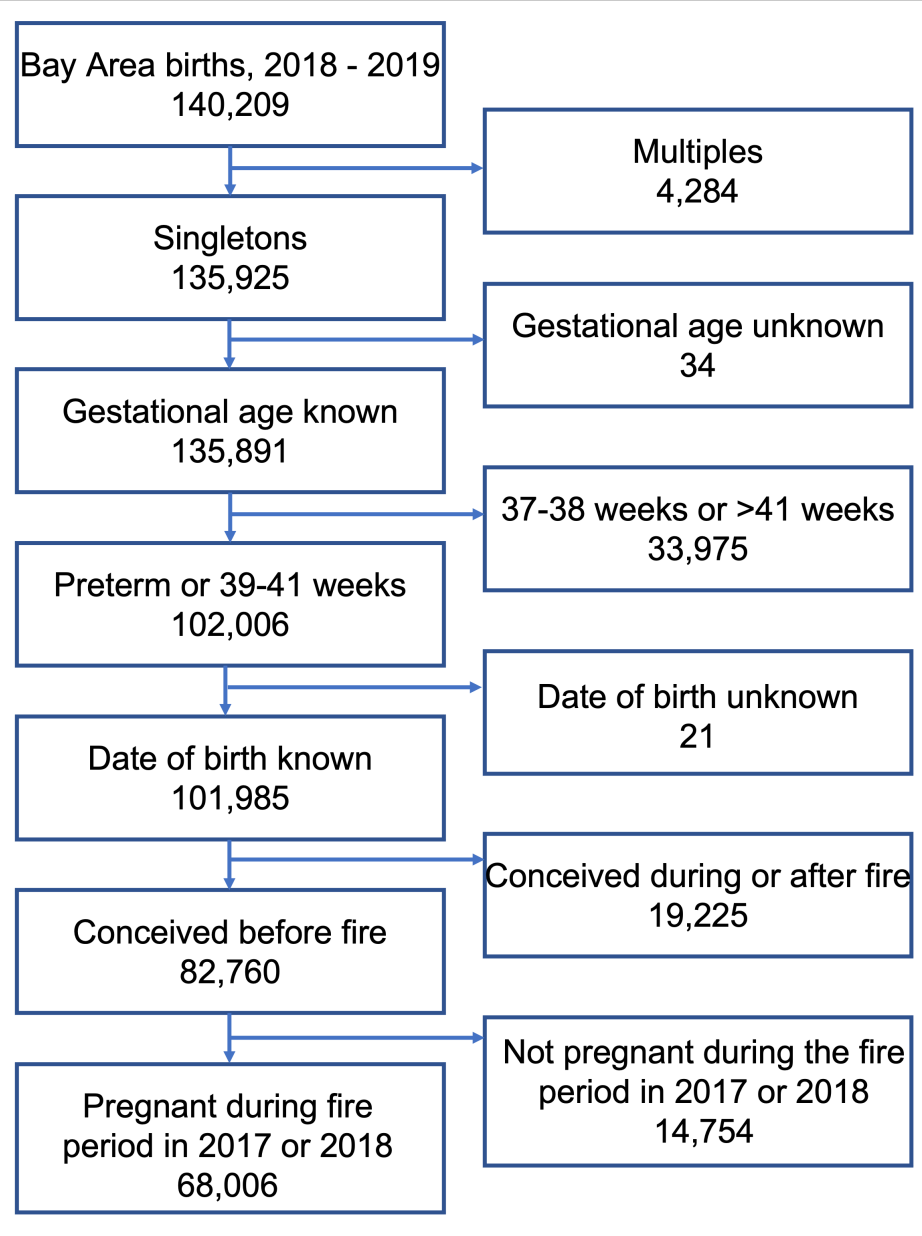

Figure 3.2: Sample selection

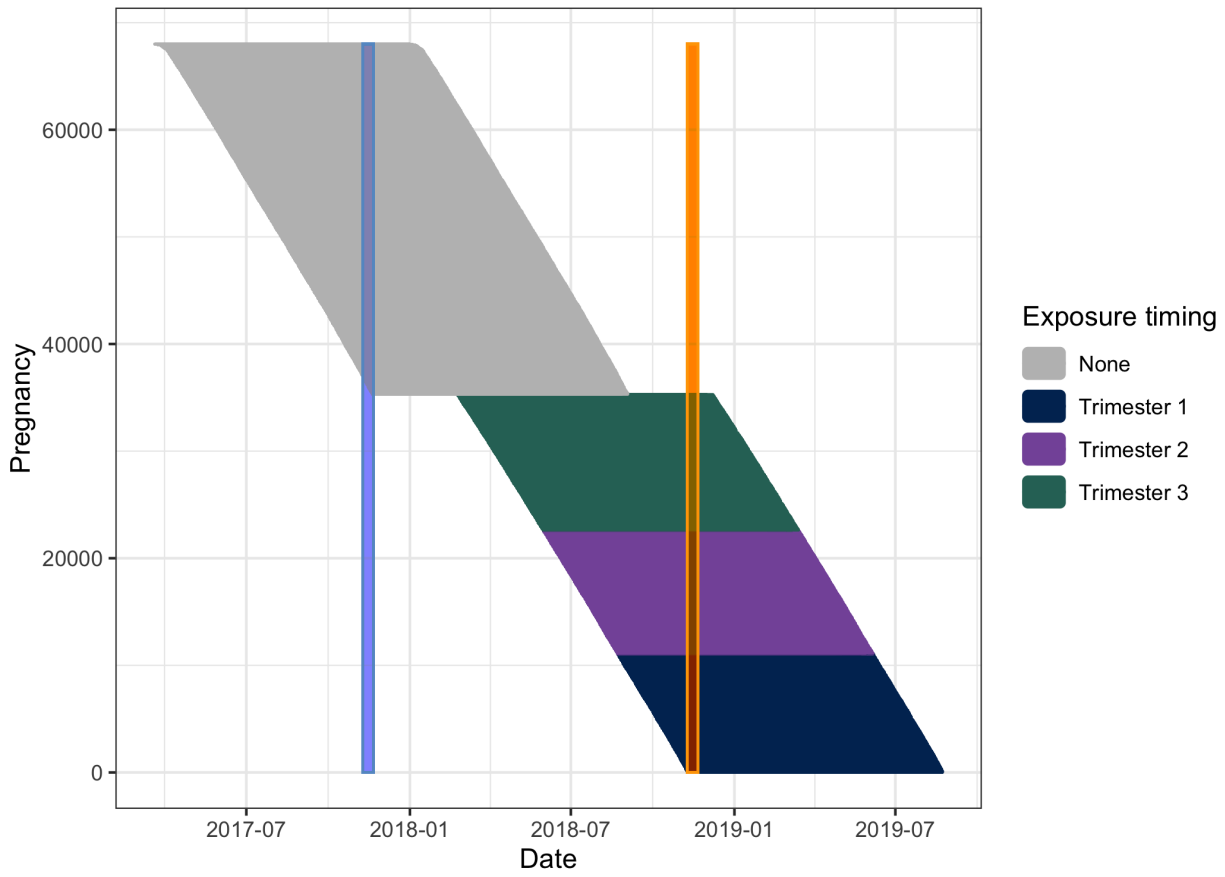

Figure 3.3: The plot is constructed of thin horizontal lines, each representing the duration of a pregnancy, from conception on the left to delivery on the right. Pregnancies are vertically ordered by start date. The orange rectangle represents the fire period (Nov 9 – 21, 2018). The blue rectangle represents a year prior to the fire period (Nov 9 – 21, 2017). Those who were exposed to the fire are colored by the trimester of exposure.

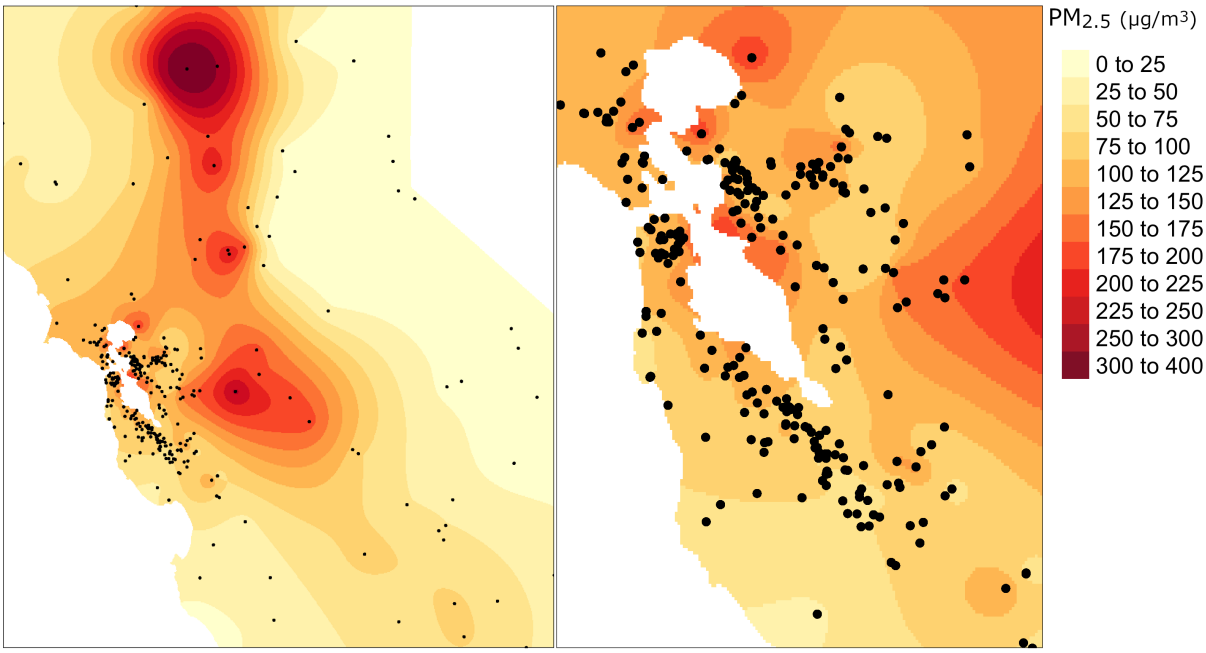

Figure 3.4: Interpolated $PM_{2.5}$ concentrations using kriging on corrected PurpleAir sensors and EPA data.

There is a dramatic spike in the ownership of PurpleAir sensors shortly before and during the wildfire. Prior to November 1, 2018, there were 128 eligible sensors; our total number of eligible sensors (including all those added before and after November 1, 2018) was 234. Approximately 8% of measurements during the fire period were removed for falling outside of our allowable range (the mean of the EPA sensors $\pm 4\sigma$).

Pregnancy mean $PM_{2.5}$ was 2.5% higher for those whose pregnancies overlapped the fire period (mean of 9.56µg/m^3) versus those whose pregnancies did not (mean of 9.33µg/m^3) (**Figure 3.5**). There was a positive association between experiencing the fire (as a binary variable) and PTB (as a binary outcome), with crude relative risk equal to 1.15 (95% CI 1.08, 1.23). These associations remained statistically significant after adjustment for demographic and clinical variables, just slightly attenuating to 1.13 (95% CI 1.06, 1.20) and 1.10 (95% CI 1.03, 1.17), respectively. Exposure to the fire was not significant for the Very Early PTB (aRR = .92 (95% CI .68, 1.26)) and Early PTB (aRR = 1.15 (95% CI 0.95, 1.40)) subgroups.

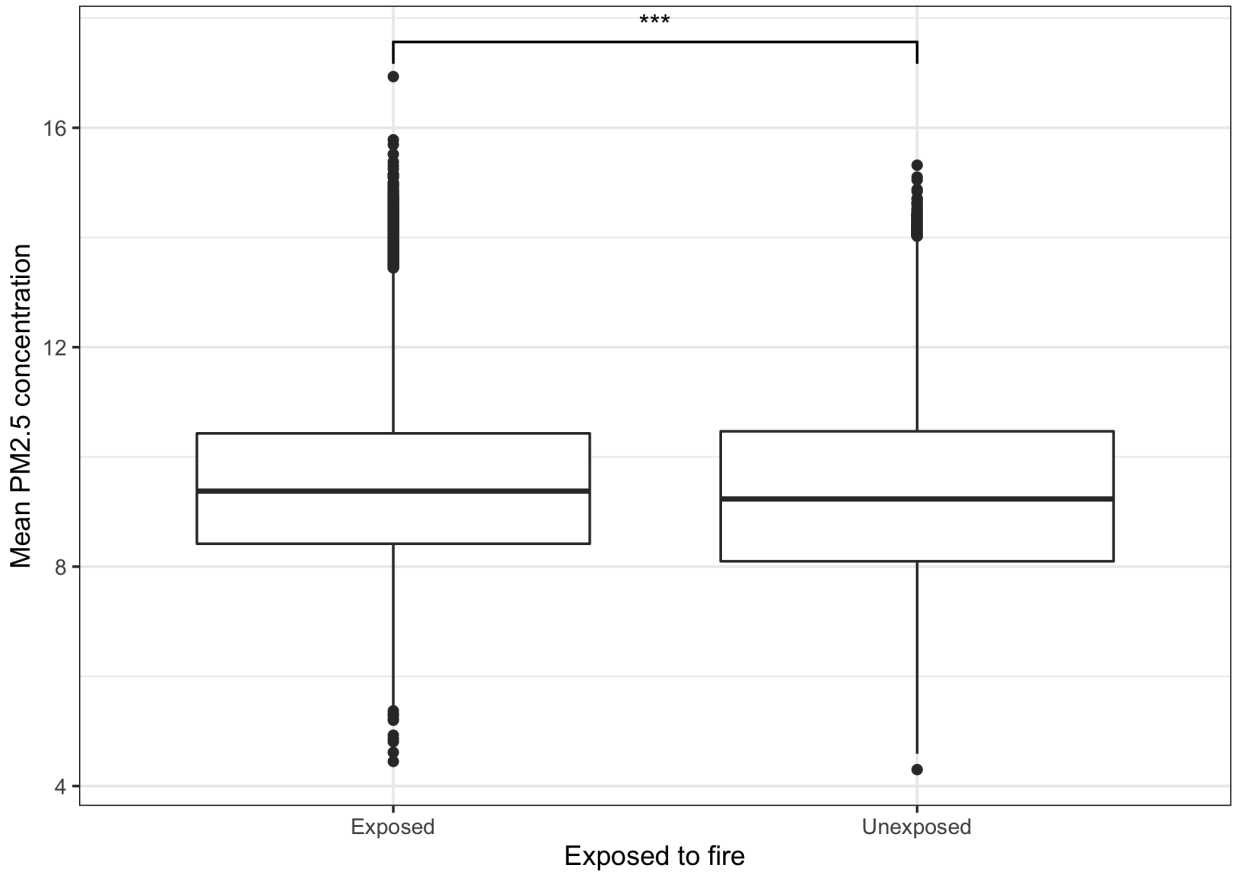

Figure 3.5: Boxplot comparing pregnancy mean $PM_{2.5}$ between the unexposed and exposed groups. Mean of unexposed is 9.33µg/m^3; mean of exposed is 9.57µg/m^3.

Results are shown in Figure **Figure 3.6** and **Table 3.2**. Mean $PM_{2.5}$ concentrations over the full pregnancy were also significant in all models; in the fully adjusted model, the aRR is 1.17 (95% CI 1.15, 1.19). The aRR of pregnancy mean $PM_{2.5}$ per 1µg/m^3 increase in the control group was 1.11 (95% CI 1.09, 1.14), and in the exposed group was 1.26 (95% CI 1.23, 1.29).

In the control group, there was no significant difference in PTB rates between the census tracts in the highest fire tertile and the lowest fire tertile. Similarly, there was no significant difference in PTB rates between the high fire tertiles to the high baseline tertiles in the control group.

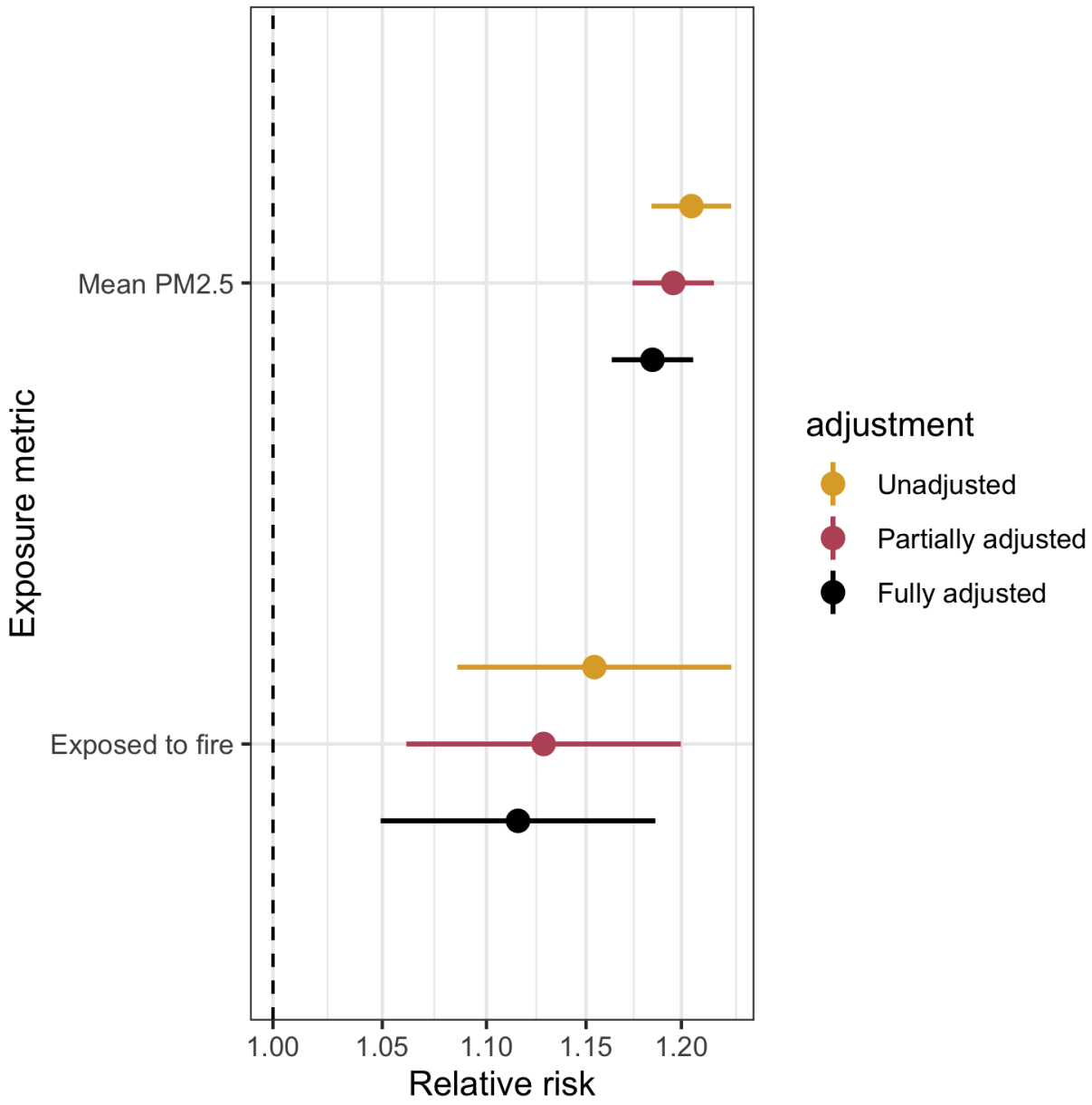

Figure 3.6: Forest plot showing crude, partially adjusted, and fully adjusted relative risk from separate log-linear models. The exposures shown are the pregnancy mean $PM_{2.5}$ as a continuous variable and exposure to fire as a binary variable. The partially adjusted models are adjusted for demographic covariates (race/ethnicity, interpregnancy interval, education, insurance type, prenatal care, and maternal age); the fully adjusted models include demographic and clinical variables (mental disorder, preexisting diabetes, preexisting hypertension, parity, drug abuse, alcohol abuse, smoking, and asthma).

Table 3.2: Table showing crude, partially adjusted, and fully adjusted relative risk from separate log-linear models by timing of preterm birth. The exposures shown are the pregnancy mean $PM_{2.5}$ as a continuous variable and exposure to fire as a binary variable. The partially adjusted models are adjusted for demographic covariates (race/ethnicity, interpregnancy interval, education, insurance type, prenatal care, and maternal age); the fully adjusted models include demographic and clinical variables (mental disorder, preexisting diabetes, preexisting hypertension, parity, drug abuse, alcohol abuse, smoking, and asthma). Preterm is birth <37 weeks, Very Early Preterm is birth <28 weeks, Early Preterm is birth 28 – 32 weeks, and Late Preterm is birth 36 – 37 weeks.

Timing	Exposure	adjustment	RR	95% CI	p value
Preterm	Exposed to fire	Unadjusted	1.15	(1.09, 1.23)	4.29E-06
Preterm	Exposed to fire	Fully adjusted	1.12	(1.05, 1.19)	4.71E-04
Preterm	Exposed to fire	Partially adjusted	1.13	(1.06, 1.20)	1.11E-04
Preterm	Mean $PM_{2.5}$	Unadjusted	1.21	(1.18, 1.23)	4.45E-94
Preterm	Mean $PM_{2.5}$	Fully adjusted	1.18	(1.16, 1.21)	1.35E-74
Preterm	Mean $PM_{2.5}$	Partially adjusted	1.20	(1.17, 1.22)	8.39E-83
Early Preterm	Exposed to fire	Unadjusted	1.17	(0.97, 1.41)	1.08E-01
Early Preterm	Exposed to fire	Fully adjusted	1.11	(0.92, 1.34)	2.66E-01
Early Preterm	Exposed to fire	Partially adjusted	1.13	(0.93, 1.36)	2.13E-01
Early Preterm	Mean $PM_{2.5}$	Unadjusted	1.44	(1.36, 1.52)	2.09E-40
Early Preterm	Mean $PM_{2.5}$	Fully adjusted	1.42	(1.34, 1.50)	2.33E-35
Early Preterm	Mean $PM_{2.5}$	Partially adjusted	1.44	(1.36, 1.52)	4.22E-38
Late Preterm	Exposed to fire	Unadjusted	1.18	(1.08, 1.30)	2.56E-04
Late Preterm	Exposed to fire	Fully adjusted	1.16	(1.06, 1.27)	1.39E-03
Late Preterm	Exposed to fire	Partially adjusted	1.16	(1.06, 1.27)	1.12E-03
Late Preterm	Mean $PM_{2.5}$	Unadjusted	1.13	(1.10, 1.16)	3.72E-19
Late Preterm	Mean $PM_{2.5}$	Fully adjusted	1.13	(1.10, 1.16)	8.54E-18
Late Preterm	Mean $PM_{2.5}$	Partially adjusted	1.13	(1.10, 1.17)	3.09E-19
Very Early Preterm	Exposed to fire	Unadjusted	1.02	(0.76, 1.38)	8.82E-01
Very Early Preterm	Exposed to fire	Fully adjusted	0.96	(0.71, 1.30)	7.94E-01
Very Early Preterm	Exposed to fire	Partially adjusted	0.98	(0.73, 1.33)	9.05E-01
Very Early Preterm	Mean $PM_{2.5}$	Unadjusted	1.80	(1.65, 1.96)	5.00E-42
Very Early Preterm	Mean $PM_{2.5}$	Fully adjusted	1.80	(1.65, 1.96)	1.89E-38
Very Early Preterm	Mean $PM_{2.5}$	Partially adjusted	1.82	(1.67, 1.99)	4.40E-40

3.4.1 Differential exposure during the fire

Among those whose pregnancies overlapped the fire, the range of peak $PM_{2.5}$ concentrations experienced was 76.4 to 202.6µg/m^3. The mean $PM_{2.5}$ over the fire period ranged from 42.3 to 91.0µg/m^3. For the fire period, the air quality tertile cutoffs were $[42.3, 60.0)$, $[60.0, 65.9)$,

and $[65.9, 91.0]$ µg/m^3.

Mean exposure over the full pregnancy was statistically significantly associated with PTB in the fully adjusted model, with an aRR of 1.25 (95% CI 1.22, 1.28) per 1µg/m^3 increase. In the unadjusted model, the risk associated with the comparison of the middle tertile to the lowest was not significant (RR = 1.07 (95% CI 0.97, 1.18)); however, this change was significant in the partially and fully adjusted models (aRR = 1.17 (95% CI 1.05, 1.29) and aRR = 1.15 (95% CI 1.04, 1.28), respectively). The risk associated with the comparison of the highest tertile to the lowest was significant in all three models, with an aRR of 1.17 (95% CI 1.05, 1.29) in the fully adjusted model. Tertile changes were not significant among the Very Early PTB or Early PTB subgroups; pregnancy mean $PM_{2.5}$ was significant among all subgroups. Results are shown in Figure **Figure 3.7** and **Table 3.3**.

Table 3.3: Table showing crude, partially adjusted, and fully adjusted relative risk from separate log-linear models by timing of preterm birth. The exposures shown are the pregnancy mean $PM_{2.5}$ as a continuous variable and the air quality tertile during the fire. The partially adjusted models are adjusted for demographic covariates (race/ethnicity, interpregnancy interval, education, insurance type, prenatal care, and maternal age); the fully adjusted models include demographic and clinical variables (mental disorder, preexisting diabetes, preexisting hypertension, parity, drug abuse, alcohol abuse, smoking, and asthma). Preterm is birth <37 weeks, Very Early Preterm is birth <28 weeks, Early Preterm is birth 28 – 32 weeks, and Late Preterm is birth 36 – 37 weeks.

Timing	**Exposure**	**adjustment**	**RR**	**95% CI**	**p value**
Preterm	Mean $PM_{2.5}$	Unadjusted	1.28	(1.25, 1.31)	8.06E-84
Preterm	Mean $PM_{2.5}$	Fully adjusted	1.25	(1.22, 1.28)	3.47E-67
Preterm	Mean $PM_{2.5}$	Partially adjusted	1.27	(1.23, 1.30)	8.76E-74
Preterm	Medium fire tertile	Unadjusted	1.02	(0.92, 1.12)	7.67E-01
Preterm	High fire tertile	Unadjusted	1.09	(0.98, 1.20)	1.04E-01
Preterm	Medium fire tertile	Fully adjusted	1.11	(1.01, 1.23)	3.86E-02
Preterm	High fire tertile	Fully adjusted	1.15	(1.03, 1.27)	1.00E-02
Preterm	Medium fire tertile	Partially adjusted	1.12	(1.01, 1.24)	3.15E-02
Preterm	High fire tertile	Partially adjusted	1.17	(1.05, 1.29)	3.57E-03
Early Preterm	Mean $PM_{2.5}$	Unadjusted	1.59	(1.47, 1.72)	1.68E-31
Early Preterm	Mean $PM_{2.5}$	Fully adjusted	1.57	(1.44, 1.70)	2.12E-26
Early Preterm	Mean $PM_{2.5}$	Partially adjusted	1.58	(1.45, 1.71)	1.20E-27
Early Preterm	Medium fire tertile	Unadjusted	1.09	(0.78, 1.52)	6.12E-01
Early Preterm	High fire tertile	Unadjusted	1.18	(0.85, 1.64)	3.30E-01

Timing	Exposure	adjustment	RR	95% CI	p value
Early Preterm	Medium fire tertile	Fully adjusted	1.28	(0.91, 1.79)	1.57E-01
Early Preterm	High fire tertile	Fully adjusted	1.24	(0.89, 1.75)	2.05E-01
Early Preterm	Medium fire tertile	Partially adjusted	1.26	(0.90, 1.76)	1.76E-01
Early Preterm	High fire tertile	Partially adjusted	1.27	(0.90, 1.78)	1.71E-01
Late Preterm	Mean $PM_{2.5}$	Unadjusted	1.20	(1.16, 1.25)	2.75E-21
Late Preterm	Mean $PM_{2.5}$	Fully adjusted	1.19	(1.14, 1.24)	1.48E-18
Late Preterm	Mean $PM_{2.5}$	Partially adjusted	1.19	(1.15, 1.24)	2.47E-19
Late Preterm	Medium fire tertile	Unadjusted	1.07	(0.92, 1.24)	3.58E-01
Late Preterm	High fire tertile	Unadjusted	1.10	(0.95, 1.28)	2.00E-01
Late Preterm	Medium fire tertile	Fully adjusted	1.16	(1.00, 1.35)	5.24E-02
Late Preterm	High fire tertile	Fully adjusted	1.16	(1.00, 1.35)	5.21E-02
Late Preterm	Medium fire tertile	Partially adjusted	1.16	(1.00, 1.35)	4.85E-02
Late Preterm	High fire tertile	Partially adjusted	1.18	(1.01, 1.37)	3.20E-02
Very Early Preterm	Mean $PM_{2.5}$	Unadjusted	2.22	(1.98, 2.49)	1.29E-42
Very Early Preterm	Mean $PM_{2.5}$	Fully adjusted	2.33	(2.05, 2.65)	9.36E-39
Very Early Preterm	Mean $PM_{2.5}$	Partially adjusted	2.28	(2.02, 2.58)	1.57E-39
Very Early Preterm	Medium fire tertile	Unadjusted	0.87	(0.51, 1.47)	6.04E-01
Very Early Preterm	High fire tertile	Unadjusted	1.23	(0.76, 2.00)	3.94E-01
Very Early Preterm	Medium fire tertile	Fully adjusted	1.07	(0.63, 1.82)	8.06E-01
Very Early Preterm	High fire tertile	Fully adjusted	1.37	(0.83, 2.25)	2.14E-01
Very Early Preterm	Medium fire tertile	Partially adjusted	1.01	(0.60, 1.71)	9.72E-01
Very Early Preterm	High fire tertile	Partially adjusted	1.30	(0.79, 2.13)	3.05E-01

3.5 Comment

3.5.1 Principal Findings

In this study, we examined the impact of one of the largest and deadliest fires in northern California history on PTB. After adjusting for demographic and clinical covariates, we found an approximate 10% increased risk of PTB by wildfire exposure.

The risk associated with mean pregnancy $PM_{2.5}$ and PTB was higher in the exposed group alone than it was in the full analysis. We also observed a difference in risk comparing those above our mean exposure to those below (regardless of exposure status), suggesting it is at least partially driven by a non-linear association between mean $PM_{2.5}$ and PTB risk. In addition, the composition of wildfire $PM_{2.5}$ may be particularly bad for human health.[37]

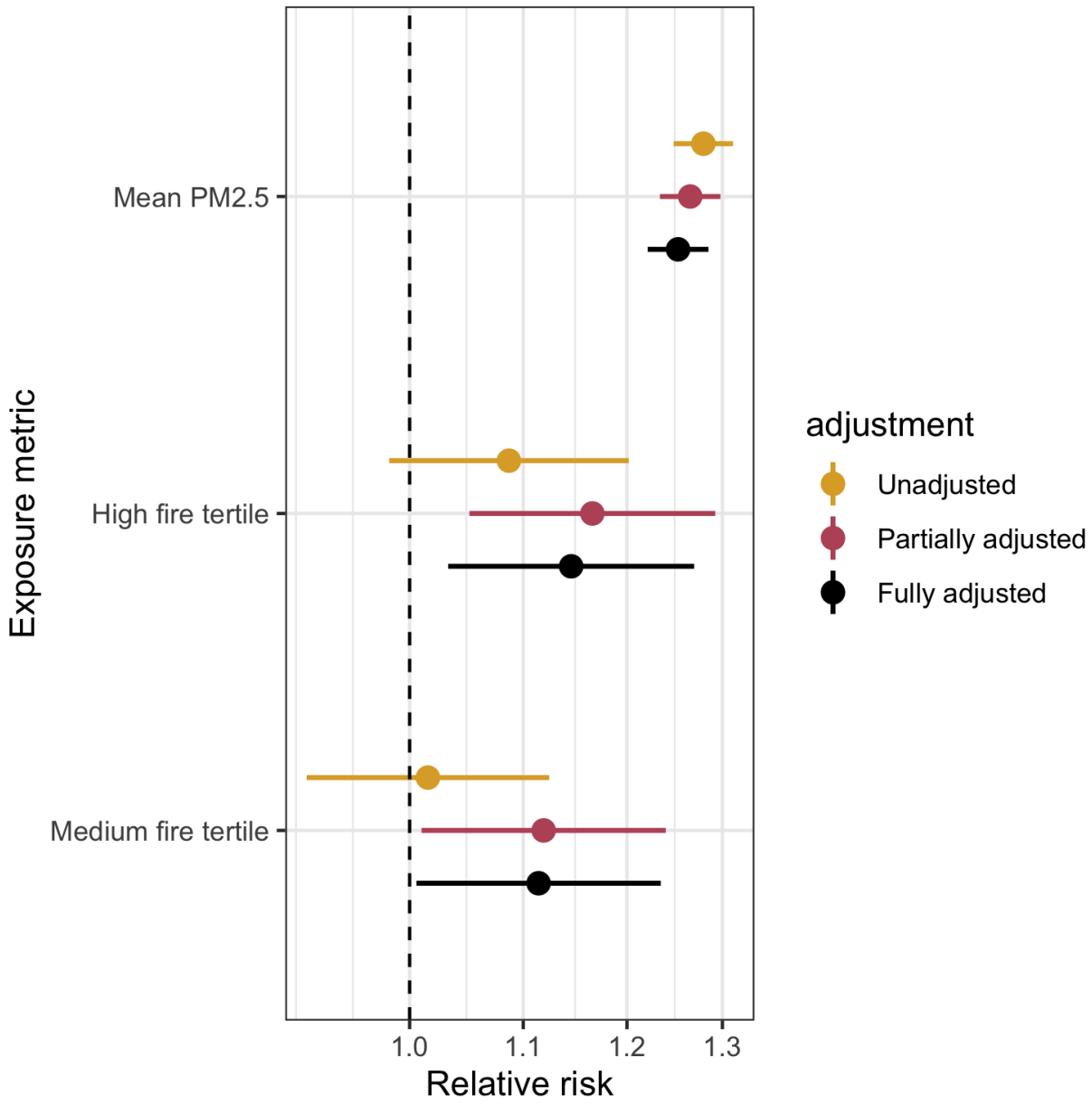

Figure 3.7: Forest plot showing crude, partially adjusted, and fully adjusted relative risk from separate log-linear models. The exposures shown are the pregnancy mean $PM_{2.5}$ as a continuous variable and the air quality tertile during the fire. The partially adjusted models are adjusted for demographic covariates (race/ethnicity, interpregnancy interval, education, insurance type, prenatal care, and maternal age); the fully adjusted models include demographic and clinical variables (mental disorder, preexisting diabetes, preexisting hypertension, parity, drug abuse, alcohol abuse, smoking, and asthma).

It could also be that $PM_{2.5}$ in our exposed group is serving as an indicator of stress (a risk factor for PTB[38]), or some other measure associated with the fire. Among those exposed to the wildfire, mean $PM_{2.5}$ had a higher risk than tertiles of exposure during the wildfire. This may suggest that it is cumulative, rather than acute exposures driving the association between mean $PM_{2.5}$ and PTB. It may also be driven by limited variation in the exposed group. We did not see major differences in timing of PTB.

Of interest is that in their work, Heft-Neal and colleagues found that baseline $PM_{2.5}$ levels were higher in lower-income ZIP codes, but these ZIP codes faced less wildfire smoke. On the other hand, higher-income ZIP codes had lower baseline $PM_{2.5}$ concentrations, but experienced more wildfire smoke. The authors proposed this situation may be due to wealthy rural landowners.[8] In our own study, the mean median income of low-baseline tracts was higher than that of high-baseline tracts. However, there was no significant difference in income between the low-fire tracts and the high-fire tracts.

3.5.2 Clinical Implications

Air chemistry is an important component that is not yet fully understood. Here we measure $PM_{2.5}$, which serves as a proxy for the chemistry of the pollutants. Prior work suggests that wildfire smoke may be particularly bad for human health.[37] The Califoria Air Resources Board performed a chemical analysis of the smoke showing that dangerous levels of toxic metals including lead were present in the wildfire smoke of the Camp Fire, having burned most of downtown Paradise, California. This further supports that this wildfire may have had unique health effects and indicates that future wildland-urban interface fires could have similar increased threats to human health over vegetation-based wildfires.[12] While different types of particulate matter may have different effects on the body, one commonly-proposed pathway from fine particulate matter to PTB is via oxidative stress. Studies have shown that wildfire combustion has high oxidative potential[39] and that wildland firefighters may

show increased biomarkers of oxidative stress.[40] However, it is important to note that biomarkers of increased oxidative stress in non-firefighters have yet to be confirmed during a real wildfire event.[41] Biomarkers of oxidative stress have been linked to PTB;[42]. In a systematic review of the literature, Moore and colleagues concluded that an imbalance of oxidants and antioxidants may be linked to PTB.[43] Further research is required to better understand potential pathways on influence in fire—PTB relationships.

3.5.3 Strengths and Limitations

A major strength of our study is the inclusion of clinical covariates, which is not typical of large population-based studies focused on air pollution and PTB.[15] In addition, we used a spatially and temporally dense map of air sensors to compute a well-resolved map of air quality. This allowed us to look not only at the effects of experiencing the fire, but also the degree to which the air was affected.

One limitation of our study is our inability to differentiate the effects of smoke and non-wildfire particulate matter. We addressed this by computing baseline $PM_{2.5}$ means by census tract and using those for comparisons. By comparing individuals from the same geographic area, we are able to control for some of the impact of background particulate matter. In addition, we compared demographics and PTB rates between census tracts in the highest fire tertile and the lowest fire tertile in our control group (those who delivered prior to the fire). We found that while the tertiles were significantly different in demographics, the PTB rates were not significantly different between the high and the low $PM_{2.5}$ tracts.

We were also not able to differentiate the effects of bad air and those of stress. Other studies have found increases in PTB following environmental disasters where pollution was not a major concern.[44] This raises the question of whether stress or exposure is a driving factor in our results. We restricted our study area to a relatively small region well away from the fire, which was never under threat of evacuation (some people may have been able to leave;

we adjusted for socioeconomic status, which may not address this fully). It is important to note that we did see a range of exposures in our area, and as such, it is possible that those facing higher $PM_{2.5}$ concentrations also experienced increased stress. It will be important that future studies more fully investigate stress as a driver of PTB in pregnant people facing fire-related smoke exposure.

We also based exposure calculations based on residential address and did not have information on residential moves or workplace. While a past study has found that moves during pregnancy did not result in dramatic differences in exposure [45], this may not hold true in acute, high-pollution events and as such, represents another important area for future inquiry.

3.5.4 Research Implications

Wildfires pose an increasing threat to human health, and further work in this field is critical. The methods presented here allow for near real-time air pollution surface construction. The field will benefit from up-to-date studies on recent fire seasons. California is seeing increasing numbers of large wildfires since the Camp Fire of 2018, validating this “new normal" of climate change. One study projects that massive wildfires will continue in the west for roughly the next 10 years.[46] The 2020 California fire season was particularly prolonged and should be examined, especially since residential addresses may better reflect the usual location of residents, due to individuals working from home during COVID-19. The wildfires of 2020 (including August complex, North Complex, CZU complex) and the 2021 fires to date (Dixie, Monument, Caldor, French, and many others) provide ample opportunity to test how much sustained exposure affects pregnant people and infants and to what extent.

In addition, the effects of wildfire smoke on respiratory outcomes in infants is an important question that may help develop policy and health care procedures. Given that preterm infants are at higher risk for adverse respiratory outcomes than term infants, this is vital to

consider in the context of extremely short-term results of smoke exposure, which we did not consider in this study.

This study has implications for influencing policies around mitigating smoke in homes or to build public health programs to provide air purifiers for sustained smoke exposure.

3.5.5 Conclusion

This study found that exposure to wildfire smoke from the 2018 Camp Fire is associated with an increased risk of preterm birth in the SF SF Bay Area, even after accounting for chronic clinical conditions of the pregnant person. Since this study found that $PM_{2.5}$ exposure posed a higher risk in the exposed group, future research should investigate wildfire smoke composition and mechanistic pathways. In addition, further work should examine the relationship of stress and smoke as combined drivers behind the association.

3.6 Acknowledgements

This work was supported by the California PTB Initiative and the WIFIRE Commons funded by National Science Foundation grants 2040676 and 2134904.

The authors wish to thank Karoline Barkjohn, Andrea Clements, and Amara Holder of the U.S. Environmental Protection Agency, Office of Research and Development and the Center for Environmental Measurement and Modeling.

This study was made possible in part due to the weather station data made available by NOAA, National Weather Service, and PG&E via Synopticdata.com

Bibliography

[1] MJ Platt. Outcomes in preterm infants. *Public health*, 128(5):399–403, 2014.

[2] Dag Moster, Rolv Terje Lie, and Trond Markestad. Long-term medical and social consequences of preterm birth. *New England Journal of Medicine*, 359(3):262–273, 2008.

[3] Howard Berger, Nir Melamed, Beth Murray Davis, Haroon Hasan, Karizma Mawjee, Jon Barrett, Sarah D McDonald, Michael Geary, and Joel G Ray. Impact of diabetes, obesity and hypertension on preterm birth: Population-based study. *PloS one*, 15(3): e0228743, 2020.

[4] Rebecca J Baer, Monica R McLemore, Nancy Adler, Scott P Oltman, Brittany D Chambers, Miriam Kuppermann, Matthew S Pantell, Elizabeth E Rogers, Kelli K Ryckman, Marina Sirota, et al. Pre-pregnancy or first-trimester risk scoring to identify women at high risk of preterm birth. *European Journal of Obstetrics & Gynecology and Reproductive Biology*, 231:235–240, 2018.

[5] Casey Crump, Elizabeth A Howell, Annemarie Stroustrup, Mary Ann McLaughlin, Jan Sundquist, and Kristina Sundquist. Association of preterm birth with risk of ischemic heart disease in adulthood. *JAMA pediatrics*, 173(8):736–743, 2019.

[6] Marshall Burke, Anne Driscoll, Sam Heft-Neal, Jiani Xue, Jennifer Burney, and Michael Wara. The changing risk and burden of wildfire in the united states. *Proceedings of the National Academy of Sciences*, 118(2), 2021. ISSN 0027-8424. doi: 10.1073/pnas.2011048118. URL `https://www.pnas.org/content/118/2/e2011048118`.

[7] David M Holstius, Colleen E Reid, Bill M Jesdale, and Rachel Morello-Frosch. Birth weight following pregnancy during the 2003 southern california wildfires. *Environmental health perspectives*, 120(9):1340–1345, 2012.

[8] Sam Heft-Neal, Anne Driscoll, Wei Yang, Gary Shaw, and Marshall Burke. Associations between wildfire smoke exposure during pregnancy and risk of preterm birth in california. *Environmental Research*, 203:111872, 2022.

[9] Mona Abdo, Isabella Ward, Katelyn O'Dell, Bonne Ford, Jeffrey R Pierce, Emily V Fischer, and James L Crooks. Impact of wildfire smoke on adverse pregnancy outcomes in colorado, 2007–2015. *International journal of environmental research and public health*, 16(19):3720, 2019.

[10] Alexander Maranghides, Eric Link, Christopher Brown, William Mell, Steven Hawks, Mike Wilson, Will Brewer, Robert Vihnanek, and William Walton. A case study of the camp fire - fire progression timeline, 2021-02-08 2021.

[11] Cal Fire. Camp incident green sheet. California Department of Forestry and Fire Protection: Sacramento, CA, USA, 2018.

[12] California Air Resources Board. Camp fire air quality data analysis. Technical report, 07 2021. URL `https://ww2.arb.ca.gov/resources/documents/camp-fire-air-quality-data-analysis`.

[13] Harold Gilliam. *Weather of the San Francisco Bay Region*, volume 63. Univ of California Press, 2002.

[14] Purple Air, Inc. Our technology - purpleair laser particle counters & arduinos. URL `https://www2.purpleair.com/pages/technology`.

[15] Sana Amjad, Dagmara Chojecki, Alvaro Osornio-Vargas, and Maria B Ospina. Wildfire exposure during pregnancy and the risk of adverse birth outcomes: A systematic review. *Environment International*, 156:106644, 2021.

[16] MH O'Donnell and AM Behie. Effects of bushfire stress on birth outcomes: A cohort

study of the 2009 victorian black saturday bushfires. *International Journal of Disaster Risk Reduction*, 5:98–106, 2013.

[17] Esri Inc. Arcgis desktop 10.7.1, 2019. URL `http://www.arcgis.com`.

[18] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2021. URL `https://www.R-project.org/`.

[19] Edzer J. Pebesma and Roger S. Bivand. Classes and methods for spatial data in R. *R News*, 5(2):9–13, November 2005. URL `https://CRAN.R-project.org/doc/Rnews/`.

[20] Jonathan Asher Greenberg and Matteo Mattiuzzi. *gdalUtils: Wrappers for the Geospatial Data Abstraction Library (GDAL) Utilities*, 2020. URL `https://CRAN.R-project.org/package=gdalUtils`. R package version 2.0.3.2.

[21] Laura August. Calenviroscreen 3.0. *Text. OEHHA. December*, 29:2016, 2016.

[22] Hongtai Huang, Tracey J Woodruff, Rebecca J Baer, Komal Bangia, Laura M August, Laura L Jellife-Palowski, Amy M Padula, and Marina Sirota. Investigation of association between environmental and socioeconomic factors and preterm birth in california. *Environment international*, 121:1066–1078, 2018.

[23] Q. Di, Y. Wei, A. Shtein, C. Hultquist, X. Xing, H. Amini, L. Shi, I. Kloog, R. Silvern, J. Kelly, and et al. Daily and annual pm2.5 concentrations for the contiguous united states, 1-km grids, v1 (2000 - 2016), 2021. URL `https://sedac.ciesin.columbia.edu/data/set/aqdh-pm2-5-concentrations-contiguous-us-1-km-2000-2016`.

[24] Qian Di, Heresh Amini, Liuhua Shi, Itai Kloog, Rachel Silvern, James Kelly, M Benjamin Sabath, Christine Choirat, Petros Koutrakis, Alexei Lyapustin, et al. An ensemble-based model of pm2.5 concentration across the contiguous united states with high spatiotemporal resolution. *Environment international*, 130:104909, 2019.

[25] Karoline K Barkjohn, Brett Gantt, and Andrea L Clements. Development and application of a united states-wide correction for pm 2.5 data collected with the purpleair sensor. *Atmospheric Measurement Techniques*, 14(6):4617–4637, 2021.

[26] Christopher Sardegna. Reagentx/purple_air_api: Purple air client v1.2.3. Feb 2021. doi: 10.5281/zenodo.4540629.

[27] Matthew J Brewer and Craig B Clements. The 2018 camp fire: Meteorological analysis using in situ observations and numerical simulations. *Atmosphere*, 11(1):47, 2020.

[28] NOAA National Centers for Environmental Information. Noaa office of satellite and product operations, (1994): Noaa geostationary operational environmental satellite (goes) i-m and n-p series imager data.

[29] David W Wong, Lester Yuan, and Susan A Perlin. Comparison of spatial interpolation methods for the estimation of air quality data. *Journal of Exposure Science & Environmental Epidemiology*, 14(5):404–415, 2004.

[30] Nynke Hofstra, Malcolm Haylock, Mark New, P. Jones, and Christoph Frei. Comparison of six methods for the interpolation of daily european climate data. *Journal of Geophysical Research. D, Atmospheres 113 (2008)*, 113, 11 2008. doi: 10.1029/2008JD010100.

[31] Nima Nikzad, Celal Ziftci, Piero Zappi, Nichole Quick, Priti Aghera, Nakul Verma, Barry Demchak, Kevin Patrick, Hovav Shacham, and Tajana Rosing. Citisense: Adaptive services for community-driven behavioral and environmental monitoring to induce change. 10 2012.

[32] Hanefi Bayraktar and F Sezer Turalioglu. A kriging-based approach for locating a sampling site—in the assessment of air quality. *Stochastic Environmental Research and Risk Assessment*, 19(4):301–305, 2005.

[33] Janaina Antonino Pinto, Prashant Kumar, Marcelo Félix Alonso, Willian Lemker Andreao, Rizzieri Pedruzzi, Sergio Ibarra Espinosa, and Taciana Toledo de Almeida Albuquerque. Kriging method application and traffic behavior profiles from local radar network database: A proposal to support traffic solutions and air pollution control strategies. *Sustainable Cities and Society*, 56:102062, 2020.

[34] John Horel, Michael Splitt, L Dunn, J Pechmann, B White, C Ciliberti, S Lazarus, J Slemmer, D Zaff, and J Burks. Mesowest: Cooperative mesonets in the western united states. *Bulletin of the American Meteorological Society*, 83(2):211–226, 2002.

[35] Milton Kotelchuck. An evaluation of the kessner adequacy of prenatal care index and a proposed adequacy of prenatal care utilization index. *American journal of public health*, 84(9):1414–1420, 1994.

[36] Len Finocchio, Matthew Newman, and Eunice Roh. Medi-cal facts and figures: Crucial coverage for low- income californians. `https://www.chcf.org/wp-content/uploads/2019/02/MediCalFactsFiguresAlmanac2019.pdf`, 2 2019.

[37] Rosana Aguilera, Thomas Corringham, Alexander Gershunov, and Tarik Benmarhnia. Wildfire smoke impacts respiratory health more than fine particles from other sources: observational evidence from southern california. *Nature communications*, 12(1):1–8, 2021.

[38] Nancy Dole, David A Savitz, Irva Hertz-Picciotto, Anna Maria Siega-Riz, Michael J McMahon, and Pierre Buekens. Maternal stress and preterm birth. *American journal of epidemiology*, 157(1):14–24, 2003.

[39] Anette Kocbach Bølling, Joakim Pagels, Karl Espen Yttri, Lars Barregard, Gerd Sallsten, Per E Schwarze, and Christoffer Boman. Health effects of residential wood smoke particles: the importance of combustion conditions and physicochemical particle properties. *Particle and fibre toxicology*, 6(1):1–20, 2009.

[40] Marta Oliveira, Solange Costa, Josiana Vaz, Adília Fernandes, Klara Slezakova, Cristina Delerue-Matos, João Paulo Teixeira, Maria Carmo Pereira, and Simone Morais. Firefighters exposure to fire emissions: Impact on levels of biomarkers of exposure to polycyclic aromatic hydrocarbons and genotoxic/oxidative-effects. *Journal of hazardous materials*, 383:121179, 2020.

[41] Carolyn Black, Yohannes Tesfaigzi, Jed A Bassein, and Lisa A Miller. Wildfire smoke exposure and human health: Significant gaps in research for a growing public health issue. *Environmental toxicology and pharmacology*, 55:186–195, 2017.

[42] Kelly K Ferguson, Thomas F McElrath, Yin-Hsiu Chen, Rita Loch-Caruso, Bhramar Mukherjee, and John D Meeker. Repeated measures of urinary oxidative stress biomarkers during pregnancy and preterm birth. *American journal of obstetrics and gynecology*, 212(2):208–e1, 2015.

[43] Tiffany A Moore, Iman M Ahmad, and Matthew C Zimmerman. Oxidative stress and preterm birth: an integrative review. *Biological research for nursing*, 20(5):497–512, 2018.

[44] Kelsey N Dancause, David P Laplante, Carolina Oremus, Sarah Fraser, Alain Brunet, and Suzanne King. Disaster-related prenatal maternal stress influences birth outcomes: Project ice storm. *Early human development*, 87(12):813–820, 2011.

[45] Philip J Lupo, Elaine Symanski, Wenyaw Chan, Laura E Mitchell, D Kim Waller, Mark A Canfield, and Peter H Langlois. Differences in exposure assignment between conception and delivery: the impact of maternal mobility. *Paediatric and perinatal epidemiology*, 24(2):200–208, 2010.

[46] Maureen C Kennedy, Ryan R Bart, Christina L Tague, and Janet S Choate. Does hot and dry equal more wildfire? contrasting short-and long-term climate effects on fire in the sierra nevada, ca. *Ecosphere*, 12(7):e03657, 2021.

Chapter 4

COVID, Air quality, and Preterm birth

4.1 Abstract

Background: Previous studies have demonstrated an association between fine particulate matter (particulate matter $\leq$2.5 µm in diameter, $PM_{2.5}$) and preterm birth (PTB). Some risk factors for PTB have been shown to modify this association. Whether the $PM_{2.5}$ – PTB association is modified by SARS-CoV-2, however, remains unknown.

Objective: Determine whether the association between $PM_{2.5}$ and PTB is modified by testing positive for SARS-CoV-2.

Methods: We performed a retrospective cohort study on nearly 200,000 pregnant people in California, USA. COVID status was obtained from birth records. Each individual was assigned a pregnancy-wide average $PM_{2.5}$ exposure by residential location at time of delivery. We applied separate Cox models, adjusted for potential confounders, to those who tested

positive and negative and compared the resulting hazard ratios (HRs).

Results: We found a non-linear relationship between $PM_{2.5}$ and PTB in our population. While the HR curves showed different shapes, they were not significantly different.

Conclusion: We found that testing positive for SARS-CoV-2 did not modify the relationship between $PM_{2.5}$ and PTB.

4.2 Introduction

Preterm birth (PTB), defined as a live birth prior to 37 weeks of gestation, is the leading cause of neonatal mortality and morbidity in the United States. The consequences of preterm birth can be both acute, such as respiratory distress syndrome[1], and long-term, including diabetes mellitus (DM)[2] and severe medical disability.[3, 4] Air pollution has been linked to increased risk of PTB in multiple studies.[5–7] While further research is needed to definitively understand the biological pathways linking air pollution and PTB, Kannan and colleagues propose that oxidative stress and inflammation, among other effects, play a role.[8] During the COVID-19 pandemic, some regions experienced improved air quality[9] and/or reduced preterm birth rates.[10] However, these changes are not universal[11, 12] and the risk between exposure to air pollution and PTB likely remains.

Pregnant people are at higher risk for developing severe COVID-19,[13] which in turn is a risk factor for preterm birth.[14] While it was initially thought that much of this risk was due to medically indicated deliveries,[14] recent studies have demonstrated that infection with SARS-CoV-2 is associated with spontaneous preterm birth as well[15]. COVID-19 causes inflammation in multiple systems,[16] increasing the risk for preterm birth.[17, 18] A COVID-19 positive status may also act as a marker of medical vulnerability. Several risk factors for SARS-CoV-2 severity, such as smoking, DM, and obesity,[19, 20] are known risk factors for preterm birth.[21] SARS-CoV-2 may also act on preterm birth by inducing stress, another

risk factor for preterm birth.[22] A positive COVID-19 result and subsequent quarantine during the prenatal period may induce anxiety,[23] further impacting birth outcomes.

Prior work has identified that certain risk factors for PTB, such as maternal DM[24] can be effect modifiers of the relationship between $PM_{2.5}$ and PTB. Since DM is linked with the severity of COVID-19, this raises the question of whether infection with SARS-CoV-2 may also act as an effect modifier on the $PM_{2.5}$ – PTB association.

In this retrospective cohort study, we examined whether testing positive for SARS-CoV-2 modifies the association between $PM_{2.5}$ exposure and PTB. Given the similar mechanisms of actions between the two risk factors, we expect the virus to amplify the existing relationship between $PM_{2.5}$ and PTB. We rely on California birth certificate data from 2019-08-15 to 2020-04-30, which cover a large and diverse population and contain individual-level data on COVID-19 status and comorbidities.

4.3 Methods

Methods and protocols for the study were approved by the Committee for the Protection of Human Subjects (CPHS) which serves as the institutional review board (IRB) for the California Health and Human Services Agency (CHHSA).

4.3.1 Study population

California birth certificate data on births conceived between 08-15-2019 and 04-30-2020 provided information on our outcome of interest (PTB) and individual covariates. Our date range was chosen to ensure we would have the full records for infants born between 20 and 41 weeks, thus avoiding fixed cohort bias. We restricted to singleton births. Multiple births to the same person were included; given our date range, this number is expected to be quite small.

4.3.2 Outcome assessment

Gestational age is recorded as best obstetric estimate, and preterm birth is defined as live birth prior to 37 weeks of gestation. Term birth was defined as birth between 39 and 41 gestational weeks.

4.3.3 Exposure assessment

$PM_{2.5}$ was calculated as a pregnancy-wide average for each individual. Latitude and longitude were pulled from birth records; 3.4% of births were missing these. We calculated $PM_{2.5}$ exposure using well-validated predictions.[25] These predictions are constructed from ground and satellite observations, in combination with a chemical transport model to represent a finely-resolved summary of $PM_{2.5}$ concentrations. This model yields highly spatially resolved predicted concentrations, at 1 km^2 resolution, and good predictive accuracy (R^2 = 0.70). The most recent available predictions are from 2018; to estimate exposures in 2019 and 2020, we obtained daily mean $PM_{2.5}$ concentrations from EPA's Air Quality System (AQS) database at all available monitoring locations in CA, which we subsequently averaged at the county level, if more than one monitoring site were present in a county. Seven counties did not have any monitoring sites; for these counties we used the concentration measured at the site nearest to the county centroid. We then multiplied the predicted concentrations within each county with the ratio of the daily county average $PM_{2.5}$ concentrations over the county average annual $PM_{2.5}$ predictions, to obtain daily finely resolved $PM_{2.5}$ concentrations in 2019 and 2020. For each person, we estimated pregnancy-wide average $PM_{2.5}$ concentrations at the delivery residential location. A small number of addresses were geocoded to locations that fell just outside the bounds of the raster; exposure for these addresses was based on the nearest raster cell.

4.3.4 COVID-19 status

Beginning in July 2020, COVID-19 status was recorded on birth certificates as "confirmed" or "presumptive." A "confirmed" diagnosis indicated that the positive test result was confirmed by a Centers for Disease Control and Prevention (CDC) laboratory, while a "presumptive" diagnosis was used for positive results from state or local laboratories. We considered both confirmed and presumptive cases as testing positive for COVID-19.

4.3.5 Covariate data

We added a binary variable to show whether any day during pregnancy exceeded 50µg/m^3; this serves as a wildfire indicator. Additional covariates included were: maternal education (12 years, <12 years, >12 years), maternal age (<18 years or >34 years compared to 18-34 years), infant sex, interpregnancy interval (IPI) (normal (24-59 months), very short (<6 months), short (6-23 months), or long (>59 months)), insurance status (public vs. non-public), race/ethnicity (white non-Hispanic, Black, Asian, Hispanic, or other), and adequacy of prenatal care: adequate (care begun by 4th month and 80%-109% of recommended visits received[26]), inadequate (care begun after 4th month or <50% of recommended visits received), intermediate (care begun by 4th month and 50-79% of recommended visits received) or adequate plus (care begun by 2nd month and ≥110% of recommended visits received). A census tract-level poverty indicator (percent of the population living in poverty) was obtained from the American Community Survey (2015–19).

4.3.6 Statistical Analysis

To determine whether COVID status was an effect modifier of the overall association between $PM_{2.5}$ and preterm birth, we constructed two separate Cox models: one for COVID positive and one for COVID negative pregnant individuals. We stratified models by race/ethnicity, insurance, and adequacy of prenatal care. We selected covariates *a priori* based on previously

identified confounders for preterm birth. We adjusted for interpregnancy interval, maternal age, infant sex, season, whether $PM_{2.5}$ exceeded 50µg/m^3on any day during the pregnancy, and census tract-level poverty. We also included a fixed effect for county to account for differences by county in COVID spread and preventative measures. $PM_{2.5}$ and the month of conception—to adjust for seasonality—were included as natural splines with three degrees of freedom. We selected the degrees of freedom based on the Akaike information criterion (AIC) and simplicity of the model.

We then compared the hazard ratios for $PM_{2.5}$ between the two by assessing the difference in hazard ratio at specific $PM_{2.5}$ intervals.

Analysis was performed in R version 4.0.4.[27]

4.3.7 Sensitivity Analysis

Both $PM_{2.5}$ sources and concentrations, as well as SARS-CoV-2 spread, differ by urbanicity. We tested adjustment by multiple measures of ruralness and the availability of medical care. These were: rural vs urban census tract, urban area compared to large rural city/town and small & isolated small rural town (as defined by the Rural Urban Commuting Area Codes[28]), distance from county centroid to the nearest Neonatal Intensive Care Unit (NICU), and physician-to-population ratio based on Medical Service Study Area.

4.4 Results

Our study population was largely Hispanic (48.1%). White was the second largest race/ethnicity grouping at 26.6%, and 5.1% of our sample identified as Black. Most of our sample (57.1%) had more than 12 years of education and the majority (74.5%) were within 18–34 years of age (**Table 4.1**). Our final sample consisted of 69,141 births conceived between 2020-02-12 and 2020-04-30 (**Figure 4.1**).

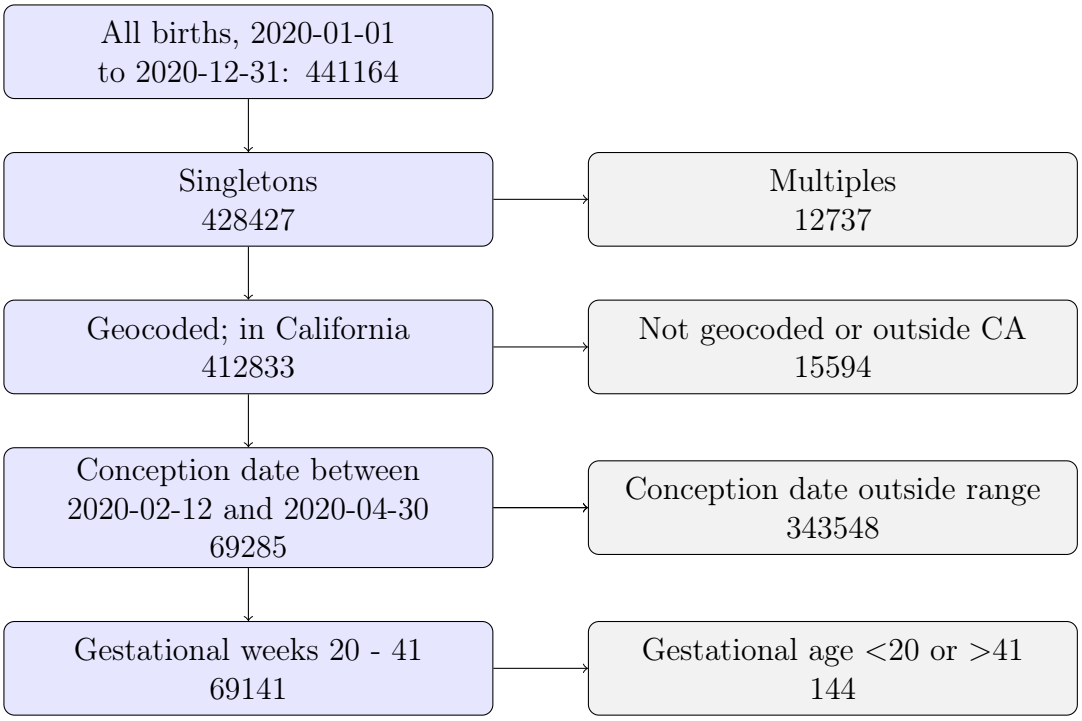

Figure 4.1: Sample selection

Table 4.1: Sample overview

Variable	Lower $PM_{2.5}$ quartiles	Highest $PM_{2.5}$ quartile	p-value
N	**29596**	**39541**	
Race/ethnicity (%)			<0.001
White	9115 (30.8)	9305 (23.5)	
Hispanic	12351 (41.7)	20932 (52.9)	
Black	1240 (4.2)	2259 (5.7)	
Asian	4217 (14.2)	4826 (12.2)	
Other	2673 (9.0)	2219 (5.6)	
Prenatal care (%)			<0.001
Adequate	12536 (42.4)	17460 (44.2)	
Adequate+	7315 (24.7)	10363 (26.2)	
Intermediate	6747 (22.8)	7597 (19.2)	
Inadequate	2998 (10.1)	4121 (10.4)	
MediCal	10388 (35.1)	18775 (47.5)	<0.001
Interpregnancy interval			<0.001
Normal	28538 (96.4)	38342 (97.0)	
Very short	525 (1.8)	579 (1.5)	
Short	438 (1.5)	487 (1.2)	
Long	95 (0.3)	133 (0.3)	
Maternal age (%)			<0.001
18 - 34	21424 (72.4)	30092 (76.1)	

	Lower $PM_{2.5}$ quartiles	Highest $PM_{2.5}$ quartile	p-value
>34	7931 (26.8)	9053 (22.9)	
<18	241 (0.8)	396 (1.0)	
Male infant (%)	15248 (51.5)	20117 (50.9)	0.095
Education (%)			<0.001
12 years	8745 (29.5)	13355 (33.8)	
< 12 years	2817 (9.5)	4715 (11.9)	
> 12 years	18034 (60.9)	21471 (54.3)	
$PM_{2.5}$ >50	20432 (69.0)	36033 (91.1)	<0.001
COVID+	1636 (5.5)	2324 (5.9)	0.052
Poverty (mean (SD))	12.24 (8.86)	16.92 (10.80)	<0.001

A map of the $PM_{2.5}$ predictions for December 31, 2020 is shown in **Figure 4.2**. There were 141 individuals along the coast who fell just outside the raster area; these were mapped to the nearest raster pixel. 81.7% people in our sample experienced a day with mean $PM_{2.5}$ > 50µg/m^3 during their pregnancy. Of those, the mean number of days above 50µg/m^3 was 8.4. The mean pregnancy-wide $PM_{2.5}$ average was 10.8 µg/m^3, with a minimum of 1.7 µg/m^3 and a maximum of 39.8µg/m^3. A map of average $PM_{2.5}$ exposure by county is shown in **Figure 4.3**.

The overall preterm birth rate was 8.4%; among those in the highest $PM_{2.5}$ quartile, the rate was 9.4%. PTB rates by county are displayed in **Figure 4.3**. There were 3960 people in our sample with a presumed or confirmed COVID-19 diagnosis, representing 5.7% of the study population.

The exposure-response curves for the COVID- and COVID+ populations are presented in **Figure 4.4**. The lowest points on the HR curves for both the COVID- and COVID+ populations occurred between 10 µg/m^3 and 15 µg/m^3. The average exposure was lower for the COVID- group than for the COVID+ group, but the range of the COVID- group extended higher. Setting our reference $PM_{2.5}$ level at 10µg/m^3 (**Figure 4.5**), we found that a $PM_{2.5}$ level of 5µg/m^3 resulted in a hazard ratio of 1.26 (95% CI 0.99, 1.60) in the COVID- group and 2.49 (95% CI 0.86, 7.20) in the COVID+ group. At a $PM_{2.5}$ level of 15µg/m^3,

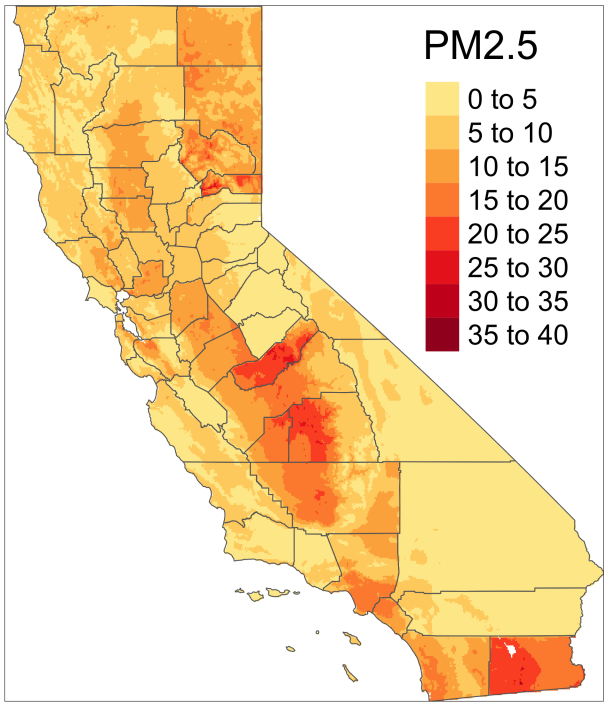

Figure 4.2: Predicted $PM_{2.5}$ surface for 2020-12-31

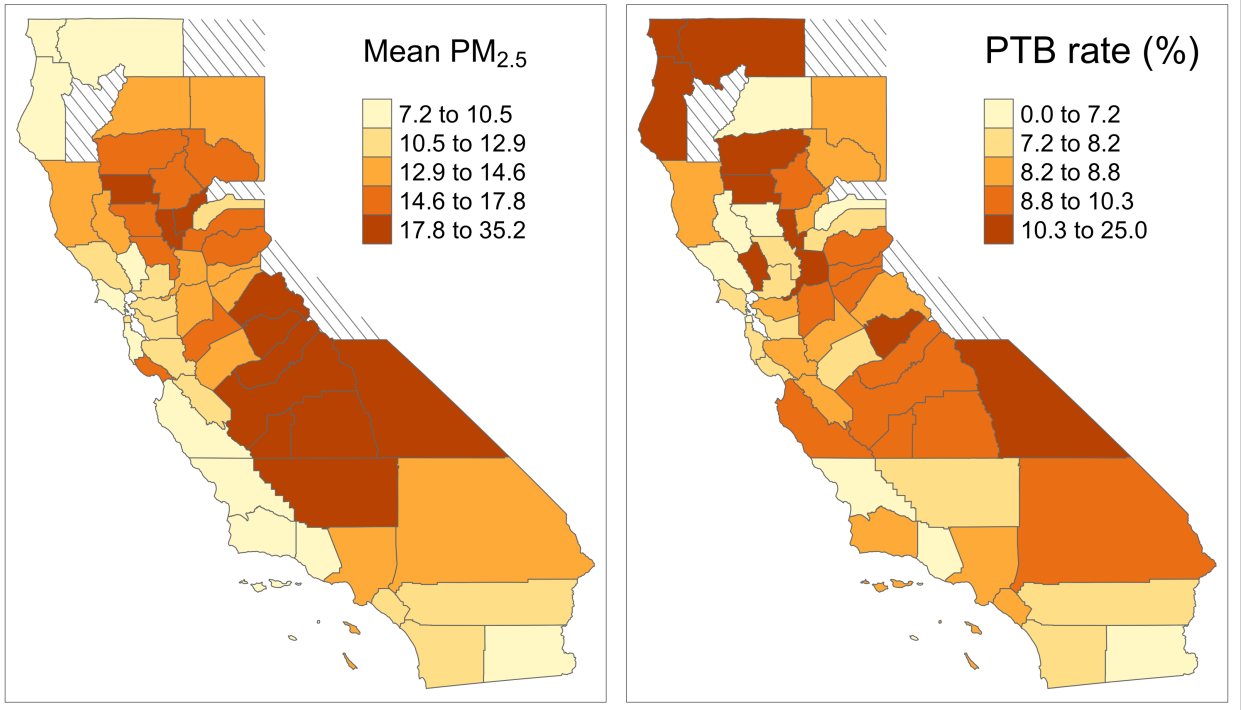

Figure 4.3: Counties are colored by quantile. Counties showing NA had fewer than 20 births.

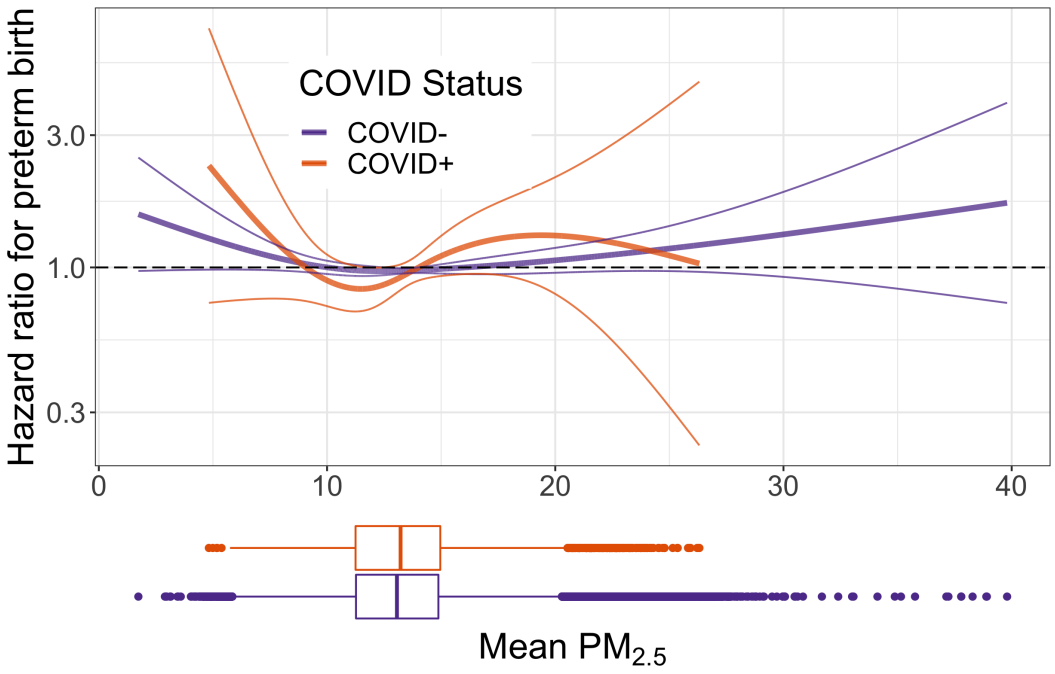

Figure 4.4: Comparison between those who tested positive for COVID-19 (orange), and those who did not (purple). The Cox model was stratified by race, insurance, and prenatal care, and adjusted for interpregnancy interval, age, infant sex, education level, month of conception, a binary indicator for $PM_{2.5}$ above 50µg/m^3, census tract poverty, and region.

the COVID- group had a hazard ratio of 0.99 (95% CI 0.90, 1.08), while the COVID+ group had a hazard ratio of 1.23 (95% CI 0.85, 1.77).

Table 4.2: The reference point is at $PM_{2.5} = 10$µg/m^3. These values are plotted in **Figure 4.5**.

PM2.5	COVID status	Hazard ratio	95% CI
5	COVID-	1.26	(0.99, 1.6)
5	COVID+	2.49	(0.86, 7.2)
10	COVID-	1	(1, 1)
10	COVID+	1	(1, 1)
15	COVID-	0.99	(0.90, 1.08)
15	COVID+	1.23	(0.85, 1.77)
20	COVID-	1.06	(0.91, 1.24)
20	COVID+	1.45	(0.76, 2.76)
25	COVID-	1.17	(0.94, 1.46)
25	COVID+	1.23	(0.33, 4.55)

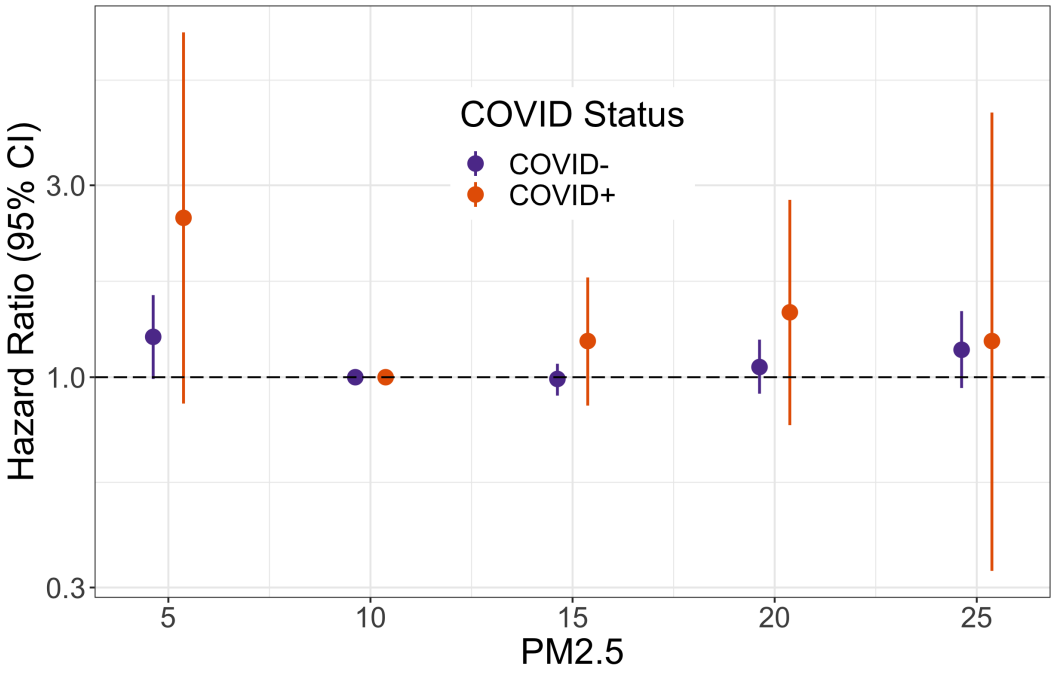

Figure 4.5: Comparison between those who tested positive for COVID-19 (orange), and those who did not (purple). The reference point is at $PM_{2.5} = 10\mu g/m^3$. Values are shown in **Table 4.2**.

4.5 Discussion

Overall, we observed a nonlinear relationship between $PM_{2.5}$ and risk of PTB in California. We also found that testing positive for COVID-19 did not modify the association between $PM_{2.5}$ exposure and PTB.

Few other studies have used splines to investigate the association between $PM_{2.5}$ and PTB. One study examining the associations between nitrogen dioxide, sulfur dioxide, and carbon monoxide (CO) and PTB found non-linear patterns emerge for each pollutant. In particular, CO exhibited a J-shape similar to our COVID- curve, with an elevated odds ratio for PTB at the low-exposure end. Nonetheless, the hazard ratio curve we observed for $PM_{2.5}$ in 2020 was unexpected in that the lowest HR was between $10\mu g/m^3$ and $15\mu g/m^3$. In our sensitivity analyses, we tested adjustment by several variables of ruralness we thought may act as confounders, but none of these changed the HR curve shape or the low point. This

may be due to the higher than expected ranges of $PM_{2.5}$. Our pregnancy-wide mean $PM_{2.5}$ range extended very high, to nearly 40µg/m^3. 6.1% of our sample had a mean exposure ≥20µg/m^3. These high numbers may have affected the shape of our expected distribution of the exposure-response curve. Most of California experienced a long wildfire period, causing weeks of increased $PM_{2.5}$, which in some cases overlapped with a heat wave, both of which are associated with preterm birth.[29, 30]

Previous studies investigating effect modification on the relationship between air quality and preterm birth have found that some risk factors are effect modifiers, while for some there is no evidence. Lavigne and colleagues found that DM, a risk factor for PTB, modified the association between $PM_{2.5}$ and PTB.[24] Similarly, a study found that advanced maternal age acted as an effect modifier.[31] However, when investigating whether hypertension, another PTB risk factor, modified the effect between air pollution and PTB, Weber and colleagues found that it generally did not.[32] Given these results, we expected that testing positive for SARS-CoV-2 would modify the association between $PM_{2.5}$ and PTB. However, our results showed no such effect modification.

A major strength of our study is that both of our primary exposures (COVID-19 status and $PM_{2.5}$ exposure) are measured at the individual-level. We had a large sample size of individuals across the state of California who were universally screened for SARS-CoV-2 at delivery. Many other studies examining the COVID-19 impacts and outcomes have relied on county-level data, which does not fully account for intra-county variation.

There are several important limitations to note. In assessing exposure, we assume that pollution spatial patterns did not change significantly, which may result in exposure measurement error. Our outcome is based on best obstetric estimate; errors are unlikely to differ by exposure. Unmeasured confounding may be present, which may have affected our results. Testing and reporting practices vary by location and evolved during the pandemic. While COVID-19 reporting inconsistencies, errors, and incompleteness may have affected our results, we have

no reason to believe this misclassification would have been differential with respect to either exposure or outcome. Information regarding the timing of testing positive for COVID-19 was limited and information regarding infection severity was unavailable. Since long-term air pollution exposure has been linked with SARS-CoV-2 severity,[33] future studies with available data could stratify by severity to assess whether the relationship is modified for only severe or only mild cases. In particular, the relationship may change depending on whether the person experiences respiratory symptoms.

4.6 Conclusion

In this study of nearly 70,000 births that took place in California during the COVID-19 pandemic we found that testing positive for COVID-19 did not amplify the association between $PM_{2.5}$ and preterm birth. Given that those who experience the highest long-term levels of air pollution likely face the compound risk of lower socioeconomic status and increased risk of contracting COVID-19, this result, while preliminary until evaluated in additional populations and time periods, provides important evidence around the interaction between air pollution and COVID-19.

Bibliography

[1] Jenny Fraser, Moira Walls, and William McGuire. Abc of preterm birth: respiratory complications of preterm birth. *Bmj*, 331(Suppl S1), 2005.

[2] Casey Crump, Jan Sundquist, and Kristina Sundquist. Preterm birth and risk of type 1 and type 2 diabetes: a national cohort study. *Diabetologia*, 63(3):508–518, 2020.

[3] Dag Moster, Rolv Terje Lie, and Trond Markestad. Long-term medical and social consequences of preterm birth. *New England Journal of Medicine*, 359(3):262–273, 2008.

[4] MJ Platt. Outcomes in preterm infants. *Public health*, 128(5):399–403, 2014.

[5] Bruce Bekkar, Susan Pacheco, Rupa Basu, and Nathaniel DeNicola. Association of air pollution and heat exposure with preterm birth, low birth weight, and stillbirth in the us: a systematic review. *JAMA network open*, 3(6):e208243–e208243, 2020.

[6] Mary Huynh, Tracey J Woodruff, Jennifer D Parker, and Kenneth C Schoendorf. Relationships between air pollution and preterm birth in california. *Paediatric and perinatal epidemiology*, 20(6):454–461, 2006.

[7] Jun Wu, Cizao Ren, Ralph J Delfino, Judith Chung, Michelle Wilhelm, and Beate Ritz. Association between local traffic-generated air pollution and preeclampsia and preterm delivery in the south coast air basin of california. *Environmental health perspectives*, 117(11):1773–1779, 2009.

[8] Srimathi Kannan, Dawn P Misra, J Timothy Dvonch, and Ambika Krishnakumar. Exposures to airborne particulate matter and adverse perinatal outcomes: a biologically plausible mechanistic framework for exploring potential effect modification by nutrition. *Environmental health perspectives*, 114(11):1636–1642, 2006.

[9] Qian Liu, Jackson T Harris, Long S Chiu, Donglian Sun, Paul R Houser, Manzhu Yu, Daniel Q Duffy, Michael M Little, and Chaowei Yang. Spatiotemporal impacts of covid-19 on air pollution in california, usa. *Science of the Total Environment*, 750:141592, 2021.

[10] Gitte Hedermann, Paula L Hedley, Marie Baekvad-Hansen, Henrik Hjalgrim, Klaus Rostgaard, Porntiva Poorisrisak, Morten Breindahl, Mads Melbye, David Hougaard, Michael Christiansen, et al. Changes in premature birth rates during the danish nationwide covid-19 lockdown: a nationwide register-based prevalence proportion study. *MedRxiv*, 2020.

[11] L-W Antony Chen, Lung-Chang Chien, Yi Li, and Ge Lin. Nonuniform impacts of covid-19 lockdown on air quality over the united states. *Science of the Total Environment*, 745:141105, 2020.

[12] Elliott K Main, Shen-Chih Chang, Andrew M Carpenter, Paul H Wise, David K Stevenson, Gary M Shaw, and Jeffrey B Gould. Singleton preterm birth rates for racial and ethnic groups during the coronavirus disease 2019 pandemic in california. *American Journal of Obstetrics & Gynecology*, 224(2):239–241, 2021.

[13] Laura D Zambrano, Sascha Ellington, Penelope Strid, Romeo R Galang, Titilope Oduyebo, Van T Tong, Kate R Woodworth, John F Nahabedian III, Eduardo Azziz-Baumgartner, Suzanne M Gilboa, 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. *Morbidity and Mortality Weekly Report*, 69(44):1641, 2020.

[14] José Villar, Shabina Ariff, Robert B Gunier, Ramachandran Thiruvengadam, Stephen Rauch, Alexey Kholin, Paola Roggero, Federico Prefumo, Marynéa Silva Do Vale, Jorge Arturo Cardona-Perez, et al. Maternal and neonatal morbidity and mortality

among pregnant women with and without covid-19 infection: the intercovid multinational cohort study. *JAMA pediatrics*, 2021.

[15] Deborah Karasek, Rebecca J Baer, Monica R McLemore, April J Bell, Bridgette E Blebu, Joan A Casey, Kimberly Coleman-Phox, Jean M Costello, Jennifer N Felder, Elena Flowers, et al. The association of covid-19 infection in pregnancy with preterm birth: A retrospective cohort study in california. *The Lancet Regional Health-Americas*, page 100027, 2021.

[16] Anjum Hajat, Matthew Allison, Ana V Diez-Roux, Nancy Swords Jenny, Neal W Jorgensen, Adam A Szpiro, Sverre Vedal, and Joel D Kaufman. Long-term exposure to air pollution and markers of inflammation, coagulation, and endothelial activation: a repeat-measures analysis in the multi-ethnic study of atherosclerosis (mesa). *Epidemiology (Cambridge, Mass.)*, 26(3):310, 2015.

[17] Roberto Romero, Jimmy Espinoza, Luís F Gonçalves, Juan Pedro Kusanovic, Lara Friel, and Sonia Hassan. The role of inflammation and infection in preterm birth. In *Seminars in reproductive medicine*, volume 25, pages 021–039. Copyright© 2007 by Thieme Publishers, Inc., 333 Seventh Avenue, New York, NY ..., 2007.

[18] Chaolin Huang, Yeming Wang, Xingwang Li, Lili Ren, Jianping Zhao, Yi Hu, Li Zhang, Guohui Fan, Jiuyang Xu, Xiaoying Gu, et al. Clinical features of patients infected with 2019 novel coronavirus in wuhan, china. *The lancet*, 395(10223):497–506, 2020.

[19] Zhaohai Zheng, Fang Peng, Buyun Xu, Jingjing Zhao, Huahua Liu, Jiahao Peng, Qingsong Li, Chongfu Jiang, Yan Zhou, Shuqing Liu, et al. Risk factors of critical & mortal covid-19 cases: A systematic literature review and meta-analysis. *Journal of infection*, 81(2):e16–e25, 2020.

[20] Feng Gao, Kenneth I Zheng, Xiao-Bo Wang, Qing-Feng Sun, Ke-Hua Pan, Ting-Yao

Wang, Yong-Ping Chen, Giovanni Targher, Christopher D Byrne, Jacob George, et al. Obesity is a risk factor for greater covid-19 severity. *Diabetes care*, 43(7):e72–e74, 2020.

[21] Rebecca J Baer, Monica R McLemore, Nancy Adler, Scott P Oltman, Brittany D Chambers, Miriam Kuppermann, Matthew S Pantell, Elizabeth E Rogers, Kelli K Ryckman, Marina Sirota, et al. Pre-pregnancy or first-trimester risk scoring to identify women at high risk of preterm birth. *European Journal of Obstetrics & Gynecology and Reproductive Biology*, 231:235–240, 2018.

[22] Nancy Dole, David A Savitz, Irva Hertz-Picciotto, Anna Maria Siega-Riz, Michael J McMahon, and Pierre Buekens. Maternal stress and preterm birth. *American journal of epidemiology*, 157(1):14–24, 2003.

[23] Qian Guo, Yuchen Zheng, Jia Shi, Jijun Wang, Guanjun Li, Chunbo Li, John A Fromson, Yong Xu, Xiaohua Liu, Hua Xu, et al. Immediate psychological distress in quarantined patients with covid-19 and its association with peripheral inflammation: a mixed-method study. *Brain, behavior, and immunity*, 88:17–27, 2020.

[24] Eric Lavigne, Abdool S Yasseen III, David M Stieb, Perry Hystad, Aaron Van Donkelaar, Randall V Martin, Jeffrey R Brook, Daniel L Crouse, Richard T Burnett, Hong Chen, et al. Ambient air pollution and adverse birth outcomes: differences by maternal comorbidities. *Environmental research*, 148:457–466, 2016.

[25] Aaron Van Donkelaar, Randall V Martin, Chi Li, and Richard T Burnett. Regional estimates of chemical composition of fine particulate matter using a combined geoscience-statistical method with information from satellites, models, and monitors. *Environmental science & technology*, 53(5):2595–2611, 2019.

[26] Milton Kotelchuck. An evaluation of the kessner adequacy of prenatal care index and a proposed adequacy of prenatal care utilization index. *American journal of public health*, 84(9):1414–1420, 1994.

[27] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2021. URL `https://www.R-project.org/`.

[28] ERS. Rural-urban commuting area codes, 2019. URL `https://www.ers.usda.gov/data-products/rural-urban-commuting-area-codes.aspx#.U9lO7GPDWHo`.

[29] Jiajia Wang, Gail Williams, Yuming Guo, Xiaochuan Pan, and Shilu Tong. Maternal exposure to heatwave and preterm birth in brisbane, australia. *BJOG: an international Journal of Obstetrics & Gynaecology*, 120(13):1631–1641, 2013.

[30] Marshall Burke, Anne Driscoll, Sam Heft-Neal, Jiani Xue, Jennifer Burney, and Michael Wara. The changing risk and burden of wildfire in the united states. *Proceedings of the National Academy of Sciences*, 118(2), 2021.

[31] Yingying Han, Panhua Jiang, Tianyu Dong, Xinliang Ding, Ting Chen, Gro Dehli Villanger, Heidi Aase, Lu Huang, and Yankai Xia. Maternal air pollution exposure and preterm birth in wuxi, china: effect modification by maternal age. *Ecotoxicology and environmental safety*, 157:457–462, 2018.

[32] Kari A Weber, Wei Yang, Frederick Lurmann, S Katharine Hammond, Gary M Shaw, and Amy M Padula. Air pollution, maternal hypertensive disorders, and preterm birth. *Environmental epidemiology (Philadelphia, Pa.)*, 3(5), 2019.

[33] Anne Bozack, Stanley Pierre, Nicholas DeFelice, Elena Colicino, Darby Jack, Steven N Chillrud, Andrew Rundle, Alfred Astua, James W Quinn, Laura McGuinn, et al. Long-term air pollution exposure and covid-19 mortality: A patient-level analysis from new york city. *American journal of respiratory and critical care medicine*, (ja), 2021.

Publishing Agreement

It is the policy of the University to encourage open access and broad distribution of all theses, dissertations, and manuscripts. The Graduate Division will facilitate the distribution of UCSF theses, dissertations, and manuscripts to the UCSF Library for open access and distribution. UCSF will make such theses, dissertations, and manuscripts accessible to the public and will take reasonable steps to preserve these works in perpetuity.

I hereby grant the non-exclusive, perpetual right to The Regents of the University of California to reproduce, publicly display, distribute, preserve, and publish copies of my thesis, dissertation, or manuscript in any form or media, now existing or later derived, including access online for teaching, research, and public service purposes.

DocuSigned by:
Joan Costello
696959F2FED3412...
Author Signature

3/14/2022
Date