Exposure to PM$_{2.5}$ during Pregnancy and Fetal Growth in Eastern Massachusetts, USA

Michael Leung,$^{1,2}$ Marc G. Weisskopf,$^{1,2}$ Francine Laden,$^{1,2,3}$ Brent A. Coull,$^4$ Anna M. Modest,$^{5,6,7}$ Michele R. Hacker,$^{1,5,8}$ Blair J. Wylie,$^{5,6}$ Yauang Wei,$^8$ Joel Schwartz,$^{1,2}$ and Stefania Papatheodorou$^1$

$^1$Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, USA
$^2$Department of Environmental Health, Harvard T.H.Chan School of Public Health, Boston, Massachusetts, USA
$^3$Channing Division of Network Medicine, Department of Medicine, Brigham and Women’s Hospital and Harvard Medical School, Boston, Massachusetts, USA
$^4$Department of Obstetrics, Gynecology, and Reproductive Biology, Beth Israel Deaconess Medical Center, Boston, Massachusetts, USA
$^5$Department of Obstetrics, Gynecology, and Biostatistics, Harvard Medical School, Boston, Massachusetts, USA
$^6$Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, USA
$^7$Department of Obstetrics, Gynecology, and Reproductive Biology, Beth Israel Deaconess Medical Center, Boston, Massachusetts, USA
$^8$Department of Obstetrics, Gynecology, and Reproductive Biology, Beth Israel Deaconess Medical Center, Boston, Massachusetts, USA

BACKGROUND: Prior studies have examined the association between fine particulate matter [PM ≤2.5 μm in aerodynamic diameter (PM$_{2.5}$)] and fetal growth with either limited spatial or temporal resolution.

OBJECTIVES: In this study, we examined the association between PM$_{2.5}$ exposure during pregnancy and fetal growth measures (ultrasound parameters and birth weight) in a pregnancy cohort using spatiotemporally resolved PM$_{2.5}$ in Eastern Massachusetts, USA.

METHODS: We used ultrasound measures of biparietal diameter (BPD), head circumference, femur length, and abdominal circumference (AC), in addition to birth weight, from 9,446 pregnancies that were delivered at the Beth Israel Deaconess Medical Center from 2011–2016. We used linear mixed-effects models to examine the associations of PM$_{2.5}$ in two exposure windows (the first 16 wk of pregnancy and the cumulative exposure up until the assessment of fetal growth) with anatomic scans (ultrasound measures at <24 wk), growth scans (ultrasound measures at ≥24 wk), and birth weight. All models were adjusted for sociodemographic characteristics, long-term trends, and temperature.

RESULTS: Higher PM$_{2.5}$ exposure in the first 16 wk was associated with smaller fetal growth measures, where associations were particularly strong for BPD, AC, and birth weight. For example, a 5–μg/m$^3$ increase in PM$_{2.5}$ was associated with a lower mean BPD$z$-score of −0.19 (95% CI: −0.31, −0.06) before 24 wk, a lower mean AC$z$-score of −0.15 (95% CI: −0.28, −0.01) after 24 wk, and a lower mean birth weight$z$-score of −0.11 (95% CI: −0.20, −0.01). Analyses examining the association with cumulative PM$_{2.5}$ exposure up until the assessment of fetal growth produced attenuated associations.

CONCLUSIONS: Higher gestational exposure to PM$_{2.5}$ was associated with smaller fetal growth measures at levels below the current national standards. https://doi.org/10.1289/EHP9824

Introduction

Exposure to fine particulate matter [PM ≤2.5 μm in aerodynamic diameter (PM$_{2.5}$)] is a considerable threat to health worldwide (Burnett et al. 2018; Cohen et al. 2017; GBD 2015 Risk Factors Collaborators 2016). Pregnant individuals and their fetuses are likely vulnerable to the effects of PM$_{2.5}$ owing to changes in maternal physiology and the rapid speed of fetal organ formation and development (Wilcox 2011). A recent systematic review showed compelling evidence for an association between maternal exposure to PM$_{2.5}$ and impaired fetal growth as characterized by birth weight (Bekkar et al. 2020), a key developmental indicator for perinatal morbidity and mortality, as well as for later life cardiometabolic outcomes (Barker 2004). Several biological mechanisms have been proposed to explain these associations, including inflammation, oxidative stress, endocrine disruption, coagulation changes, and placental dysfunction (Glinianaia et al. 2004; Liu et al. 2016; Slama et al. 2008; Veras et al. 2008).

Because the sequence of events during fetal development is very specific (e.g., cell differentiation, organ development, changes in fetal metabolism), the timing of exposure to PM$_{2.5}$ during pregnancy is likely to manifest in distinct effects on fetal growth parameters. However, the use of newborn anthropometry does not allow for the identification of these developmental windows. Routine ultrasound measurements would make these internal (and therefore hidden) processes observable. However, only a few studies have used fetal ultrasound parameters to examine the timing of when the growth-restricting effects of PM$_{2.5}$ exposure manifest (Cao et al. 2019; Clemens et al. 2017; Lin et al. 2020). Although all three studies found that increased prenatal PM$_{2.5}$ exposure was associated with reduced fetal growth from mid-gestation onward, they suffer from some limitations with regard to exposure assessment: a) limited spatial resolution, in that two of the three used the nearest land-based monitors for exposure assessment (Cao et al. 2019; Zhao et al. 2018); and b) assessment of long-term cumulative PM$_{2.5}$ exposure without considering other potentially relevant developmental windows [two of the three studies assessed PM$_{2.5}$ concentrations from conception to the date of ultrasound (Cao et al. 2019; Zhao et al. 2018), whereas the third assessed annual mean PM$_{2.5}$ concentrations (Clemens et al. 2017)]. Therefore, we aimed to overcome these limitations by examining the association between PM$_{2.5}$ in two exposure windows (the first 16 wk of pregnancy and the cumulative exposure up until the assessment of fetal growth) and fetal growth measures in a pregnancy cohort with routine ultrasound and spatiotemporally resolved PM$_{2.5}$ in Eastern Massachusetts, USA.

Methods

Study Population

We used prenatal and obstetric data from the Beth Israel Deaconess Medical Center (BIDMC), one of the large private tertiary-care hospitals in Eastern Massachusetts. In the present study, we included all pregnancies delivered by the practices in
which all obstetric ultrasounds were performed through the BIDMC. We restricted the analyses to live births at ≥20 wk of gestation from 2011 through 2016, which is the period for which both ultrasound and PM$_{2.5}$ data were available. Of these, we excluded individuals with multifetal gestations because of different growth trajectories and residential addresses outside of Massachusetts. Full addresses were available for each delivery and were geocoded to latitude and longitude using the Google Maps Application Programming Interface. This study was approved by the institutional review boards of the Harvard T.H. Chan School of Public Health and the Beth Israel Deaconess Medical Center. The data were previously collected from medical and administrative records; thus, informed consent was not required.

**PM$_{2.5}$ Exposure**

We assigned PM$_{2.5}$ exposure based on where the pregnant individual resided at birth in Massachusetts from a state-of-the-art spatiotemporal model that estimates daily PM$_{2.5}$ concentration for each 1 × 1 km grid across the continental United States (Di et al. 2019). Briefly, the model uses an ensemble of three machine learning algorithms (artificial neural network, random forest, and gradient boosting) that incorporates satellite-based aerosol optical depth, simulation outputs from three chemical transport models, land-use predictors, and meteorological predictors to estimate daily concentrations of PM$_{2.5}$ at each grid. The predictive model is calibrated using data from 1,928 monitoring stations that belong to the U.S. Environmental Protection Agency (EPA) Air Quality System plus additional monitoring from the National Park Service’s Interagency Monitoring of Protected Visual Environments (IMPROVE) network, the Southeastern Aerosol Research and Characterization Study (SEARCH) network in the Southeastern United States, and the Multiple Air Toxics Exposure Study (MATES) III and IV networks in California. Tenfold cross-validation revealed good model performance with a total $R^2$ of 0.86 for the entire United States (Di et al. 2019). These predicted PM$_{2.5}$ data have been used in previous work that have examined associations with birth outcomes in Massachusetts (Fong et al. 2018, 2019; Qiu et al. 2020).

We considered two overlapping exposure windows during which daily PM$_{2.5}$ levels were averaged: a) the first 16 wk of pregnancy, because this is the period where organ formation takes place and most major functional defects occur in the fetal anatomy (Moore and Persaud 1993); and b) the cumulative PM$_{2.5}$ exposure from the date of conception (calculated by subtracting gestational age at delivery from the date of birth) up until the assessment of fetal growth (i.e., from conception to the ultrasound for fetal ultrasound parameters, and from conception to birth for birth weight). For example, a pregnancy with ultrasound scans at weeks 18, 28, and gestational age at delivery 37 wk has three fetal growth measures, each with two exposure values assigned. That is, ultrasound parameters from the 18–wk scan would be assigned average PM$_{2.5}$ in the first 16 wk and average PM$_{2.5}$ in the first 18 wk; ultrasound parameters from the 28–wk scan would be assigned average PM$_{2.5}$ in the first 16 wk and average PM$_{2.5}$ in the first 28 wk; and birth weight would be assigned average PM$_{2.5}$ in the first 16 wk and average PM$_{2.5}$ during the entire 37–wk pregnancy.

We did not consider splitting exposures by trimesters because that has no biological basis; that is, gestation is a continuum, and the concept of trimesters implies three biologically distinct periods of development (e.g., it assumes a constant effect for one trimester, and a different constant effect for another trimester). Furthermore, the definition of trimesters differs across organizations (NICHD 2017; ACOG 2020), which could make interpreting the results from trimester-specific analyses challenging.

**Fetal Ultrasounds and Birth Weight**

We used repeated ultrasound measurements of biparietal diameter (BPD), head circumference (HC), femur length (FL) and abdominal circumference (AC), all of which were reported in millimeters. Ultrasounds were interpreted by maternal–fetal medicine specialists or radiologists. Gestational age at the time of ultrasound was based on the best obstetric estimate, combining information from the last menstrual period and the earliest ultrasound performed in pregnancy (ACOG 2017).

Ultrasounds were classified a priori into two groups based on the timing of the scan because those conducted later in pregnancy are more likely to include diagnostic scans. The standard of care is one scan at ∼18–20 wk to evaluate fetal anatomy, but that can sometimes be delayed to 21–23 wk if the participant is late for prenatal care, whereas scans conducted later in pregnancy are typically ordered for pregnancies at risk for impaired fetal growth or other complications. Thus, we henceforth refer to scans before 24 wk as “anatomic scans,” and those at 24 wk and beyond as “growth scans”.

We excluded individuals with only one growth scan but no corresponding anatomic scan (n = 732) because this subgroup likely did prenatal care elsewhere and transferred to the BIDMC late in their pregnancy. We also excluded measurements that were considered implausible, defined as values below or above 4 standard deviations (SDs) from the mean of the cohort at a given gestational age (736 measurements for BPD, 839 measurements for HC, 827 measurements for FL, and 922 measurements for AC were considered implausible and were excluded). That is, we excluded implausible measures, but not individuals with implausible values, and so an individual with one plausible and one implausible measurement would remain in the study. Furthermore, to enable comparisons across gestational weeks, we generated age-specific z-scores for each of the fetal growth measures by applying the International Fetal and Newborn Growth Consortium for the 21st Century (INTERGROWTH-21st) standards for fetal growth, which are globally validated standards based on healthy women with no risk factors for intrauterine growth restriction (Papageorghiou et al. 2014). Because these standards are available only up to 40 wk of gestation, ultrasound scans conducted after 40 wk were excluded. Finally, we also abstracted birth weight from the medical records, which was reported in grams, and generated age- and sex-specific z-scores using the INTERGROWTH-21st standards for newborn size (Papageorghiou et al. 2014).

**Covariates**

Individual-level covariates were abstracted from the medical records and included maternal age (continuous), education (college or higher; lower than college or not specified for those who declined or were unavailable to answer), insurance type (private or public/uninsured; given that <1% were uninsured, we combined them with individuals who had public insurance), parity (nulliparous or parous), and fetal sex (male or female). We also had information on race, which was self-reported at registration for care. Responses were categorized into five groups for the analyses: White, Black, Asian, Hispanic, and Other, where individuals who identified as more than one race or did not specify their race were placed in the “Other” category (original self-reported categories are reported in Table S1). We also adjusted for ambient temperature because it has been previously recommended that variables that are predictors of the exposure but are also confounders of the exposure–outcome association should be
controlled for in the analyses (Cefalu and Dominici 2014). Temperature data were obtained from the Land Data Assimilation Systems at the National Aeronautics and Space Administration Goddard Earth Sciences Data and Information Services Center with a 12×12 km spatial resolution (National Aeronautics and Space Administration; https://ldas.gsfc.nasa.gov/data). For each delivery, we assigned the 12-km grid to the address that the pregnant individual reported to have resided at the time of birth. Finally, for area-level socioeconomic status (SES), we used the national percentile rankings of the Area Deprivation Index (ADI) (Kind and Buckingham 2018), which is a neighborhood disadvantage metric derived from 17 census variables on income, education, employment, and housing quality from the American Community Survey (ACS). The ADI from the 2009–2013 ACS was linked to each pregnancy at the census block group level.

**Statistical Analysis**

We fitted generalized additive mixed models (GAMMs) to examine the association between PM$_{2.5}$ in the two exposure windows (the first 16 wk of pregnancy and the cumulative exposure to fetal growth assessment) with anatomic scans (<24 wk), growth scans (≥24 wk), and birth weight. We used a penalized spline for PM$_{2.5}$ to allow for a nonlinear exposure–response function in each window (Wood 2006). To adjust for confounding by long-term and seasonal trends, we included the date of conception and modeled it using a natural spline with 24 degrees of freedom (df) for the entire 6-y study period (as opposed to a year indicator and a natural spline with 4 df for each year) to capture long-term trends more flexibly (i.e., we allowed seasonal trends to vary by year). All covariates previously described were also included in the model, where linear and quadratic terms were used for continuous variables. Furthermore, we included a random intercept for each pregnancy because ultrasounds within each pregnancy were likely to be correlated. Our GAMMs showed that the associations were approximately linear (Figures S1–S4), thus we present estimates from the linear mixed-effects models. We also fitted models stratified by the number of ultrasound measures (1 and ≥2 ultrasounds) because there may be systematic differences in individuals; for example, those with more ultrasound measures may be at a higher risk of pregnancy complications.

We conducted three sensitivity analyses to examine the robustness of our findings. First, we adjusted for self-reported smoking during pregnancy, which was abstracted from the medical records and was categorized as either smoker or nonsmoker, given that prenatal smoking has been shown to interfere with fetal growth at various stages of development (Lampl et al. 2003). Second, we restricted our analyses of birth weight to full-term births (≥37 wk) to disentangle the associations with growth restriction from those with prematurity. Third, we included in our models a fixed effect for the pregnant individual to account for potential correlations among those who gave birth multiple times during the study period.

We also conducted several subgroup analyses to assess for potential effect modification. Prior analyses have identified several subgroups in which the association between PM$_{2.5}$ and fetal growth appeared stronger, such as male fetuses (Jedrychowski et al. 2009) and those of Black women and individuals with lower SES (e.g., less education, lower median household income) (Bekkar et al. 2020). Male fetuses may be more susceptible to the effects of PM$_{2.5}$ for several possible reasons, including their increased vulnerability to placental inflammation because their immune systems take longer to develop (Perni et al. 2005). People of color and those with low education are likely to live in disadvantaged communities, which not only have higher levels of pollution (i.e., pollution sources tend to be located near these communities) but which may also have more neighborhood-level stressors (e.g., lack of access to health care, poorer job opportunities), which can contribute to chronic stress and increased individual susceptibility to environmental exposures (McGuinn et al. 2019; Morello-Frosch and Shenassa 2006). Thus, for each type of fetal growth measure (anatomic scan, growth scan, and birth weight), we assessed whether the association was separately modified by fetal sex, maternal race (the “Other” group was excluded from these analyses), maternal education (the “Not specified” group was excluded from these analyses), and ADI. To do so, we examined the $p$-value for the product term between PM$_{2.5}$ and the modifier. For each potential modifier, we tested whether the association differed from that of the reference level, which was “female” for fetal sex, “White” for maternal race, “lower than college” for maternal education, and the first quartile for ADI. All analyses were performed in R (version 3.6.1; R Development Core Team).

### Table 1. Maternal and fetal characteristics of deliveries at the Beth Israel Deaconess Medical Center, Boston, Massachusetts, in 2011–2016 ($n=9,446$).

| Characteristics | n (%) | Mean ± SD, or median (25th–75th percentile) |
|-----------------|-------|------------------------------------------|
| Age (y)         |       | Mean ± SD 31.4 ± 5.4 | Median (25th–75th percentile) 31.7 (28–35) |
| Education       |       | College or higher 4,608 (49) | Lower than college 3,193 (34) | Not specified 1,645 (17) |
| Race            |       | White 4,916 (52) | Black 1,626 (17) | Asian 899 (10) |
| Hispanic        |       | Hispanic 882 (9) | Other 1,123 (12) |
| Parity          |       | Nulliparous 4,655 (49) | Parous 4,793 (51) |
| Insurance       |       | Private 7,436 (79) | Public or uninsured 2,010 (21) |
| Smoking         |       | Smoker 226 (2) | Nonsmoker 9,220 (98) |
| Child sex       |       | Male 4,791 (51) | Female 4,655 (49) |
| Gestational age at delivery |       | 38.9 ± 2.0 | Median (25th–75th percentile) 39.3 (38.4–40.1) |
| Term (≥37 wk)   |       | 805 (9) | 8,641 (91) |
| ADI (percentile)|       | Mean ± SD 22 ± 40 | Median (25th–75th percentile) 17 (9–28) |
| Number of ultrasounds |       | Quartile 1 (1, 9) 2,611 (28) | Quartile 2 (9, 17) 2,209 (23) |
|                 |       | Quartile 3 (17, 28) 2,343 (25) | Quartile 4 (28, 100) 2,283 (24) |
|                 |       | Mean ± SD 2.8 ± 1.6 | Median (25th–75th percentile) 2 (2–4) |
|                 |       | 1 ultrasound 2,106 (22) | ≥2 ultrasounds 7,340 (78) |

Note: $n=9,446$ refers to the number of deliveries, and not the number of unique individuals (there were 8,241 unique individuals); the ADI is a composite measure of neighborhood socioeconomic disadvantage derived from 17 census block group variables on income, employment, housing from the American Community Survey, where higher values indicate neighborhoods with more disadvantage. ADI, Area Deprivation Index; SD, standard deviation.
Results

Maternal and fetal characteristics for the 9,446 pregnancies included in our study are shown in Table 1, where this sample of eligible participants was formed as in Figure S5. Participants were, on average, 31 years of age at the time of conception and were majority White (52%), with the racial composition of our sample closely resembling that of Suffolk County, which represents the Boston Metropolitan Area and is where the BIDMC is located (Table S2). Most individuals had private insurance (79%), about half had a college education or higher (49%), and few reported to have smoked during pregnancy (2%). Furthermore, the mean and median ADI was 22 and 17, respectively, suggesting that participants in this cohort lived in more advantaged neighborhoods compared with the rest of the United States, given that an ADI of 50 is the median ADI nationwide.

Distributions of PM$_{2.5}$ and temperature are presented in Table 2 and Figure S6. In our sample, the mean and median PM$_{2.5}$ exposure in the first 16 wk (7.41 and 7.40 μg/m$^3$), from conception to the time of ultrasound (7.41 and 7.43 μg/m$^3$), and during the entire pregnancy (7.32 and 7.35 μg/m$^3$) were all below the current annual ambient standard of 12 μg/m$^3$ for PM$_{2.5}$ (U.S. EPA 2021). Exposure to PM$_{2.5}$ across the overlapping exposure windows was highly correlated, and correlations were strongest between PM$_{2.5}$ in the first 16 wk and PM$_{2.5}$ during the entire pregnancy (Table S3). Furthermore, the mean and median temperature in the first 16 wk, from conception to the time of ultrasound, and during the entire pregnancy was about 10°C and remained stable over the years of the study period (Table 2; Figure S6).

Most participants (78%) had two or more ultrasound measurements during pregnancy (Table 1; Figure S7), and summary statistics for the fetal ultrasound parameters are displayed in Table 3. The mean and median z-scores for all ultrasound parameters were similar for anatomic and growth scans, where fetuses had, on average, smaller BPD, but larger HC, FL, and AC measures during pregnancy compared with the international standards (Papageorghiou et al. 2014). At delivery, the mean and median birth weight z-score was 0.33 and 0.32, respectively, which suggests that newborns in our sample were slightly heavier than the international norm.

The associations between PM$_{2.5}$ exposure in the first 16 wk of pregnancy and fetal growth measures (ultrasound parameters and birth weight) are shown in Table 4. For both anatomic and growth scans, we observed that increased PM$_{2.5}$ was associated with smaller ultrasound parameters, and associations were strongest for BPD and AC. For example, a 5-μg/m$^3$ increase in PM$_{2.5}$ was associated with a lower mean z-score among anatomic scans (<24 wk) of −0.19 (95% confidence interval [CI]: −0.31, −0.06 for BPD), −0.10 (95% CI: −0.22, 0.02) for HC, −0.08 (95% CI: −0.21, 0.05) for FL, and −0.15 (95% CI: −0.28, −0.01) for AC and a lower mean z-score among growth scans (≥24 wk) of −0.15 (95% CI: −0.29, −0.02) for BPD, −0.06 (95% CI: −0.19, 0.08) for HC, −0.06 (95% CI: −0.19, 0.08) for FL, and −0.16 (95% CI: −0.31, −0.02) for AC. Furthermore, PM$_{2.5}$ exposure in the first 16 wk was also associated with lower birth weight, where a 5-μg/m$^3$ increase was associated with a lower mean z-score of −0.11 (95% CI: −0.20, −0.01). Analyses with cumulative PM$_{2.5}$ exposure produced attenuated associations for the ultrasound parameters, with most 95% CIs including the null. However, the association with birth weight was similar to our primary analysis (Table 4). Our stratified analyses based on the number of ultrasound (1 and ≥2 ultrasounds) showed that the associations were stronger for those with two or more ultrasounds, whereas associations were attenuated for those with a single scan (Table S4). Furthermore, the results of our sensitivity analyses produced findings similar to our primary analyses. The inclusion of smoking status in our models yielded similar estimates of the associations (Table S5), as did analyses restricted to full-term births (Table S6) and those that included a fixed effect for the pregnant individual (Table S7).

Results from our effect modification analyses can be found in Tables S8–S15. We did not find any evidence for effect modification by fetal sex (Tables S8–S9). However, we found that maternal race, education, and ADI potentially modified the association during specific windows. For maternal race, we found that higher PM$_{2.5}$ exposure during the first 16 wk produced stronger negative associations with all ultrasound parameters at ≥24 wk for Hispanic women compared with White women (Table S10). On the other hand, the negative associations were not observed after 24 wk (i.e., among growth scans and birth weight) in Asian women (Table S10). We also found that the negative associations with birth weight were not observed for Black individuals when using cumulative exposures (Table S11). For maternal education, we found that the negative associations appeared to be stronger

Table 2. Summary statistics for PM$_{2.5}$ and temperature data among deliveries at the Beth Israel Deaconess Medical Center, Boston, Massachusetts, in 2011–2016 (n = 9,446).

| Exposure window          | Mean  | SD   | Median | 25th percentile | 75th percentile | IQR   | Minimum | Maximum |
|--------------------------|-------|------|--------|-----------------|-----------------|-------|---------|---------|
| PM$_{2.5}$ (μg/m$^3$)    |       |      |        |                 |                 |       |         |         |
| First 16 weeks           | 7.41  | 1.49 | 7.40   | 6.39            | 8.43            | 2.04  | 2.57    | 13.76   |
| Conception to ultrasound | 7.41  | 1.38 | 7.43   | 6.45            | 8.34            | 1.90  | 2.63    | 14.68   |
| Entire pregnancy         | 7.52  | 1.29 | 7.35   | 6.38            | 8.22            | 1.84  | 3.05    | 12.34   |
| Temperature (°C)         |       |      |        |                 |                 |       |         |         |
| First 16 weeks           | 9.85  | 7.44 | 9.28   | 3.10            | 17.08           | 13.98 | −4.88   | 21.63   |
| Conception to ultrasound | 10.50 | 3.21 | 10.35  | 8.41            | 12.42           | 4.01  | −3.42   | 21.39   |
| Entire pregnancy         | 10.48 | 2.92 | 10.49  | 8.08            | 13.09           | 5.01  | 0.20    | 19.87   |

Note: IQR, interquartile range; SD, standard deviation.

Table 3. Summary statistics for fetal ultrasound parameters and birth weight from deliveries at the Beth Israel Deaconess Medical Center, Boston, Massachusetts, in 2011–2016 (n = 9,446).

| Growth measure       | n       | Mean ± SD | Median (25th–75th percentile) |
|----------------------|---------|-----------|-------------------------------|
| Anatomic scans       |         |           |                               |
| BPD z-score (at <24 wk) | 10,859  | −0.68 ± 1.10 | −0.67 (−1.39–0.05)           |
| HC z-score           | 10,813  | 0.38 ± 1.09  | 0.39 (−0.32–1.07)             |
| FL z-score           | 10,787  | 0.83 ± 1.14  | 0.81 (0.11–1.51)              |
| AC z-score           | 10,755  | 0.52 ± 1.17  | 0.49 (−0.24–1.25)             |
| Growth scans (at ≥24 wk) | 15,677  | −0.85 ± 1.13 | −0.83 (−1.57 to −0.10)        |
| BPD z-score          | 15,606  | 0.34 ± 1.15  | 0.32 (−0.39–1.09)             |
| HC z-score           | 15,629  | 1.11 ± 1.10  | 1.15 (0.45–1.80)              |
| FL z-score           | 15,608  | 0.57 ± 1.10  | 0.53 (−0.14–1.25)             |
| Birth weight z-score | 9,428   | 0.33 ± 0.99  | 0.32 (−0.33–0.98)             |

Note: AC, abdominal circumference; BPD, biparietal diameter; FL, femur length; HC, head circumference; SD, standard deviation.
Table 4. Linear mixed effects model estimates and 95% CIs for the association between PM$_{2.5}$ in two exposure windows (the first 16 weeks of pregnancy and the cumulative exposure to fetal growth assessment) and fetal growth measures (ultrasound parameters and birth weight) from deliveries at the Beth Israel Deaconess Medical Center, Boston, Massachusetts, in 2011–2016 ($n=9,446$).

| Growth measure          | First 16 weeks Estimate (95% CI) | p-value | Cumulative Estimate (95% CI) | p-value |
|-------------------------|----------------------------------|---------|-------------------------------|---------|
| Anatomic scans (at <24 wk) |                                  |         |                               |         |
| BPD z-score             | $-0.19 (-0.31, -0.06)$            | 0.003   | $-0.15 (-0.26, -0.04)$        | 0.009   |
| HC z-score              | $-0.10 (-0.22, 0.02)$             | 0.11    | $-0.09 (-0.20, 0.02)$         | 0.12    |
| FL z-score              | $-0.08 (-0.21, 0.05)$             | 0.22    | $-0.05 (-0.17, 0.07)$         | 0.42    |
| AC z-score              | $-0.15 (-0.28, -0.01)$            | 0.03    | $-0.12 (-0.24, -0.00)$        | 0.04    |
| Growth scans (at ≥24 wk) |                                  |         |                               |         |
| BPD z-score             | $-0.15 (-0.29, -0.02)$            | 0.03    | $-0.08 (-0.22, 0.06)$         | 0.27    |
| HC z-score              | $-0.06 (-0.19, 0.08)$             | 0.40    | $0.02 (-0.12, 0.17)$          | 0.74    |
| FL z-score              | $-0.06 (-0.19, 0.08)$             | 0.42    | $-0.04 (-0.18, 0.10)$         | 0.57    |
| AC z-score              | $-0.16 (-0.31, -0.02)$            | 0.03    | $-0.06 (-0.21, 0.09)$         | 0.44    |
| Birth weight            | $-0.11 (-0.20, -0.01)$            | 0.04    | $-0.11 (-0.23, 0.00)$         | 0.05    |

Note: Estimates represent the difference in mean z-score with a 5-µg/m$^3$ increase in PM$_{2.5}$ after adjusting for maternal age, race, education, insurance type, parity, fetal sex, date of conception, temperature, and Area Deprivation Index. AC, abdominal circumference; BPD, biparietal diameter; CI, confidence interval; FL, femur length; HC, head circumference.

Discussion

This large retrospective cohort study showed that higher PM$_{2.5}$ exposure was associated with smaller ultrasound parameters, as well as lower birth weight at levels below the current annual ambient standard of 12 µg/m$^3$ for PM$_{2.5}$ (U.S. EPA 2021). Associations were particularly strong for BPD, AC, and birth weight; for example, a 5-µg/m$^3$ increase in PM$_{2.5}$ was associated with a lower mean BPD z-score of $-0.19$ (95% CI: $-0.31$, $-0.06$) before 24 wk. These findings have implications for later health and childhood development. For example, head size, as measured by BPD and HC, has been associated with brain development and cognitive achievement in childhood (Jaekel et al. 2019; Kirkegaard et al. 2020; Villar et al. 2021); height, as approximated in utero by FL, has been associated with educational attainment and economic productivity in adulthood (Magnusson et al. 2006; McGovern et al. 2017); abdominal circumference, which is an indicator for the size of the fetal liver and the amount of subcutaneous fat deposition, has been associated with later cardiometabolic conditions (Rückinger et al. 2010); and birth weight, a summary measure of in utero growth, is a key indicator for later life morbidity and mortality (Barker 2004).

Our findings are consistent with prior studies that have found that increased prenatal PM$_{2.5}$ exposure was associated with reduced fetal ultrasound parameters, even though our estimates are not directly comparable (Clemens et al. 2017; Lin et al. 2020). One study conducted in Beijing did not assess the same fetal growth parameters, and they found that a 10-µg/m$^3$ increase in PM$_{2.5}$ from conception to ultrasound was associated with a lower mean z-score of 0.3 in estimated fetal weight; however, they did not examine associations with either HC, FL, or AC, all of which were used to compute estimated fetal weight (Hadlock et al. 1985). Furthermore, another study conducted in Shanghai used raw ultrasound measurements and did not standardize their parameters by gestational age, but still aggregated all ultrasound measurements together as a single outcome (Cao et al. 2019). They found that a 10-µg/m$^3$ increase in PM$_{2.5}$ from conception to ultrasound was associated with smaller BPD, FL, and AC by about 5.5 mm each. Finally, the last study, conducted in Scotland, internally standardized their fetal ultrasound parameters and thus their z-scores are not directly comparable to ours, but they found that PM$_{2.5}$ exposure was associated with smaller BPD measures (Clemens et al. 2017). Despite the heterogeneity in analytic treatments, the current body of evidence (including our contribution) suggests a robust signal for the association between higher prenatal PM$_{2.5}$ exposure and smaller ultrasound parameters of fetal growth. This is further corroborated by the negative associations we found between PM$_{2.5}$ and birth weight, a finding which is concordant with the findings from 25 of the 29 studies included in a recent systematic review examining this association (Bekkar et al. 2020).

Among studies of fetal growth and PM exposure [both PM$_{2.5}$ and PM ≤10 µm in aerodynamic diameter (PM$_{10}$)], there is still inconsistency with regard to the critical window of exposure. The three studies focusing on PM$_{2.5}$ examined only long-term exposure, either entire pregnancy average (Clemens et al. 2017) or the cumulative exposure from conception to the date of ultrasound (Cao et al. 2019; Lin et al. 2020), and they did not, or were not able to, consider other exposure windows. Studies examining PM$_{10}$ have also identified that PM$_{10}$ exposure in the first 16 wk of pregnancy (Hansen et al. 2008), and also in the third trimester (Lamichhane et al. 2018; van den Hooven et al. 2012), was associated with smaller ultrasound parameters. Here, we found that the two overlapping exposure windows we considered (the first 16 wk and cumulative exposure) produced negative associations, but those using cumulative PM$_{2.5}$ produced weaker associations for all fetal growth measures (ultrasound parameters and birth weight). This pattern indicates that the effects of PM$_{2.5}$ potentially interfere with fetal tissue development or the developing placenta in early pregnancy, rather than with the period of rapid growth in late pregnancy. Thus, it is important that future studies also assess exposures early in pregnancy in addition to cumulative exposures; otherwise, harmful exposures may be overlooked. That is, if the potential critical window is indeed in early pregnancy, then only assessing cumulative exposures by averaging values in the critical window with other periods later in pregnancy would add measurement error such that the estimate would be biased toward the null.

Our analyses stratified by the number of ultrasounds (1 or ≥2) showed that the associations were stronger for pregnancies with two or more ultrasounds compared with those with a single scan. It is possible that this is because those with two or more ultrasounds included more high-risk pregnancies, among which associations with PM$_{2.5}$ may be stronger owing to increased susceptibility; for example, those with hypertensive disorders may be more susceptible to the effects of PM$_{2.5}$ (Koman et al. 2018). Thus, exposure to PM$_{2.5}$ can not only contribute to a pregnancy...
becoming a high-risk pregnancy—given that past work has shown that it is associated with pregnancy conditions, such as gestational diabetes and preeclampsia (Tang et al. 2020; Yu et al. 2020)—but also, based our findings, may predispose such high-risk individuals to its adverse health effects.

In our effect modification analyses, we found that the associations did not differ by fetal sex. Although this is contrary to previous literature that suggests that stressful exposures during pregnancy may be more harmful to male births, those investigations mostly focused on acute exposures (Bruckner et al. 2010; Fukuda et al. 1998; Jedrychowski et al. 2009; Mocarelli et al. 1996). We, however, found suggestive evidence that race, education, and ADI could modify the association during specific developmental windows. Associations with fetal ultrasound parameters were stronger for individuals who were Hispanic and were null for those who were Asian. These findings, however, should be interpreted with caution given that the 95% CIs were wide due to the few Asian and Hispanic women in our cohort (about 900 deliveries each over the 6-year study period); future studies should examine these racial disparities in other settings with different demographic distributions. Furthermore, we also found stronger associations for individuals who were more educated and lived in areas with the least disadvantage. This finding, which was conditional on the other covariates, such as maternal age, was not expected. One possible explanation is that although these individuals are presumably, on average, healthier (as there may be fewer neighborhood-level stressors), they may be less resilient to the effects of PM$_{2.5}$. Potential effect heterogeneity among fetal ultrasound measurements were also observed mostly in birthweight; however, these two sets of measurements are not directly comparable for several reasons. First, birth weight poorly reflects fetal growth during early pregnancy; that is, it is hard to distinguish a healthy newborn from one that experiences fetal growth restriction early in pregnancy but later catches up to achieve population growth standards by birth (Bloomfield et al. 2006). Furthermore, birth weight is a summary measure of in utero growth (e.g., fetal weight is computed on the basis of HC, FL, and AC using Hadlock’s formula (Hadlock et al. 1985)), and it perhaps may be more sensitive to symmetric growth restriction, which is not the case in the present study given that the associations of PM$_{2.5}$ exposure with different organs were distinct (i.e., the reductions in the size of the structures were not equal).

This study has several strengths. We assessed PM$_{2.5}$ exposure at home addresses at birth with high spatial and temporal resolution using a state-of-the-art model, where two exposure windows were considered. We used repeated ultrasound measurements to longitudinally assess growth trajectories, which enabled the assessment and identification of critical windows of development. We controlled for an extensive panel of confounders, the most important being SES, for which we had both individual and neighborhood-level variables. Furthermore, the characteristics of our sample closely resembled that of the Boston Metropolitan Area, and so our results are generalizable to the population residing in the study area. Finally, we conducted several sensitivity analyses to assess the robustness of our findings. We found that the additional adjustment for smoking status during pregnancy did not change our results, likely because too few women smoked to substantially alter our estimates and because smoking status is unlikely to be related to ambient PM$_{2.5}$ exposure conditional on individual- and area-level SES (Weisskopf and Webster 2017). In addition, the sensitivity analysis restricted to full-term births produced associations that were similar to the primary analysis, where these results showed that higher PM$_{2.5}$ exposure was still associated with smaller ultrasound parameters even if the pregnancy did not end in a premature birth. Finally, models with a fixed effect for the pregnant individual produced almost identical estimates, likely because correlations within pregnant individuals were low and only a small proportion of individuals gave birth multiple times during the study period.

Some limitations also should be acknowledged. Our PM$_{2.5}$ exposure assessment was based on modeled outdoor values at the home address, which may not reflect personal PM$_{2.5}$ exposure. This measurement error will bias the estimate toward the null, but this could be offset by better control for confounding because more proxy measures of exposure are less vulnerable to individual-level confounding (Weisskopf and Webster 2017). For example, we did not have information on several risk factors for fetal growth restriction, such as maternal height, prepregnancy weight, passive smoking, or household income, but similar to the results of the sensitivity analysis controlling for smoking status, these variables are unlikely to confound our associations. Furthermore, we used maternal residence at delivery to assign exposures, which could result in further misclassification owing to residential mobility during pregnancy. Yet, past simulations in other settings have shown that ignoring residential mobility had only a minor impact on point estimates of association and that the identification of critical windows was robust to this type of exposure misclassification (Warren et al. 2018).

We also ignored the time–activity pattern of participants during pregnancy, where participants are less likely to spend time at home during early pregnancy (Nethery et al. 2009). However, residential mobility and time–activity patterns during pregnancy were likely nondifferential conditional on the covariates included, with the magnitude of the bias being contingent on the degree of misclassification in our sample. Furthermore, our analyses were restricted to live-born children and so our estimates may be biased upward (Goin et al. 2021; Leung et al. 2021; Liew et al. 2015; Raz et al. 2018). That is, if pregnancy loss is driven by both the exposure and other unmeasured factors that also affect the outcome, then the PM$_{2.5}$–growth associations estimated in the subpopulation of “healthier” live births are likely biased. These considerations suggest that perhaps the associations are even stronger than they are shown here, which further supports the need to reduce PM$_{2.5}$ exposure during pregnancy. With our data structure, we were unable to examine other critical windows using a distributed lag model framework (Gasparini 2016; Schwartz 2000; Zanobetti et al. 2000); the timing of the ultrasounds varied across pregnancies, and so individuals would need to contribute a different number of lags, where the lags would not correspond to the same gestational week. Finally, this study was retrospective, and so analyses of growth scans should be interpreted with caution because high-risk pregnancies could be overrepresented. That is, those who had scans at ≥24 wk were likely coming in for diagnostic purposes, whereas low-risk pregnancies with normal fetal growth were less likely to have growth scans.

In conclusion, we show that gestational exposure to PM$_{2.5}$ was associated with fetal ultrasound parameters and birth weight at exposure levels below the current national standards (U.S. EPA 2021). These findings add to the growing body of literature documenting the harmful effects of PM$_{2.5}$ not only during pregnancy, but also for overall health. Future work should explore this topic further in other settings and different populations, at exposure levels that are higher than used in this study, while also investigating critical windows of exposure during pregnancy. Follow-up studies are warranted to examine the clinical relevance of our observed associations. In terms of policy implications, our findings suggest the need to focus efforts on reducing exposures even at “safe” concentrations.

**Acknowledgments**

This work was supported through a Pilot Project grant from the Harvard Chan National Institute of Environmental Health Sciences Center for Environmental Health (National Institutes of Health grant P30 ES000002).
