Published in final edited form as:
Environ Int. 2021 March ; 148: 106378. doi:10.1016/j.envint.2021.106378.

Application of the Navigation Guide Systematic Review Methodology to Evaluate Prenatal Exposure to Particulate Matter Air Pollution and Infant Birth Weight

Inyang Uwak[a], **Natalie Olson**[b], **Angelica Fuentes**[b], **Megan Moriarty**[a], **Jairus Pulczinski**[c], **Juleen Lam**[d], **Xiaohui Xu**[e], **Brandie D. Taylor**[f], **Samuel Taiwo**[a], **Kirsten Koehler**[c], **Margaret Foster**[g], **Weihsueh A. Chiu**[b], **Natalie M. Johnson**[a]

[a]Department of Environmental and Occupational Health. Texas A&M University, College Station, TX, USA.

[b]Department of Veterinary Integrative Biosciences. Texas A&M University, College Station, TX, USA.

[c]Department of Environmental Health and Engineering. Johns Hopkins University, Baltimore, MD, USA.

[d]Department of Health Sciences, California State University, East Bay, Hayward, CA USA.

[e]Department of Epidemiology and Biostatistics. Texas A&M University, College Station, TX, USA.

[f]Department of Epidemiology and Biostatistics. Temple University, Philadelphia, PA, USA.

[g]Medical Sciences Library. Texas A&M University, College Station, TX, USA.

Abstract

Address correspondence to N.M. Johnson, Department of Environmental and Occupational Health, Texas A&M University, 1266 TAMU, College Station, TX 77843 USA. Tel: 979-436-9325. nmjohnson@tamu.edu.

CRediT authorship contribution statement
Inyang Uwak: Investigation, Formal analysis, Validation, Writing - Original Draft
Natalie Olson: Investigation, Formal analysis, Validation, Writing - Review & Editing
Angelica Fuentes: Investigation, Validation
Megan Moriarty: Investigation, Validation
Jairus Pulczinski: Investigation, Validation
Juleen Lam: Methodology, Validation, Writing - Review & Editing
Xiaohui Xu: Validation
Brandie D. Taylor: Validation, Writing - Review & Editing
Samuel Taiwo: Validation
Kirsten Koehler: Validation, Writing - Review & Editing
Margaret Foster: Methodology
Weihsueh A. Chiu: Conceptualization, Methodology, Formal analysis, Validation, Project administration, Supervision, Writing - Review & Editing
Natalie M. Johnson: Conceptualization, Validation, Project administration, Supervision, Writing - Review & Editing

Declaration of interests
The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Low birth weight is an important risk factor for many co-morbidities both in early life as well as in adulthood. Numerous studies report associations between prenatal exposure to particulate matter (PM) air pollution and low birth weight. Previous systematic reviews and meta-analyses report varying effect sizes and significant heterogeneity between studies, but did not systematically evaluate the quality of individual studies or the overall body of evidence. We conducted a new systematic review to determine how prenatal exposure to $PM_{2.5}$, PM_{10}, and coarse PM ($PM_{2.5–10}$) by trimester and across pregnancy affects infant birth weight. Using the Navigation Guide methodology, we developed and applied a systematic review protocol [CRD42017058805] that included a comprehensive search of the epidemiological literature, risk of bias (ROB) determination, meta-analysis, and evidence evaluation, all using pre-established criteria. In total, 53 studies met our inclusion criteria, which included evaluation of birth weight as a continuous variable. For $PM_{2.5}$ and PM_{10}, we restricted meta-analyses to studies determined overall as "low" or "probably low" ROB; none of the studies evaluating coarse PM were rated as "low" or "probably low" risk of bias, so all studies were used. For $PM_{2.5}$, we observed that for every 10 μg/m^3 increase in exposure to $PM_{2.5}$ in the 2nd or 3rd trimester, respectively, there was an associated 5.69g decrease (I^2: 68%, 95% CI: −10.58, −0.79) or 10.67g decrease in birth weight (I^2: 84%, 95% CI: −20.91, −0.43). Over the entire pregnancy, for every 10 μg/m^3 increase in $PM_{2.5}$ exposure, there was an associated 27.55g decrease in birth weight (I^2: 94%, 95% CI: −48.45, −6.65). However, the quality of evidence for $PM_{2.5}$ was rated as "low" due to imprecision and/or unexplained heterogeneity among different studies. For PM_{10}, we observed that for every 10 μg/m^3 increase in exposure in the 3rd trimester or the entire pregnancy, there was a 6.57g decrease (I^2: 0%, 95% CI: −10.66, −2.48) or 8.65g decrease in birth weight (I^2: 84%, 95% CI: −16.83, −0.48), respectively. The quality of evidence for PM_{10} was rated as "moderate," as heterogeneity was either absent or could be explained. The quality of evidence for coarse PM was rated as very low/low (for risk of bias and imprecision). Overall, while evidence for $PM_{2.5}$ and course PM was inadequate primarily due to heterogeneity and risk of bias, respectively, our results support the existence of an inverse association between prenatal PM_{10} exposure and low birth weight.

Keywords

air pollution; particulate matter; prenatal; low birth weight; Navigation Guide; systematic review; risk of bias; meta-analysis

1. Introduction

Prenatal exposure to particulate matter (PM) air pollution has been linked with adverse birth outcomes, namely infant low birth weight (LBW) (Dadvand et al. 2013). A number of studies have investigated the association between prenatal PM exposure and infant LBW, which is attributable to the regular collection and large-scale availability of birth weight data through birth records (Blencowe et al. 2019). Additionally, LBW has been associated with an increased risk of certain long-term health outcomes later in life (Belbasis et al. 2016). LBW is defined as infants born weighing less than 2,500 grams (Cutland et al. 2017). A primary cause of LBW is preterm birth (PTB), delivery of a live born infant <37 weeks of gestation (Cutland et al. 2017). Thus, mean birthweights across gestational age ranges has been used as a proxy for fetal growth. Specifically, small for gestational age (SGA) is defined as infants that fall within the smallest 10th percentile of infants of the same gestational age (Cutland et al. 2017). A subset of LBW and SGA infants require intensive neonatal care for immediate health issues and may have chronic health outcomes later in life (Belbasis et al. 2016).

In spite of the substantial evidence on the association between developmental PM exposure to PM and outcomes, including LBW, PTB, and SGA births, there have been inconsistencies in the conclusions on the magnitude of the effect (Lamichhane et al. 2015; Stieb et al. 2012). We applied the Navigation Guide systematic review methodology to assess the quality and strength of evidence on the effect of prenatal PM exposure on infant birth weight. The Navigation Guide was developed in 2011 to strengthen approaches for assessing evidence in environmental health sciences (Woodruff et al. 2011). The Navigation Guide is a systematic and transparent approach that draws from best practices in the clinical arena while accounting for differences in evidence and decision context involved in environmental health risk assessments, such as the reliance on human observational studies versus randomized controlled trials (Cumpston et al. 2019; Guyatt et al. 2008; Woodruff and Sutton 2014). To date, the Navigation Guide methodology has been applied in numerous reviews of environmental exposures, including the human evidence for effects of airborne pollutants on the diagnosis of autism spectrum disorder (Lam et al. 2016) and both the human and non-human evidence for effects of Perfluorooctanoic acids (PFOAs) on fetal growth (Johnson et al. 2014; Koustas et al. 2014; Lam et al. 2014). The results of these studies and others demonstrate the utility of this approach in applying rigor and transparency in support of evidence-based decisions to environmental health problems.

In this review, we evaluated the human evidence regarding prenatal PM exposure and infant birth weight. We assessed each study for the risk of bias and conducted a meta-analysis on a subset of studies to estimate the overall magnitude of effect. We focused on birth weight as a continuous outcome variable to determine the impact of bias on effect size estimates. Consideration as a continuous variable allows for the assessment of the effect on population distributions. Case studies have illustrated the importance of considering a continuous scale to provide added information about the exposure-disease continuum, inform population variability, and increase the predictive power of risk assessment (Woodruff et al. 2008).

2. Methods

2.1. Systematic Review Methodology

While systematic review methods have been used for decades in the clinical sciences, specific techniques for conducting a systematic review directly applicable to the decision context and evidence streams in environmental health have only recently been developed and utilized in the field of environmental health sciences (Rooney et al. 2014; Woodruff and Sutton 2014). We conducted our review using the Navigation Guide approach, which is based on the Cochrane Collaboration and Grading of Recommendations Assessment Development and Evaluation (Guyatt et al. 2008). We developed a protocol before initiating the study and registered it in PROSPERO [CRD42017058805].

2.2. Study question

Our ultimate objective was to evaluate whether ambient air pollution is "toxic" to the developing fetus in the sense of reducing birth weight with increasing exposure, since lower birth weight is a risk factor for both short- and long-term morbidities (Belbasis et al. 2016). The "Population," "Exposure," "Comparator," and "Outcome" (PECO) statement, is briefly outlined below with additional specifics available in our protocol. Population: Pregnant women. Exposure: Gestational exposure to ambient particulate air pollution. "Particulate air pollution" is defined as outdoor sources of inhaled airborne matter classified as $PM_{2.5}$ (mass concentration of particles with diameter smaller than 2.5 μm), PM_{10} (mass concentration of particles with diameter smaller than 10 μm), or $PM_{2.5–10}$ (mass concentration of particles with diameters 2.5–10 μm), excluding active and passive smoking. Comparator: Pregnant women exposed to lower levels of PM than the more highly exposed humans. Outcome: Birth weight measured as a continuous variable.

2.3. Data Sources

We searched the databases Ovid Medline, Embase, and Global Health on November 23, 2015, using the search terms developed in collaboration with librarian (MF), shown in the Supplemental Materials, Table S1. Our search was not limited by publication date. We limited our search to English language and used the Medical Subject Headings (MeSH) database to compile synonyms for ambient particulate air pollution and birth weight (details in our protocol). We updated the search on February 27, 2020, to identify any new studies, applying the same strategies used in the original search. We also supplemented these results by hand-searching references of all included studies.

2.4. Study Selection

We included original studies that evaluated ambient particulate air exposure and reported associations with birth weight. Three reviewers (MM, JP, IU) independently screened titles and abstracts of each reference in RefWorks to determine eligibility. In the event of a discrepancy between reviewers, the default was to move the article forward for full-text screening. We excluded studies if: 1) the article did not report birth weight outcomes; 2) the article did not report ambient particulate air pollution exposure; 3) the article contained no original data; 4) the article did not involve human subjects; or 5) other reason, with an explanation required. All duplicate articles were removed. At the full-text screening stage, the same reviewers (MM, JP, IU) independently screened references in RefWorks for inclusion using the same criteria as above. Additionally, at this stage, studies were excluded if the article did not report birth weight as a continuous variable. Studies reporting birth weight as z-scores were excluded.

2.5. Data Extraction

Two reviewers (NO, AF) independently extracted data related to study characteristics and outcome measures into the Health Assessment Workspace Collaborative (HAWC) database (Supplemental Materials, Table S2). In the case of missing data, the protocol was to contact study authors; however, all relevant data was able to be extracted from the full text articles. Data extracted by each author was independently reviewed (WC, NMJ, IU) for quality assurance/quality control on all the studies to resolve any discrepancies between the two independent extractors and further ensure accuracy. We extracted all characteristics of the study population, including location and sample size, exposure period duration, pollutant class, methods used to estimate exposures, and all relevant estimates of association relating particulate air pollution exposure with birth weight, specifically recording estimates as related to exposure assessment technique or by spatial scale (i.e., city- or county-level versus <5km radius). For the meta-analysis, we extracted adjusted regression estimates and standard errors or 95% confidence interval limits and standardized to a continuous increment in exposure (i.e., per 10 $\mu g/m^3$ unit increase in pollutant). For instance, if change in birthweight was originally reported in grams per 1 ug/m^3 exposure, the effect and confidence interval limits were multiplied by 10. Some studies reported the change in birthweight per IQR increase of exposure. For these the values were standardized by multiplying the (change in birthweight per IQR) by (10 ug/m^3 divided by the value of the IQR). For studies where a 95% confidence interval was not reported one was calculated from available p-values or standard errors assuming a normal distribution. For articles reporting multiple models adjusting for different sets of covariates, we selected estimates from the fully-adjusted model including the most confounders.

2.6. Assessing the risk of bias

We evaluated the risk of bias for each of the studies across the following domains: recruitment strategy, blinding, confounding, exposure assessment, incomplete outcome data, selective outcome reporting, conflicts of interest, or other problems that could put the study at risk of bias (Table 1). Ratings for each domain were "low," "probably low," "probably high," or "high" risk of bias, with customized instructions for each domain based on the type of evidence anticipated (Supplemental Materials, Table S3). For example, we determined for a study to be rated "low" risk of bias in the confounding domain, all five pre-determined potential confounders were accounted for. These included socioeconomic status, race/ethnicity, maternal tobacco use, maternal age, and season of conception/birth. Likewise, to determine if exposure assessment measurements were robust, reviewers took into consideration the validity and reliability of the monitoring or modeling methods employed. Review authors with subject-matter expertise from our team (NMJ, JL, XX, BT, MM, IU, ST, WC) independently determined the risk of bias across all domains. An additional QA/QC author was matched with each study to solve any discrepancies between ratings. An overall risk of bias rating was assigned as "low," "probably low," "probably high," or "high" risk of bias by evaluating the individual domain ratings. If any of the ratings were "high" or "probably high," the overall rating was automatically rated as "high" or "probably high," respectively. If the majority of domains were rated as "low" or "probably low," the overall rating was determined to be "low" or "probably low," respectively.

2.7. Meta-analysis

Details of the meta-analysis approach are in the study protocol. In brief, the analysis separately considered each of the three pollutant classes of $PM_{2.5}$, PM_{10}, and $PM_{2.5–10}$ and the four exposure windows, first, second, and third trimester, as well as entire pregnancy. The primary analysis utilized study results for the entire population in each study, using the exposure metric at the smallest spatial scale, analyzed using single pollutant models, adjusted for covariates. Additionally, due to the sufficient number of studies, the primary analyses for $PM_{2.5}$ and PM_{10} utilized only studies with "low" or "probably low" risk of bias. Studies were pooled using random-effects models with the Knapp-Hartung modification (Knapp and Hartung 2003). This approach accounts for uncertainty in the estimate of τ^2 in the standard error estimates, generally resulting in wider confidence intervals. Heterogeneity was evaluated using the I^2 metric. Sources of heterogeneity explored using subgrouping included the following: ethnicity (non-Hispanic White only, Hispanic only, Black only), geographic locale (Americas, Europe, Asia), spatial scale of exposure assessment, and risk of bias rating. Additionally, influence analysis was conducted by removing individual studies one at a time.

2.8. Rating the quality of evidence across studies

We rated the quality of the overall body of evidence as "high," "moderate," "low," or "very low." An initial rating of "moderate" quality was assigned based on the previously described rationale for rating human evidence according to the Navigation Guide approach (Johnson et al. 2014). We considered "downgrades" to the quality rating based on five categories of considerations: risk of bias, indirectness, inconsistency, imprecision, and potential for publication bias (Table 3). We considered "upgrades" to the quality rating due to a large magnitude of effect, dose-response, and whether residual confounding would minimize the overall effect estimate (Balshem et al. 2011). Possible downgrades or upgrades were: 0 (no change from initial quality rating), −1 (1 level downgrade) or − 2 (2 level downgrade), +1 (1 level upgrade) or +2 (2 level upgrade). Review authors evaluated the quality of the evidence according to our protocol (Supplemental Materials, Table S4) and then compared ratings as a group to reach the final decision.

2.9. Rating the strength of the evidence across studies

We assigned an overall strength of evidence rating based on a combination of 4 considerations, outlined in Table 3 and detailed in Supplemental Materials, Table S4: (1) Quality of body of evidence (i.e., the rating from the previous step), (2) Direction of effect, (3) Confidence in effect (likelihood that a new study could change our conclusion), and (4) Other compelling attributes of the data that may influence certainty. Possible ratings were "sufficient evidence," "limited evidence," "inadequate evidence," or "evidence of lack of toxicity."

3. Results

3.1. Included studies

Figure 1A depicts the screening of eligible articles: the original November 2015 search retrieved 532 unique records, of which 103 were screened at the full-text review stage. Of these, 32 met our pre-defined criteria for inclusion. Figure 1B illustrates the February 2020 search which retrieved 223 additional studies, of which 50 were screened at full-text review stage and 21 studies met our pre-defined criteria for inclusion into the final analysis, totaling 53 articles. A summary of the characteristics of these studies is detailed in Table 2. The included studies were largely cohort studies, with 44 using similar retrospective methods to investigate the relationship between air pollution and birth weight. Nine of the studies used prospective methods, enrolling pregnant mothers and collecting information to determine air pollution exposure during pregnancy. Studies varied in the pollutant measured, type of exposure assessment method, and exposure window (i.e., entire pregnancy or trimester specific) reported. Overall, 20 studies measured $PM_{2.5}$ exposure alone, 17 studies measured PM_{10} exposure alone, and only 1 study measured $PM_{2.5-10}$ alone. Several studies measured pollutants in combination, either all three (3 studies) or two of the three pollutant classes (12 studies). Exposure assessment methods included ambient monitoring as the primary technique (30 studies), followed by modeling (20 studies), a combination of monitoring and modeling (2 studies) or in one case personal modeling for a 48h duration in the second trimester of pregnancy. In general, studies reported effect estimates for trimester-specific and entire pregnancy exposure windows (28 studies). In some cases, only estimates were reported for the entire pregnancy and not by trimester (15 studies) or just by trimester and not entire pregnancy (10 studies). Study locations ranged globally, and geographic location was taken into consideration in the meta-analysis.

3.2. Risk of bias for individual studies

Risk of bias designations generally were rated as “low” or “probably low” for most domains (Figure 2). Individual study determinations are summarized in Figure S11 and individual study ratings are also available in HAWC (https://hawcproject.org/assessment/227/) and Figure S12. In a few cases, recruitment across study groups were determined to be “high” risk. For instance, Pedersen et al. 2013 investigated low birth weight in a large European cohort study, wherein study participants were recruited from different populations in varying proportions. Confounding was predominantly rated as “probably low” (58% of studies). In some cases, studies were rated as “high” or “probably high” risk in addressing confounding. In these cases, investigators only accounted for two or fewer of the pre-determined important potential confounders, which could have introduced bias into analyses. In a few cases reviewers determined a “probably high” risk of bias in the “other” category, defined as if the study appeared to be free of other problems that could put it at a risk of bias. For instance, regarding Mannes et al. 2005, reviewers determined a risk of residual confounding and over adjustment bias in the linear regression model, as authors adjusted for an intermediate on the pathway between exposure and outcome. In addition, authors also did not account for extreme values in birthweight for gestational age. In general, the domain with a considerable number of studies rated as “probably high” (43%) was related to the robustness of exposure assessment. This was mainly due to reliance on county-level monitoring data without adequate temporal coverage or spatial resolution. Overall, for $PM_{2.5,}$ 12 studies (out of a total of 30 studies measuring $PM_{2.5}$) were rated overall as “low” or “probably low” risk of bias. For PM_{10}, 10 studies (out of a total of 29 studies measuring PM_{10}) were rated overall as “low” or “probably low” risk of bias and used for subsequent meta-analysis. For studies on coarse PM, none of the 5 studies were given an overall rating of “low” or “probably low.” This was largely the result of risk of exposure misclassification based on county-level measurements employed in most these studies (Darrow et al. 2011; Ebisu et al. 2016; Morello-Frosch et al. 2010; Parker and Woodruff 2008). Complete descriptions of risk of bias evaluations and their justifications are provided online in the HAWC workspace (https://hawcproject.org/assessment/227/).

3.3. Meta-analysis

We conducted a primary meta-analysis on studies rated as “low” or “probably low” risk of bias for exposures to $PM_{2.5}$ and PM_{10}. This included 18 total studies for $PM_{2.5}$ and 10 total studies for PM_{10}. For $PM_{2.5–10}$, there were a limited number of studies overall that measured this pollutant class, and none were rated as “low” or “probably low.” Thus, we used the existing 5 studies rated as “high” or “probably high” in our primary meta-analysis. A summary of the meta-analysis results using a random effects model is shown in Table 4, separated by pollutant class and exposure window (trimester or entire pregnancy). For $PM_{2.5}$, the overall random effects estimates ranged from 5.69g to 27.55g decrease in birth weight per 10 μg/m^3 increase in $PM_{2.5}$ (Supplemental Figure S1A–D). The meta-estimate for the 1st trimester was not statistically significant, but those for the other exposure windows were. Substantial heterogeneity was evident in each exposure window (I^2 ranged from 68% to 94%). In each exposure window, at least one study reported a positive relationship (increase in birth weight with increasing PM). Subgrouping based on ethnicity, spatial scale, or geographic location did not explain the observed heterogeneity (Supplemental Figures S6A, S7A–D, S8A–D); the only statistically significant subgroup differences were by geographic location for entire pregnancy (Supplemental Table S5). Including “high” and “probably high” risk of bias studies further increased heterogeneity (Supplemental Figure S4A–D), though subgroup differences by risk of bias were not in of themselves statistically significant (Table S5). Influence analysis showed that for the second trimester, heterogeneity is explained by a single study (Hyder et al. 2014) with a large effect size (Supplemental Figure S9B). Omitting this study reduced I^2 from 68% to 40% and reduced the meta-estimate from −5.69g (−10.58, −0.79) to −3.81g (−7.88, 0.25). For other exposure windows, heterogeneity could not be attributed to any single study (Supplemental Figure S9A, C–D). No evidence of publication bias (all p-values > 0.05) was found as assessed using funnel plots and tests for asymmetry (Begg and Mazumdar 1994; Egger et al. 1997; Sterne et al. 2011) (Supplemental Figures S10A–D).

For PM_{10}, the overall random effects estimates ranged from a 3.22g increase to an 8.65g decrease in birth weight per 10 μg/m^3 increase in PM_{10} (Supplemental Figure S2A–D). The meta-estimates for the 1st and 2nd trimesters were not statistically significant (effect estimate 3.22g, 95% CI: −3.13, 9.58 and −3.37g, 95% CI: −8.22, 1.48, respectively), but estimates for the other exposure windows were statistically significant. Low heterogeneity was seen in the trimester-based exposure windows (I^2 0–14%). However, substantial heterogeneity was evident for the entire pregnancy (I^2 84%). Subgrouping based on ethnicity was not possible due to too few studies, and subgrouping by spatial scale, or geographic location did not explain the observed heterogeneity (Supplemental Figures S7E–H, S8E–H); the only statistically significant subgroup differences were by geographic location for first trimester and entire pregnancy (Supplemental Table S5). Including “high” and “probably high” risk of bias studies increased heterogeneity in all cases (Supplemental Figure S5A–D), and subgroup differences by risk of bias were statistically significant for first and third trimesters (Table S5). Influence analysis showed that for the entire pregnancy, heterogeneity was explained largely by a single study (Geer et al. 2012) that reported a positive association, whereas all the other studies consistently showed an inverse association (Supplemental Figure S9H). Omitting this study reduced the I^2 from 84% to 0%, and changed the meta-estimate from −8.65g (−16.83, −0.48) to −11.22g (−13.17, −9.26). For the other exposure windows, similar results in terms of both heterogeneity and meta-estimates were obtained under influence analyses (Supplemental Figures S9E–G). No evidence of publication bias (all p-values > 0.05) was found as assessed using funnel plots and tests for asymmetry (Begg and Mazumdar 1994; Egger et al. 1997; Sterne et al. 2011) (Supplemental Figures S10E–H).

A smaller number of studies examined “coarse” PM ($PM_{2.5-10}$). None of these studies were rated as having “low” or “probably low” risk of bias, as discussed previously. Thus, when including all studies, overall random effects estimates ranged from a 2.70g to 8.81g decrease in birth weight per 10 μg/m^3 increase in $PM_{2.5-10}$ (Supplemental Figure S3A–D). The meta-estimates for the 2nd and 3rd trimesters were not statistically significant, −2.90g (−10.04, 4.23) and −4.93g (−10.82, 0.96) respectively, and each of these had high heterogeneity (I^2 70–76%). Due to the small number of studies, subgrouping based on ethnicity, spatial scale, or geographic location were not possible did not explain this observed heterogeneity (Supplemental Figures S6E–F, S7I–L, S8I–L). Heterogeneity was reduced to 55% for the 2nd trimester when omitting the most influential study (Parker and Woodruff 2008), though this left only two studies remaining with a pooled estimate that remained statistically non-significant (Supplemental Figure S9J). Similarly, in the 3rd trimester, omitting the most influential study ((Ebisu et al. 2016)) reduced heterogeneity to 64%, but the pooled estimate remained statistically non-significant (Supplemental Figure S9K). For the 1st trimester and the entire pregnancy, the meta-estimates were statistically significant, −2.70g (−3.90, −1.49) and −8.81g (−10.32, −7.31) respectively, with no observed heterogeneity in both cases (I^2 0%). For the 1st trimester, omitting any one study lead to meta-estimates that were either statistically non-significant or that were only barely significant (p=0.0498) (Supplemental Figure S9I). For the entire pregnancy, meta-estimates remained statistically significant under influence analyses, with no heterogeneity (Supplemental Figure S9L). Insufficient studies were available to examine publication bias.

3.4. Quality of the body of evidence

In all cases, the initial rating for the quality of evidence was “moderate” based on Navigation Guide methods (Johnson et al. 2014). Using the factors for rating the quality of evidence (Table 3), we determined the following evaluations (Supplemental Table S6, Table 4). For $PM_{2.5}$ exposure in the first trimester, a downgrade of 2 levels was supported, based on “imprecision” due to the lack of a statistically significant meta-estimate, as well as a wide confidence interval indicating potential impact of random error. Moreover, a downgrade for “inconsistency” was due to the substantial heterogeneity that could not be explained. The resulting quality of evidence rating was “very low.” For $PM_{2.5}$ exposure in the second trimester, a downgrade of 1 level was supported based on “imprecision”. Heterogeneity was explained by a single study, and omitting this study lead to an effect estimate no longer statistically significant. The resulting quality of evidence rating was “low.” Last, for $PM_{2.5}$ exposure in the third trimester, as well as exposure throughout entire pregnancy, a downgrade of 1 level was supported, based on “inconsistency” due to the substantial heterogeneity that could not be explained. Thus, the resulting quality of evidence rating was “low.”

For PM_{10} exposure during the first trimester, a downgrade of 2 levels was supported based on "imprecision" due to a wide confidence interval and the lack of a statistically significant meta-estimate with low heterogeneity. The resulting quality of evidence rating was "low." For PM_{10} exposure during the second trimester, a downgrade of 1 level was supported based on "imprecision" due to the lack of a statistically significant meta-estimate with low heterogeneity. The resulting quality of evidence rating was "low." For PM_{10} exposure during the third trimester, no change in the quality of evidence was indicated, as the meta-estimate was statistically significant with low heterogeneity. The resulting quality of evidence rating was "moderate." Last, for PM_{10} exposure during the entire pregnancy, no change in the quality of evidence was indicated. Heterogeneity was explained by a single study and omitting that study lead to a precise, statistically significant meta-estimate. The resulting quality of evidence rating was "moderate." Meta-analysis results with "moderate" quality of evidence ratings are displayed in Figure 3.

For exposure to coarse PM ($PM_{2.5–10}$) during the first trimester, a downgrade of 2 levels was supported based on "risk of bias" (all studies were rated "high" or "probably high"), "imprecision" due to few studies (n=3), and a high degree of influence of any one study had on statistical significance. The resulting quality of evidence rating was "very low." For $PM_{2.5–10}$ exposure during the second and third trimesters, downgrades of 3 levels are supported based on "risk of bias" (all studies were rated "high" or "probably high"), "imprecision" due to the lack of a statistically significant meta-estimate, and "inconsistency" due to high, unexplained heterogeneity. The resulting quality of evidence rating was "very low." Last, for $PM_{2.5–10}$ exposure throughout the entire pregnancy, a downgrade of 1 level was supported based on "risk of bias" (all studies were rated "high" or "probably high"). The resulting quality of evidence rating was "low."

3.5. Strength of the body of evidence

Using the considerations for rating the strength of evidence in Table 3, the following evaluations were made (Table 5). For $PM_{2.5}$, there is "inadequate evidence" for all exposure windows due to "low" or "very low" quality of evidence, based on either imprecision of the estimate or high and unexplained heterogeneity (none of the other considerations were influential in this evaluation). For PM_{10}, there is "limited evidence" that increasing exposure during the third trimester or during the entire pregnancy will lead to a reduction in birth weight. The quality of evidence for these exposure windows was rated as "moderate." Although the direction of the effect estimate was in the "adverse" direction, confidence in the effect estimate is limited because chance, bias, and confounding cannot be ruled out with reasonable confidence, and additional data could alter this conclusion. No other compelling attributes of the data exist that would influence this evaluation. For other exposure windows, evidence for PM_{10} is "inadequate" due to "low" or "very low" quality of evidence, based on either imprecision of the estimate and/or the presence of a relationship in the opposite (non-adverse) direction (none of the other considerations were influential in this evaluation). For $PM_{2.5–10}$, there is "inadequate evidence" that increasing exposure is during any exposure window leads to a reduction in birth weight. The available evidence is insufficient to assess the effects of exposure, mainly due to high risk of bias in individual studies and the reliance on a small set of often heterogeneous studies. None of the other considerations from Table 3 were influential to this evaluation.

4. Discussion

Numerous case-control and cohort studies demonstrate an association between prenatal exposure to ambient air pollution and reduced fetal growth or infant birthweight. An early systematic review found an association between $PM_{2.5}$ exposure and LBW and SGA births, as well as PM_{10} exposure and SGA (Shah et al. 2011). Despite these observed associations, there have been inconsistencies in the conclusions about the association and magnitude of the effect. Initial systematic reviews based on a relatively small number of studies (n=4), were not able to draw conclusions on effect size (Bonzini et al. 2010; Bosetti et al. 2010; Ghosh et al. 2007). More recent systematic reviews, which performed a meta-analysis on a larger number of studies (>30) showed that pooled estimates of effect size for LBW for a 10 $\mu g/m^3$ increase in $PM_{2.5}$ exposure during entire pregnancy ranged from −15.9g (−26.8, −5.0) (Sun et al. 2016) to −22.17g (−37.93, −6.41) (Lamichhane et al. 2015). Steib et al. also reported estimates per 10 $\mu g/m^3$ increase in $PM_{2.5}$ exposure to be −23.4g (−45.5, −1.4) (Stieb et al. 2012), all of which are consistent with our pooled estimate of −27.55g (−48.45, −6.65) per 10 $\mu g/m^3$. This agreement is likely due to several of the same studies used across these meta-analyses. For PM_{10}, Lamichhane et al. reported estimates for a 10 $\mu g/m^3$ increase at −10.31g (−13.57 to −3.13 g), whereas Stieb et al. published estimates for a 20 $\mu g/m^3$ increase at −16.8g (−20.2 to −13.3) (Lamichhane et al. 2015; Stieb et al. 2012), both of which are also consistent with our pooled estimate of −8.65g (−16.83, −0.48) per 10 $\mu g/m^3$. These previous investigators cited that they were not able to rule out the consequences of specific biases that may be as a result of differences in study methodology, study design, population demographics, exposure period, characterization of confounding and data collection.

In our analysis, there was substantial heterogeneity across the different pollutant classes. Also, the spatial scale employed, large scale (at the city or county level or >/= 10km) in comparison to medium scale (census tract, zip code, postal code, nearest monitor, <10km and >/=5km) or small scale (<5km) led to greater heterogeneity. These findings underscore the complexity of estimating exposure across gestation. While one study (Jedrychowski et al. 2009) employed personal monitoring during pregnancy, the cost of adequate temporal coverage is great since it is infeasible for participants to carry monitors over time. Despite the significant heterogeneity, we still observed a decrease in birthweight for every 10 $\mu g/m^3$ increase in $PM_{2.5}$ across all trimesters (except the 1^{st}) and entire pregnancy, as well as for every10 $\mu g/m^3$ increase in PM_{10} across the third trimester and entire pregnancy. The “inadequate” evidence rating for $PM_{2.5}$ reflects the quality, which received downgrades for inconsistency, driven mainly by heterogeneity. Similar conclusions were drawn by Lam et al. for the association between early-life exposure to air pollution as a whole and diagnosis of autism spectrum disorder (Lam et al. 2016).

Some limitations that may be associated with our study include the reliance on expert evaluation in the process used for the risk of bias, quality and strength ratings. However, this limitation was overcome by creating a diverse team of experts from relevant fields to participate in this process. Moreover, by publishing a pre-specified protocol and employing two independent reviewers for each study, our analysis includes a degree of transparency and robustness that is absent when using less structured approaches. Additionally, the rating of the quality of evidence across studies was dependent on the available data. For instance, PM_{10} and $PM_{2.5}$ are typically reported separately, but also likely occur in combination. Thus, models that consider multi-pollutant exposures may better represent gestational PM exposure. Furthermore, most studies fail to consider secondary/co-exposures like ultrafine particulate matter, gas phase pollutants, or heat, which can also affect birth weight. A recent systematic review including cohort and cross-sectional studies in U.S. populations demonstrated a significant association of air pollutant and heat exposure with adverse birth outcomes, such as preterm birth and low birth weight (Bekkar et al. 2020). There is also the potential for additional unmeasured confounding. For instance, (Wilson et al. 2017) noted that associations between infant health and with air pollution during individual trimesters may be biased unless all trimesters are included in the same model to fully address confounding and seasonal trends. Less than a quarter of the studies we identified addressed this issue, though subgrouping analyses revealed no statistically significant differences between studies that treated trimesters separately versus together in a single model. Recent studies also include measures of more temporality refined exposure windows, for instance, monthly or weekly averages. These studies may yield important insight into the critical windows of exposure ((Arroyo et al. 2019; Liu et al. 2019; Yuan et al. 2020). However, our analyses did not include enough studies to evaluate weekly exposure.

A major strength of our study is the transparency and thoroughness of the Navigation Guide systematic review process, which incorporates the GRADE system for assessing the quality of synthesized human evidence in environmental health research in the absence of randomized clinical trials (Woodruff et al. 2014). Overall, our results support the vast evidence that prenatal PM exposure is associated with reduced infant birth weight. These implications on infant mortality burden were included for the first time in the State of Global Air report, which highlighted air pollution accounts for 20% of newborn deaths worldwide, mostly related to complications of low birth weight and preterm birth (Health Effects Institute). Thus, public health interventions to address infant birth weight suppression from PM may have a substantial impact on infant health, especially those at high risk for exposure. Future research and implementation strategies are recommended to help optimize interventions and policies to mitigate infant health effects.

5. Conclusions

Overall, we conclude that the existing evidence supports an association between prenatal exposure to ambient particulate matter air pollution and a decrease in birth weight, particularly for PM_{10}. However, our findings reveal the need to standardize and improve exposure assessment methods in air pollution research because the various forms of exposure measurement utilized in the studies contributed to the heterogeneity seen in the meta-analysis. Furthermore, some of the unexplained heterogeneity found in our study may be resolved with additional studies which could also strengthen the evidence.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. N. Olson and A. Fuentes received funding from the Texas A&M University Tier One Program, which supported their work on this project. I. Uwak and K. Koehler received support from the U.S. Department of Transportation funded University Transportation Center (UTC), the Center for Advancing Research in Transportation Emissions, Energy, and Health (CARTEEH). N. Johnson received support from the National Institute of Environmental Sciences (NIEHS) (ES028866). N. Olson, N. Johnson and W. Chiu were also supported, in part, by the NIEHS (P30 ES029067).

References

Arroyo V, Diaz J, Salvador P, Linares C. 2019 Impact of air pollution on low birth weight in spain: An approach to a national level study. Environ Res 171:69–79. [PubMed: 30660920]

Balshem H, Helfand M, Schunemann HJ, Oxman AD, Kunz R, Brozek J, et al. 2011 Grade guidelines: 3. Rating the quality of evidence. J Clin Epidemiol 64:401–406. [PubMed: 21208779]

Basu R, Harris M, Sie L, Malig B, Broadwin R, Green R. 2014 Effects of fine particulate matter and its constituents on low birth weight among full-term infants in california. Environ Res 128:42–51. [PubMed: 24359709]

Begg CB, Mazumdar M. 1994 Operating characteristics of a rank correlation test for publication bias. Biometrics 50:1088–1101. [PubMed: 7786990]

Bekkar B, Pacheco S, Basu R, DeNicola N. 2020 Association of air pollution and heat exposure with preterm birth, low birth weight, and stillbirth in the us: A systematic review. JAMA Netw Open 3:e208243. [PubMed: 32556259]

Beland L-P, Oloomi S. 2019 Environmental disaster, pollution and infant health: Evidence from the deepwater horizon oil spill Journal of Environmental Economics and Management 98:102265.

Belbasis L, Savvidou MD, Kanu C, Evangelou E, Tzoulaki I. 2016 Birth weight in relation to health and disease in later life: An umbrella review of systematic reviews and meta-analyses. BMC Med 14:147. [PubMed: 27677312]

Bell ML, Ebisu K, Belanger K. 2007 Ambient air pollution and low birth weight in connecticut and massachusetts. Environ Health Perspect 115:1118–1124. [PubMed: 17637932]

Bell ML, Belanger K, Ebisu K, Gent JF, Lee HJ, Koutrakis P, et al. 2010 Prenatal exposure to fine particulate matter and birth weight: Variations by particulate constituents and sources. Epidemiology 21:884–891. [PubMed: 20811286]

Bijnens EM, Derom C, Gielen M, Winckelmans E, Fierens F, Vlietinck R, et al. 2016 Small for gestational age and exposure to particulate air pollution in the early-life environment of twins. Environ Res 148:39–45. [PubMed: 27003124]

Blencowe H, Krasevec J, de Onis M, Black RE, An X, Stevens GA, et al. 2019 National, regional, and worldwide estimates of low birthweight in 2015, with trends from 2000: A systematic analysis. Lancet Glob Health 7:e849–e860. [PubMed: 31103470]

Bonzini M, Carugno M, Grillo P, Mensi C, Bertazzi PA, Pesatori AC. 2010 Impact of ambient air pollution on birth outcomes: Systematic review of the current evidences. Med Lav 101:341–363. [PubMed: 21105590]

Bosetti C, Nieuwenhuijsen MJ, Gallus S, Cipriani S, La Vecchia C, Parazzini F. 2010 Ambient particulate matter and preterm birth or birth weight: A review of the literature. Arch Toxicol 84:447–460. [PubMed: 20140425]

Cumpston M, Li T, Page MJ, Chandler J, Welch VA, Higgins JP, et al. 2019 Updated guidance for trusted systematic reviews: A new edition of the cochrane handbook for systematic reviews of interventions. Cochrane Database Syst Rev 10:ED000142. [PubMed: 31643080]

Cutland CL, Lackritz EM, Mallett-Moore T, Bardaji A, Chandrasekaran R, Lahariya C, et al. 2017 Low birth weight: Case definition & guidelines for data collection, analysis, and presentation of maternal immunization safety data. Vaccine 35:6492–6500. [PubMed: 29150054]

Dadvand P, Parker J, Bell ML, Bonzini M, Brauer M, Darrow LA, et al. 2013 Maternal exposure to particulate air pollution and term birth weight: A multi-country evaluation of effect and heterogeneity. Environ Health Perspect 121:267–373. [PubMed: 23384584]

Darrow LA, Klein M, Strickland MJ, Mulholland JA, Tolbert PE. 2011 Ambient air pollution and birth weight in full-term infants in atlanta, 1994–2004. Environ Health Perspect 119:731–737. [PubMed: 21156397]

Ebisu K, Berman JD, Bell ML. 2016 Exposure to coarse particulate matter during gestation and birth weight in the u.S. Environ Int 94:519–524. [PubMed: 27324566]

Egger M, Davey Smith G, Schneider M, Minder C. 1997 Bias in meta-analysis detected by a simple, graphical test. BMJ 315:629–634. [PubMed: 9310563]

Erickson AC, Ostry A, Chan LH, Arbour L. 2016 The reduction of birth weight by fine particulate matter and its modification by maternal and neighbourhood-level factors: A multilevel analysis in british columbia, canada. Environ Health 15:51. [PubMed: 27079512]

Fong KC, Di Q, Kloog I, Laden F, Coull BA, Koutrakis P, et al. 2019 Relative toxicities of major particulate matter constituents on birthweight in massachusetts. Environ Epidemiol 3:e047. [PubMed: 31342007]

Geer LA, Weedon J, Bell ML. 2012 Ambient air pollution and term birth weight in texas from 1998 to 2004. J Air Waste Manag Assoc 62:1285–1295. [PubMed: 23210220]

Ghosh R, Rankin J, Pless-Mulloli T, Glinianaia S. 2007 Does the effect of air pollution on pregnancy outcomes differ by gender? A systematic review. Environ Res 105:400–408. [PubMed: 17493608]

Giovannini N, Schwartz L, Cipriani S, Parazzini F, Baini I, Signorelli V, et al. 2018 Particulate matter (pm10) exposure, birth and fetal-placental weight and umbilical arterial ph: Results from a prospective study. J Matern Fetal Neonatal Med 31:651–655. [PubMed: 28277922]

Gouveia N, Bremner SA, Novaes HM. 2004 Association between ambient air pollution and birth weight in sao paulo, brazil. J Epidemiol Community Health 58:11–17. [PubMed: 14684720]

Gray SC, Edwards SE, Miranda ML. 2010 Assessing exposure metrics for pm and birth weight models. J Expo Sci Environ Epidemiol 20:469–477. [PubMed: 19773814]

Gray SC, Edwards SE, Schultz BD, Miranda ML. 2014 Assessing the impact of race, social factors and air pollution on birth outcomes: A population-based study. Environ Health 13:4. [PubMed: 24476365]

Guo P, Chen Y, Wu H, Zeng J, Zeng Z, Li W, et al. 2020 Ambient air pollution and markers of fetal growth: A retrospective population-based cohort study of 2.57 million term singleton births in china. Environ Int 135:105410. [PubMed: 31884132]

Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. 2008 Grade: An emerging consensus on rating quality of evidence and strength of recommendations. BMJ 336:924–926. [PubMed: 18436948]

Han Y, Ji Y, Kang S, Dong T, Zhou Z, Zhang Y, et al. 2018 Effects of particulate matter exposure during pregnancy on birth weight: A retrospective cohort study in suzhou, china. Sci Total Environ 615:369–374. [PubMed: 28988071]

Hannam K, McNamee R, Baker P, Sibley C, Agius R. 2014 Air pollution exposure and adverse pregnancy outcomes in a large uk birth cohort: Use of a novel spatio-temporal modelling technique. Scand J Work Environ Health 40:518–530. [PubMed: 24647593]

He T, Zhu J, Wang J, Ren X, Cheng G, Liu X, et al. 2018 Ambient air pollution, h19/dmr methylation in cord blood and newborn size: A pilot study in zhengzhou city, china. Chemosphere 212:863–871. [PubMed: 30193235]

Health Effects Institute. Available: www.stateofglobalair.org [accessed 11/09/2020.

Huang C, Nichols C, Liu Y, Zhang Y, Liu X, Gao S, et al. 2015 Ambient air pollution and adverse birth outcomes: A natural experiment study. Popul Health Metr 13:17. [PubMed: 26190943]

Hyder A, Lee HJ, Ebisu K, Koutrakis P, Belanger K, Bell ML. 2014 Pm2.5 exposure and birth outcomes: Use of satellite- and monitor-based data. Epidemiology 25:58–67. [PubMed: 24240652]

Jedrychowski W, Perera F, Mrozek-Budzyn D, Mroz E, Flak E, Spengler JD, et al. 2009 Gender differences in fetal growth of newborns exposed prenatally to airborne fine particulate matter. Environ Res 109:447–456. [PubMed: 19261271]

Johnson PI, Sutton P, Atchley DS, Koustas E, Lam J, Sen S, et al. 2014 The navigation guide - evidence-based medicine meets environmental health: Systematic review of human evidence for pfoa effects on fetal growth. Environ Health Perspect 122:1028–1039. [PubMed: 24968388]

Keller JP, Chang HH, Strickland MJ, Szpiro AA. 2017 Measurement error correction for predicted spatiotemporal air pollution exposures. Epidemiology 28:338–345. [PubMed: 28099267]

Kim OJ, Ha EH, Kim BM, Seo JH, Park HS, Jung WJ, et al. 2007 Pm10 and pregnancy outcomes: A hospital-based cohort study of pregnant women in seoul. J Occup Environ Med 49:1394–1402. [PubMed: 18231086]

Kirwa K, McConnell-Rios R, Manjourides J, Cordero J, Alshawabekeh A, Suh HH. 2019 Low birth weight and pm2.5 in puerto rico. Environ Epidemiol 3.

Knapp G, Hartung J. 2003 Improved tests for a random effects meta-regression with a single covariate. Stat Med 22:2693–2710. [PubMed: 12939780]

Koustas E, Lam J, Sutton P, Johnson PI, Atchley DS, Sen S, et al. 2014 The navigation guide - evidence-based medicine meets environmental health: Systematic review of nonhuman evidence for pfoa effects on fetal growth. Environ Health Perspect 122:1015–1027. [PubMed: 24968374]

Kumar N 2012 Uncertainty in the relationship between criteria pollutants and low birth weight in chicago. Atmos Environ (1994) 49:171–179.

Lam J, Koustas E, Sutton P, Johnson PI, Atchley DS, Sen S, et al. 2014 The navigation guide - evidence-based medicine meets environmental health: Integration of animal and human evidence for pfoa effects on fetal growth. Environ Health Perspect 122:1040–1051. [PubMed: 24968389]

Lam J, Sutton P, Kalkbrenner A, Windham G, Halladay A, Koustas E, et al. 2016 A systematic review and meta-analysis of multiple airborne pollutants and autism spectrum disorder. PLoS One 11:e0161851. [PubMed: 27653281]

Lamichhane DK, Leem JH, Lee JY, Kim HC. 2015 A meta-analysis of exposure to particulate matter and adverse birth outcomes. Environ Health Toxicol 30:e2015011. [PubMed: 26796890]

Lamichhane DK, Ryu J, Leem JH, Ha M, Hong YC, Park H, et al. 2018 Air pollution exposure during pregnancy and ultrasound and birth measures of fetal growth: A prospective cohort study in korea. Sci Total Environ 619–620:834–841.

Laurent O, Wu J, Li L, Chung J, Bartell S. 2013 Investigating the association between birth weight and complementary air pollution metrics: A cohort study. Environ Health 12:18. [PubMed: 23413962]

Lavigne E, Burnett RT, Stieb DM, Evans GJ, Godri Pollitt KJ, Chen H, et al. 2018 Fine particulate air pollution and adverse birth outcomes: Effect modification by regional nonvolatile oxidative potential. Environ Health Perspect 126:077012. [PubMed: 30073952]

Li Z, Yuan X, Fu J, Zhang L, Hong L, Hu L, et al. 2019 Association of ambient air pollutants and birth weight in ningbo, 2015–2017. Environ Pollut 249:629–637. [PubMed: 30933760]

Liu Y, Xu J, Chen D, Sun P, Ma X. 2019 The association between air pollution and preterm birth and low birth weight in guangdong, china. BMC Public Health 19:3. [PubMed: 30606145]

Mannes T, Jalaludin B, Morgan G, Lincoln D, Sheppeard V, Corbett S. 2005 Impact of ambient air pollution on birth weight in sydney, australia. Occup Environ Med 62:524–530. [PubMed: 16046604]

Medeiros A, Gouveia N. 2005 [relationship between low birthweight and air pollution in the city of sao paulo, brazil]. Rev Saude Publica 39:965–972. [PubMed: 16341408]

Merklinger-Gruchala A, Kapiszewska M. 2015 Association between pm10 air pollution and birth weight after full-term pregnancy in krakow city 1995–2009--trimester specificity. Ann Agric Environ Med 22:265–270. [PubMed: 26094521]

Morello-Frosch R, Jesdale BM, Sadd JL, Pastor M. 2010 Ambient air pollution exposure and full-term birth weight in california. Environ Health 9:44. [PubMed: 20667084]

Parker JD, Woodruff TJ, Basu R, Schoendorf KC. 2005 Air pollution and birth weight among term infants in california. Pediatrics 115:121–128. [PubMed: 15629991]

Parker JD, Woodruff TJ. 2008 Influences of study design and location on the relationship between particulate matter air pollution and birthweight. Paediatr Perinat Epidemiol 22:214–227. [PubMed: 18426516]

Pedersen M, Giorgis-Allemand L, Bernard C, Aguilera I, Andersen AM, Ballester F, et al. 2013 Ambient air pollution and low birthweight: A european cohort study (escape). Lancet Respir Med 1:695–704. [PubMed: 24429273]

Rahmalia A, Giorgis-Allemand L, Lepeule J, Philippat C, Galineau J, Hulin A, et al. 2012 Pregnancy exposure to atmospheric pollutants and placental weight: An approach relying on a dispersion model. Environ Int 48:47–55. [PubMed: 22836169]

Rhee J, Fabian MP, Ettinger de Cuba S, Coleman S, Sandel M, Lane KJ, et al. 2019 Effects of maternal homelessness, supplemental nutrition programs, and prenatal pm2.5 on birthweight. Int J Environ Res Public Health 16.

Rooney AA, Boyles AL, Wolfe MS, Bucher JR, Thayer KA. 2014 Systematic review and evidence integration for literature-based environmental health science assessments. Environ Health Perspect 122:711–718. [PubMed: 24755067]

Salam MT, Millstein J, Li YF, Lurmann FW, Margolis HG, Gilliland FD. 2005 Birth outcomes and prenatal exposure to ozone, carbon monoxide, and particulate matter: Results from the children's health study. Environ Health Perspect 113:1638–1644. [PubMed: 16263524]

Santos Vde P, de Medeiros AP, de Lima TA, Nascimento LF. 2014 [the effect of air pollutants on birth weight in medium-sized towns in the state of sao paulo]. Rev Paul Pediatr 32:306–312. [PubMed: 25510993]

Savitz DA, Bobb JF, Carr JL, Clougherty JE, Dominici F, Elston B, et al. 2014 Ambient fine particulate matter, nitrogen dioxide, and term birth weight in new york, new york. Am J Epidemiol 179:457–466. [PubMed: 24218031]

Schembari A, de Hoogh K, Pedersen M, Dadvand P, Martinez D, Hoek G, et al. 2015 Ambient air pollution and newborn size and adiposity at birth: Differences by maternal ethnicity (the born in bradford study cohort). Environ Health Perspect 123:1208–1215. [PubMed: 25978617]

Schwarz L, Bruckner T, Ilango SD, Sheridan P, Basu R, Benmarhnia T. 2019 A quantile regression approach to examine fine particles, term low birth weight, and racial/ethnic disparities. Environ Epidemiology 3:e060.

Sellier Y, Galineau J, Hulin A, Caini F, Marquis N, Navel V, et al. 2014 Health effects of ambient air pollution: Do different methods for estimating exposure lead to different results? Environ Int 66:165–173. [PubMed: 24598283]

Shah PS, Balkhair T, Knowledge Synthesis Group on Determinants of Preterm LBWb. 2011 Air pollution and birth outcomes: A systematic review. Environ Int 37:498–516. [PubMed: 21112090]

Sterne JA, Sutton AJ, Ioannidis JP, Terrin N, Jones DR, Lau J, et al. 2011 Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. BMJ 343:d4002. [PubMed: 21784880]

Stieb DM, Chen L, Eshoul M, Judek S. 2012 Ambient air pollution, birth weight and preterm birth: A systematic review and meta-analysis. Environ Res 117:100–111. [PubMed: 22726801]

Stieb DM, Chen L, Beckerman BS, Jerrett M, Crouse DL, Omariba DW, et al. 2016 Associations of pregnancy outcomes and pm2.5 in a national canadian study. Environ Health Perspect 124:243–249. [PubMed: 26090691]

Sun X, Luo X, Zhao C, Zhang B, Tao J, Yang Z, et al. 2016 The associations between birth weight and exposure to fine particulate matter (pm2.5) and its chemical constituents during pregnancy: A meta-analysis. Environ Pollut 211:38–47. [PubMed: 26736054]

van den Hooven EH, Pierik FH, de Kluizenaar Y, Willemsen SP, Hofman A, van Ratingen SW, et al. 2012 Air pollution exposure during pregnancy, ultrasound measures of fetal growth, and adverse birth outcomes: A prospective cohort study. Environ Health Perspect 120:150–156. [PubMed: 22222601]

Vinikoor-Imler LC, Davis JA, Meyer RE, Messer LC, Luben TJ. 2014 Associations between prenatal exposure to air pollution, small for gestational age, and term low birthweight in a state-wide birth cohort. Environ Res 132:132–139. [PubMed: 24769562]

Wilson A, Chiu YM, Hsu HL, Wright RO, Wright RJ, Coull BA. 2017 Potential for bias when estimating critical windows for air pollution in children's health. Am J Epidemiol 186:1281–1289. [PubMed: 29206986]

Winckelmans E, Cox B, Martens E, Fierens F, Nemery B, Nawrot TS. 2015 Fetal growth and maternal exposure to particulate air pollution--more marked effects at lower exposure and modification by gestational duration. Environ Res 140:611–618. [PubMed: 26056995]

Woodruff TJ, Zeise L, Axelrad DA, Guyton KZ, Janssen S, Miller M, et al. 2008 Meeting report: Moving upstream-evaluating adverse upstream end points for improved risk assessment and decision-making. Environ Health Perspect 116:1568–1575. [PubMed: 19057713]

Woodruff TJ, Sutton P, Navigation Guide Work G. 2011 An evidence-based medicine methodology to bridge the gap between clinical and environmental health sciences. Health Aff (Millwood) 30:931–937. [PubMed: 21555477]

Woodruff TJ, Sutton P. 2014 The navigation guide systematic review methodology: A rigorous and transparent method for translating environmental health science into better health outcomes. Environ Health Perspect 122:1007–1014. [PubMed: 24968373]

Xiao Q, Chen H, Strickland MJ, Kan H, Chang HH, Klein M, et al. 2018 Associations between birth outcomes and maternal pm2.5 exposure in shanghai: A comparison of three exposure assessment approaches. Environ Int 117:226–236. [PubMed: 29763818]

Xue T, Zhu T, Han Y. 2018 Association between birthweight and ambient pm2.5 in the united states: Individually-varied susceptibility and spatial heterogeneity. Environ Int 119:388–397. [PubMed: 30005187]

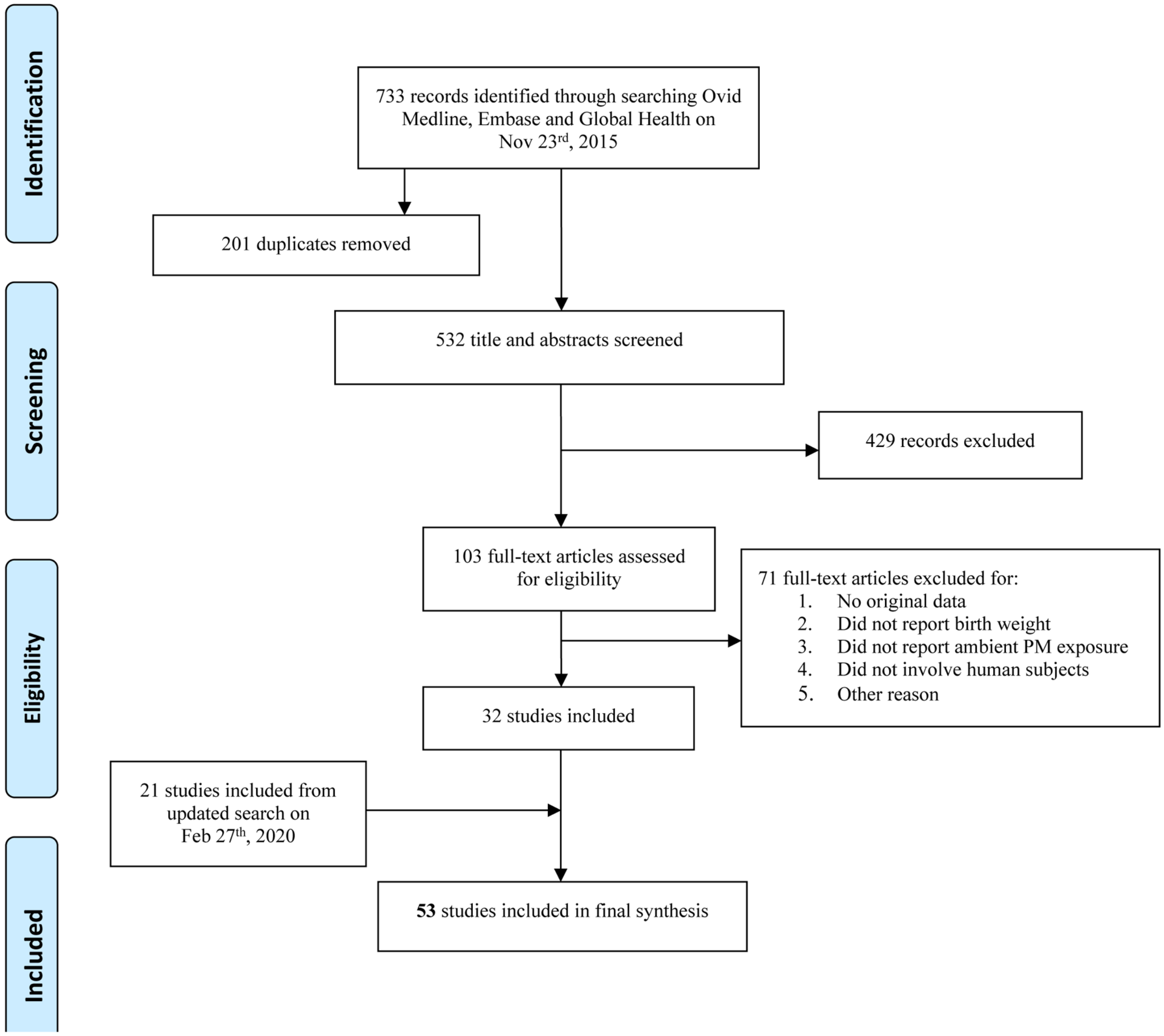
Yang CY, Tseng YT, Chang CC. 2003 Effects of air pollution on birth weight among children born between 1995 and 1997 in kaohsiung, taiwan. J Toxicol Environ Health A 66:807–816. [PubMed: 12746128]

Ye L, Ji Y, Lv W, Zhu Y, Lu C, Xu B, et al. 2018 Associations between maternal exposure to air pollution and birth outcomes: A retrospective cohort study in taizhou, china. Environ Sci Pollut Res Int 25:21927–21936. [PubMed: 29797193]

Yuan L, Zhang Y, Wang W, Chen R, Liu Y, Liu C, et al. 2020 Critical windows for maternal fine particulate matter exposure and adverse birth outcomes: The shanghai birth cohort study. Chemosphere 240:124904. [PubMed: 31550593]

Highlights:

- The Navigation Guide systematic review methodology represents a transparent and rigorous approach to reduce bias in evaluation of environmental health studies.
- Existing evidence supports an association between developmental exposure to ambient particulate matter air pollution and decreased infant birth weight.
- Heterogeneity observed in the meta-analysis supports the application of high spatial resolution air pollution exposure assessment methods in epidemiological studies.


Identification
Screening
Eligibility
Included
733 records identified through searching Ovid Medline, Embase and Global Health on Nov 23rd, 2015
201 duplicates removed
532 title and abstracts screened
429 records excluded
103 full-text articles assessed for eligibility
71 full-text articles excluded for:
1. No original data
2. Did not report birth weight
3. Did not report ambient PM exposure
4. Did not involve human subjects
5. Other reason
32 studies included
21 studies included from updated search on Feb 27th, 2020
53 studies included in final synthesis

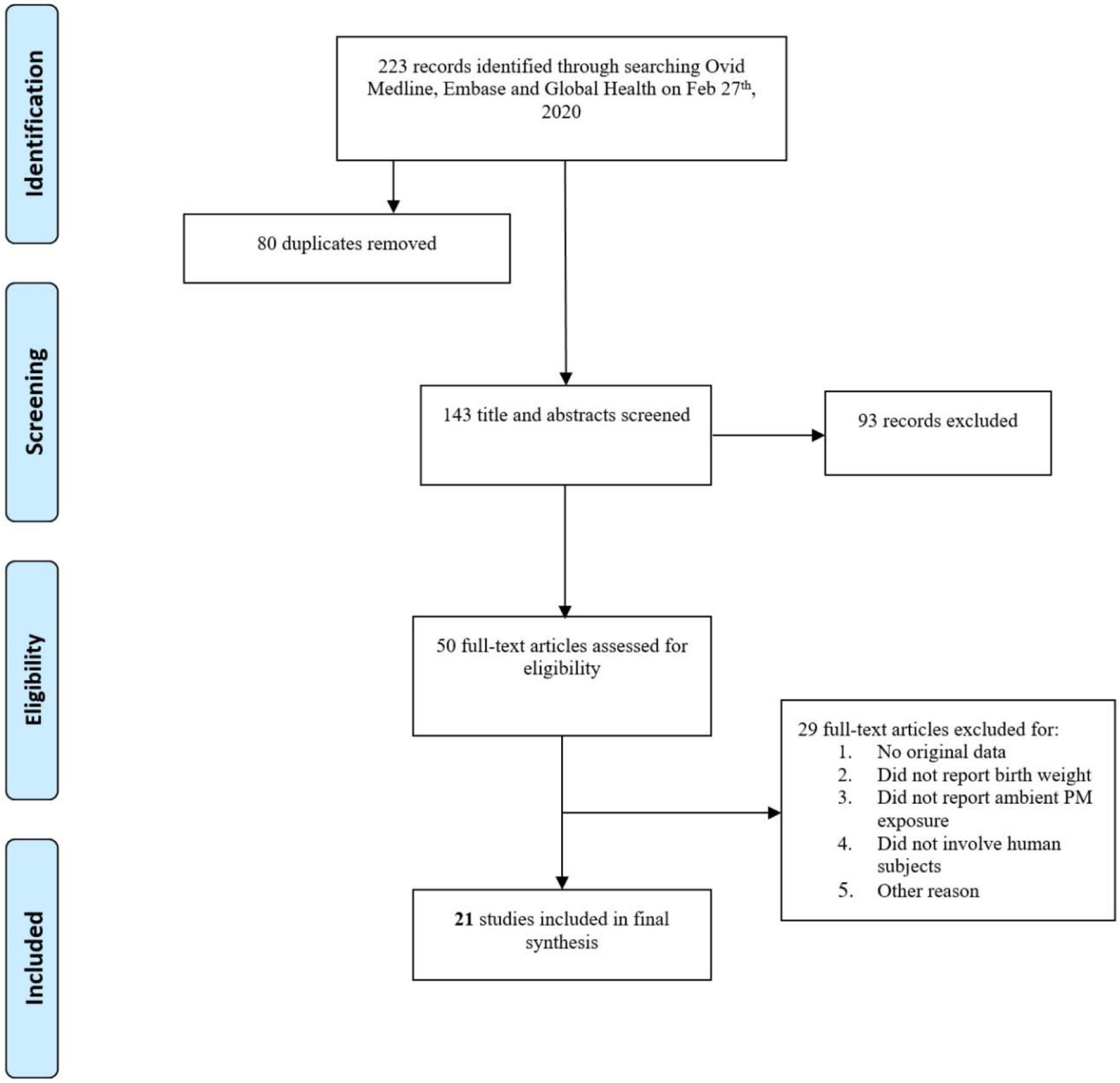


Figure 1A.

Flowchart showing the literature search and screening process for studies relevant to prenatal particulate matter exposure and birth weight measured as a continuous variable. 1B. Flowchart showing the updated search and screening process (February 27th, 2020). The search terms used are provided in Supplemental Material, Table S1.

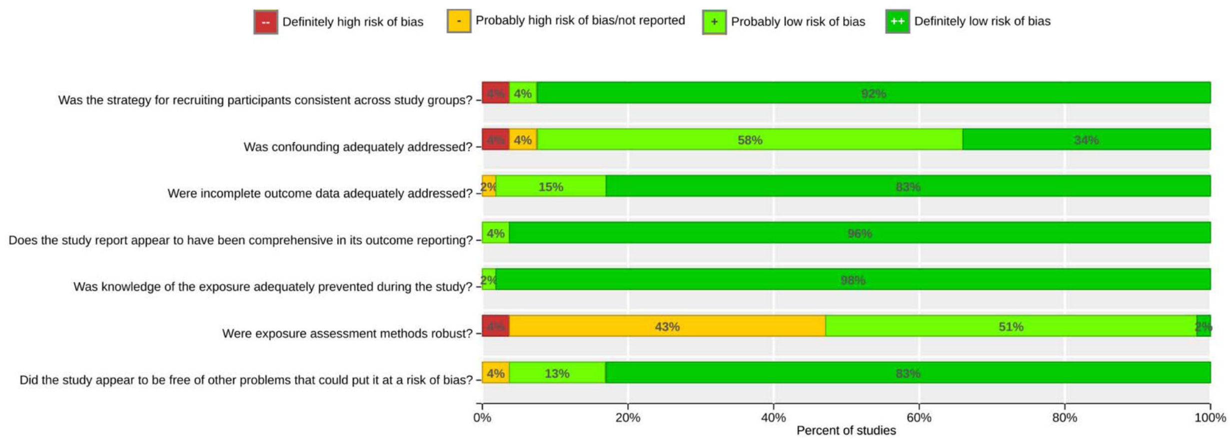
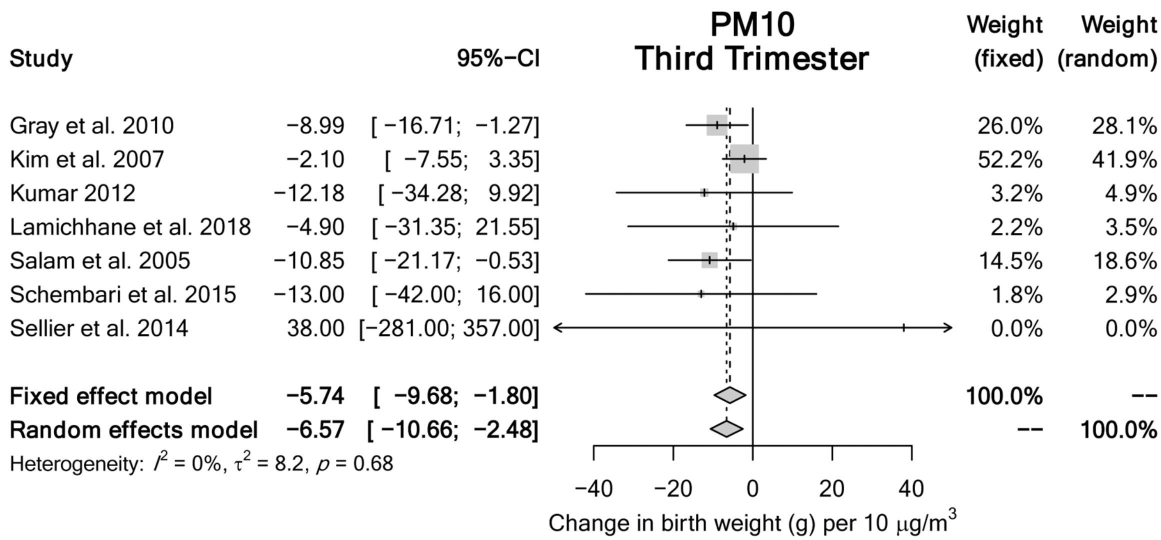


Figure 2.

Summary of risk of bias judgments. Determinations for each domain were assigned according to Supplemental Material, Table S3. In general, the domain with a considerable number of studies rated as “probably high” (43%) was related to the robustness of exposure assessment.

A. PM_{10} (third trimester) including studies rated as "low" or "probably low" risk of bias



B. PM_{10} (Full pregnancy) including studies rated as "low" or "probably low" risk of bias

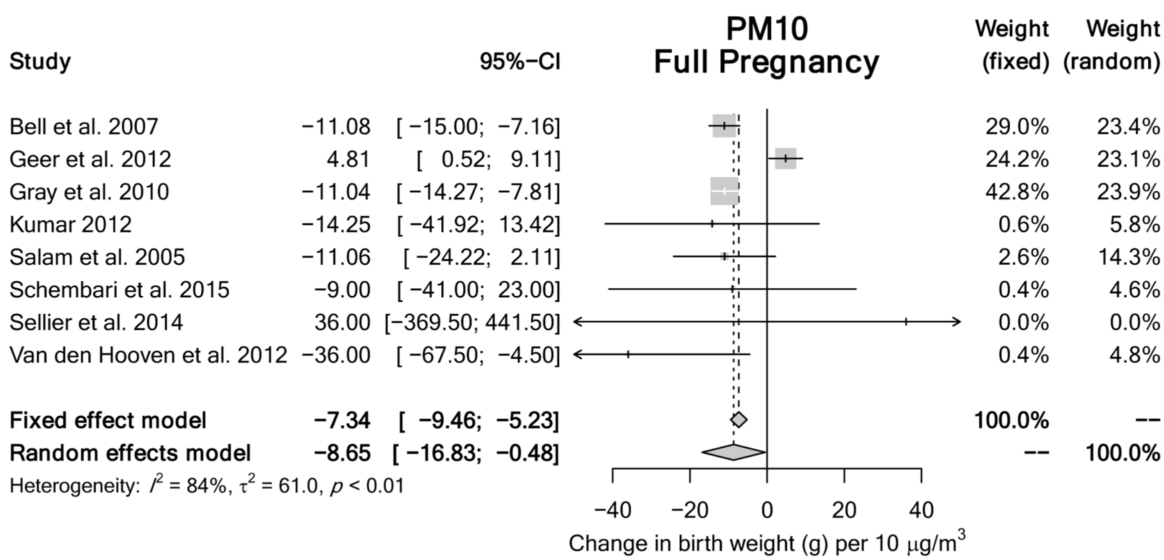


Figure 3.

Meta-analysis results for pollutants demonstrating "moderate" quality of evidence rating include (A) PM_{10} exposure during the 3rd trimester and (B) PM_{10} exposure throughout entire pregnancy.

Table 1.

Summary of risk of bias domains and criteria for low risk designation

Risk of bias domain	**Low risk of bias designation**[a]
Recruitment strategy	Protocols for recruitment and inclusion/exclusion criteria applied similarly across study groups
Blinding	Knowledge of the exposure ensured when assessing outcome, or judgement that outcome measurement not likely to be influenced by lack of blinding
Exposure assessment	Confidence in the accuracy of the exposure assessment methods that minimizes exposure misclassification, i.e., validity and reliability measures specified for monitoring and modeling
Confounding	All five important potential confounders pre-specified by reviewers are accounted for (i.e., matched, stratified, multivariate analysis or otherwise statistically controlled for)
Incomplete outcome	No missing outcome data, balanced attrition across groups, or for continuous outcome data, plausible effect size among missing outcomes not enough to have a relevant impact on the observed effect size
Selective outcome reporting	All pre-specified outcomes outlined in the protocol, methods, abstract, and/or introduction reported in the pre-specified way
Conflicts of Interest	The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study
Other bias	The study appears to be free of other sources of bias

[a]The complete criteria for determining risk of bias designations for individual studies are provided in Supplemental Material Table S3, "Instructions for Making Risk of Bias Determinations."

Table 2.

Summary of study characteristics for studies included in the meta-analysis

Reference	Study location	Study design	Sample size	Pollutant(s) (exposure assessment method)	Exposure period	Overall ROB rating
(Basu et al. 2014)	California, USA (8 counties)	Cohort R	646,296	$PM_{2.5}$ (Ambient monitoring)	Entire pregnancy, 1st, 2nd, and 3rd trimesters	Probably low
(Beland and Oloomi 2019)	southern USA	Cohort R	9,324,839	$PM_{2.5}$ (Ambient monitoring)	Entire pregnancy	Probably low
(Bell et al. 2007)	Connecticut and Massachusetts, USA	Cohort R	358,504	$PM_{2.5}$, PM_{10} (Ambient monitoring)	Entire pregnancy	Probably low
(Bell et al. 2010)	Connecticut and Massachusetts, USA (4 counties)	Cohort R	76,788	$PM_{2.5}$ (Ambient monitoring)	Entire pregnancy, 1st, 2nd, and 3rd trimesters	Probably high
(Bijnens et al. 2016)	Flanders, Belgium	Cohort R	4,760	PM_{10} (Modeling)	Entire pregnancy, 1st, 2nd, and 3rd trimesters, last month, last week	Probably high
(Darrow et al. 2011)	Atlanta, USA (5 Counties)	Cohort R	402, 627	$PM_{2.5}$, PM_{10}, $PM_{2.5–10}$ (Ambient monitoring)	Entire pregnancy, 3rd trimester only	Probably high
(Ebisu et al. 2016)	USA (224 Counties)	Cohort R	8,017,865	$PM_{2.5–10}$ (Ambient monitoring)	Entire pregnancy, 1st, 2nd, and 3rd trimesters	Probably high
(Erickson et al. 2016)	British Columbia, Canada	Cohort R	231,929	$PM_{2.5}$ (Modeling)	Entire pregnancy	Probably high
(Fong et al. 2019)	Massachusetts, USA	Cohort R	907,766	$PM_{2.5}$ (Modeling)	Entire pregnancy	Probably low
(Geer et al. 2012)	Texas, USA	Cohort R	1,548,904	$PM_{2.5}$, PM_{10} (Ambient monitoring)	Entire pregnancy	Probably low
(Giovannini et al. 2018)	Italy	Cohort R	3,614	PM_{10} (Ambient monitoring)	1st, 2nd, and 3rd trimesters	High
(Gouveia et al. 2004)	São Paulo, Brazil	Cross-sectional	179,460	PM_{10} (Ambient monitoring)	Entire pregnancy, 1st, 2nd, and 3rd trimesters	Probably high
(Gray et al. 2010)	North Carolina, USA	Cohort R	350,754	$PM_{2.5}$, PM_{10} (Ambient monitoring)	Entire pregnancy, 1st, 2nd, and 3rd trimesters	Probably low
(Gray et al. 2014)	North Carolina, USA	Cohort R	457, 642	$PM_{2.5}$ (Modeling)	Entire pregnancy	Probably low
(Guo et al. 2020)	Guangdong province, China	Cohort R	2,567,457	$PM_{2.5}$, PM_{10} (Ambient monitoring)	Entire pregnancy	Probably high
(Han et al. 2018)	Suzhou, China	Cohort R	10,915	$PM_{2.5}$, PM_{10} (Ambient monitoring)	Entire pregnancy, 1st, 2nd, and 3rd trimesters	Probably high
(Hannam et al. 2014)	United Kingdom (Northwest England)	Cohort R	203,562	$PM_{2.5}$, PM_{10} (Modeling)	Entire pregnancy, 1st, 2nd, and 3rd trimesters	Probably low

Reference	Study location	Study design	Sample size	Pollutant(s) (exposure assessment method)	Exposure period	Overall ROB rating
(He et al. 2018)	Zhengzhou, China	Cohort P	591	PM_{10} (Ambient monitoring)	Entire pregnancy, 1^{st}, $2^{nd,}$ and 3^{rd} trimesters	Probably high
(Huang et al. 2015)	Beijing, China	Cohort R	50,874	PM_{10} (Ambient monitoring)	1^{st}, $2^{nd,}$ and 3^{rd} trimesters	Probably high
(Hyder et al. 2014)	Connecticut and Massachusetts, USA	Cohort R	834,332	$PM_{2.5}$ (Ambient monitoring and modeling)	Entire pregnancy, 1^{st}, $2^{nd,}$ and 3^{rd} trimesters	Probably low
(Jedrychowski et al. 2009)	Krakow, Poland	Cohort P	481	$PM_{2.5}$ (Personal monitoring)	Entire pregnancy	Low
(Keller et al. 2017)	Georgia, USA	Cohort R	403,881	$PM_{2.5}$ (Modeling)	1^{st}, $2^{nd,}$ and 3^{rd} trimesters	Probably low
(Kim et al. 2007)	Seoul, Korea	Cohort P	1,514	PM_{10} (Ambient monitoring)	1^{st}, 2^{nd}, and 3^{rd} trimesters	Probably low
(Kirwa et al. 2019)	Puerto Rico	Cohort R	332,129	$PM_{2.5}$ (Ambient monitoring)	Entire pregnancy	Probably high
(Kumar 2012)	Chicago, USA	Cohort R	400,000	$PM_{2.5}$, PM_{10} (Ambient monitoring)	Entire pregnancy, 1^{st}, $2^{nd,}$ and 3^{rd} trimesters	Probably low
(Lamichhane et al. 2018)	South Korea	Cohort P	648	PM_{10} (Modeling)	1^{st}, $2^{nd,}$ and 3^{rd} trimesters	Probably low
(Laurent et al. 2013)	California, USA (2 counties)	Cohort R	105,092	$PM_{2.5}$, PM_{10} (Ambient monitoring and modeling)	Entire pregnancy	High
(Lavigne et al. 2018)	Ontario, Canada	Cohort R	196,171	$PM_{2.5}$ (Ambient monitoring)	Entire pregnancy, 1^{st}, 2^{nd}, and 3^{rd} trimesters	Probably low
(Li et al. 2019)	Ningbo, China	Cohort R	170,008	$PM_{2.5}$, PM_{10} (Ambient monitoring)	Entire pregnancy, 1^{st}, $2^{nd,}$ and 3^{rd} trimesters	Probably high
(Mannes et al. 2005)	Sydney, Australia	Cohort R	138,056	$PM_{2.5}$, PM_{10} (Ambient monitoring)	1^{st}, 2^{nd}, and 3^{rd} trimesters	Probably high
(Medeiros and Gouveia 2005)	São Paulo, Brazil	Cohort R	311,735	PM_{10} (Ambient monitoring)	1^{st}, $2^{nd,}$ and 3^{rd} trimesters	High
(Merklinger-Gruchala and Kapiszewska 2015)	Krakow, Poland	Cohort R	84,842	PM_{10} (Ambient monitoring)	Entire pregnancy, 1^{st}, $2^{nd,}$ and 3^{rd} trimesters	Probably high
(Morello-Frosch et al. 2010)	California, USA	Cohort R	3,545,177	$PM_{2.5}$, PM_{10}, $PM_{2.5-10}$ (Ambient monitoring)	Entire pregnancy, 1^{st}, $2^{nd,}$ and 3^{rd} trimesters	Probably high
(Parker and Woodruff 2008)	USA (excluding Alaska and Hawaii)	Cohort R	785,965	$PM_{2.5}$, $PM_{2.5-10}$ (Ambient monitoring)	Entire pregnancy, 1^{st}, $2^{nd,}$ and 3^{rd} trimesters	Probably high
(Parker et al. 2005)	California, USA	Cohort R	18,247	$PM_{2.5}$ (Ambient monitoring)	Entire pregnancy	Probably high
(Pedersen et al. 2013)	12 European countries	Cohort P	74,178	$PM_{2.5}$, PM_{10}, $PM_{2.5-10}$ (Modeling)	Entire pregnancy	High
(Rahmalia et al. 2012)	Poiters and Nancy, France	Cohort P	888	PM_{10} (Modeling)	Entire pregnancy, 1^{st}, $2^{nd,}$ and 3^{rd} trimesters	Probably high

Reference	Study location	Study design	Sample size	Pollutant(s) (exposure assessment method)	Exposure period	Overall ROB rating
(Rhee et al. 2019)	Boston, USA	Cohort R	3,366	$PM_{2.5}$ (Modeling)	Entire pregnancy, 1st, 2nd, and 3rd trimesters	Probably low
(Salam et al. 2005)	California, USA	Cohort R	3,901	PM_{10} (Ambient monitoring)	Entire pregnancy, 1st, 2nd, and 3rd trimesters	Probably low
(Santos Vde et al. 2014)	São José dos Campos, Brazil	Cross-sectional	21,591	PM_{10} (Ambient monitoring)	3rd trimester only	High
(Savitz et al. 2014)	New York, USA	Cohort R	252,967	$PM_{2.5}$ (Modeling)	Entire pregnancy, 1st, 2nd, and 3rd trimesters	Probably low
(Schembari et al. 2015)	Bradford, United Kingdom	Cohort P	9,067	$PM_{2.5}$, PM_{10} (Modeling)	Entire pregnancy, 3rd trimester only	Probably low
(Schwarz et al. 2019)	California, USA	Cohort R	2,768,898	$PM_{2.5}$ (Ambient monitoring)	Entire pregnancy	Probably low
(Sellier et al. 2014)	Poiters and Nancy, France	Cohort P	1,026	PM_{10} (Modeling)	Entire pregnancy, 1st, 2nd, and 3rd trimesters	Probably low
(Stieb et al. 2016)	Canada	Cohort R	2,781,940	$PM_{2.5}$ (Modeling)	Entire pregnancy, 1st, 2nd, and 3rd trimesters	Probably high
(van den Hooven et al. 2012)	Netherlands	Cohort P	7,772	PM_{10} (Modeling)	Entire pregnancy	Probably low
(Vinikoor-Imler et al. 2014)	North Carolina, USA	Cohort R	322,981	$PM_{2.5}$ (Modeling)	1st, 2nd, and 3rd trimesters	Probably low
(Winckelmans et al. 2015)	Flanders, Belgium	Cohort R	525,635	PM_{10} (Modeling)	Entire pregnancy, 1st, 2nd, and 3rd trimesters	Probably high
(Xiao et al. 2018)	Shanghai, China	Cohort R	132,783	$PM_{2.5}$ (Modeling)	Entire pregnancy, 1st, 2nd, and 3rd trimesters	Probably low
(Xue et al. 2018)	USA	Cohort R	18,317,707	$PM_{2.5}$ (Ambient monitoring)	Entire pregnancy	High
(Yang et al. 2003)	Kaohsiung, Taiwan	Cohort R	13,396	PM_{10} (Ambient monitoring)	1st, 2nd, and 3rd trimesters	Probably high
(Ye et al. 2018)	Taizhou, China	Cohort R	24,246	PM_{10} (Modeling)	Entire pregnancy, 1st, 2nd, and 3rd trimesters	Probably high
(Yuan et al. 2020)	Shanghai, China	Cohort R	3,692	$PM_{2.5}$ (Modeling)	Entire pregnancy, 1st, 2nd, and 3rd trimesters	Probably low

R: retrospective cohort; P: prospective cohort

Table 3.

Factors for evaluating the quality and strength of the body of evidence

Quality is rated across all studies. Evidence begins as "moderate" and may be downgraded (−1 or −2) or upgraded (+1 or +2) according to factors.		Strength is rated across all studies. The final ratings represent the level of certainty of toxicity.	
Downgrade factors	• Risk of bias across studies • Indirectness • Inconsistency • Imprecision • Publication bias	**Considerations**	• Quality of body of evidence • Direction of effect estimates • Confidence in effect estimates • Other compelling attributes of the data that may influence certainty
Upgrade factors	• Large magnitude of effect • Dose response • Confounding minimizes effect		
Quality rating	• High quality • Moderate quality • Low quality • Very low quality	**Strength rating**	• Sufficient evidence • Limited evidence • Inadequate evidence • Evidence of lack of toxicity

Table 4:

Summary of main meta-analysis results and quality of evidence rating conclusions

Exposure Window	No. of studies	Effect estimate g per 10 μg/m^3 (95% CI)	I^2 (%)	Quality of evidence rating
$PM_{2.5}$				
1st Trimester	11	−6.50 (−15.07, 2.07)	87%	Very low (downgrades for imprecision and inconsistency)
2nd Trimester	12	−5.69 (−10.58, −0.79)	68%	Low (downgrade for imprecision)
3rd Trimester	12	−10.67 (−20.91, −0.43)	84%	Low (downgrade for inconsistency)
Full Pregnancy	15	−27.55 (−48.45, −6.65)	94%	Low (downgrade for inconsistency)
PM_{10}				
1st Trimester	6	3.22 (−3.13, 9.58)	14%	Low (downgrade for imprecision)
2nd Trimester	6	−3.37 (−8.22, 1.48)	0%	Low (downgrade for imprecision)
3rd Trimester	7	−6.57 (−10.66, −2.48)	0%	Moderate (no changes)
Full Pregnancy	8	−8.65 (−16.83, −0.48)	84%	Moderate (heterogeneity explained by single study with inverse effect)
$PM_{2.5–10}$				
1st Trimester	3	−2.70 (−3.90, −1.49)	0%	Very low (downgrades for risk of bias and imprecision)
2nd Trimester	3	−2.90 (−10.04, 4.23)	70%	Very low (downgrades for risk of bias, imprecision, inconsistency)
3rd Trimester	4	−4.93 (−10.82, 0.96)	76%	Very low (downgrades for risk of bias, imprecision, inconsistency)
Full Pregnancy	5	−8.81 (−10.32, −7.31)	0%	Low (downgrade for risk of bias)

For $PM_{2.5}$, we included 18 unique studies rated as “low” or “probably low” risk of bias. For PM_{10}, we included 10 studies rated as “low” or “probably low” risk of bias. For coarse PM ($PM_{2.5–10}$), there were no studies rated as “low” or “probably low” risk of bias, thus we included 5 studies rated as “high” or “probably high.”

Table 5:

Summary of strength of evidence conclusions

Exposure Window	Quality of evidence rating[1]	Direction of effect estimates	Confidence in effect estimates	Other compelling attributes	Strength of evidence rating
$PM_{2.5}$					
1st Trimester	Very low	Adverse[2]	Low[3]	None	Inadequate
2nd Trimester	Low	Adverse[2]	Low[3]	None	Inadequate
3rd Trimester	Low	Adverse[2]	Low[3]	None	Inadequate
Full Pregnancy	Low	Adverse[2]	Low[3]	None	Inadequate
PM_{10}					
1st Trimester	Low	Adverse[2]	Low[3]	None	Inadequate
2nd Trimester	Low	Adverse[2]	Low[3]	None	Inadequate
3rd Trimester	Moderate	Adverse[2]	Limited[4]	None	Limited
Full Pregnancy	Moderate	Adverse[2]	Limited[4]	None	Limited
$PM_{2.5-10}$					
1st Trimester	Very low	Adverse[2]	Low[3]	None	Inadequate
2nd Trimester	Very low	Adverse[2]	Low[3]	None	Inadequate
3rd Trimester	Very low	Adverse[2]	Low[3]	None	Inadequate
Full Pregnancy	Low	Adverse[2]	Low[3]	None	Inadequate

[1] From Table 4.

[2] Decreasing birth weight with increasing exposure is considered an effect in the adverse direction.

[3] Results may be due to chance, bias, or confounding, so additional data are likely to alter the results.

[4] A credible association is observed, but chance, bias, and confounding cannot be ruled out with reasonable confidence, so additional data could alter the results.