COMMITTEE ON THE MEDICAL EFFECTS OF AIR POLLUTANTS

Effects of maternal exposure to ambient air pollution on adverse birth outcomes

A report by the Committee on the Medical Effects of Air Pollutants

Chair: Professor Frank Kelly (outgoing), Professor Anna Hansell

Chair of Sub-group: Professor Jonathan Grigg

© Crown copyright 2025

Produced by UK Health Security Agency for the
Committee on the Medical Effects of Air Pollutants

© Crown copyright 2025

Produced by UK Health Security Agency for the
Committee on the Medical Effects of Air Pollutants

Foreword

A healthy start to life can help promote lifelong health. A growing number of studies have suggested that current levels of air pollution may increase the risks of adverse birth outcomes – prematurity, being born underweight or stillbirth. This is a concern for both parents and society and needs careful evaluation.

In this report, COMEAP examines the evidence regarding the impact of outdoor air pollution on adverse birth outcomes. We conducted a detailed review of epidemiological research available to mid-2021. We also considered mechanistic studies and discuss how the biological impacts of air pollution might be linked to the outcomes studied. We found that many of the published systematic reviews were not of a suitable quality to draw firm conclusions regarding the likelihood that air pollutants (particulate matter – $PM_{2.5}$ and PM_{10}, nitrogen dioxide and ozone) are associated with low birth weight, pre-term birth or stillbirth.

Our findings align with conclusions on this topic in the US Environmental Protection Agency (USEPA) Integrated Science Assessments (ISAs) published in 2016, 2019, 2020, together with commentary reports from the Health Effects Institute, the World Health Organization, the UK Royal College of Physicians (RCP) and the UK Royal College of Paediatrics and Child Health. The USEPA reports concluded that the evidence was suggestive of, but not sufficient to infer, a causal relationship. Our report generally agrees with the USEPA conclusions, taking account of differences in approach to reviewing the evidence.

We concluded that new systematic reviews and meta-analyses were required that would include recently published studies. These would allow us to draw firmer conclusions as to whether air pollutants are likely to cause these adverse effects on births. We might also be able to recommend an exposure-response function that we think is sufficiently robust to estimate the numbers of births affected in the UK, and to inform policy analysis.

We note that a 2021 Global Burden of Disease study includes an estimate of the burden, attributable to fine particle air pollution ($PM_{2.5}$), of health outcomes linked to low birth weight and effects on gestational age. We evaluated the underpinning evidence synthesis by Ghosh et al, published in 2021, as part of this report. In our opinion, some aspects of the methods used – including review design and study selection – could have had an impact on the conclusions, including the estimated exposure-response relationship. We therefore recommend that all of the available evidence should be synthesised in up-to-date meta-analyses.

The preparation and publication of this report were delayed by the COVID-19 pandemic and also by other time-critical requests from government departments

for advice from COMEAP. We recognise that newer reviews, meta-analyses and mechanistic studies on air pollution and birth outcomes have been published since we undertook our literature review, and we are aware of others that are underway. These more recent analyses will require the same careful, methodological examination to assess their suitability as a basis for policy analysis and decision-making.

I would like to extend congratulations, and many thanks, to everyone in the Sub-group who produced the report, which represents a huge amount of work. I would particularly like to thank the Sub-group Chair Professor Jonathan Grigg and the co-opted members of the Sub-group who generously contributed their time and specialist expertise to this work. I would also like to thank the COMEAP Secretariat at UKHSA who ably supported the Sub-group's work.

Professor Anna Hansell

Chair of the Committee on the Medical Effects of Air Pollutants

Authors

COMEAP and Sub-group Members	
Professor Jonathan Grigg Sub-group Chair, and Chair of Mechanisms Working Group	Centre for Genomics and Child Health, The Blizard Institute, Queen Mary University of London Professor Grigg is supported by i) NERC grant: NE/W002264/1 Relating Environment-use Scenarios in Pregnancy/Infanthood and Resulting airborne material Exposures to child health outcomes (RESPIRE) ii) NIHR Senior Investigator Award Conceptualisation, Methodology, Investigation, Writing – Original Draft, Writing – Review & Editing, Supervision
Dr Rachel Smith Chair of Epidemiology Working Group	Mohn Centre for Children's Health and Wellbeing, School of Public Health, Imperial College London National Institute for Health and Care Research (NIHR) Health Protection Research Unit (HPRU) in Environmental Exposures and Health, Imperial College London Medical Research Council (MRC) Centre for Environment and Health, Imperial College London Conceptualisation, Methodology, Investigation, Writing – Original Draft, Writing – Review & Editing, Supervision
Professor Peter Brocklehurst	Birmingham Clinical Trials Unit, Institute of Applied Health Research, University of Birmingham Conceptualisation, Methodology, Investigation, Writing – Original Draft, Writing – Review & Editing
Dr Julia Fussell	Environmental Research Group, School of Public Health, Imperial College London Medical Research Council (MRC) Centre for Environment and Health, Imperial College London National Institute for Health and Care Research (NIHR) Health Protection Research Unit (HPRU) in Environmental Exposures and Health, Imperial College London Conceptualisation, Methodology, Investigation, Writing – Original Draft, Writing – Review & Editing, Supervision
Professor Anna Hansell	Centre for Environmental Health and Sustainability, University of Leicester National Institute for Health and Care Research (NIHR) Health Protection Research Unit (HPRU) in Chemical Threats and Hazards National Institute for Health and Care Research (NIHR) Leicester Biomedical Research Centre (BRC) Writing – Review & Editing, Project Administration
Professor Debbie Jarvis	National Heart and Lung Institute, Faculty of Medicine, Imperial College London Medical Research Council (MRC) Centre for Environment and Health, Imperial College London Conceptualisation, Methodology, Investigation, Writing – Original Draft, Writing – Review & Editing
Professor Eric Jauniaux	Institute for Women's Health, University College London University College London Hospitals NHS Foundation Trust

	Conceptualisation, Methodology, Investigation, Writing – Original Draft, Writing – Review & Editing, Supervision
Professor Frank Kelly	Environmental Research Group, School of Public Health, Imperial College London Medical Research Council (MRC) Centre for Environment and Health, Imperial College London National Institute for Health and Care Research (NIHR) Health Protection Research Unit (HPRU) in Environmental Exposures and Health, Imperial College London Conceptualisation, Project Administration
Professor Mireille Toledano	Mohn Centre for Children's Health and Wellbeing, School of Public Health, Imperial College London National Institute for Health and Care Research (NIHR) Health Protection Research Unit (HPRU) in Environmental Exposures and Health, Imperial College London Medical Research Council (MRC) Centre for Environment and Health, Imperial College London Conceptualisation, Methodology, Investigation, Writing – Original Draft, Writing – Review & Editing
Dr Heather Walton	Environmental Research Group, School of Public Health, Imperial College London Medical Research Council (MRC) Centre for Environment and Health, Imperial College London National Institute for Health and Care Research (NIHR) Health Protection Research Unit (HPRU) in Environmental Exposures and Health, Imperial College London Conceptualisation, Methodology, Investigation, Writing – Original Draft, Writing – Review & Editing
Scientific Secretariat	
Dr Philippa Douglas	Air Quality and Public Health, UK Health Security Agency (UKHSA) (now Principal Air Quality Scientist, Environment Agency) Conceptualisation, Methodology, Investigation, Data Curation, Writing – Original Draft, Writing – Review & Editing, Supervision, Project Administration
Eve Draper	Air Quality and Public Health, UKHSA Writing – Review & Editing, Project Administration
Dr Naomi Earl	Environmental Hazards and Emergencies, UKHSA Writing – Review & Editing, Project Administration
Dr Karen Exley	Air Quality and Public Health, UKHSA Conceptualisation, Methodology, Investigation, Writing – Original Draft, Writing – Review & Editing, Project Administration
Alison Gowers	Air Quality and Public Health, UKHSA Conceptualisation, Writing – Review & Editing, Project Administration
Dr Sarah Robertson	Air Quality and Public Health, UKHSA (now Healthcare Scientist Advanced (Environmental Public Health) Public Health Scotland) Conceptualisation
Eleanor Sykes	Air Quality and Public Health, UKHSA (now Programme Advisor – Flood and Coastal Risk Management, Environment Agency) Investigation, Data Curation, Writing – Review & Editing, Project Administration

Acknowledgements

COMEAP would like to thank all of the members of the Adverse Birth Outcomes Sub-group for contributing their time and expertise in developing this report, in particular co-opted members: Peter Brocklehurst, Julia Fussell, Eric Jauniaux, Rachel Smith and Mireille Toledano. Special thanks go to Peter Brocklehurst and Eric Jauniaux for sections on pregnancy and the function of the placenta, Julia Fussell and Jonathan Grigg for leading on and finalising the sections discussing the mechanistic evidence relating to effects of air pollutants, and Rachel Smith for leading on and finalising the substantial review of the epidemiological literature.

The COMEAP Adverse Birth Outcomes Sub-group would like to thank the following individuals for their helpful contributions during the preparation of this report:

Dr Mike Holland (EMRC) for his contribution to the section on the UK Burden of stillbirth, preterm birth, and low birth weight.

Mr John Stedman (Ricardo Energy and Environment) and Professor Mathew Heal (University of Edinburgh) for their knowledge and expert contribution to Typical UK levels of $PM_{2.5}$, PM_{10}, NO_2 and O_3 (Appendix 12).

Dr Allen Wilcox (National Institute of Environmental Health Sciences, US National Institutes of Health) for his expert opinion on a paper.

Dr Valentina Guercio (UK Health Security Agency) for re-running the stratified meta-analyses presented in Figures 8, 9 and 10.

Executive summary

Given growing interest from the public and the government about the effects of air pollution on pregnancy and early life, the Department of Health and Social Care (DHSC) requested that COMEAP undertake work on this topic. To support COMEAP's work, DHSC approached the National Institute of Health Research (NIHR) Health Protection Research Unit (HPRU) on Health Impact of Environmental Hazards at King's College London to undertake an evidence-based scoping review on air pollution and adverse birth outcomes.

Using the scoping review as a starting point, a Sub-group of the Committee undertook a programme of work to address the following questions:

- What is the strength and consistency of epidemiological evidence linking ambient air pollutants with adverse birth outcomes?
- Is there evidence of biological mechanisms by which air pollutants could affect adverse birth outcomes?
- Is the evidence sufficient to propose quantification of specific adverse birth outcomes associated with air pollution in the UK?
- Can recommendations be identified from current meta-analyses or should new meta-analyse(s) be undertaken?

We reviewed studies identified in the scoping review and by updated literature searches. The main health outcomes considered were pre-term birth (PTB), term low birth weight (TLBW) and stillbirth. We examined epidemiological studies investigating associations with particulate matter (including $PM_{2.5}$, PM_{10}), nitrogen dioxide (NO_2) and ozone (O_3). For mechanistic studies, we additionally considered evidence relating to black carbon, polycyclic aromatic hydrocarbons (PAHs, including B[a]P) and diesel engine exhaust.

For the epidemiological evidence, we reviewed and quality assessed the available meta-analyses published up until August 2020 in relation to ambient air pollutants and birth outcomes. Meta-analyses are a synthesis of the data for the same health endpoint and possible cause (for example, the same pollutant), from independent studies, to determine the overall weight of evidence. We selected as core meta-analyses those which were highest quality, to form the basis for our views on the available evidence. We also identified individual studies, (published until March 2020) that were not included in these core meta-analyses. It was not our intention to comment on all the individual studies but, rather, to identify key studies that could influence conclusions drawn on the basis of the core meta-analyses.

The meta-analyses of epidemiological studies were quality assessed using modifications of the AMSTAR 2 critical appraisal tool for systematic reviews and

the mechanistic evidence was quality assessed using the Texas Commission Environmental Quality guidelines.

Overview of the epidemiological evidence

From the available meta-analyses, there was weak evidence for associations between some air pollutants and adverse birth outcomes. We identified a large number of additional studies that were not included in the meta-analysis evidence. Therefore, for most pollutant-outcome pairs, up-to-date meta-analyses are required to synthesise the full evidence base.

In the table below, we present the overall conclusions from reviewing the epidemiological evidence on air pollutants and adverse birth outcomes, by pollutant and birth outcome pairs.

Preterm Birth (a birth occurring at a gestational age of < 37 weeks)			
Pollutant	Conclusion from the core meta-analyses	Conclusion from the key individual studies (studies that would influence, or be important to note, over and above the evidence already provided in the core meta-analyses)	Recommendation
$PM_{2.5}$	Weak evidence* for an association	Large number of studies not captured in or published since the core meta-analyses. Results from key studies were inconsistent[+]	Up-to-date meta-analyses to take into account all of the evidence are needed for $PM_{2.5}$, PM_{10}, NO_2 and O_3. This is because, whilst selected individual studies published since the meta-analyses were considered, we did not systematically review all these additional studies.
PM_{10}	Weak evidence* for an association	Large number of studies not captured in or published since the core meta-analyses. Note: the key studies that have been conducted in Europe do not support an association, but there are a number of studies with varying design and methods from Asia that need further consideration	
NO_2	Does not support an association*	Large number of studies not captured in or published since the core meta-analyses. Results from key studies were inconsistent[+]	
O_3	Weak evidence* for an association	Large number of studies not captured in or published since the core meta-analyses. Results from key studies were inconsistent[+]	
Term Low Birth Weight (a birth at term (≥37 weeks) with a weight of <2,500 g)			
Pollutant	Conclusion from the core meta-analyses	Conclusion from the key individual studies	Recommendation
$PM_{2.5}$	Weak evidence* for an association	Results of the key studies are inconsistent[+]	Up-to-date meta-analyses to take into account all of the evidence are needed for $PM_{2.5}$ and PM_{10} and recommended for NO_2 and O_3. This is because, whilst selected individual studies published since the meta-analyses were considered, we did not systematically review all these additional studies.
PM_{10}	Weak evidence* for an association	Several large key studies which have inconsistent[+] findings	
NO_2	Does not support an association[♦]	Key studies do not support an association[♦]	
O_3	Does not support an association[♦]	Key studies do not support an association[♦]. However, there are several large studies not captured in or published since the core meta-analyses, which expand geographical coverage of the evidence base.	
Stillbirth (a baby born after 24 completed weeks^ of gestation showing no signs of life)			
Pollutant	Conclusion from the core meta-analyses	Conclusion from the key individual studies	Recommendation
$PM_{2.5}$	Does not support an association[♦]	The key studies provide inconsistent evidence[+].	Up-to-date meta-analyses to take into account all of the evidence are needed for $PM_{2.5}$, PM_{10}, O_3, and recommended for NO_2.
PM_{10}	Does not support an association[♦]	The key studies have inconsistent results[+].	

NO_2	Does not support an association♦	The key studies do not suggest an association♦.	Whilst selected individual studies published since the meta-analyses were considered, we did not systematically review all these additional studies. However, for PM_{10} and NO_2, up to date meta-analyses are unlikely to alter the conclusion presented.
O_3	Does not support an association♦	The key studies provide inconsistent+ evidence for an association. However, we note that the well conducted studies from the USA and the UK provide some evidence of an association of stillbirth with O_3 exposure in pregnancy although the small number of studies preclude robust evaluation of whether this is related to a trimester- specific exposure or to whole pregnancy exposure.	

*For context, a conclusion of weak evidence for an association from the meta-analyses, is defined in section 2.1.3 as where the evidence from different meta-analyses (or different exposure windows within the same meta-analysis) is inconsistent (either in direction or statistical significance), but there is evidence from some portion of the meta-analyses indicating an association and/or associations observed in the meta-analyses are subject to considerable heterogeneity in direction in forest plots and/or considerable potential for residual confounding.

♦ For context, a conclusion of the evidence does not support an association from the meta-analyses is defined in section 2.1.3 as where summary estimates from meta-analyses are consistently either null (odds ratio 1 or very close to 1) or fail to reach conventional levels of statistical significance.

+ For context, inconsistent evidence/results is defined in section 2.1.3 as effect estimates from different studies point in both positive and negative directions (odds ratios above and below 1) and/or effect estimates from some studies indicate an association (odds ratio above or below 1 and reaching conventional statistical significance) whilst effect estimates from other studies are null (odds ratio 1 or very close to 1) and/or fail to reach conventional levels of statistical significance.

^ UK definition provided. Different timing cut-offs, typically ranging from 20 to 28 weeks' gestation, are used for defining stillbirth in different countries

Overview of the mechanistic evidence

When reviewing the mechanistic evidence, published until June 2021, the primary focus was placed on human studies because animal studies are less useful for understanding the effects of pollutants on adverse birth outcomes in humans than for some other health endpoints.

Overall, there is some support, from the rather limited mechanistic evidence available, that it is biologically plausible that exposure to ambient levels of air pollution, particularly particulate matter, could cause adverse birth outcomes. Many of the mechanisms identified appear relevant to TLBW, and stillbirth.

Nitrosative/oxidative stress (for example, through the nitration of placental proteins involved in trophoblast invasion and regulation of placental vascular reactivity), inflammation (for example, through the expression of molecules involved in the formation of a dysfunctional placenta), PAH exposure (for example, through possible anti-estrogenic effects) and epigenetic changes (for example, by altering the expression of genes involved in key placental cellular processes) all have potential to disturb placental function and, as a consequence, appear relevant to compromised fetal growth and development as well increasing the risk of stillbirth.

In addition, given that inflammatory events take place in preparation for delivery and there is an increased inflammatory response during preterm delivery, inflammation is a plausible mechanism mediating the effects of air pollution on PTB.

Quantification of specific adverse birth outcomes associated with air pollution in the UK

There is some evidence for associations between air pollution and birth outcomes. However, we do not consider it sufficient to make recommendations for quantification of adverse birth outcomes associated with air pollutants at this stage, given the need for up-to-date meta-analyses.

Contextualising with reports from other authoritative bodies

We have examined our findings in the context of reports from other authoritative bodies. The conclusions reached in our report are broadly compatible with conclusions in reports from other authoritative bodies.

Recommendations for future work

In the course of this work, the Sub-group identified several areas for future research:

- New meta-analyses are required to synthesise all of the available evidence and enable recommendations for quantification to be made, if appropriate
- A new systematic review of the epidemiological evidence for PAHs

- More studies investigating the mechanisms by which air pollutants impact on birth outcomes are required

We also make a number of suggestions regarding methods, reporting and publication of epidemiological studies, systematic reviews and meta-analyses that we think would help improve the evidence base in the future.

Contents

Foreword iii
Authors v
Acknowledgements vii
Executive summary viii

Chapter 1 Introduction 11
1.1 Background 11
1.2 Approach 11
1.3 Structure of this report 12
1.4 Description of a 'standard' pregnancy and possible influences of air pollution 13
1.5 Birth outcomes of interest 24
1.6 UK burden of stillbirth, preterm birth, and low birth weight 25
1.7 Exemplar of an effect of environmental toxicity: exposure to tobacco smoke during pregnancy 26

Chapter 2 Epidemiological evidence 29
2.1 Approach 29
2.2 Results 43
2.3 Conclusions from epidemiological studies 122

Chapter 3 Mechanistic evidence 126
3.1 Approach 126
3.2 Identified literature 126
3.3 Evidence from human mechanistic studies 137
3.4 Studies from animal models 157

Chapter 4 Discussion 162
4.1 Overview of evidence 162
4.2 Strengths and limitations of the report 165
4.3 Limitations of the evidence base 167
4.4 Bringing the epidemiological and mechanistic evidence together 170
4.5 Use of single pollutant models 176
4.6 Is there a key exposure window? 177
4.7 Is there a key air pollutant? 178
4.8 Implications 178

Chapter 5 Conclusions 192

Chapter 6 Recommendations for future work 195

References 198

Appendices 227

Appendix 1. Full list of Sub-group and Working Group Members 228

Appendix 2. Terms of reference 230

Appendix 3. UK Burden of stillbirth, preterm birth, and low birthweight 232

Appendix 4. Search terms used in the updated literature search 234

Appendix 5. Details of whether included meta-analyses concerned term low birth weight (TLBW) or low birth weight (LBW) 235

Appendix 6. AMSTAR 2 scoring tool 242

Appendix 7. AMSTAR 2 modifications 250

Appendix 8. Modified AMSTAR 2 weighting 253

Appendix 9. Population overlaps in the meta-analyses 255

Appendix 10. The approach utilised to determine the level of commentary written for Preterm birth (PTB), term low birth weight (TLBW) and stillbirth and the pollutants ($PM_{2.5}$, PM_{10}, NO_2 and O_3) 340

Appendix 11. Detailed results and summary characteristics from selected meta-analyses per pollutant and outcome pair 341

Appendix 12. Selected key studies of interest not included in core meta-analyses 362

Appendix 13. Summary tables of studies investigating underlying mechanisms 399

Appendix 14. BaP concentrations in the UK 423

Glossary of terms and abbreviations

Term (Abbreviation)	Definition
3-nitrotyrosine (3-NT)	A product of the attack of peroxynitrite (a reactive nitrogen species [RNS]) on tyrosine groups of proteins AND/OR: a measure of protein modification by RNS
Placental 3-NT (3-NTp)	See above
8-iso-prostaglandin F2α (8-iso-PGF2α)	A F2-isoprostane. These are markers of oxidative stress and are generated by the oxidation of arachidonic acid esterified in phospholipids
15-F2t-isoprostane (15-F2t-IsoP)	A F2-isoprostane (see above)
8-Oxo-2'-deoxyguanosine (8-OHdG or 8-oxodG))	The most widely studied product of DNA oxidation. It is an oxidised derivative of deoxyguanosine (a nucleoside that is used for DNA synthesis) A DNA base modification and the major DNA lesion induced by oxidative stress
A priori	Rather than empirical observation, a priori is based on theoretical deduction. In the context of this report, setting the method in advance before knowing the results
Active Transport	The movement of ions or molecules across a cellular membrane from an area or region of lower concentration to an area or region of higher concentration via transporters and requiring the consumption of energy
Ambient air	Outdoor air
A MeaSurement Tool to Assess Systematic Reviews (AMSTAR 2)	A tool to assess the quality of meta-analyses
Angiogenic	The formation of new blood vessels involving the migration, growth and migration of endothelial cells
Apoptosis	Process of programmed cell death
Air Quality Index (AQI)	A tool used to communicate how polluted the air currently is, or predicted to be (forecast)
Arachidonic acid	A polyunsaturated fatty acid
Aryl hydrocarbon receptor (AhR)	A protein present in the cells that regulates the metabolism of xenobiotics (including chemicals, such as polycyclic aromatic hydrocarbons, that are foreign to the body)
Ascorbate (AA)	A salt, ester, or the anion of ascorbic acid (vitamin C). It is a naturally occurring organic compound with antioxidant properties.
Atherosclerosis	Vascular disease caused by the build up and deposition of plaque of fatty material in the inner lining of an artery
Benzo[a]pyrene (BaP)	A polycyclic aromatic hydrocarbon
Benzo[b]fluoranthene (BbF)	A polycyclic aromatic hydrocarbon
Bias	In this report, bias refers to the relationship observed between exposure and disease that may not reflect the true relationship, due to a systematic error (bias) in the selection of the study participants or in the measurement/assessment of exposure or disease.
Black carbon (BC)	A form of particulate air pollution produced from incomplete combustion of fuels and is a major part of soot.
Blastocyst	5-6 day post-fertilisation mammal embryo formed of an inner mass of rapidly dividing cells and outer layer of trophoblast.

Body Mass Index (BMI)	The body mass divided by the square of the body height, normally expressed in units of kg/m^2
Bulk flow/solvent drag	The movement of ions across cell membranes by bulk transport following the movement of water rather than being facilitated by ion channels or cellular pumps
Cadmium	The chemical element of atomic number 48, a silvery-white metal resembling zinc
Caliber	The diameter of the inside of a cylinder
Carbon monoxide (CO)	A poisonous gas produced when carbon-containing fuels burn without an adequate supply of oxygen
Carcinogenesis	The formation of a malignant growth (cancer)
Cardiac output	The effective volume of blood pumped by the heart per unit of time (generally litres per minute); it usually refers to left ventricle output
Case-control studies	A type of observational study in which patients who have developed a disease are compared to patients who do not have the disease, with research into their past exposure to decipher a supposed causal attribute
Case-cross over studies	An epidemiological technique involving comparing ambient conditions during a period when an individual suffered from some effect on health with a period when no such effect occurred (often the same day of the week within the same month).
Chemokines	A large group of small proteins that belong to the family of cytokines. They stimulate the migration of cells, particularly white blood cells, and play a central role in the development and homeostasis of the immune system.
Cleft	In this context, usually refers to an abnormal fissure or opening especially when resulting from failure of parts to fuse during embryonic development
Copy Number (Cn) variation	A phenomenon in which sections of the genome (an organisms complete set of DNA) are repeated; the number of repeats in the genome varies between individuals
Coefficient	The quantification of the association between an explanatory variable (for instance NO_2) and the response (for instance health outcome) in a regression model. Positive coefficients (>0) indicate that increases in the explanatory variable are associated with increases in the health outcome
Cohort study	A longitudinal observational study that follows groups of individuals (cohorts) who are free of disease at the outset, but who vary according to exposures of interest (for example environmental exposures, lifestyle, behaviours). The individuals are followed over time to see who develops disease. Disease rates are compared between groups according to exposure (either presence/absence (for example smokers vs non-smokers) or groups with different levels of exposure (for example for air pollution)), to evaluate whether the exposure influences risk of a particular disease outcome.
Committee on the Medical Effects of Air Pollutants (COMEAP)	A Department of Health and Social Care (DHSC) expert advisory committee that advises the UK government on the health effects of air pollutants.
Concentration response function (CRF)	A function that expresses the quantitative relationship between concentrations (for instance of an air pollutant) and the (health) outcome of interest. This term is often used to refers to the function adopted to estimate effects associated with pollutants
Confidence interval	If it is possible to define 2 statistics t1 and t2 (functions of sample values only) such that, θ being a parameter underestimate, $P(t1 \leq \theta > t2) = \alpha$

	where α is some fixed probability (for example 0.95 or 95%), the interval between t1 and t2 is called a confidence interval. The assertion that θ lies in this interval will be true, on average, in a proportion α of the cases when the assertion is made. For example, 95% confidence intervals are calculated in such a way that, in the absence of bias, 95% of such intervals will include the parameter that is being estimated
Confounding	Confounding is the masking or distortion of the relationship between the exposure and health outcome of interest by another variable. To be a confounding factor the other variable must be correlated with both the exposure and health outcome (but not on the causal pathway between the two). In this situation the effect of the exposure and the effect of the confounding factor are mixed together, masking or distorting the true relationship between the exposure and health outcome. This is overcome by adjusting for the confounding factor in the analysis, to separate the independent effects of both the exposure and the confounding factor.
Correlation	The quantification of the relationship between 2 variables; often used to refer to how close 2 variables are to having a linear relationship with each other. A correlation of zero indicates no relationship with values of +1 / -1 corresponding to perfect positive / negative correlation
Cotinine	A chemical substance formed by the breakdown of nicotine in the body. It is widely used as an indicator of recent exposure to nicotine.
Covariance	A measure of the joint variability of 2 variables, covariance describes the way that 2 variables deviate from their expected values (means). If 2 variables are independent, then the covariance is zero. If the 2 variables are positively correlated, then the covariance is >0. Covariance is closely related to correlation, which is the normalised (unitless values between -1.0 and 1.0) version of covariance
C-reactive protein (CRP)	A protein made by the liver. The concentration of CRP increases when there is inflammation in the body.
Crown Rump Length (CRL)	First-trimester measurement of the length of an embryo or fetus from the top of the head (the crown) to the base of the buttocks (rump)
Cross-sectional studies	A type of observational design that measures the prevalence of health outcomes or determinants of health, or both, in a population at a point in time or over a short period.
Cytokine	Proteins, secreted by cells, that carry signals to neighbouring cells. They have an effect on the immune system; examples include interleukins.
Cytotrophoblast	The inner cellular layer of the trophoblast covering the placental villi that gives rise to the syncytiotrophoblast and extravillous trophoblast
Diesel Exhaust (DE)	A mixture of gaseous and particulate. The emissions produced by a diesel-fuelled internal combustion engine
Decidua	Transformed endometrial layer that lines the uterine cavity during pregnancy
Decidual Stroma	The cellular component of the decidua tissue following decidualisation of the endometrial stromal cells
Dibenz[a,h]anthracene (DBA)	A polycyclic aromatic hydrocarbon
Differentially methylated regions (DMR)	A region of the genome with multiple consecutive methylated CpG sites (regions of DNA where a cytosine nucleotide is followed by a guanine nucleotide).possibly involved in the regulation of gene transcription (the first step in gene expression)
Diffusion	The movement of molecules from an area of greater concentration to an area of lesser concentration

Displacement Loop (D-loop)	A DNA structure where the 2 strands of a double-stranded DNA molecule are separated and held apart by a third strand of DNA. D-loops occur during DNA repair and in telomeres and, relevant to a paper reviewed in this report, in mitochondrial circular DNA molecules.
Deoxyribonucleic acid (DNA)	The genetic material of the cell. A double-strand of complex molecules which contains the genetic information for the development and functioning of an organism.
DNA adduct	A form of DNA damage caused by covalent attachment of a chemical to DNA.
DNA methylation	The addition of a methyl group to DNA (specifically at a cytosine base followed by a guanine base that is called a CpG site)
Doppler velocity ultrasound	Non-invasive technique used during pregnancy to evaluate blood flow and assess the perfusion of the placenta, and fetal brain, heart and liver perfusion.
EDD	Estimated date of delivery of a baby ('due date') calculated from the first day of the last menstrual period
Endometrium	The mucus membrane lining of the uterus during the menstrual cycle
eNOS	An enzyme, endothelial nitric oxide synthase. In the endometrium, it plays a role in implantation of the blastocyst (the early stage of an embryo)l
Epidemiological studies	Investigations of diseases conducted at a population level
Epigenetics	The study of how cells control gene activity without changing the DNA sequence.
Estrogen Receptor (ER)	A protein found inside cells that is activated by the hormone estrogen.
European Study of Cohorts for Air Pollution Effects(ESCAPE)	A collaboration of more than 30 European cohort studies, aimed at quantifying health impacts of air pollution.
Extravillous trophoblast (EVT)	Differentiated trophoblastic cells that detach from cytotrophoblast columns of the anchoring villus and migrate through the decidua down to the superficial myometrium.
Exocytosis	The release of the contents of a cell vacuole to the exterior
Facilitated diffusion	The process of passive movement of molecules or ions along a concentration gradient across a biological membrane via specific transmembrane integral proteins
Fetus	An unborn offspring in a mammal. In humans, more than 8 weeks after conception
Fibrinolysis	The breakdown of fibrin within blood clots
Free radical	An especially reactive atom or group of atoms that has one or more unpaired electron. They can damage cells, proteins, and DNA by altering their chemical structure
FT3	Unbound, free fraction of triiodothyronine – one of 2 major hormones produced by the thyroid gland (the other being thyroxine, T4)
FT4	Unbound, free fraction of thyroxine – one of 2 major hormones produced by the thyroid gland (the other being triiodothyronine, T3)
Gestational Days (GD)	The days after conception
Genotype	The genetic constitution of an individual organism
Germinal	The earliest stage of development
Gestational adjusted birth weight	The birth weight of an infant, accounting (adjusted for) for the estimated gestational age at the time the infant was born
Gestational age	The clinical term given in weeks to describe human development timed from the first day of the last menstrual period (LMP).
Glomerular filtration	The filtration of waste products out of the blood into the urine by the kidney.
Glutathione peroxidase (GPX)	An antioxidant enzyme that catalyses the reduction of hydrogen peroxide to water, providing protection from oxidative damage.

Glutathione (GSH)	An antioxidant capable of preventing damage to important cellular components caused by reactive oxygen species
Glutathione transferase (GST)	A family of detoxification enzymes that catalyse a broad range of reactions that involve the addition of glutathione to compounds
Haemodynamics	The dynamics of blood flow
Heart rate	The number of heart beats per minute
Health Effects Institute (HEI)	An independent, non-profit corporation specialising in research on the health effects of air pollution, typically receiving balanced funding from the USEPA and the motor vehicle industry.
Heterogeneity	Diversity in character or content
Histone	A basic protein that provides structural support for a chromosome. DNA is wrapped around complexes of histone proteins.
Histotrophic	In early placenta development the initial transfer of nutrition from maternal tissue (mainly the decidual glands) to the developing fetus (histotrophic nutrition) compared to later intraplacental blood-borne nutrition (haemotrophic nutrition)
Homeostasis	The maintenance of a stable equilibrium, especially through physiological processes
Hazard Ratio (HR)	Hazard ratio is a type of relative risk: a measure of whether the risk of a particular event occurring differs between groups, from analysis which takes into account not just occurrence of events, but also time to events (survival analysis).
Health risks of air pollution in Europe (HRAPIE)	The World Health Organization's health risks of air pollution in Europe – HRAPIE – project
High sensitivity CRP (hsCRP)	Detected by a CRP test that is more sensitive than a normal CRP test. See C-reactive protein (CRP) definition
Hydrocarbon	A compound that contains carbon and hydrogen.
Hyperdynamic	Marked by excessive muscular activity
Hypertension	Persistently high blood pressure
I^2	A measure of heterogeneity. It is the percentage of variability in effect estimates in a meta-analysis that is due to heterogeneity between studies rather than sampling error (chance)
Iatrogenic preterm birth	Deliberate preterm birth due to intervention, either to prevent intrauterine death or prevent severe disease in the mother, for example worsening preeclampsia
Immunoglobulin	Glycoproteins produced by plasma cells that function as antibodies
Inflammation	The body's response to protect itself from an irritant, injury or infection
Interleukins (IL)	A group of proteins that regulate inflammatory and immune responses. An interleukin is a type of cytokine
Interquartile range (IQR)	A measure of the spread of the middle 50% of a set of data. It represents all the data greater than or equal to the 25th percentile and less than or equal to the 75th percentile
Intracellular	Located or occurring within a cell or cells
Intrauterine	Within the uterus
Intrauterine fetal death (IUFD) or Stillbirth	In the UK, stillbirths are legally defined as being 'born after 24 completed weeks gestation showing no signs of life' and can only be registered as stillbirths if they meet this definition. Deaths occurring prior to 24 weeks' gestation are not registered and are legally defined as miscarriages. Different timing cut-offs, typically ranging from 20-28 weeks' gestation, are used for defining stillbirth in different countries. The definition recommended by WHO for international comparison is a baby born with no signs of life at or after 28 weeks' gestation.
Intrauterine growth restriction (IUGR)	Drop in fetal growth as evaluated by ultrasound measurements with advancing gestation
Intrauterine inflammation (IUI)	Inflammation of the fetal membranes

Introns	A sequence within a gene that does not code for protein
Isoprostanes	Compounds formed from the free radical oxidation of arachidonic acid (a polyunsaturated fatty acid).
King's College London (KCL)	King's College London is a public research university located in London
Lamina	A thin layer of tissue
Land use regression (LUR)	The land use regression model is an exposure assessment tool frequently used in air pollution epidemiological studies to estimate the concentration of air pollution at unmonitored locations
Leptin (LEP)	A protein hormone that is involved in the regulation of appetite
Linear regression	Linear regression is a statistical method that attempts to model the relationship between 2 variables by fitting a linear equation to observed data. A linear regression line has an equation of the form Y = a + bX, where X is the explanatory variable and Y is the dependent variable. The slope of the line is b, and a is the intercept (the value of y when x = 0).
Long interspersed nuclear element-1 (LINE-1)	A family of DNA repeat sequences with about 500,000 copies. As LINE-1 constitutes approximately 17% of the human genome, the extent of LINE-1 methylation is regarded as a surrogate marker of global DNA methylation
Longitudinal studies	Epidemiological techniques that involve following populations or individuals over a period of time
Low birth weight (LBW)	Birth with a weight of <2,500g
Lymphocyte	A type of white blood cell that is part of the body's immune system
Macrophage	A large white blood cell that specialises in engulfing and digesting cellular debris, pathogens and other foreign substances in the body
Mechanistic studies	A study or test designed to provide information concerning the molecular, cellular, or physiological mechanisms by which substances exert their effects on living cells and organisms
Mesenchyme	Loosely organised undifferentiated mesodermal (middle layer) cells that give rise to such structures as connective tissues, blood, lymphatics, bone, and cartilage
Meta-regression	An extension to subgroup analyses that allows the effect of continuous, as well as categorical, characteristics to be investigated, and in principle allows the effects of multiple factors to be investigated simultaneously (although this is rarely possible due to inadequate numbers of studies)
Meta-analyses	A statistical method used to combine the results of a number of individual epidemiological studies
Metabolite	A substance produced by a metabolic process
microRNA (MiRNA)	Single-stranded, short (approximately 21 nucleotides) non-coding RNA molecules that play important roles in regulating gene expression
Miscarriage	The UK defines miscarriage as the loss of a pregnancy up to 24 weeks of gestation. Other countries may use different timing cut-offs. The World Health Organization uses 28 weeks.
Mitochondria	A rod-shaped structure found in the cytoplasm of cells that generate most of the energy needed to power the cell's biochemical reactions
Mitochondrial DNA (MtDNA)	The circular chromosome found inside mitochondria - MtDNA can be used as a biomarker of mitochondrial response and dysfunction in the presence of oxidative damage
Multi/two pollutant model	A regression model containing 2 (or more) pollutants, plus other covariates (for example possible confounders such as socioeconomic status). Aims to separate out the independent effects of each pollutant.
Myometrium	The muscular layer of the wall of the uterus
Necrosis	Death of cells of a living tissue

Neonate	A newborn baby (typically up to 28 days (4 weeks) after birth)
National Health Service (NHS)	The publicly funded healthcare system in England, and one of the 4 National Health Service systems in the United Kingdom.
Nitrogen Dioxide (NO_2)	A gas formed when high temperatures (for example inside a car engine) cause oxygen and nitrogen in the air to react together
Newcastle Ottawa Scale (NOS)	A tool used for assessing the quality of non-randomised studies included in a systematic review and/or meta-analyses
Oxides of nitrogen (NO_x)	A mixture of gases composed of nitrogen and oxygen (typically nitric oxide (NO) and nitrogen dioxide (NO_2))
Nanoparticle (NP)	A very small particle, less than 0.1 µm (100nm) in diameter
Oxidative potential of ascorbate (OP^{AA})	See oxidative potential definition.
Oxidative potential of glutathione (OP^{GSH})	See oxidative potential definition.
Odds ratio (OR)	An odds ratio (OR) is a measure of association between an exposure and an outcome. The OR represents the odds that an outcome will occur given a particular exposure, compared to the odds of the outcome occurring in the absence of that exposure.
Organogenesis	Phase of the fetal development after the embryogenesis during which fetal organs form
Oxidative potential (OP)	The capacity of particles to cause oxidative reactions, measured by the depletion of specified target reagents
Oxidative stress	The imbalance between reactive oxygen species and antioxidant defences, leading to potential damage
Ozone (O_3)	An oxidant gas. Ozone is a secondary pollutant produced by the interaction of oxides of nitrogen and volatile organic compounds in the presence of sunlight.
Particulate matter (PM)	A mixture of solid particles and liquid droplets found in the air that are so small they can be inhaled
Phagolysosome	A digestive vesicle formed by the fusion of a phagosome (for example, a macrophage containing ingested material) and a lysosome (a vesicle containing hydrolytic enzymes)
Pinocytosis	The ingestion of liquid droplets into a cell
Placenta	Fetal organ attached by the umbilical cord to the fetal abdomen, that provides the mammal fetus with oxygen, essential nutrients and water during pregnancy
Placental growth factor (PlGF)	A protein mainly produced by the trophoblast which is essential for placental angiogenesis (formation of new blood vessels)
Plasma	In the studies in this report: the colourless fluid part of the blood in which corpuscles (blood cells) are suspended
PM_{10}, $PM_{2.5}$	The mass concentrations (expressed in µg/m^3) of particles of generally less than 10 µm and 2.5 µm in the ambient air. More formally, PM_{10} is the mass concentration of particles passing through the inlet of a size selective sampler with a transmission efficiency of 50% at an aerodynamic diameter of 10 micrometres. $PM_{2.5}$ refers to the mass concentration of particles passing through the inlet of a size selective sampler with a transmission efficiency of 50% at an aerodynamic diameter of 2.5 micrometres
Polycyclic aromatic hydrocarbon (PAH)	A group of solid or semi-volatile organic compounds. PAHs form when complex organic substances are exposed to high temperatures or pressures.
Postnatal	Happening or existing after birth
Parts per billion (ppb)	This term describes the concentration of a pollutant in air in terms of volume ratio. A concentration of 1 ppb means that for every billion (10^9) units of air, there is one unit of pollutant present

Parts per million (ppm)	This term describes the concentration of a pollutant in air in terms of volume ratio. A concentration of 1 ppm means that for every million units of air, there is one unit of pollutant present
Preeclampsia	Multisystemic placental-related disorder of the second half of human pregnancy characterised by the onset of high blood pressure and proteinuria (increased levels of protein in the urine).
Prenatal	Happening or existing before birth
Preterm birth (PTB)	A live birth or stillbirth at a gestational age of < 37 weeks. There are also sub-categories of preterm birth, based on gestational age; extremely preterm (<28 weeks); very preterm (28 to 32 weeks), moderate to late preterm (32 to 37 weeks). Sub-category timescales can vary between studies
Prospective cohort study	A cohort study is prospective when measurements of exposures and other relevant characteristics are made before the cases of disease occur
Phosphatase and tensin homolog (PTEN)	A protein that plays an important role in cell growth, proliferation and differentiation
Public Health England (PHE)	An executive agency of the Department of Health and Social Care. UK Health Security Agency from 1 October 2021
Pulsatility index (PI)	Calculation of Doppler measurements of systolic and diastolic velocities in the uterine, umbilical, or fetal circulations.
Quantitative real-time polymerase chain reaction (qPCR)	A laboratory technique of molecular biology that is used to measure gene expression
Quantification of Air Pollution Risks in the UK (QUARK)	A subgroup of COMEAP (see COMEAP definition)
Quinone	A member of a class of cyclic organic compounds containing 2 carbonyl groups (a carbon atom double-bonded to an oxygen atom; C=O); quinones are highly redox active molecules that can lead to the formation of reactive oxygen species
Randomised control trial	An experimental design in which treatments (or exposures in the case of toxicology) are allocated to a subject on a random basis. Commonly, a number of subjects are randomly allocated to a small number of treatment groups
Reactive Oxygen Species (ROS)	Collective term for chemically reactive species.that can generate free radicals and/or cause oxidative damage.
Red cell	Red blood cell: contains haemoglobin, which transports oxygen around the body .
Redox signalling pathway	Specific, and usually reversible, oxidation-reduction reactions, by a reactive species, within a group of molecules in a cell. Redox signalling is important in many physiological and pathological processes
Relative risk (RR)	A relative risk is a ratio of the probability of an event (for example a birth outcome) occurring in the exposed group (for example exposure to a pollutant) versus the probability of that same event occurring in the non-exposed group.
Retrospective cohort study	A cohort study is retrospective when measurements of exposures and other relevant characteristics are made after the cases of disease occur. The cohort study is constructed retrospectively (for example from routine records), but critically (and this is what distinguishes it from the case-control design) the individuals who form the cohort(s) are not selected on the basis of whether a disease outcome has occurred or not. They are selected on the basis of some relevant characteristic at the start of the follow-up time period, for example all workers in factory X between 1970-1980.
Standard Deviation (SD)	The measure of dispersion of a set of data from its mean

Sensitivity analysis	The study of how the uncertainty in the output of a mathematical model or system (numerical or otherwise) can be apportioned to different sources of uncertainty in its inputs, for example methods, models, assumptions
Socio-economic status (SES)	The position of an individual or group on the socioeconomic scale, which is determined by a combination of social and economic factors such as income, amount and kind of education, type and prestige of occupation, place of residence,
Simple Diffusion	The process by which solutes are moved along a concentration gradient in a solution or across a semipermeable membrane without the need for transportors and use of energy.
Single Pollutant Model	In epidemiology: A regression model containing a single pollutant, plus other covariates (for instance possible confounders such as socioeconomic status)
Small for gestational age (SGA)	A birth weight below the 10th percentile of the sex-specific birth weight for gestational age
Sulphur Dioxide (SO_2)	An acidic gas formed by the oxidation of sulphur (found in fossil fuel) during combustion
Soluble fms-like tyrosine kinase 1 (sFlt-1)	A protein that binds to placental growth factor and thereby inhibits its activity
Stillbirth	In the UK: A fetal death occurring after 24 weeks of gestation (see Intrauterine fetal death IUFD). Different timing cut-offs, typically ranging from 20-28 weeks' gestation, are used for defining stillbirth in different countries. The definition recommended by WHO for international comparison is a baby born with no signs of life at or after 28 weeks' gestation.
Stroke volume	The volume of blood pumped from a ventricle of the heart in one beat
Syncytiotrophoblast	Layer of multinucleated trophoblastic cells covering the surface of the villous trees and is the main site of exchange for drugs and metabolites, nutrients, waste products, and gases between the maternal and fetal circulations
Syncytium	Layer of cells containing several nuclei.
Term low birth weight (TLBW)	Birth at term (≥37 weeks) with a weight <2,500 g
Time-series study	An epidemiological method that focuses on the relationship between outcome (for example number of deaths or hospital admissions in a population) and explanatory variable (for example pollutant concentrations) using measures of these variables at regular time intervals, for example days
Transcription factors	Proteins involved in the process of transcribing genetic information from DNA to messenger RNA
Trimester	A division of the duration of pregnancy, which is divided into 3 trimesters (each lasting approximately 3 months): trimester 1 (T1): 1-12 weeks' gestation trimester 2 (T2): 13-27 weeks' gestation trimester 3 (T3): 28-40 weeks' gestation
Trophoblast	Specialised placental cells formed from the outer layer of the blastocyst trophoectoderm that differentiate into cytotrophoblast, syncytiotrophoblast and extravillous trophoblast.
Thyroid stimulating hormone (TSH)	A glycoprotein hormone produced by the pituitary gland which stimulates thyroid hormone production by the thyroid gland
Tumor necrosis factor-alpha (TNF-α)	An inflammatory cytokine produced by the immune system
United States Environmental Protection Agency (USEPA)	An independent executive agency of the United States federal government tasked with environmental protection matters
Umbilical cord (UC)	A flexible cordlike structure containing normally 2 arteries and 1 vein connecting the mammalian fetus to the placenta during gestation

UK Health Security Agency (UKHSA)	A government agency in the United Kingdom, responsible since 1 October 2021 for UK-wide public health protection and infectious disease capability, and replacing Public Health England.
Ultrafine particles (UFP)	Particles of less than 100nm (in any dimension) in diameter
Vasoactive	Affecting the blood vessels, by changing their relaxation (dilation) or contraction (constriction)
Ventricle	The 2 lower chambers of the heart, which pump blood to the rest of the body.
Very low birth weight (VLBW)	Birth with a weight <1500g
Villi	Elongated projections set closely together on a surface that increase the surface exchange area between the maternal blood and fetal circulations
Villous Tree	Villous trees are the main anatomical structure of the placenta
Volatile organic compounds (VOCs)	A very large group of organic compounds with a low boiling point and therefore easily become vapours or gases.
World Health Organization (WHO)	A specialised agency of the United Nations responsible for international public health
Whole Pregnancy (WP) exposure	Exposure across the whole period of pregnancy (however long it was)
Xenobiotic	Relating to or denoting a substance that is foreign to the body, typically referring to synthetic chemicals

Chapter 1
Introduction

1.1 Background

There is increasing concern, both from the medical community and members of the public, of the potential for exposure to air pollution to negatively affect adverse birth outcomes. In 2016, the Committee on the Medical Effects of Air Pollutants (COMEAP) identified early life effects and adverse birth outcomes as a topic for consideration. Since then, there have been a number of epidemiological and toxicological studies on this topic and the Royal College of Physicians and Royal College of Paediatrics and Child Health published its report 'Every breath we take: the lifelong impact of air pollution' (RCP, 2016). As there has been growing interest from the public and the government about the effects of air pollution on early life, the Department of Health and Social Care (DHSC) requested that COMEAP begin work on this topic. To support COMEAP's work, DHSC approached the National Institute of Health Research (NIHR) Health Protection Research Unit (HPRU) on the Health Impact of Environmental Hazards at King's College London to undertake an evidence-based scoping review on air pollution and adverse birth outcomes. The scope of the review was discussed and agreed with Public Health England (PHE)[1]. This review is available online (Fussell and Walton, 2018). PHE[1] proposed that COMEAP consider the scoping review in more detail and carry out further work needed for preparing a statement on the current evidence linking maternal exposure to ambient air pollutants with adverse birth outcomes.

1.2 Approach

To deliver this work, a Sub-group was established, comprising current COMEAP Members, with additional Members co-opted to cover the specialist expertise required. The Sub-group received support and scientific input from the Scientific Secretariat. Members of the Sub-group were divided into 2 working groups: one focusing on epidemiological evidence and the other on the mechanistic and toxicological evidence. The full list of Sub-group Members is provided in Appendix 1.

Sub-Group Members agreed that the focus would be on preterm birth (PTB), term low birth weight (TLBW) and stillbirth as these are the most prevalent

[1] UK Health Security Agency from 1 October 2021 (Public Health England was replaced by UK Health Security Agency and Office for Health Improvement and Disparities).

adverse birth outcomes, have a considerable burden (as discussed in section 1.6) and are where most of the evidence lies. Members agreed that for the epidemiological studies, the main air pollutants of focus were particulate matter (including $PM_{2.5}$, PM_{10}), nitrogen dioxide (NO_2), and ozone (O_3). For the mechanistic studies, the air pollutants considered were particulate matter (including $PM_{2.5}$, PM_{10}, black carbon), NO_2, polycyclic aromatic hydrocarbons (PAHs) (including benzo[a]pyrene (BaP), diesel engine exhaust, and O_3. It was agreed that the focus would be on single pollutants and not mixtures. This was partly because the regulatory framework for air quality is on a specific air pollutant basis and partly for consistency with other quantification recommendations that use single-pollutant model recommendations. Consideration of multi-pollutant models is complex (see section 4.5) and the methodological implications of this forms a separate ongoing COMEAP workstream.

Sub-Group members addressed the following questions (Appendix 2):

- What is the strength and consistency of epidemiological evidence linking ambient air pollutants with adverse birth outcomes?
- Is there evidence of biological mechanisms by which air pollutants could affect adverse birth outcomes?
- Is the evidence sufficient to propose quantification of specific adverse birth outcomes associated with air pollution in the UK?
- Can recommendations be identified from current meta-analyses or should new meta-analyse(s) be undertaken?

1.3 Structure of this report

Chapter 1 of the report provides an overview of a ‘standard’ pregnancy, the general mechanisms of transfer across the placenta and the mechanisms of toxicity of air pollutants. The report is then divided by epidemiological evidence (Chapter 2) and the mechanistic evidence (Chapter 3).

Chapter 2 reviews the process of selecting (section 2.1.1) and analysing core-meta-analyses (Figure 5) to provide a commentary on the evidence for each pollutant, outcome and exposure window (section 2.1.2). Additional selected studies that had either been recently published or were not included in the chosen meta-analyses were identified (section 2.1.4), and commentaries made on the evidence from these (section 2.2). Overall conclusions on the epidemiological evidence, based on the core-meta-analyses and key studies are summarised.

Chapter 3 provides an overview of the mechanisms by which air pollutants may elicit one or more adverse birth outcome including oxidative stress, inflammation, DNA damage, epigenetic modifications, thyroid hormone synthesis alteration and metabolomic changes, in addition to effects on placental circulations and haemodynamics. The mechanistic evidence reviewed in the report primarily focuses on summarising the evidence on ‘human’

mechanistic studies since it is recognised that animal studies are less useful for understanding the underlying effects of air pollution on adverse birth outcomes in humans (section 3.2).

In the discussion (Chapter 4), we have brought the epidemiological and mechanistic evidence together and made comment on whether there is a key exposure window and/or a key air pollutant. The strengths and limitations of both the report and the evidence base are also discussed.

Chapter 5 outlines the conclusions and recommendations based on the report findings.

1.4 Description of a 'standard' pregnancy and possible influences of air pollution

The beginning of pregnancy is variably defined as either the moment of conception, or the moment of implantation of the fertilised egg into the endometrium. Counting of the duration of pregnancy is conventionally in weeks and days and is calculated from the first day of the last menstrual period (LMP), approximately 2 weeks before conception occurs, until the estimated date of delivery (EDD) (Farchi et al, 2014) which is at 40 weeks. In practice, in many middle- and high-income countries, ultrasound measurement of the fetal crown-rump length (CRL) in the first trimester is now used to confirm the gestational age of the pregnancy, and to calculate the EDD. Definitions of key terms relating to pregnancy and birth outcomes are summarised in Table 1.

Starting from very early in pregnancy, there are marked physiological changes in pregnant women. Increases in plasma volume (up by 40 to 50%) and red cell mass (up by 20 to 30%) and increases in cardiac output, stroke volume and heart rate result in a hyperdynamic state with an increased volume of distribution for both fat-soluble and water-soluble compounds which may, in turn, affect the distribution of inhaled air pollutants (Vinturache et al, 2021). Renal function also increases in pregnancy, with glomerular filtration rate increasing by up to 50%, which may enhance excretion of water-soluble toxins.

Table 1. UK definitions of key terms relating to pregnancy and birth outcomes

Term	Definition	Synonyms
Pregnancy		
Trimester	The duration of pregnancy is divided into 3 time periods or trimesters (each lasting approximately 3 months): Trimester 1: 1-12 weeks gestation Trimester 2: 13-27 weeks gestation Trimester 3: 28-40 weeks gestation	T1, T2, T3
Prenatal	Happening or existing before birth	
Birth outcomes		
Miscarriage	The UK defines miscarriage as a loss of a pregnancy prior to 24 weeks of gestation. Other countries may use different timing cut-offs (see definition for stillbirth below)	Early pregnancy failure Spontaneous abortion
Stillbirth	In the UK, stillbirths are legally defined as being 'born after 24 completed weeks gestation showing no signs of life' and can only be registered as stillbirths if they meet this definition. Deaths occurring prior to 24 weeks' gestation are not registered and are legally defined as miscarriages. Different timing cut-offs, typically ranging from 20-28 weeks' gestation, are used for defining stillbirth in different countries. For example, the definition used by WHO for international comparison is a baby born with no signs of life at or after 28 weeks' gestation[1].	Intrauterine fetal death (IUFD)

[1] Why we need to talk about losing a baby (who.int)

Preterm Birth (PTB)	A live birth or stillbirth at a gestational age of < 37 weeks (NICE, 2019). There are also sub-categories of PTB, based on gestational age; extremely preterm (<28 weeks); very preterm (28 to 32 weeks), moderate to late preterm (32 to 37 weeks).	Preterm Delivery Premature birth
Term	**Definition**	**Synonyms**
Birth Weights		
Low Birth Weight (LBW)*	Birth with a weight of <2,500g (Cutland et al, 2017;NICE, 2014)	
Very Low Birth Weight (VLBW)	Birth with a weight <1,500g (Cutland et al, 2017;)	
Term Low Birth Weight (TLBW)	Birth at term (≥37 weeks) with a weight of <2,500g (Cutland et al, 2017;NICE, 2014)	
Gestational adjusted birth weight	The birth weight of an infant, accounting (adjusting) for the estimated gestational age at the time the infant was born (Oken et al, 2003)	Birth weight corrected for gestational age
Small for Gestational Age (SGA)	A birth weight below the 10th percentile of the sex-specific birth weight for all fetuses at that gestational age (Cutland et al, 2017)	

Intrauterine growth restriction (IUGR)	A fetus of any size who has not achieved its genetically determined potential size regardless of growth centile (Peleg et al, 1998). For example, a fetus that is growing on the 40th centile on ultrasound at 20 weeks, with both parents born with weight and size around the 50th centile, where the growth falls below the 15th centile at 34 weeks. This could be associated with maternal preeclampsia or a viral infection after 20 weeks leading to utero-placental insufficiency (non-genetically determined). IUGR implies evidence of a dynamic process (unlike SGA which is fixed in time that is, a fetal weight below the 10th centile at any gestational age without evidence that the fetus was within the normal 10-90th centile range earlier (during) in pregnancy).	Fetal growth restriction

*Low birth weight included any baby born <2500g regardless of the gestational age at birth. A baby born very preterm which is normally grown for its age may be <2500g and included in this definition. Conversely no baby born at term who is <2500g will be normally grown. Before widespread use of ultrasound in early pregnancy to accurately date pregnancies, gestational age was imprecisely estimated, and many studies relied solely on weight at the time of birth to categorise growth.

1.4.1 The development of the human placenta and fetus

The development of a placenta is characteristic of those mammalian species called the 'placentals'. The majority of mammals, including humans are placentals, meaning that their fetuses are nourished by the placenta during the entire pregnancy and develop(s) entirely inside the placental membranes. One of the principal functions of the placenta is diffusional exchange, and the physical requirements for this impose the greatest influence on the structure of the organ. The definitive placenta is the essential interface between the maternal circulation carrying O_2-rich blood and nutrients, and the fetal circulation during the second and third trimesters of pregnancy. The placenta also provides the fetus with a route for clearance of excretory products, and produces a vast array of proteins, steroid hormones and factors necessary for the maintenance of pregnancy (Jauniaux et al, 2020).

When the utero-placental circulation is fully established and the umbilico-placental circulation is fully functional at the end of the first trimester, materno-fetal exchanges increase rapidly. This means that the placenta, and possibly the fetus, may be exposed to particulate matter (PM) and/or adhered constituents such as metals and PAHs present in the maternal bloodstream. Particulate components such as lead, arsenic, titanium, nickel, carbonaceous material and vanadium may enter the villous mesenchyme (the dominant villous type during the early stages of pregnancy) via pinocytosis.

1.4.2 The role of the placenta

Any adverse birth outcomes caused by air pollution could be a consequence of effects on the mother, fetus and/or placenta. These effects could be the result of a cascade of events occurring in the lung and general circulation of the mother and/or the translocation of ultrafine particles into the placenta towards the fetus. Emerging evidence suggests that pollutants can reach the placenta (Bové et al, 2019; Liu et al, 2021) (section 3.3.1). Such pollutants may then exert direct effects by transferring across the placenta into the fetal tissues and circulation or could act indirectly on the placenta by compromising transplacental oxygen and nutrient transport, inducing maternal/placental inflammatory conditions and/or disrupting hormonal balance. We now discuss the development of, and mechanisms of transfer across the placenta.

Placental organogenesis

Human placentation is almost unique amongst mammals in that it is invasive, and the gestational sac embeds itself completely within the uterine endometrium and superficial myometrium. Human placentation is also characterised by remodelling of the spiral arteries, during which the vessels lose their elastic lamina and smooth muscle coat, and consequently their responsiveness to circulating vasoactive compounds. At term, the spiral arteries that, in the non-pregnant state, transport just a few millilitres of blood per minute need to carry up to 600 ml per minute (Jauniaux et al, 2020).

The trophoblastic cells are the germinal cells forming on the outer layer of the blastocyst. Soon after implantation, the trophoblast differentiates into 2 layers: the syncytiotrophoblast and the cytotrophoblast that will cover the placental villi and form the placental membrane. The terminal villi are the main functional units of the villous

tree (Figure 1A). They are highly vascularised, but by capillaries alone, and are highly adapted for diffusional exchange (Figure 1B). The most distal cytotrophoblast cells break through the syncytium and spread laterally to form the cytotrophoblastic shell separating the placenta from the decidua. Cells on the outer surface of the shell differentiate into non-proliferative, cytotrophoblast cells that migrate into the decidual stroma, collectively called extravillous trophoblast (EVT).

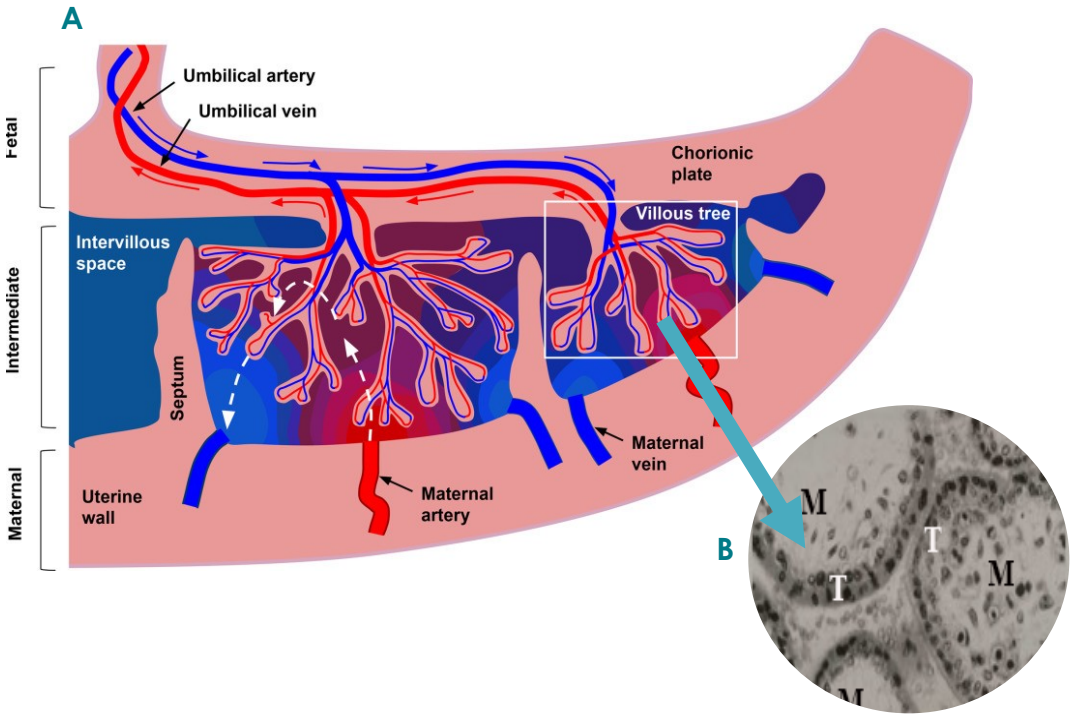


Figure 1. (A) Schematic representation of blood flow through the placenta and surrounding tissue (reproduced from Slator et al, 2018) (left) and (B) microscopic view of first trimester placental villi (right) showing from outside to inside: the trophoblastic (T) layers and the villous mesenchyme (M) containing the fetal capillaries with nucleated red cells (credit: Eric Jauniaux). Figure 1(A) is reproduced from Slator PJ, Hutter J, McCabe L, Gomes ADS, Price AN, Panagiotaki E, Rutherford MA, Hajnal JV, Alexander DC. 'Placenta microstructure and microcirculation imaging with diffusion MRI', Magnetic Resonance in Medicine volume 80(2), pages 756-766, Copyright 2018. Available under Public Licence Deed - Attribution 4.0 International - Creative Commons

The EVT cells migrate down to the superficial myometrium and the wall of the final portion of the maternal circulation down to the basal arteries that connect the radial and spiral arteries. The EVT also form a continuous cytotrophoblastic shell at the interface with the decidua, so that the tips of the spiral arteries are effectively plugged by endovascular EVT cells (Figure 2). This shell and the trophoblastic plugs act like a labyrinthine interface permitting a slow seepage of maternal plasma, but no true blood flow, inside most of the intervillous space during the first 3 months of pregnancy (Burton et al, 2007; Jauniaux et al, 2003). This seepage is supplemented by secretions from the decidual glands and the materno-fetal exchanges are described as histotrophic (Figure 3). The absence of a continuous maternal circulation inside most of the primitive placenta during the first trimester placenta is

essential to control the oxygen level inside the gestational sac during organogenesis (Figure 2). It also adds to the natural defence of the fetus against parasites and viruses when at its most vulnerable (Jauniaux et al, 2020). During most of the first trimester, the secondary yolk sac and the extra-embryonic coelom play important roles as an additional transport route inside the gestational sac (Burton et al, 2007).

The tip of the spiral arteries in the area destined to become the definitive placenta are plugged by EVT plugs, only allowing plasma to filter, whilst in the periphery there is a free flow entering the primitive placenta. This creates focal oxidative stress with degeneration of the above villous tissue and leads to the formation of the placental membranes. At the end of the first trimester, the EVT plugs of the cytotrophoblastic shell are progressively dislocated, allowing maternal blood to flow more freely and continuously within the intervillous space (Figure 2). This transforms the placenta into a haemochorial organ. During this transitional phase, which takes place between 10 to 12 weeks gestation, two-thirds of the primitive placenta disappears, and the chorionic cavity is obliterated by the growth of the amniotic sac (Jauniaux et al, 2003). These events bring the maternal erythrocytes closer to the fetal tissues, facilitating nutrient and gaseous exchange between the maternal and fetal circulations across the villous membrane.

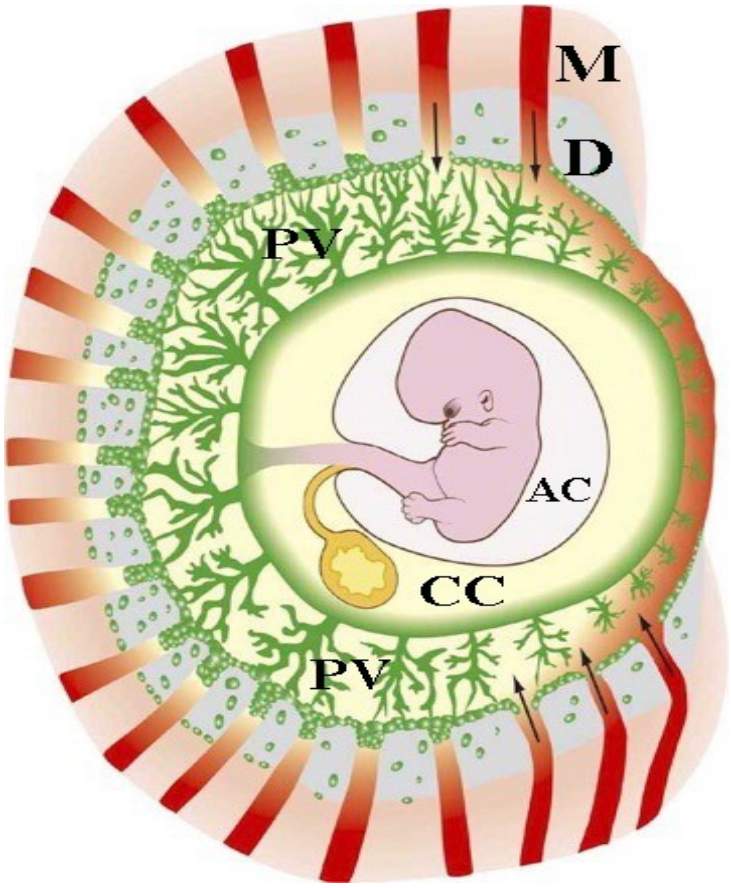


Figure 2. Utero-placental circulation in the first trimester of pregnancy. The tip of the spiral arteries in the area destined to become the definitive placenta are plugged by cytotrophoblastic plugs (shown in green) only allowing plasma to filter whereas in the periphery there is a free flow entering the primitive placenta (arrows) creating focal oxidative stress with degeneration of the above placental villi (PV) and leading to the formation of the placental membranes. M= myometrium; D= Decidua; CC= Chorionic cavity (extraembryonic coelome); AC= Amniotic cavity. Modified from Jauniaux et al (2004). Used with permission of Oxford University Press Journals, from 'Distribution and transfer pathways of antioxidant molecules inside the first trimester human gestational sac', Jauniaux E, Cindrova-Davies T, Johns J, Dunster C, Hempstock J, Kelly FJ and Burton GJ, Journal of Clinical Endocrinology and Metabolism volume 89(3), Copyright 2004; permission conveyed through Copyright Clearance Centre, Inc

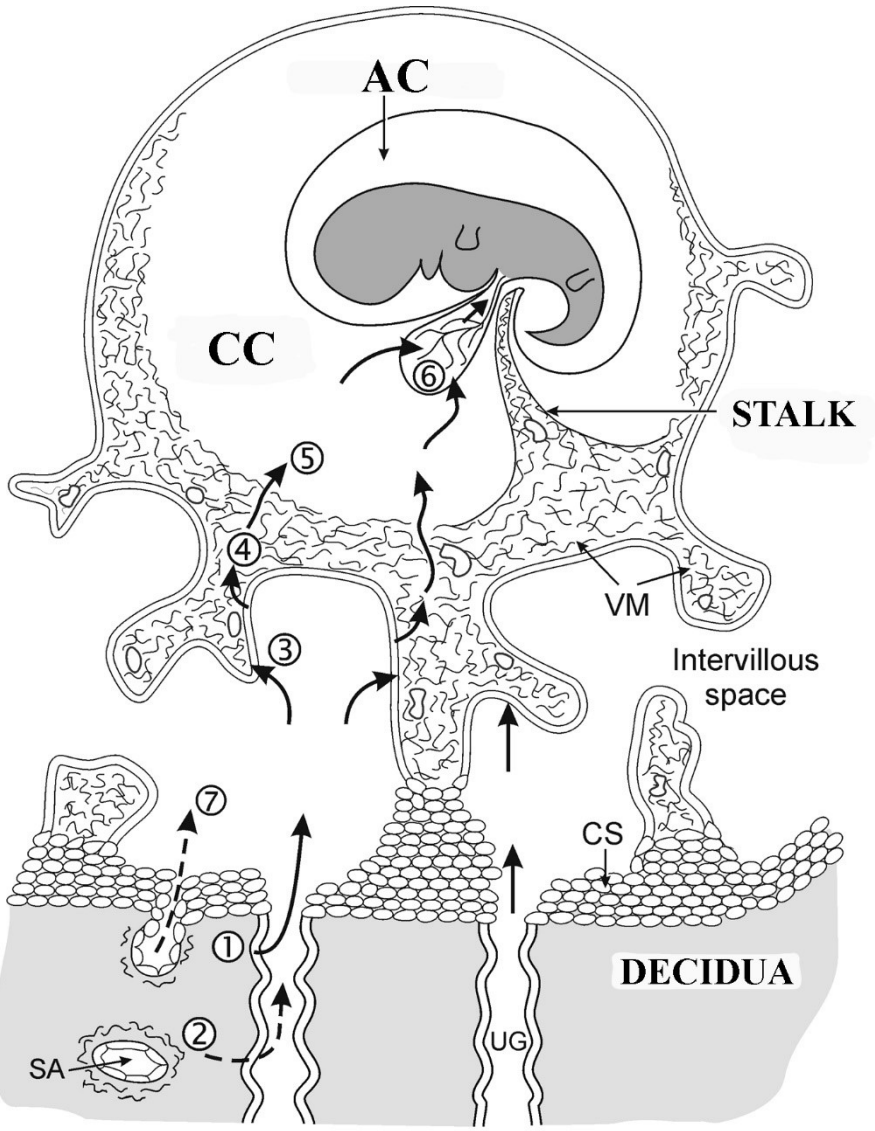


Figure 3. Histiotrophic pathway between the mother and the fetal compartments during the first trimester of pregnancy. AC=Amniotic cavity CC= Chorionic cavity; VM= Villous mesenchyme; CS= Cytotrophoblastic shell; UG= Uterine gland; SA= Spiral Arteries. 1)= Uterine gland epithelial secretion; 2) = Trans- epithelial transfer; transfer from the intervillous space into the fetal circulation via: 3) the villous trophoblast, 4) the villous mesenchyme, 5) the chorionic cavity and 6) secondary yolk sac. 7) Maternal plasma filtered by the cytotrophoblastic plugs entering the intervillous space. This figure shows the absence of a continuous maternal circulation inside most of the primitive placenta during the first trimester. Modified from Burton et al, 2001. Used with permission of Elsevier Science and Technology Journals, from 'Nutrition of the Human Fetus during the First Trimester—A Review', Burton GJ, Hempstock J and Jauniaux E, Placenta, volume 22 Supplement A, copyright 2001; permission conveyed through Copyright Clearance Centre, Inc.

General Mechanisms of transfer across placenta

Many of the transport mechanisms are present in the early or primitive placenta and these may be up- or down-regulated throughout the rest of pregnancy to meet the requirements of fetal growth and homeostasis. For a molecule to reach the fetal plasma from the maternal plasma, and vice versa, it must cross the villous exchange membrane, which is made up of the trophoblast, the villous mesenchyme and the

endothelium of the fetal capillary (Figure 1B). The villous membrane becomes thinner as pregnancy advances bringing the villous capillaries closer to the intervillous circulation and thus facilitating transport between the maternal and fetal circulation. There is also some limited transfer between maternal blood in the decidua and the fluid of the amniotic sac.

Exchange across the definitive placenta occurs by one of the 4 following mechanisms (Hay, 1994): simple diffusion, facilitated diffusion, active transport, and pinocytosis (Figure 4). Water, and with it, dissolved solutes, is transferred by bulk flow and is driven by differences in hydrostatic and osmotic pressures between the maternal and fetal circulations across the exchange barrier (Stulc, 1997). The dissolved solutes are filtered as they move through the components of the villous barrier. Water movement may also be via paracellular channels or through pores in the plasma membranes.

Diffusion of any molecule occurs in both directions across the villous barrier. When there is a concentration gradient and/or, for charged species, an electrical gradient, one of these unidirectional fluxes (rates of transfer) is greater in one direction than it is in the other, so that there is a net flux in one direction. Small molecules such as oxygen diffuse rapidly whereas hydrophilic molecules such as glucose will not diffuse easily. Rates of diffusion of many molecules are flow-limited, and thus may vary over gestation with changes in utero-placental and feto-placental blood flow. Changes in concentration and, for charged molecules, electrical gradients between maternal and fetal plasma will also affect rates of transfer.

Some amino acids and glucose are actively transported across the villous membrane at faster rates than would occur by diffusion down concentration gradients (Illsley, 2000; Regnault et al, 2002). Transporter proteins are a large and diverse group of molecules. In the placenta, transporter proteins are found most abundantly in the microvillous and basal plasma membranes of the syncytiotrophoblast. Large proteins such immunoglobulin cross the placenta via pinocytosis. These molecules are entrapped in invaginations of the microvillous plasma membrane of the syncytiotrophoblast, which detach and pass across the intracellular compartment. These vesicles eventually fuse with the basal plasma membrane and undergo exocytosis, releasing their contents into the fetal circulation.

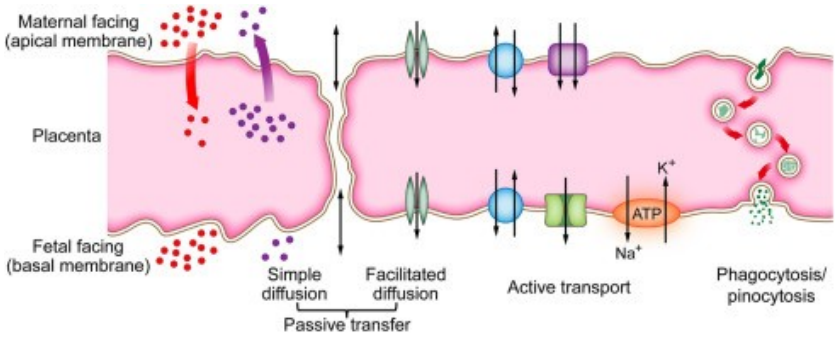


Figure 4. Diagram showing the main mechanisms of placental transfer: passive transfer (simple and facilitated diffusion), active transport, and phagocytosis/pinocytosis. Reprinted from 'Concepts and Models for Drug Permeability Studies (Second Edition), Cell and Tissue Based in Vitro Culture Models', Li J, Hu B, Bai, C, Chen D and Yu Y 'Cell-based in vitro models for placental barrier permeability studies' pages 251-272, Copyright 2024, with permission from Elsevier.

1.4.3 General mechanisms of toxicity of air pollutants

The mechanistic pathways by which airborne gaseous and particulate pollutants exert their wide-ranging detrimental effects on human health are undoubtedly multiple, complex, and interdependent in nature. Indeed, it is highly likely that associations with acute and chronic health effects are the result of multiple components of the ambient pollutant mix acting via different physiological mechanisms.

For example, evidence exists that the effects on the cardiovascular system, and specifically haemostasis, may involve platelet activation, a combined insult of oxidative stress and inflammation, impaired fibrinolysis, circulating microvesicles, epigenetic modifications as well as variations in an individual's sensitivity to air pollution by genetic polymorphisms (COMEAP, 2018a).

A chain of events involving pollutant-induced pulmonary and systemic oxidative stress and inflammation, mediated via redox sensitive signalling pathways, is considered a central mechanism underlying the health effects of air pollution (Borm et al, 2007; Kelly 2003). An excessive oxidative challenge (as opposed to the degree that is essential in physiological redox signalling and for steady-state maintenance of stress response systems) is a serious imbalance between the generation of reactive oxygen species (ROS) and antioxidant protection in favour of the former, causing excessive oxidative damage (Kelly, 2003). Reactive oxygen species is the collective term for oxygen free radicals (any species containing one or more unpaired electrons) and other non-radical derivatives of oxygen that can easily generate free radicals and/or cause oxidative damage. Reactive nitrogen species (RNS) also exist and when overproduced or under eliminated can create damage via nitrosative stress. Biomolecules that are targeted by ROS/RNS include DNA, proteins, lipids, and carbohydrates, giving rise to DNA mutations, lipid peroxidation and protein oxidations. This can lead to appreciable impairment of cellular function and in the worst case to cell death, organ dysfunction and severe disease phenotypes.

The individual pollutants that make up the ambient mix are free radicals (for example NO_2) or have the ability to drive free oxidative reactions (for example PM and O_3).

Once inhaled, O_3 and NO_2 react with proteins and lipids present in the lung lining fluid compartment to produce secondary oxidant species (Langford et al, 1995; Kelly and Tetley, 1997).

As such, cellular responses to concentrations of gaseous pollutants that overcome endogenous antioxidant defences are not the result of a direct reaction with epithelial cell surface components, but instead, are mediated via a cascade of secondary, free radical derived products.

Several more complicated and inter-related pathways exist by which inhaled PM can generate oxidative stress at the air-lung interface: an ability of the particle surface per se to elicit oxidative stress or the introduction of oxidising species such as redox active transition metals present on the particle surface that can assist in the production of hydroxyl free radicals via the Fenton reaction or surface absorbed PAHs that can undergo bio-transformation in vivo into quinone species (Kelly et al, 2007).

On reaching the lung surface, secondary oxidation products arising from all forms of oxidant pollutants, via the activation of transcription factors, initiate cytokine and chemokine generation, adhesion molecule expression and tight junction modification. These responses lead to the influx of activated inflammatory cells from the vasculature into the lung, which in turn generate and release large quantities of free radicals, a second wave of pulmonary oxidative stress accompanied by systemic inflammation and oxidative imbalances. The latter, plus any translocation of ultrafine particles and/or particle constituents across the alveolar membrane into the general circulation paves the way for toxicity within the vasculature and at organ sites distant from the lung.

The biological mechanisms through which gestational exposure to air pollution may cause adverse birth outcomes are not well-understood but are expected to be multifactorial and dependent upon the pollutant mixture, its concentration and composition as well as the exposure window owing to changes in physiologic maturity of the placenta, circulations and fetus as pregnancy advances.

Several mechanisms have however been hypothesised and investigated, including DNA adduct formation, epigenetic changes, inflammation, oxidative stress and hormonal disruption. Such experimental work includes human mechanistic studies and animal and in vitro research. The former is vital to further our understanding in this area owing to specific inadequacies of animal models. For example, whilst spontaneous PTB is rare in most species, placentation as well as labour in humans differs from those in rodents and non-human primates.

1.5 Birth outcomes of interest

Whilst there are a number of birth outcomes that can be studied, the main health outcomes considered were PTB, TLBW and stillbirth as these are the most prevalent and are where most of the evidence lies. Adverse birth outcomes covered in this report, as well as wider definitions for context, are defined in Table 1. PTB remains the largest cause of infant mortality around the world (WHO, 2023). PTB can be spontaneous or iatrogenic (deliberate preterm delivery by induction of labour or

caesarean section, either to prevent intrauterine death or prevent severe disease in the mother, for example worsening preeclampsia). The risk of mortality or childhood impairment increases with decreasing gestational age at the time of birth (Hack and Fanaroff, 1999; Husby et al, 2023).

Intrauterine growth restriction (IUGR) occurs when the transfer of nutrients (including oxygen), between the mother and fetus, via the placenta, is impaired due to utero-placental malperfusion. If this occurs the fetus fails to reach its growth potential and this, when extreme, can result in intrauterine death. In the UK, if this death occurs after 24 weeks of gestation, this is called a stillbirth (Fairburn, 2018). The sequelae of growth restriction vary depending on when the growth restriction occurred and how severe this is prior to birth. Growth restriction is often seen with pre-eclampsia, a multisystem disease of pregnancy, diagnosed by sudden-onset hypertension (>20 weeks of gestation) and at least one other associated complication, including proteinuria, maternal organ dysfunction or uteroplacental dysfunction (Dimitriadis et al, 2023; NHS, 2021).

1.5.1 Inter-relationship between birth outcomes

Many adverse outcomes of pregnancy are inter-related. If a single pollutant exposure early in pregnancy could adversely affect the development of the placenta by altering the physiological transformation of the spiral arteries (terminal branches of the uterine circulation providing maternal blood to the placenta), it could lead to an increase in the risk of all of the adverse birth outcomes described in Table 1. Even if a pollutant only acted directly on fetal metabolism in the third trimester, it could still lead to an increase in most of the adverse outcomes described in Table 1, with the exception of miscarriage.

Preterm birth, both spontaneous and iatrogenic, and growth restriction are clinical manifestations of a wide variety of pathologies that affect pregnancy. They are not outcomes that are specific to environmental exposures. Epidemiological studies can only explore whether there is an increased incidence of these outcomes in women exposed to air pollutants compared with unexposed women.

1.6 UK burden of stillbirth, preterm birth, and low birth weight

Data from the Office of National Statistics (ONS) for England and Wales in 2019, showing number of births, stillbirths (total and categorised by gestational age) and birthweight are presented in Table 2. It shows that stillbirths account for 0.39% of all births, 7.8% of live births are preterm, 6.8% of live births are LBW and 2.63% of live births are low birth weight at term (TLBW). A clear trend between rates of stillbirth and deprivation index is shown in Appendix 3, Table 22 with stillbirth rates for the most deprived decile being over 2.2 times higher than for the least deprived for England and 2.5 times higher for Wales.

Table 2. Summary of data on live births, stillbirths, births by gestational age and birthweight for England and Wales in 2019

Birth Outcomes for 2019 England and Wales	**Total Number**	**Percentage (%)** (For stillbirths, the denominator is total live births and stillbirths. For all other %s the denominator is total live births)
Total births (live and stillbirths)	**642,855**	
Total Stillbirths	**2,485**	**0.39**
Stillbirths by gestational age		
24 to 27 weeks	765	0.12
28 to 31 weeks	455	0.07
32 to 36 weeks	605	0.09
37 to 41 weeks	651	0.10
42 weeks and above	9	0.0014
Total Live births	**640,370**	
Live births by gestational age		
Under 24 weeks	946	0.15
24 to 27 weeks	2,072	0.32
28 to 31 weeks	5,092	0.80
32 to 36 weeks	41,839	6.53
37 to 41 weeks	575,417	89.86
42 weeks and above	12,696	1.98
Low birth weight		
<1,500g	5,640	0.88
1,500g to 1,999g	7,854	1.23
2,000g to 2,499g	29,801	4.65
Low birth weight at term	16,837	2.63

Data sourced from:

https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/livebirths/bulletins/birthsummarytablesenglandandwales/2019

Note that this includes data down to local authority area and for 'small geographic areas'

https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/livebirths/datasets/birthcharacteristicsinenglandandwales

For illustration, an overview of the cost per case and the UK costs of stillbirth and PTB are presented in Appendix 3, Table 23. Associated costs of adverse birth outcomes are high, with the impact of PTB on health services costs estimated to be £3b per year in 2009 (Mangham et al, 2009). For 2013/2014 the annual costs to the health service and through lost productivity resulting from stillbirths has been estimated at £694 million (Campbell et al, 2018). Neither cost estimate can be considered complete as they do not account for some important elements, including grief and stress to parents, other family members and healthcare professionals.

1.7 Exemplar of an effect of environmental toxicity: exposure to tobacco smoke during pregnancy

The effects of exposure to tobacco smoke by active and passive smoking in pregnancy have been well documented. Understanding the effect of exposure to smoking on the placenta may give indications as to how air pollution might also

affect the placenta and is an example for how passive inhalation of an environmental toxin can adversely affects birth outcomes. Although tobacco smoke contains high concentrations of some substances that are not found at similar concentrations in ambient air pollution, it is a source of airborne PM and, to some extent, can be regarded as a high-dose parallel to exposure to air pollution.

1.7.1 Main findings on exposure to passive smoking during pregnancy

There is good supporting evidence from epidemiological studies that passive smoking during pregnancy:

- increases the risks of stillbirth, miscarriages and preterm delivery
- has a limited and dose-dependent effect on fetal growth

Mechanistic studies suggest that these effects are due to excessive placental oxidative stress/chronic inflammation.

The boxes below collate some of the evidence from cohort and case-control studies and systematic reviews on passive smoking in pregnancy. Cohort studies with uncertain/weak methodology or small numbers have been identified with an asterisk (*).

Literature findings on impacts of passive smoking in pregnancy

Miscarriage/Stillbirth

- In a cohort study, the Odds Ratio (OR) of stillbirth in relation to exposure to household members' smoke was 1.321 (95% Confidence Interval (CI) 1.150, 1.517). Sub-group analysis found statistically significant associations among certain age groups only (Yaya and Bishwajit, 2019)

- A systematic review (Pineles et al, 2014) found that second-hand smoke exposure during pregnancy increase the risk of miscarriage by 11% (95% CI: 0.95, 1.31; n = 17 studies)

- A population-based study found that co-exposure to passive smoking and vitamin D deficiency increased the risk of miscarriage (OR = 2.50, 95%CI: 1.63, 3.84) (Lin, 2022)

SGA/IUGR/Birth weight

- Infants of mothers exposed to tobacco smoke (by active or passive smoking) with a plasma cotinine level of ≥3.03 ng/mL showed an OR for SGA of 2.02 (95% CI 1.45 to 2.83) and 2.44 (95% CI 1.73 to 3.44) in the full population and term infants respectively, in a cohort study (Kobayashi et al, 2019)

- Second-hand smoke exposure during pregnancy was associated with reduced birth weight (-56 g, 95% CI: -97 to -16 g) in a cohort study, but no increase in the risk of SGA was found (Ion et al, 2015)
- Hoyt et al (2018) found no evidence for an association with SGA births for second-hand smoke exposure during pregnancy in a cohort study but a case-control study (Niu et al 2015) did find an association
- Birth weight had a significant inverse correlation with the maternal exposure to passive smoking (number of cigarettes smoked by household members) in a cohort study * (Wadi and Al-Sharbatti 2011)
- Lower birth weight due to significant reduction of fat mass and most anthropometric measurements was found following passive smoke exposure during pregnancy in a study of full-term newborn infants (Luciano et al, 1998)
- A case-control study reported an association with severe small for gestational age (SGA) with adjusted OR of 2.51 (95% CI 1.59-3.95) and a dose-response relationship to the number of smokers in the home (Fantuzzi et al, 2008)

Preterm birth (PTB)

- Maternal passive smoking during pregnancy was associated with higher risk of PTB in a case-control study (adjusted odds ratio (aOR)=2.20 (95% CI 1.56;3.12)*) (Luo et al, 2012)
- Second-hand smoke exposure during pregnancy was associated with earlier delivery (-0.19 weeks; 95% CI -0.32 to -0.05) but no increase in the risk of PTB in a cohort study (Ion et al, 2015)
- A study reported associations of maternal exposure to second-hand smoke during pregnancy with PTB: aOR of 1.99; 95% CI 1.13-3.50 for PTB in general, and aOR 2.17; 95% CI 1.22-3.88 for moderate PTB (32 to 36 weeks) (Hoyt et al, 2018)
- A study reported a non-significant association of environmental tobacco smoke exposure with early preterm delivery (aOR 1.56; 95% CI 0.99, 2.46) with an exposure-response relationship with the number of smokers in the home (Fantuzzi et al, 2007)
- A birth cohort study found an increased risk of very PTB (<32 weeks) with OR of 1.98 (95% CI 1.41, 2.76) but not moderate PTB (32 to 36 weeks). The risk of very PTB increased with the duration of exposure (P for trend = 0.0014) (Qiu et al, 2014)

Chapter 2
Epidemiological evidence

2.1 Approach

Evidence from the epidemiological literature on potential associations of air pollutants (particulate matter, carbon monoxide (CO), ozone, nitrogen dioxide, polycyclic aromatic hydrocarbons (PAHs) or traffic-related air pollutants) on 3 key adverse birth outcomes (preterm birth, term low birthweight and stillbirth) was assessed by the epidemiological working group. We have not considered other birth outcomes, such as congenital anomalies, in this report.

The first step for the epidemiological working group was to identify relevant epidemiological evidence. The scoping review conducted by King's College London (KCL) (Fussell and Walton, 2018) was used as a basis for identifying key literature. This was accompanied by an updated literature search, using revised search terms used in the KCL scoping review, as described further in Appendix 4.

The epidemiological evidence was then summarised in 2 steps:

1. All relevant existing meta-analyses were identified, quality assessed and reviewed (section 2.1.2). Meta-analyses were excluded from further consideration if they contained major errors, population overlaps which we regarded as 'intolerable', or if all individual eligible studies included in the meta-analysis were already included in more recent and/or better-quality meta-analyses. From the core meta-analyses (those which were not excluded), a commentary on the evidence was prepared for each pollutant, outcome, and exposure window (section 2.2).

2. Key studies that had either been recently published, or were not included in the selected meta-analyses, were identified (section 2.1.4), and commentaries made on the evidence from these (section 2.2).

It was not the intention of the epidemiological working group to conduct a full systematic review and meta-analysis, but rather to consider the KCL scoping review in more detail, and carry out any further work needed to prepare a statement on the current evidence linking maternal exposure to ambient air pollutants with adverse birth outcomes, as per the terms of reference
(Appendix 2).

2.1.1 Identifying relevant epidemiological evidence

The KCL scoping review was used a basis for identifying key epidemiological evidence. The KCL scoping review summarised meta-analyses conducted to date (publication dates ranging from 2011 to 2017), and identified and summarised studies that had been published since the most recent meta-analysis (publication dates ranging from 2016 to 2018). The studies already included in the KCL scoping review were considered. As this review was published in October 2018, and the literature search was conducted in December 2017, the search terms used were re-run using PubMed to identify epidemiological studies published between January 2018 and March 2019 (search terms are provided in Appendix 4). Following this, alerts from PubMed from this search were set up, and evidence was continuously collated from March 2019 to March 2020 (with the exception of relevant systematic reviews with meta-analyses, which we continued to consider for inclusion until August 2020, as it had become clear that the epidemiological evidence considered in the report would be mainly focused on meta-analyses). Additional studies were identified by reviewing 'grey' literature, being informed of key publications by experts, and by conducting hand-searches (for example, checking the reference lists in recently published reviews). In summary, we considered:

- 18 meta-analyses published from 2011 to August 2020
- approximately 150 individual studies published from 2016 – March 2020

2.1.2 Identifying and choosing meta-analyses to comment on

Identifying relevant meta-analyses

The KCL scoping review included 13 meta-analyses (Dadvand et al, 2013; Estarlich et al, 2011; Estarlich et al, 2016; Klepac et al, 2018; Lamichhane et al, 2015; Li et al, 2017; Pedersen et al, 2013; Sapkota et al, 2012; Siddika et al 2012; Stieb et al, 2012; Sun et al, 2015; Sun et al, 2016; Zhu et al, 2015). A further 5 meta-analyses published up to August 2020 were identified (Giorgis-Allemand et al, 2017; Guo et al, 2019; Liu et al, 2017; Li et al, 2020;Ji Y et al, 2019). Of these 18 meta-analyses, 4 were excluded as they were multi-centre studies using meta-analytical methods to conduct the epidemiological analysis, and did not contain a systematic review element (Dadvand et al, 2013; Estarlich et al, 2016; Pedersen et al, 2013; Giorgis-Allemand et al, 2017). Another study was excluded as it did not examine an outcome of focus (Estarlich et al, 2011), instead reporting on birth weight considered as a continuous variable. Therefore, 13 meta-analyses were examined.

Figure 5 is a flow chart describing the approach taken to identify the 'core meta-analyses' which we regarded as being of good quality. Table 3 summarises which birth outcomes and pollutants each of the 13 meta-analyses examined.

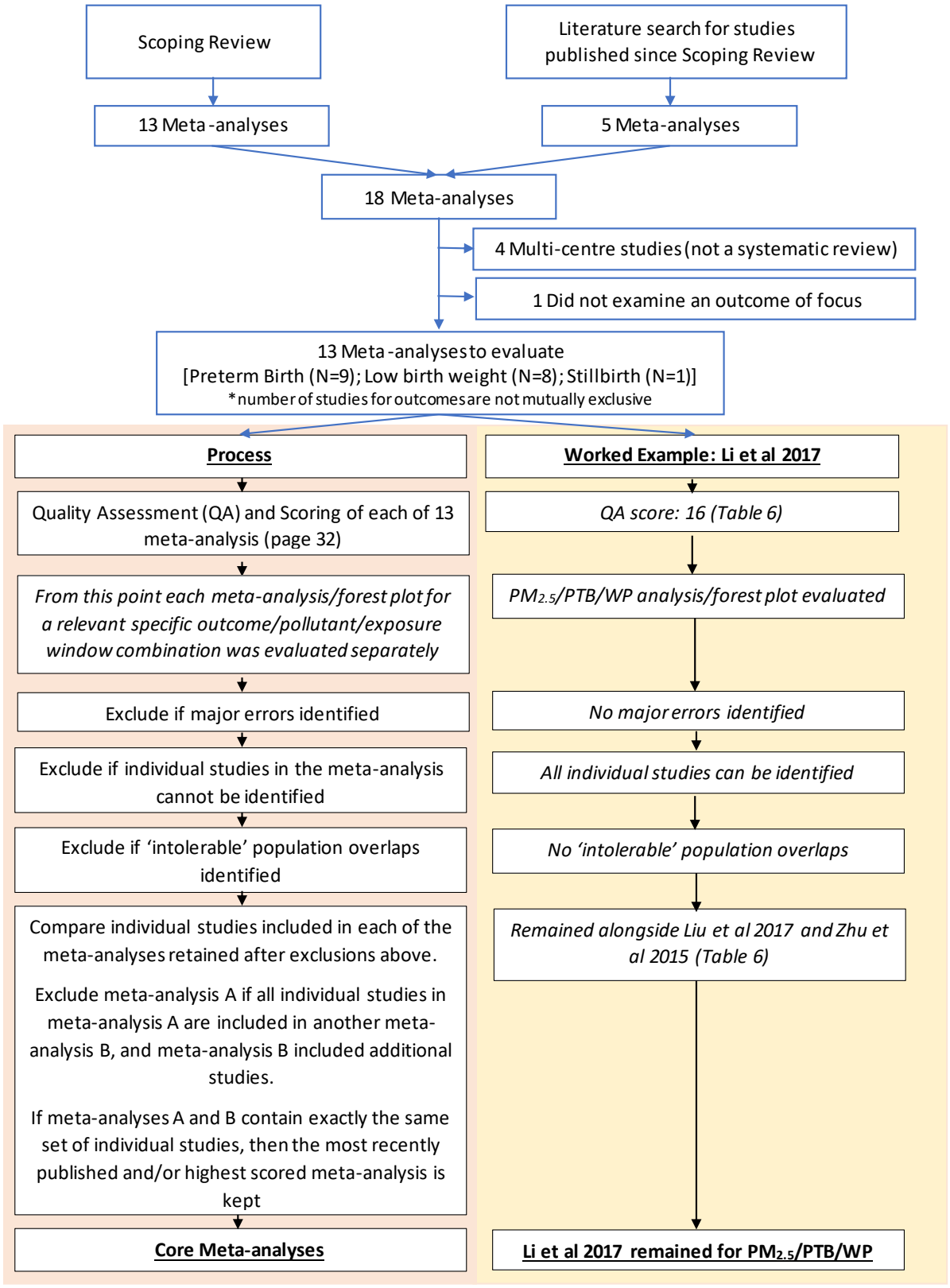


Figure 5. Flow chart describing the approach taken to get to the 'core meta-analyses'. Example provided with Li et al (2017)

Text version of Figure 5

Figure 5 is a flow chart describing the approach taken to identify the 'core meta-analyses' which we regarded as being of good quality.

Eighteen papers reporting meta-analyses were initially identified. Of these, the scoping review had previously identified 13. A literature search for studies published since the scoping review identified 5 additional meta-analyses.

Of these 18 meta-analyses, 5 were excluded: 4 were meta-analyses undertaken as part of multi-centre studies (not systematic reviews) and the birth outcome examined by 1 paper was not within the focus of our review. Therefore, 13 meta-analyses were evaluated.

The process of evaluation involved:

Step 1: Quality assessment (QA)and scoring of each of the 13 meta-analyses.

From this point, each meta-analysis/forest plot for a relevant specific outcome/pollutant/exposure window combination was evaluated separately

Step 2: A meta-analysis was excluded if:

- Major errors were identified
- Individual studies in the meta-analysis could not be identified
- 'Intolerable' population overlaps were identified

Step 3: For the meta-analyses remaining after these exclusions, the individual studies included were compared and the following process adopted:

- Exclude meta-analysis A if all individual studies in meta-analysis A are included in another meta-analysis B, and meta-analysis B included additional studies
- If meta-analyses A and B contain exactly the same set of individual studies, then the most recently published and/or highest scored meta-analysis is kept

Final evaluation: Any meta-analyses remaining after this process were regarded as 'core meta-analyses' for the purposes of our review.

A worked example is given for Li et al (2017):

Step 1: The paper was given a QA score of 16

For the forest plot examining associations of pre-term birth with whole pregnancy exposure to $PM_{2.5}$:

Step 2 found

- No major errors
- All individual studies could be identified
- No 'intolerable' population overlaps

In Step 3, this meta-analysis was compared (in terms of the individual studies it contained) with all the other meta-analyses for pre-term birth and whole pregnancy exposure to $PM_{2.5}$ retained after Step 2. Li et al (2017) included some studies which

were not contained within the other retained meta-analyses, thus it was not excluded.

Final evaluation: Therefore, Li et al (2017) was regarded as a 'core meta-analysis' for associations of pre-term birth with whole pregnancy exposure to $PM_{2.5}$.

End of text version of Figure 5

Table 3. The birth outcomes and pollutants examined in each of the meta-analyses

Pollutant	Birth Outcome		
	PTB	LBW	Stillbirth
$PM_{2.5}$	Klepac et al 2018 Lamichhane et al 2015 Li et al 2017 Liu et al 2017 Sapkota et al 2012 Stieb et al 2012 Sun et al 2015 Zhu et al 2015	Ji Y et al 2019 Li et al 2017 Li et al 2020 Sapkota et al 2012 Stieb et al 2012 Sun et al 2016 Zhu et al 2015	Siddika et al 2016
PM_{10}	Guo et al 2019 Klepac et al 2018 Lamichhane et al 2015 Sapkota et al 2012 Stieb et al 2012	Guo et al 2019 Ji Y et al 2019 Li et al 2020 Sapkota et al 2012 Stieb et al 2012	Siddika et al 2016
NO_2	Guo et al 2019 Klepac et al 2018) Stieb et al 2012	Guo et al 2019 Li et al 2020 Stieb et al 2012)	Siddika et al 2016
O_3	Klepac et al 2018 Stieb et al 2012	Li et al 2020 Stieb et al 2012	Siddika et al 2016

Two meta-analysis studies were excluded in their entirety on the basis that they had major errors (Guo et al, 2019; Ji Y et al 2019). Guo et al (2019) [3] was excluded in its entirety because it contained major transformation errors. For example, in the forest plot for PM_{10} and PTB, estimates from Zhao et al (2015) and Hansen et al (2006) appear to be untransformed and not on the 20 μg/m^3 increment that Guo et al (2019) state. Zhao et al (2015) is on a 10 μg/m^3 increment and Hansen et al (2006) a per interquartile range (IQR) increment. Similarly, Ji Y et al (2019) [4] was excluded as it appeared not to have converted the original ORs onto a comparable scale before meta-analysing.

[3] The major transformation errors were reported to the journal editors. The editors agreed to retreact the paper from the journal.

[4] The editors were contacted highlighting the fact that the original ORs had not been converted onto a comparable scale before meta-analysing. This resulted in the retraction of the paper.

It should be noted that not all the meta-analyses examining LBW, applied a criterion restricting to studies specifically examining LBW at term, for example Stieb et al (2012) (see Appendix 5). However, it was noted that: a) the majority of the individual studies included in such meta-analyses had specifically restricted to TLBW, and conversely b) in some meta-analyses where the meta-analysis examined TLBW, estimates for LBW not restricted to term had, on occasion, been included. It was decided that a meta-analysis would not be removed simply because it included estimates for LBW not restricted to term, as this would result in loss of much valuable information regarding TLBW. It was agreed that the epidemiological working group would comment on the extent to which the meta-analysis results reflected studies examining TLBW versus LBW, bearing in mind that our focus is on the evidence in relation to TLBW.

Quality scoring

The second iteration of 'A MeaSurement Tool to Assess Systematic Reviews' (AMSTAR 2), was used to assess the quality of the meta-analyses (Shea et al, 2017). AMSTAR 2 includes 16 questions, each of which are answered with a 'yes' or 'no' answer. The tool comes with guidance to give an indication on how the criteria should be applied (Shea et al, 2017) (see Appendix 6 for full details).

AMSTAR 2 was designed for systematic reviews of well-defined (usually clinical) interventions, and not designed for environmental epidemiological studies, therefore the AMSTAR 2 score was adapted as follows (see Appendix 7 for full details):

- for some questions, 'partial' was also allowed to be an answer as well as 'yes' or 'no'
- additional guidance was agreed for some questions
- question 9 was divided into 2 parts to be answered separately for the systematic review element, and the meta-analysis element
- addition of our own question 'have the authors addressed potential confounding by smoking in the meta-analysis, and if yes how' to reflect the importance of accounting for maternal smoking in the analysis

We also designed our own scoring system to provide a quantitative score for each meta-analysis, to make it easier to compare the quality across meta-analyses:

- if the answer to a question was 'yes', it scored 1
- if the answer to a question was 'partial', it scored 0.5
- if the answer to a question was 'no', it scored 0

In addition, each question was given a numerical weighting:

- a weighting of 1 was given if the question was not considered to be very important
- a weighting of 2 was given if the question was considered to be somewhat important

- a weighting of 3 was given if the question was considered to be very important

The weighting assigned to each question is summarised in Appendix 8. The numerical answer to the question was multiplied by the weighting to give a modified AMSTAR 2 score. The maximum possible score was 37.

Each meta-analysis was quality assessed by one author (RS, DJ, JF, PD or HW) using the modified AMSTAR 2 score 'blind', that is, without considering the effect estimates reported in each meta-analysis, so as not to bias the score that each meta-analysis was assigned. Some meta-analyses did not report, evaluate, or discuss the elements evaluated by AMSTAR 2 as fully or completely for all the different pollutants and outcomes. In these cases, a modified AMSTAR 2 score was calculated for each pollutant and outcome individually.

We regarded scores of <40% as Low; 40% to 60% as Moderate; >60% as High.

Identifying population overlaps

Some of the meta-analyses contained individual studies that were conducted in the same, or similar, geographical regions and were conducted over the same or overlapping time periods. This would introduce some bias into the meta-analysis, as this contradicts the assumption in meta-analysis that effect sizes (and thus their underlying populations) are independent.

Identification of meta-analyses that contained population overlaps was completed on a per pollutant, outcome and exposure window basis. This was conducted 'blind', that is without considering the effect estimates reported in the meta-analyses. Secondly, if there was a population overlap, we assessed it to determine whether the overlap was 'tolerable' or 'intolerable', as follows:

- an example of a 'tolerable' population overlap

  In Stieb et al (2012), the forest plot examining $PM_{2.5}$, PTB and Whole Pregnancy, includes studies conducted by Huynh et al (2006) (which was conducted in California, USA, between 1999 and 2000) and Wu et al (2009) (which was conducted in Los Angeles and Orange counties, in California, USA, between 1997-2006). The extent of geographical and time period overlap is quite limited. Both studies add considerable value to the analysis, as Huynh et al (2006) adds more subjects from a much larger geographical area, and Wu et al (2009) adds an additional 8 years of data. The population overlap was therefore considered to be tolerable.

- an example of an 'intolerable' population overlap

  In Klepac et al (2018), the forest plot examining $PM_{2.5}$, PTB and Whole Pregnancy includes studies conducted by Brauer et al (2008) and Stieb et al (2016a). Brauer et al (2008) was conducted in Vancouver, Canada, between 1999 and 2002 and Stieb et al (2016a) was a national study conducted across Canada between 1999 and 2008. Therefore, all of the subjects included in Brauer et al (2008) will also be included in Stieb et al (2016a). This 'complete' population overlap was therefore considered to be 'intolerable'.

Some of the overlaps which we regarded as 'intolerable' did not concern complete population overlaps. For example, in Lamichhane et al (2015), the forest plot examining $PM_{2.5}$, PTB and Whole Pregnancy includes studies conducted by Hyder et al (2014) and Kloog et al (2012). Hyder et al (2014) was conducted in Massachusetts and Connecticut, USA, between 2000 and 2006 and Kloog et al (2012) was conducted in Massachusetts, USA, between 2000 and 2008. Potentially therefore all the Massachusetts births in Hyder et al (2014) are also included in the Kloog et al (2012) study. As Kloog et al (2012) covers a smaller geographical area, and only adds an additional 2 years of data, we consider there was insufficient added value from the Kloog et al (2012) study to justify including both studies in a meta-analysis, and this was, therefore, considered to be an 'intolerable' overlap.

A meta-analysis was excluded for an individual pollutant, outcome, and exposure window, if it contained one or more 'intolerable' population overlaps. The exception to this rule was when assessing population overlaps in Li et al (2020), as the authors conducted sensitivity analyses to assess the robustness of the pooled effect by excluding the included individual studies one-at-a-time, and therefore the impact of one 'intolerable' overlap, on the meta-analysis could be assessed. However, if there were 2 or more 'intolerable' overlaps for a pollutant, outcome, exposure window combination, that particular meta-analysis was excluded, as it is impossible to tell from the sensitivity analyses what effect the combination of the overlapping studies may have had.

A summary of the population overlaps within the meta-analyses, per each pollutant, outcome and exposure window, and whether the overlap was determined to be 'tolerable' or 'intolerable' is summarised in Appendix 9.

Other exclusions

Some meta-analyses were excluded either partially, or wholly, for other reasons. If it was not possible to identify the individual studies included in a meta-analysis, the meta-analysis was excluded. For example, Sapkota et al (2012) and Sun et al (2016) did not present forest plots for trimester-specific estimates, and only presented pooled estimates. Therefore, it was not possible to assess population overlaps for these estimates, as the individual studies included in these estimates were not cited in the study. We did write to authors for additional information, but this was not provided. However, it was possible to infer some of the individual studies included for some trimester-specific estimates in Sapkota et al (2012) by cross checking the information in Tables 1 and 2 in Sapkota et al (2012) (which includes a column indicating whether each individual study examined Whole Pregnancy, specific trimesters, or both) with information in Table 3 in Sapkota et al (2012) (which details the number of studies included in each pollutant, outcome and exposure window specific risk estimate); for example if there were only 5 studies that examined a particular pollutant and outcome combination (as inferred by Table 1 or 2), and all 5 of those studies were included in a trimester-specific risk estimate (as inferred by Table 3), then we could deduce the 5 individual studies behind the estimate.

After all the exclusions described above, there were still more than one meta-analysis remaining for some pollutant, outcome, and exposure window combinations. In this case, the individual studies included in each of the meta-analyses were

compared. If all of the individual studies in meta-analysis A were included in another meta-analysis B, and meta-analysis B included additional studies, then meta-analysis B was kept, and meta-analysis A was excluded. If meta-analyses A and B contained exactly the same set of individual studies, then the modified AMSTAR 2 score and the publication date of the meta-analysis was considered (and the most recently published and/or highest scored meta-analysis was kept).

2.1.3 Interpreting evidence from meta-analyses

When we are commenting on summary OR/RRs and CIs from meta-analyses specifically, conventional statistical significance of these summary estimates holds greater weight in our interpretation of the strength of evidence, than it might do if we were examining results from individual studies.

Table 7 to Table 15 describe the results from the meta-analyses and estimates, rounded to 2 decimal places, for consistency throughout the report.

Definition of conclusions on the evidence from meta-analyses:

Each conclusion on the overall evidence from core meta-analyses states a) the strength of evidence and, if the evidence was considered to be limited, is qualified by b) a statement regarding the sufficiency of the evidence. The latter was incorporated as it reflects the common situation where limited evidence means that only a conditional and provisional conclusion can be made. However, it is important to note that comment on the sufficiency of evidence was only made where the evidence was limited. Definitions are given in the box below.

Overall strength of evidence considering core meta-analyses – definitions used in this report

a. The evidence does not support an association – where summary estimates from meta-analyses are consistently either null (odds ratio / risk ratio 1 or very close to 1) or fail to reach conventional levels of statistical significance.

b. Weak evidence for an association – where the evidence from different meta-analyses (or different exposure windows within the same meta-analysis) is inconsistent (either in direction or statistical significance), but there is evidence from some portion of the core meta-analyses indicating a statistically significant association

   AND/OR

   associations observed in the core meta-analyses are subject to considerable heterogeneity in direction in forest plots and/or considerable potential for residual confounding.

c. Strong evidence for an association – where the evidence from meta-analyses is consistent in direction of association, and is robust that is, not subject to

heterogeneity of direction within forest plots, is without considerable potential for residual confounding, and consistently reaches conventional levels of statistical significance.

Sufficiency of evidence[5] – definition

d. Limited evidence – where the evidence is based on meta-analyses analysing a small number of studies (around 5 studies or fewer, as an approximate guide), and/or studies of limited quality or size, and/or where the core meta-analyses may not have fully captured the evidence base available at the time.

2.1.4 Commentaries on key individual studies not included in the core meta-analyses

Identifying key individual studies not included in the core meta-analyses

Potentially relevant studies were identified as described in section 2.1.1.

The aim was to identify key studies that would influence, or be important to note, over and above the evidence already provided in the selected meta-analyses. It was not the intention of the epidemiological working group to comment on all relevant studies that were not included in the meta-analyses that had been summarised.

This process was conducted separately for each outcome/pollutant combination. To make this process manageable, studies already included in the core meta-analyses for a given outcome/pollutant combination (either in full, or in part for specific exposure windows only) were not considered further. For example, if for a given outcome/pollutant combination, an individual study was included in the core meta-analyses that had been evaluated for trimester 3, but not for Whole Pregnancy exposure, it would not be considered further here (even if the individual study had also analysed Whole Pregnancy exposure). This means there may be effect estimates for specific exposure time windows from individual studies that would add to the evidence from the meta-analyses for a specific exposure window which we do not consider here).

A systematic review from Bekkar et al (2020) was identified that evaluated evidence for association between prenatal exposure to $PM_{2.5}$, ozone, and heat and PTB, LBW and stillbirth, in US populations. However, no meta-analysis was performed so it was not eligible for inclusion in the epidemiological section of the report. The individual studies included within this review were cross checked, to identify whether they had been included in the core meta-analyses or had been captured within the list of 'new' studies.

The number of additional studies examined for each pollutant-outcome pair is summarised in Table 4.

[5] To note: Comment on the sufficiency of evidence will only be made where the evidence is limited.

Table 4. The number of studies examined (which did not appear in any of the core meta-analyses)

	PTB	Term LBW	Stillbirth
$PM_{2.5}$	34	18	6
PM_{10}	24	15	4
NO_2	31	12	4
O_3	26	12	5

Approach to writing commentaries on key individual studies not included in the core meta-analyses

Key information was extracted from each of the individual studies for each pollutant and outcome pair, as follows:

- exposure windows examined
- country and city/region the study was conducted in
- the study period
- the number of subjects included in the study
- whether or not the study adjusted for maternal smoking
- the study design
- brief description of the exposure assessment used in the study (for example, dispersion modelling, distance to nearest air pollution monitor, personal monitoring)
- whether a multi-pollutant model was used
- key results (adjusted OR/Risk Ratios (RR) and CIs)
- where relevant, how PTB was defined and whether the study examined LBW or TLBW

We followed an agreed decision tree to determine the level of commentary to be written for each pollutant outcome pair. This was based on whether there was a minimum number of 'acceptable and informative' individual studies that were not included in the core meta-analyses, and whether there was a substantial proportional increase in the number of studies compared to the number of studies included in the core meta-analyses. We define 'acceptable' studies as those which do not contain any obvious major errors and 'informative' studies as those which present an estimate of the association between the exposure (analysed on a quantitative scale) and the outcome (that is those that could contribute to quantifying an exposure-response relationship). This process is summarised in Figure 6.

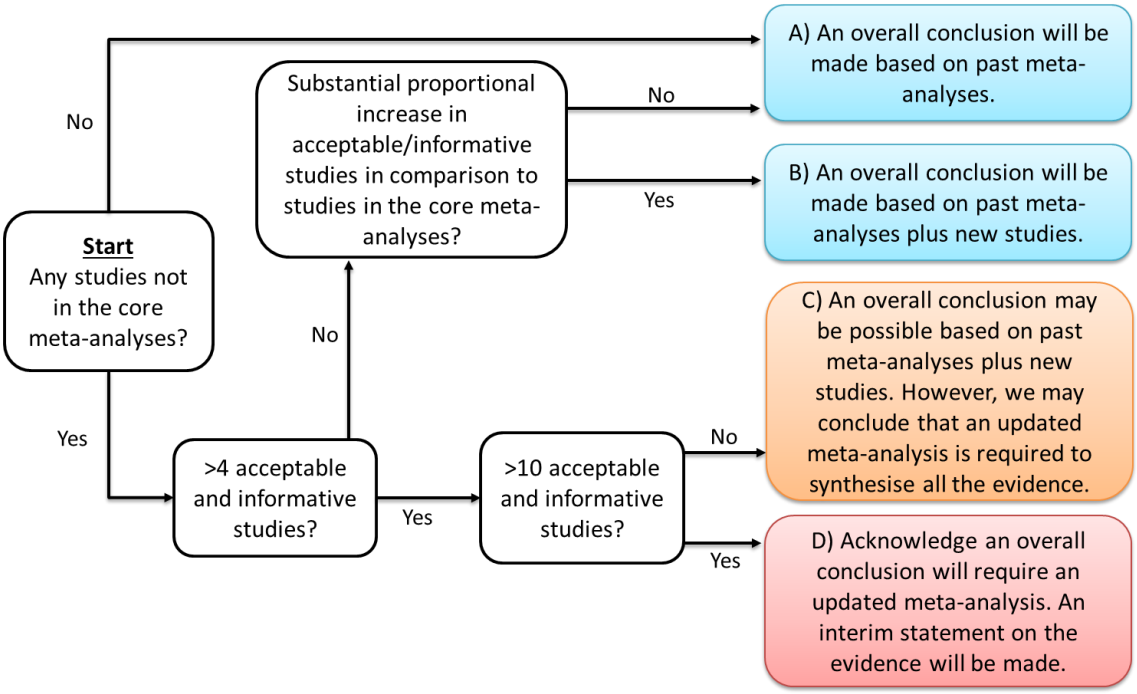


Figure 6. Process used to determine the level of commentary written for each pollutant outcome pair. *Acceptable and informative studies in this context are those that do not contain any obvious major errors and present an estimate of the association between the exposure (analysed on a quantitative scale) and the outcome

Text version of Figure 6:

Figure 6 shows a decision tree used to determine the level of commentary to be written for each pollutant-outcome pair. This decision is dependent on the number of 'acceptable and informative' studies identified in a literature review.

Start/Question 1: The starting point is to assess whether there are any additional studies that were not included in any of the core meta-analyses.

- If the answer to Question 1 is 'No', Outcome A is reached: An overall conclusion will be made based on past meta-analyses
- If the answer to Question 1 is 'Yes', then Question 2 is relevant

Question 2: is whether 4 or more additional acceptable and informative studies have been found

- If the answer to Question 2 is 'No', Question 3 is relevant
- If the answer to Question 2 is 'Yes', then Question 4 is relevant

Question 3: Do the additional studies identified represent a substantial proportional increase in acceptable/informative studies in comparison to studies in the core meta-analyses?

- If the answer to Question 3 is 'No', then Outcome A is reached: An overall conclusion will be made based on past meta-analyses

- If the answer to Question 3 is 'Yes', then Outcome B is reached: An overall conclusion will be made based on past meta-analyses plus new studies

Question 4: is whether more than 10 additional acceptable and informative studies were identified

- If the answer to Question 4 is 'No', Outcome C is reached: An overall conclusion may be possible based on past meta-analyses plus new studies. However, we may conclude that an updated meta-analysis is required to synthesise all the evidence

- If the answer to Question 4 is 'Yes', then Outcome D is reached: Acknowledge an overall conclusion will require an updated meta-analysis. An interim statement on the evidence will be made

End of text version of Figure 6.

For each pollutant outcome pair:

- if endpoint a) was reached, an overall conclusion on the strength of the evidence was made based on the evidence from the core meta-analyses
- if endpoint b) was reached, a commentary would be written based on the results from all of the acceptable and informative studies - an overall conclusion on the strength of the evidence was made considering these, and the evidence from the core meta-analyses
- if endpoint c) was reached, a brief description of the available studies would be written, with a detailed commentary written on all key acceptable and informative studies - it may have been possible to make an overall conclusion on the strength of the evidence based on these, and the evidence from the core meta-analyses, assuming there is not a lot of heterogeneity and there are no obvious population overlaps, and that meta-analyses and key highlighted studies are pointing to the same conclusion - however, this may not have been possible, in which case we concluded that an up to date meta-analysis is required to synthesise all the available evidence
- if endpoint d) was reached, a brief description of the available studies would be written, and a more detailed commentary would be written to highlight a few key acceptable and informative studies of interest (for example extremely large studies, pooled cohorts) - it would not have been possible to make an overall conclusion on the strength of the evidence based on these and the evidence from the core meta-analyses, given the volume of newer studies needing to be synthesised - we made an interim statement on the strength of the evidence based on the evidence from the core meta-analyses (and key highlighted studies, if these are pointing to the same conclusions as the meta-analyses), and concluded that an up to date meta-analysis is required to synthesise all the available evidence

General definitions used when interpreting epidemiological evidence from key highlighted studies

Where we choose to comment on consistency of study results our meaning is as follows:

a. **Consistent evidence/results:** all, or the vast majority, of effect estimates point in the same direction of association and reach statistical significance **or** all, or the vast majority, of effect estimates are either null (odds ratio 1 or very close to 1) or fail to reach conventional levels of statistical significance.

b. **Inconsistent evidence/results:** effect estimates from different studies point in both positive and negative directions (odds ratios above and below 1) **and/or** effect estimates from some studies indicate an association (odds ratio above or below 1 and reaching conventional statistical significance) whilst effect estimates from other studies are null (odds ratio 1 or very close to 1) and/or fail to reach conventional levels of statistical significance.

Specific considerations for studies looking at Preterm birth (PTB) and Stillbirth

The temporal resolution of the exposure assessment and statistical approach used in studies concerning PTB and stillbirth outcomes is important. Length of Whole Pregnancy, trimester 3 and trimester 2 is a function of gestational age calculated from the LMP and confirmed by ultrasound measurements of the fetus, preferably the crown rump length (CRL). By definition, preterm cases (<37 weeks) have shorter gestation than term births (>=37 weeks). In developed countries (in which the majority of air pollution and birth outcomes research is conducted), population average gestational age at delivery is artificially shorter for stillbirths versus live births, due to intervention to deliver stillborn babies soon after diagnosis. If there is a strong seasonal pattern in the air pollutant exposures (common, but whether this translates in exposure estimates depends on the temporal resolution of the study's exposure assessment, for example, if using annual averages that vary only spatially this is not an issue) this means that average exposures for those with shorter gestations (that is, shorter averaging period) are more strongly influenced by seasonal peaks and troughs. Thus, they are more likely to be further above or further below the mean in the exposure distribution, than those with longer gestation periods (that is average exposure may be a function of gestation length (case or non-case status), rather than a cause of it). This can result in artefactual U-shaped functional relationships between exposure and outcome when using a logistic modelling approach, as observed for Whole Pregnancy and trimester 3 exposures to air pollutants in Smith et al (2020) and for temperature in Giorgis-Allemand et al (2017). A survival-modelling approach may be more appropriate for analysis of PTB and stillbirth in relation to Whole Pregnancy, trimester 3, and in some instances trimester 2, exposure, where exposures are seasonally varying, to avoid such bias. Alternatively, a matched case-control approach where the average exposure is calculated for the case for the full gestation period, and then for the corresponding control on the basis of the gestational length of the matched case, regardless of the control's gestational length (that is, truncating to ensure equal length exposure windows), would overcome this issue.

2.2 Results

Meta-analyses

A summary of the included and excluded meta-analyses for each pollutant, outcome and exposure window period, are presented in Table 5. Some meta-analyses were excluded completely or partially (for example, for some pollutant, outcome, and exposure window combinations). Overall, 8 meta-analyses were included (Klepac et al, 2018; Li et al, 2017; Sapkota et al, 2012; Siddika et al, 2016; Stieb et al, 2012; Zhu et al, 2015; Liu et al, 2017 Li et al, 2020).

Please note, as Li et al (2017) conducted analyses for PTB and $PM_{2.5}$ separately per interquartile range (per IQR), and per 10 $\mu g/m^3$ (per 10) (depending on which increment the underlying individual studies used), they are presented separately in Table 5. Estimates from the per IQR analyses in Li et al (2017) give general qualitative evidence but are difficult to compare to results from other meta-analyses for which effect estimates represent a given absolute increment.

Table 5. Core meta-analyses per pollutant, outcome, and exposure window period

Pollutant	Exposure window	Birth Outcome		
		PTB	LBW/TLBW	Stillbirth
$PM_{2.5}$	WP	(16) Li et al 2017 (per 10 $\mu g/m^3$) (25) Liu et al 2017 (30) Zhu et al 2015	(24.5) Stieb et al 2012 (30) Zhu et al 2015	(21.5) Siddika et al 2016
	T1	(25) Liu et al 2017 (24.5) Stieb et al 2012 (30) Zhu et al 2015)	(24.5) Li et al 2020	(21.5) Siddika et al 2016
	T2	(11.5) Klepac et al 2018 (24.5) Stieb et al 2012 (30) Zhu et al 2015	(24.5) Li et al 2020	(21.5) Siddika et al 2016
	T3	(11.5) Klepac et al 2018 (24.5) Stieb et al 2012 (30) Zhu et al 2015	(24.5) Li et al 2020	(21.5) Siddika et al 2016
PM_{10}	WP	(12) Sapkota et al 2012 (24.5) Stieb et al 2012	(12) Sapkota et al 2012 (24.5) Stieb et al 2012	(21.5) Siddika et al 2016
	T1	(11.5) Klepac et al 2018 (24.5) Stieb et al 2012	(24.5) Li et al 2020 (24.5) Stieb et al 2012	(21.5) Siddika et al 2016
	T2	(11.5) Klepac et al 2018 (24.5) Stieb et al 2012	(24.5) Li et al 2020 (24.5) Stieb et al 2012)	(21.5) Siddika et al 2016
	T3	(11.5) Klepac et al 2018 (24.5) Stieb et al 2012	(24.5) Li et al 2020 (12) Sapkota et al 2012)	(21.5) Siddika et al 2016

Pollutant	Exposure window	Birth Outcome		
		PTB	LBW/TLBW	Stillbirth
NO_2	**WP**	(11.5) Klepac et al 2018 (24.5) Stieb et al 2012	(24.5) Stieb et al 2012	(21.5) Siddika et al 2016
	T1	(24.5) Stieb et al 2012	(24.5) Li et al 2020 (24.5) Stieb et al 2012	(21.5) Siddika et al 2016
	T2	(11.5) Klepac et al 2018 (24.5) Stieb et al 2012	(24.5) Li et al 2020 (24.5) Stieb et al 2012	(21.5) Siddika et al 2016
	T3	(11.5) Klepac et al 2018 (24.5) Stieb et al 2012	(24.5) Li et al 2020 (24.5) Stieb et al 2012	(21.5) Siddika et al 2016
O_3	**WP**	(11.5) Klepac et al 2018 (24.5) Stieb et al 2012	(24.5) Stieb et al 2012	(21.5) Siddika et al 2016
	T1	(24.5) Stieb et al 2012	(24.5) Li et al 2020 (24.5) Stieb et al 2012	(21.5) Siddika et al 2016
	T2	(11.5) Klepac et al 2018 (24.5) Stieb et al 2012	(24.5) Li et al 2020 (24.5) Stieb et al 2012	(21.5) Siddika et al 2016
	T3	(11.5) Klepac et al 2018 (24.5) Stieb et al 2012	(24.5) Li et al 2020 (24.5) Stieb et al 2012	(21.5) Siddika et al 2016

The modified AMSTAR 2 quality scores assigned to each meta-analysis are presented in brackets before the first author and year. It should be noted that not all included the meta-analyses examined TLBW (some examined LBW); details are provided in Appendix 5.

Table 6. Excluded meta-analyses per pollutant, outcome, and exposure window period

Pollutant	Exposure window	Birth Outcome		
		PTB	LBW/TLBW	Stillbirth
$PM_{2.5}$	**WP**	(11.5) Klepac et al 2018* (23) Lamichhane et al 2015* (16) Li et al 2017 (per IQR)*[1] (12) Sapkota et al 2012* (24.5) Stieb et al 2012** (18) Sun et al 2015*	Ji Y et al 2019***[5] (16) Li et al 2017 (per 10 μg/m^3)* (16) Li et al 2017 (per IQR)* (24.5) Li et al 2020*[6] (12) Sapkota et al 2012* (18) Sun et al 2016*	
	T1	(11.5) Klepac et al 2018* (23) Lamichhane et al 2015* (16) Li et al 2017 (per IQR)*[1] (12) Sapkota et al 2012*[2] (18) Sun et al 2015*	Ji Y et al 2019***[5] (16) Li et al 2017 (per IQR)** (12) Sapkota et al 2012***[2] (18) Sun et al 2016***[3]	
	T2	(23) Lamichhane et al 2015* (16) Li et al 2017 (per IQR)*[1] (12) Sapkota et al 2012***[2] (18) Sun et al 2015*)	Ji Y et al 2019***[5]) (16) Li et al 2017 (per IQR)** (12) Sapkota et al 2012***[2] (18) Sun et al 2016***[3]	
	T3	(23) Lamichhane et al 2015* (16) Li et al 2017 (per IQR)*[1] (12) Sapkota et al 2012***[2] (18) Sun et al 2015*	Ji Y et al 2019***[5] (16) Li et al 2017 (per IQR)** (12) Sapkota et al 2012***[2] (18) Sun et al 2016***[3]	
PM_{10}	**WP**	Guo et al 2019***[5] (30) Lamichhane et al 2015***[4]	Guo et al 2019***[5] Ji Y et al 2019***[5] (24.5) Li et al 2020*/***[6, 7]	
	T1	Guo et al 2019***[5] (30) Lamichhane et al 2015***[4] (12) Sapkota et al 2012***[2]	Guo et al 2019***[5] Ji Y et al 2019***[5] (12) Sapkota et al 2012***[2])	
	T2	Guo et al 2019***[5] (30) Lamichhane et al 2015***[4] (12) Sapkota et al 2012***[2]	Guo et al 2019***[5] Ji Y et al 2019***[5] (12) Sapkota et al 2012***[2]	
	T3	Guo et al 2019***[5] (30) Lamichhane et al 2015***[4] (12) Sapkota et al 2012***[2]	Guo et al 2019***[5] Ji Y et al 2019***[5] (24.5) Stieb et al 2012**	
NO_2	**WP**	Guo et al 2019***[5]	Guo et al 2019***[5] (24.5) Li et al 2020*/***[6, 7]	
	T1	Guo et al 2019***[5] (11.5) Klepac et al 2018*	Guo et al 2019***[5]	
	T2	Guo et al 2019***[5]	Guo et al 2019***[5]	

Pollutant	Exposure window	Birth Outcome		
		PTB	LBW/TLBW	Stillbirth
	T3	Guo et al 2019***[5]	Guo et al 2019***[5]	
O_3	**WP**		(24.5) Li et al 2020*[6]	
	T1	(11.5) Klepac et al 2018*		
	T2			
	T3			

The modified AMSTAR 2 quality scores assigned to each meta-analysis are presented in brackets before the first author and year. The reason for exclusion is indicated by the symbol after the first author and year (*= Contained ≥1 'intolerable overlaps' (see footnotes for Li et al (2020), **=Individual studies were included in other included meta-analyses, ***=Other reason (see footnotes)). It should be noted that not all included the meta-analyses examined TLBW (some examined LBW); details are provided in Appendix 5.

1) Concerned a 'Dust source' which was treated as $PM_{2.5}$. 2) Sapkota et al (2012) did not present forest plots for trimester-specific estimates and therefore it was not possible to assess potential population overlaps. It was possible to infer this information for some exposure windows. 3) Sun et al (2016) did not present forest plots for trimester-specific estimates and therefore it was not possible to assess potential population overlaps. 4) Lamichhane et al (2015) did not transform the estimates from Zhao et al (2015) correctly . 5) Guo et al (2019) and Ji Y et al (2019) were excluded in their entirety as they contained major errors (as described in section 2.1.2) and were not quality assessed. 6) Contained ≥2 'intolerable' overlaps (as described in section 2.1.2). 7) Contained a major error in the quality assessment; the meta-analysis included Arroyo et al (2019b) and it is not clear if the published estimates coming from a meta-analysis within Arroyo et al are correctly derived (as the figure suggests they are based only on the provinces in which statistically significant positive associations were observed, rather than all provinces).

Additional individual studies

We determined the appropriate level of commentary on key individual studies not included in the core meta-analyses by following the process set out in Figure 6. The results of that process for each pollutant-outcome pair are shown in Appendix 10. Approach D was adopted for PTB and TLBW in relation to each pollutant, and approach C was adopted for stillbirth in relation to each pollutant. These commentaries were then combined with the strength of evidence assessments from the core meta-analyses to produce a synthesis of the available evidence

General summaries of the core meta-analyses

Li et al 2020

Li et al (2020) aimed to "evaluate the effects of major air pollutant exposure during different trimesters on LBW" and conducted meta-analyses of associations of LBW with $PM_{2.5}$, PM_{10}, NO_2 and O_3 during Whole Pregnancy and each trimester.

Li et al (2020) searched for studies up to June 2020 that reported associations of ambient air pollution and LBW in English from PubMed and Web of Science. Inclusion criteria were based on study design (cohort studies, nested case-control, or case-cohort studies), where they had quantitatively estimated risk and where estimates were presented as risk estimates for LBW odds ratio (OR), relative risk (RR), or hazard ratio (HR) and 95% confidence interval (CI) in trimesters and/or the entire pregnancy period. Studies were included if they defined LBW as birth weight less than 2,500g and gestational age of 37 weeks or more (that is, the aim was to examine TLBW).

After title and abstract review, there were 108 studies and of these 54 went on to be included in at least one meta-analysis. When there were multiple publications on the same population or based on overlapped data, the latest or the largest study was included but no list of excluded studies was provided.

A table presenting main characteristics of included studies was published but this did not always state which pollutant was reported nor whether effect estimates were OR, RR or HR. The table includes information on whether maternal smoking was included in the adjustments and, overall, authors state 'they mostly do'. Study quality was assessed using the Newcastle Ottawa Scale (NOS) and study quality was overall considered high with scores ranging from 7 to 9. Forest plots of Whole Pregnancy associations were published, and we communicated with the authors who then provided forest plots for the trimester specific associations. Meta-analysis was repeated with stratification by region, by the units of increase in pollutant, by whether studies reported detailed birthweights, and by the type of effect estimate (OR, RR or HR). The supplemental material included repeated analysis with one study removed each time (this table was used to infer which studies were included in trimester specific estimates) and relevant funnel plots for each meta-analysis reported.

This meta-analysis was assigned a high modified AMSTAR 2 score of 24.5 (out of 37, 66.2%) because the overall conduct of the meta-analysis was good. However, it was not registered in Prospero or other relevant database and did not provide a list of excluded studies. Although some sensitivity analyses were presented these were not based on whether studies had adjusted for maternal smoking nor on the quality of the study.

Klepac et al 2018

Klepac et al (2018) aimed to "comprehensively review the studies on the association of (ambient air pollution) AAP and pregnancy outcomes." This study conducted meta-analyses of PTB and exposure to $PM_{2.5}$, PM_{10}, NO_2, CO and O_3 during Whole Pregnancy, each trimester, and the first and last months of pregnancy.

Klepac et al (2018) searched for studies up to June 2016 examining the associations between maternal exposure to ambient air pollution and pregnancy outcomes, in English, from one bibliographic database (PubMed). Following title/abstract review of 126 studies, 96 remained for qualitative synthesis. Of these, 48 studies (of the association between the 4 most commonly investigated pregnancy outcomes and the 5 most commonly observed ambient air pollutants in the 6 most commonly studied gestational time windows of exposure) were selected for descriptive quantitative synthesis, and 28 of these were selected for meta-analyses of PTB and a range of air pollutants. Klepac et al (2018) state that "For pregnancy outcomes other than PTB, pooled effect estimates were not calculated due to smaller number of studies included in each gestational window of exposure" (Klepac et al, 2018).

Klepac et al (2018) provided no table of primary study characteristics, instead focusing on grouping studies by exposure assessment methodology. It is not possible to discern from the report by Klepac et al (2018) which confounders were adjusted for in specific individual studies included in the review/meta-analyses. Klepac et al (2018) state that smoking during pregnancy was adjusted for in 42 out of

the 96 studies included in the qualitative synthesis. Klepac et al (2018) did not perform any sensitivity analyses, that is, no stratification/exclusion based on adjustment for smoking, or any other covariate. We checked whether original studies included in Klepac et al (2018) meta-analyses were adjusted for smoking, to allow more detailed critique of these meta-analyses (section 2.2.2).

This meta-analysis was assigned a low modified AMSTAR 2 score of 11.5 (out of 37; 31.1%) because there were several omissions when judged against standard systematic review/meta-analysis methodology and reporting. In particular; only a single bibliographic database (PubMed) was searched; it lacked an explicit statement on whether study selection/data extraction was performed in duplicate (in its absence it has to be assumed this was not performed in duplicate); the list of excluded studies was not reported; it lacked adequate description of included studies (for example, study location, sample size, study design, covariates and so on); it lacked any sensitivity analyses, particularly lack of stratification/exclusion based on adjustment for smoking; and there was no evaluation of publication bias. However, all studies included in meta-analyses were reported to have a Newcastle Ottawa Score (NOS)>=8 (range = 8-9, median = 9), so there is relatively low risk of bias in the meta-analysis results. It is thus difficult to judge whether the meta-analysis estimates presented by Klepac et al (2018) comprehensively represent the body of evidence available, as the initial search may not have been sufficiently comprehensive and the decision-making process as to how studies were excluded from the systematic review (and thus ultimately from the meta-analysis) is not transparently reported (2018). Notwithstanding, Klepac et al (2018) was retained in this commentary, as it included some individual studies that were not included in any of the other core meta-analyses.

Li et al 2017

Li et al (2017) stated they had “collected all data presently available” in conducting their meta-analyses of PTB and exposure to $PM_{2.5}$ (per 10 μg m^3 and per IQR) during Whole Pregnancy and each trimester.

Li et al (2017) searched for studies published in English up to July 2016, examining the associations between $PM_{2.5}$ and PTB from 4 bibliographic databases. These included the China National Knowledge Infrastructure and Wanfang Data Knowledge Service Platform to collect local data from Chinese studies. Following title/abstract review of 1,750 studies, and full text review of 105 studies, 23 studies were selected for inclusion in the meta-analyses.

Studies that adjusted effect estimates for smoking during pregnancy and exposure to environmental tobacco smoke during pregnancy were not identified in the table of primary study characteristics. Comment was made that not all studies included/acknowledged smoking as a confounding variable. The meta-analysis does not have any sensitivity analyses for stratification/exclusion based on adjustment for smoking, or any other covariate. We have therefore checked whether original studies included in Li et al (2017) meta-analyses were adjusted for smoking, to allow more detailed critique of these meta-analyses (Section 2.2.2).

This meta-analysis was assigned a moderate modified AMSTAR 2 score of 16 (out of 37; 43.2%) because there were several omissions when judged against standard systematic review/meta-analysis methodology and reporting. In particular: it lacked an explicit statement on whether study selection was performed in duplicate (in its absence it has to be assumed this was not performed in duplicate); the list of excluded studies was not reported; it lacked any sensitivity analyses, particularly lack of stratification/exclusion based on adjustment for smoking; and it lacked satisfactory explanation for, and discussion of, any heterogeneity observed in the results. All studies included in the meta-analyses were reported to have a NOS ≥ 7 (range=7-9, median=7), so there is moderately low risk of bias in the meta-analysis results.

Liu et al 2017

Liu et al (2017) conducted a meta-analysis of the associations between $PM_{2.5}$ exposure (during Whole Pregnancy and trimester 1) and PTB from cohort and nested case-control studies. Their rationale for restricting by study design was that "results from long-term cohort studies are of substantially superior quality to those obtained from case-control/cross-sectional studies."

Liu et al (2017) searched for cohort studies, in English or Chinese, relating to associations between $PM_{2.5}$ and PTB from 5 bibliographic databases (Embase, PubMed, Web of Science, Cochrane databases, and Wanfang database), and via manually searching reference lists of previous relevant published systematic reviews/meta-analyses. The date bounds of the literature search are not reported. Following title/abstract review of 1,382 studies, and full text review of 96 studies, 11 studies were identified for final inclusion in the meta-analysis. Of the final studies included, 7 used a retrospective cohort design using birth record data, 3 were prospective cohort studies and one was a nested case-control study. One of the study exclusion criteria was that "exposure and health effect in first trimester or entire pregnancy were not mentioned." Liu et al (2017) present results for $PM_{2.5}$ and PTB and it was not clear if this was an a priori decision, despite contacting the author. Liu et al (2017) only present meta-analyses for Whole Pregnancy and trimester 1 exposure. The rationale for these decisions is not given in the published paper, so we wrote to Liu et al (2017) to ascertain the rationale and whether this was decided a priori. Liu et al (2017) responded "because the collected literature mainly studies the exposure in early pregnancy and the Whole Pregnancy, there are few articles in other stages. So, we did not analyse other exposure windows." The included studies were published from 2008-2016 and covered a wide geographical range: USA, Australia, Europe, China. Exposure assessment in 10 of the studies was based on monitoring station data, and in the remaining study based on land use regression modelling.

To be included in the meta-analyses individual studies had to have adjusted for maternal age, ethnicity, parity, education level and infant sex as a minimum. Seven out of the 11 included studies also adjusted for smoking. Eight out of 11 studies adjusted for season (either of birth or conception). Liu et al (2017) stratified their meta-analyses by mean exposure level ($PM_{2.5}$ >= 15 $\mu g/m^3$ vs. <15 $\mu g/m^3$) but did not perform any sensitivity analyses according to confounder adjustment, that is, no stratification/exclusion based on adjustment for smoking, or any other covariate.

This meta-analysis was assigned a high modified AMSTAR 2 score of 25 (out of 37; 67.6%), although there were a few notable omissions: the list of excluded studies was not reported; there were no sensitivity analyses performed (in particular, no stratification/exclusion based on adjustment for smoking); there was no rigorous discussion of risk of bias when interpreting results; there was no comprehensive explanation/discussion of heterogeneity observed in results. However, all studies included in the meta-analyses were reported to have a Newcastle Ottawa Score>=7 (range = 7 to 9, median = 8), so there is relatively low risk of bias in the meta-analysis results.

Siddika et al 2016

Siddika et al (2016) performed a systematic review and meta-analysis examining associations of stillbirth with air pollutant exposure. It reported on 4 exposure periods (Whole Pregnancy and each of the 3 trimesters) against 6 pollutants (PM_{10}, $PM_{2.5}$, NO_2, O_3, sulphur dioxide (SO_2) and CO). This was the only review identified in which stillbirth was the outcome.

They collated all epidemiological studies up to mid-April 2015 which reported relevant associations from 3 appropriate bibliographic databases. Following abstract review of 369 studies, and full text review of 14 studies, 13 studies (2 on short term associations: 6 cohorts, 2 semi-ecological, one case control and 2 cross-sectional studies on long term exposure) were finally identified. After quality assessment (NOS > 7 considered suitable for inclusion) 4 studies Faiz et al (2012) – California, USA, cohort; Green et al (2015) – New Jersey, USA, cohort; Hwang et al (2011) - Taiwan, case control; Kim et al (2007) - Seoul, South Korea, cohort) were included in at least one of the meta-analyses reported. Of these studies only Faiz et al (2012) included adjustment for smoking.

This meta-analysis was assigned a moderate modified AMSTAR 2 score of 21.5 (out of 37; 58.1%), although there were a few notable omissions. There was no explicit statement that the methods used were agreed in advance and no evidence of prior registration of the review on Prospero; there was no clear definition of stillbirth provided at the outset, although all studies included in the meta-analysis considered that stillbirth could only occur after 20 weeks' gestation; a list of excluded studies was not provided and no clear reasons for their exclusion given. There was no clear statement in the published article as to why only 4 studies were included in the meta-analysis, however readers could reasonably infer that this was because they exceeded a quality threshold score on the NOS. Publication bias and heterogeneity were assessed, but as the authors pointed out these are relatively uninformative when only 2 or 3 studies are included in a meta-analysis (as was the case here). The authors acknowledge that smoking is a major risk factor for stillbirth but do not explore or discuss the potential role of confounding by smoking in their analyses.

Zhu et al 2015

Zhu et al (2015) performed a review and meta-analysis of $PM_{2.5}$ and the change in birthweight, risk of low birthweight, pre-term birth and small for gestational age during Whole Pregnancy and for each trimester. Stillbirth was included in the review, but as only one stillbirth study was included no pooled meta-analysis was possible.

PubMed, Web of Science, Embase and Google Scholar were searched for studies in English. No start date was specified but the end date was 1 March 2014. The earliest study publication date was 2005. Twenty-five studies remained after sifting and exclusions.

This meta-analysis was assigned a high modified AMSTAR 2 score of 30 (out of 37; 81.1%). There were a few missing quality aspects: there was no *a priori* protocol; it lacked a list of excluded studies; and there were some missing details in the table of original studies (subjects, exposure levels and assessment method, but other details such as study design, adjusted covariates and time period were given, and there was a general summary of exposure assessment methods used in the discussion). There were some good quality aspects that were included: there was an explanation for which study designs were included; the study selection and data extraction was conducted in duplicate; the factors related to risk of bias were considered in the inclusion and exclusion criteria, although no formal scoring system was used; sensitivity analyses were conducted (excluding studies one by one); publication bias was examined (none found); the authors specified whether constituent studies adjusted for smoking; and heterogeneity was analysed by meta-regressions. The qualitative conclusions from these meta-regressions were discussed, although quantitative information was not given.

Zhu et al (2015) explicitly discussed the issue of population overlap and the selection of estimates for inclusion in the meta-analysis took this into account.

Sapkota et al 2012

Sapkota et al (2012) performed a review and meta-analysis of $PM_{2.5}$ and PM_{10} and the risk of low birthweight and pre-term birth during Whole Pregnancy and for each trimester.

Sapkota et al (2012) searched ISI Web of Knowledge and PubMed. No date range is specified but the searches were performed between February and July 2009. No identified papers were earlier than 2000. They identified 20 peer-reviewed articles providing quantitative estimates of exposure and outcome that met their selection criteria. Numbers of studies by pollutant and outcome are given in the relevant sections below but, from a meta-analysis quality perspective, it is worth noting here that there was a lack of clarity as to which studies were included in the trimester-specific pooled estimates.[6]

This meta-analysis was assigned a low modified AMSTAR 2 score of 12 (out of 37; 32.4%), for a variety of reasons: There was no a priori protocol; it lacked a list of excluded studies (although exclusion numbers in different categories were given); there was no discussion of whether constituent studies adjusted for smoking or not

[6] In addition, there is a lack of clarity regarding the derivation of the whole pregnancy summary effect estimates. The authors explain that "When only trimester-specific risk estimates were provided, a combined estimate for the entire pregnancy was also calculated using meta-analytic approach, and this new estimate was used in the overall analysis". However, some of the studies included in the whole pregnancy meta-analyses did not report odds ratios for all three trimesters. Therefore, the extent to which some of the "whole pregnancy" risk estimates included in the main meta-analyses represent associations over the whole period of pregnancy is unclear.

(we have therefore checked whether original studies included in Sapkota et al (2012) meta-analyses were adjusted for smoking, to allow comment to be made on the results of these meta-analyses (section 2.2.2); there were inadequate details in the tables of original studies (subjects, location, exposure assessment method and outcomes were included, with some mean pollutant levels provided in discussion, but lacked details of study covariates or time period); tables indicated which study had trimester specific information but not which trimester; and there was no discussion of risk of bias in individual studies (although some discussion of study size). On the other hand, some aspects were good: study selection and data extraction was conducted in duplicate; sensitivity analyses were completed (excluding studies one by one, although there was no analysis of effect of stratification by study quality); publication bias was examined (none found); and heterogeneity was analysed and discussed (although there was no testing of the potential effect of the suggested reasons for heterogeneity by stratifying results by those factors).

Stieb et al 2012

Stieb et al (2012) aimed to "present a systematic review and meta-analysis of over 60 studies examining associations between ambient air pollution and birth weight and PTB."

Stieb et al (2012) searched for studies published in English between January 1980-January 2011, examining the associations between various ambient air pollutants (NO_2, SO_2, O_3, CO, PM_{10}, $PM_{2.5}$) and birth weight and PTB from a number of bibliographic databases (EMBASE, MEDLINE, Scopus, Current Contents, Global Health, Cochrane, TOXLINE, Canadian Research Index). Following title/abstract review of 2,026 unique studies, and full text review of 301 studies, 61 studies were selected for inclusion in the meta-analyses and one additional study was identified through review of the reference lists of relevant papers. This made a total of 62 papers. Studies were carried out in 39 locations. It reported on 4 exposure periods (Whole Pregnancy and each of the 3 trimesters) against 6 pollutants (NO_2, SO_2, O_3, CO, PM_{10}, $PM_{2.5}$). Stieb et al (2012) did not perform meta-regressions against potential confounding factors for PTB and TLBW. However meta-regressions of estimates of change in birth weight were conducted against maternal smoking, alcohol consumption, education, socio-economic status, mean pollutant concentration, and restriction to singleton or term pregnancies. The study quantified heterogeneity and evaluated publication bias.

This meta-analysis was assigned a high modified AMSTAR 2 score of 24.5 (out of 37; 66.2%), although there were a few small omissions such as: the list of excluded studies was not reported; a duplicate process for study selection was employed but no mention of the consensus process was given should disagreements have arisen; and there was no comprehensive explanation/ discussion of heterogeneity observed in results provided. Publication bias was assessed through funnel plots and both Begg's test p-values and Egger's bias coefficients were calculated.

Commentary on results from studies examining Preterm Birth (PTB)

$PM_{2.5}$ and PTB

Commentary on results from meta-analyses

After exclusions, the core meta-analyses that examined associations between $PM_{2.5}$ exposure and PTB were Klepac et al (2018) for trimesters 2 and 3, Li et al (2017) for Whole Pregnancy, Liu et al (2017) for Whole Pregnancy and trimester 1, Stieb et al (2012) for trimesters 1, 2 and 3 and Zhu et al (2015) for Whole Pregnancy and trimesters 1, 2 and 3. A summary of the results is provided in Table 7, with more detailed information provided in Appendix 11, Table A 36.

Table 7. Summary of the results from the core meta-analysis examining preterm birth and $PM_{2.5}$ exposure

Meta-analyses (Increment)	OR (95% CI)			
	WP	T1	T2	T3
Klepac et al 2018 (Per 10 µg/m³)	Excluded	Excluded	1.10 (0.96, 1.27)	1.05 (1.02, 1.09)
Li et al 2017 (Per 10 µg/m³)	1.02 (0.93, 1.12)	Excluded	Excluded	Excluded
Liu et al 2017 (Per 10 µg/m³)	1.15 (0.99, 1.33)	1.15 (1.05, 1.24)	Not performed	Not performed
Zhu et al 2015 (Per 10 µg/m³)	1.10 (1.03, 1.18)	0.96 (0.77,1.21)	0.90 (0.79, 1.03)	0.97 (0.89, 1.05)
Stieb et al 2012 (Per 20 µg/m³)	Excluded	0.85 (0.60, 1.20)	0.66 (0.57, 0.77)	1.05 (0.98, 1.13)
Individual studies in common between meta-analyses?	Li & Liu=1 Li & Zhu=3 Liu & Zhu=4	Liu & Stieb=1 Liu & Zhu=4 Stieb & Zhu=3	Klepac & Zhu=1	Klepac & Zhu=1 Stieb & Zhu=2

Klepac et al 2018

Klepac et al (2018) report summary estimates for PTB per 10 µg/m³ $PM_{2.5}$ by exposure period as follows: trimester 2 OR 1.10 (95% CI 0.96, 1.27) and trimester 3 OR 1.05 (95% CI 1.02, 1.09) (Table 7). These meta-analyses included 6 and 8 individual studies respectively.

Heterogeneity was considerable (I^2 97%, p<0.00001) for trimester 2 and moderate (I^2 57%, p=0.04) for trimester 3. For trimester 2 odds ratios in the individual studies were both greater than 1 and below 1, that is, the heterogeneity was both in size and in direction of the association. For trimester 3 the odds ratios in the individual studies were all greater than 1, that is, the heterogeneity was in size of effect. Publication bias was not examined.

The majority of the weight of the meta-analysis estimates (67% and 84% for trimester 2 and 3 respectively) came from studies that had adjusted for smoking. The summary estimate for trimester 3, but not trimester 2, reached conventional levels of statistical significance. Given the consistency in direction of effect for individual studies (for trimester 3), and that the majority of the weight in the summary estimate reflects studies which have adjusted for smoking, and taking into account statistical significance, we conclude that Klepac et al (2018) provides **some evidence to support an association of PTB with $PM_{2.5}$ exposure in trimester 3**. However, we conclude that Klepac et al (2018) provides **no evidence for an association with $PM_{2.5}$ exposure in trimester 2,** due to considerable heterogeneity in underlying studies and lack of statistical significance.

Li et al 2017

Li et al (2017) report a Whole Pregnancy summary estimate for PTB per 10 µg/m^3 $PM_{2.5}$ with OR 1.02 (95% CI 0.93, 1.12) (Table 7). The meta-analysis included 6 studies.

Heterogeneity among studies (I^2 97%, p<0.00001) was considerable, with the odds ratios in the individual studies being both greater than 1 (sometimes substantially) and below 1, that is, the heterogeneity was both in size and in direction of the association. 53% of the weight of the summary estimate came from studies with no adjustment for smoking. No evidence of publication bias, based on a Begg's test and Egger's test, was found.

The summary estimate is null and did not reach conventional levels of statistical significance, and we conclude that Li et al (60) provide **no evidence to support an association between PTB and $PM_{2.5}$ exposure across Whole Pregnancy**.

Liu et al 2017

Liu et al (2017) report summary estimates for PTB per 10 µg/m^3 $PM_{2.5}$ by exposure period as follows: Whole Pregnancy OR 1.15 (95% CI 0.99, 1.33) and trimester 1 with OR 1.15 (95% CI 1.05, 1.24) (Table 7). The meta-analyses included 7 and 9 individual studies, respectively.

Heterogeneity was considerable (I^2 85%, p<0.00001) for Whole Pregnancy, with the odds ratios in the individual studies being both greater than 1 and below 1 (sometimes substantially), that is, the heterogeneity was both in size and in direction of the association. There was no indication of an important degree of heterogeneity (I^2 33%, p=0.15) for trimester 1.

Publication bias was only evaluated for sub-groups of studies (n= 3-5) characterised by low versus high $PM_{2.5}$ exposure levels. Liu et al (2017) interpreted the funnel plot as asymmetrical in the high-level Whole Pregnancy exposure group, suggesting possible publication bias, however Egger's test did not reach statistical significance for any sub-group (p>0.05). Given the small numbers of studies in the funnel plot tests, statistical power may be insufficient to identify real asymmetry.

All studies included adjustment for smoking. The summary estimate for trimester 1 reached conventional levels of statistical significance, and for Whole Pregnancy is borderline significant. Given lack of heterogeneity (for trimester 1), and that all of the weight in the summary estimate reflects studies which have adjusted for smoking, and taking into account statistical significance, we conclude that Liu et al (2017) **provides good evidence to support an association of PTB with $PM_{2.5}$ exposure in trimester 1, and weak evidence suggesting an association with $PM_{2.5}$ exposure across Whole Pregnancy**.

Zhu et al 2015

Zhu et al (2015) report summary estimates for PTB per 10 μg/m^3 $PM_{2.5}$ by exposure period as follows: Whole Pregnancy OR 1.10 (95% CI 1.03, 1.18), trimester 1 OR 0.96 (95% CI 0.77, 1.21), trimester 2 OR 0.90 (95% CI 0.79, 1.03) and trimester 3 OR 0.97 (95% CI 0.89, 1.05) (Table 7). The meta-analyses included 8, 6, 3, and 6 studies respectively.

Heterogeneity was moderate for Whole Pregnancy (I^2=52.0%, p<0.042), with odds ratios in the individual studies being predominantly greater than 1 (sometimes substantially), that is, the heterogeneity was more about size than direction of the association. Heterogeneity was considerable for trimester 1 (I^2 87.2%, p<0.001), with the odds ratios in the individual studies being both greater than 1 and below 1, that is, the heterogeneity was both in size and in direction of the association. There was no indication of an important degree of heterogeneity for trimesters 2 (I^2 0%, p=0.700) or 3 (I^2 31.4%, p=0.200). No evidence of publication bias, based on a Begg's funnel plots and Egger's test, was found.

Almost all included studies adjusted for smoking, meaning that 76%, 81%, 100% and 100% of the weight of the Whole Pregnancy, trimesters 1, 2 and 3 meta-analysis estimates respectively came from studies which had adjusted for smoking. Only the summary estimate for Whole Pregnancy reached conventional levels of statistical significance. Given the relatively consistent direction of effect for individual studies (for Whole Pregnancy), and that most of the weight in the summary estimate reflects studies which have adjusted for smoking, and taking into account statistical significance, we conclude that Zhu et al (2015) provides **good evidence to support an association of PTB with $PM_{2.5}$ exposure in Whole Pregnancy, but no evidence for an association in specific trimesters**. However, for trimester 2 this conclusion is based on limited evidence as this meta-analysis contained only 3 studies.

Stieb et al 2012

Stieb et al (2012) report summary estimates for PTB per 20 μg/m^3 $PM_{2.5}$ by exposure period as follows: trimester 1 OR 0.85 (95% CI 0.60, 1.20) and trimester 3 OR 1.05 (95% CI 0.98, 1.13) (Table 7). The meta-analyses each included 4 studies. The forest plot for trimester 2 included only a single study and therefore is not a meta-analysis and is not commented upon further.

Heterogeneity was considerable (I^2 94.5%, p<0.001) for trimester 1, with the odds ratios in the individual studies being both greater than 1 and below 1 (sometimes substantially) that is, the heterogeneity was both in size and in direction of the association. However, with such small numbers of studies included this statistic may not be particularly informative. There was no indication of an important degree of heterogeneity (I^2 33.2%; p=0.213) for trimester 3. Evidence of publication bias was not reported among studies that examined PTB and $PM_{2.5}$.

Two of the 4 individual studies per meta-analysis adjusted for smoking, meaning that 54% and 92% of the weight of the trimester 1 and trimester 3 meta-analysis estimates respectively, came from studies with no adjustment for smoking. Neither summary estimate reached conventional levels of statistical significance, although that for trimester 3 was borderline significant. Given the considerable potential for residual confounding by smoking in these estimates, and that none of the summary estimates reached conventional levels of significance we conclude that Stieb et al (2012) provides **no evidence to support an association of PTB with $PM_{2.5}$ exposure in trimesters 1 or 3**, but this is on the basis of limited evidence as the meta-analyses contained small numbers of studies (n=4).

Synthesis

For each of the exposure windows, there is inconsistency in the findings from the meta-analyses – both null and positive (adverse) associations are observed.

For Whole Pregnancy $PM_{2.5}$ exposure, summary estimates reported by Liu et al (2017) and Zhu et al (2015) were of similar magnitude and indicative of elevated risk, although only reaching statistical significance for Zhu et al (2015). The results were consistent despite the 4 studies they had in common (Brauer et al, 2008; Gehring et al, 2011; Hannam et al, 2014; Pereira et al, 2014a) contributing only 20% and 9% weight respectively to Liu et al (2017) and Zhu et al (2015) meta-analyses. All studies included in Liu et al (2017), and 7 out of 8 studies included in Zhu et al (2015) adjusted for smoking. The one study which did not adjust for smoking (Huynh et al, 2006) contributed 24% weight to the Zhu et al (2015) meta-analysis, and whilst the size of the odds ratio was not notably out of line with the other studies, had this been excluded it is likely that the summary effect size would have been smaller and perhaps not reached statistical significance. In contrast, the meta-analysis by Li et al (2017) did not support an association between PTB and Whole Pregnancy $PM_{2.5}$ exposure. Only 3 of the 6 included studies adjusted for smoking, but there was no notable difference in the range of odds ratios from studies which had (ORs 0.98, 1.00, 1.06) and had not (ORs 0.96, 0.96, 1.15) adjusted for smoking, and all studies contributed relatively equal weight. Had studies which did not adjust for smoking been excluded it is unlikely that the summary estimate would have been changed much, it would simply have had a wider confidence interval. Lack of consistency in results is most likely explained by limited crossover in terms of included studies: Li et al (2017) include 3 studies which are not included in either of the other 2 meta-analyses, which have null (Ritz et al, 2007; Fleischer et al, 2014) or negative (Stieb et al, 2016b) associations, and together contribute 51.6% weight to the Li et al (2017) meta-analysis. Stieb et al (2016b) is an extremely large study, containing almost 3 million births.

There were some studies that examined very PTB (vPTB) (<30 weeks) rather than PTB (<37 weeks). It appears that Zhu et al (2015) included an estimate from the Brauer et al (2008) study which related to very PTB (vPTB) rather than PTB. It is not possible to discern which estimate Liu et al (2017) included from Brauer et al (2008); back calculation was attempted but it did not tally exactly with any of the estimates published in Brauer et al (2008). If the published result for PTB (<37 weeks) from Brauer et al (2008) had been included in these meta-analyses instead, then this may have attenuated the summary estimate in Zhu et al (2015), but increased the summary estimate in Liu et al (2017), however the Brauer et al (2008) study was not particularly influential in terms of % weight in either.

For trimester 1 $PM_{2.5}$ exposure, the meta-analysis by Liu et al (2017 reported a statistically significant elevated risk, in contrast to Zhu et al (2015) and Stieb et al (2012) which did not provide evidence of association. These inconsistencies are unlikely to be explained by adjustment for smoking as all or most of the weight the meta-analyses by Liu et al (2017) and Zhu et al (2015) came from studies which adjusted for this. There was only one study in common (108) between Liu et al (2017) and Stieb et al (2012) contributing 6% and 22% weight respectively. There were 4 studies in common (Gehring et al, 2011; Hannam et al, 2014; Pereira et al, 2014a, Lee et al, 2013) between Liu et al (2017) and Zhu et al (2015) together contributing 32% and 65% weight respectively. There were 3 studies in common (Gehring et al, 2011; Jalaludin et al, 2007; Wilhelm et al, 2005) between Zhu et al (2015) and Stieb et al (2012), together contributing 49% and 72% weight respectively. Lack of consistency is most likely explained by limited crossover in terms of included studies. Liu et al (2017) include 5 studies (Ritz et al, 2007; DeFranco et al, 2016; Ha et al, 2014; Pereira et al, 2014b; Pereira et al, 2016) which are not included in either of the other 2 meta-analyses which together contribute 68% weight to this meta-analysis. Three of these 5 studies have strong statistically significant positive associations.

For trimester 2 $PM_{2.5}$ exposure, the summary estimate from Klepac et al (2018) suggested elevated risk (not statistically significant), whereas that from Zhu et al (2015) did not. Adjustment for smoking is unlikely to explain the difference as all or most of the weight in these 2 meta-analyses came from studies which adjusted for this. Lack of consistency is most likely explained by minimal crossover in terms of included studies, with only one study in common (Pereira et al, 2014a) which contributed 45% weight to Zhu et al (2015), but only 8.3% weight to Klepac et al (2018).

For trimester 3 $PM_{2.5}$ exposure, Klepac et al (2018) reported a statistically significant elevated risk, and the summary estimate from Stieb et al (2012) was suggestive of a positive association (but of a much smaller magnitude), whereas the summary estimate from Zhu et al (2015) was null. These inconsistencies are unlikely to be explained by adjustment for smoking as all or most of the weight in meta-analyses by Klepac et al (2018) and Zhu et al (2015) came from studies which adjusted for this. Lack of consistency is most likely explained by minimal crossover in terms of included studies, with only one study (Pereira et al, 2014a) common to Klepac et al (2018) and Zhu et al (2015) contributing 2% and 24% weight respectively; none common to Klepac et al (2018) and Stieb et al (2012); and 2

studies (Gehring et al, 2011; Jalaludin et al, 2007) common to Zhu et al (2015) and Stieb et al (2012) contributing 15% and 8% weight respectively.

Conclusion regarding evidence for $PM_{2.5}$ and PTB from core meta-analyses

The 5 meta-analyses provide weak evidence (section 2.1.3) for an association between PTB and $PM_{2.5}$ exposure. Specifically, 3 out of the 5 meta-analyses above provide some evidence (ranging in strength) for associations between PTB and $PM_{2.5}$ exposure during Whole Pregnancy, trimester 1, and trimester 3. Limited overlap in terms of studies included in the 3 meta-analyses points to the need for an updated comprehensive meta-analysis, in order to quantitatively synthesise the evidence.

Commentary on evidence from key individual studies regarding $PM_{2.5}$ and Preterm Birth (PTB) (not included in the core meta-analyses)

Following the flowchart process as outlined in Figure 6, Endpoint D was reached (Appendix 10).

Brief description of available studies:

Thirty-four studies examining the association between $PM_{2.5}$ exposure during pregnancy and PTB were captured by our literature search and were not included in any of the core meta-analyses above (Giorgis-Allemand et al, 2017; Chen et al, 2018; Hyder et al, 2014; Mendola et al, 2016; Rappazzo et al, 2014, Zhu et al, 2019; Trasande et al, 2013; Smith et al, 2020; Sheridan et al, 2019; Laurent et al, 2016a; Basu et al, 2017; Chang et al, 2015; Huang et al, 2018; Kinglsey et al, 2017; Padula et al, 2019; Rudra et al, 2011; Salihu et al, 2012; Chang et al, 2013; Poirier et al, 2015; Siddika et al, 2019; Liang et al, 2019a,b; Liu Y et al, 2019; Guo et al, 2018; Sun et al, 2019; Ye et al, 2018; Yuan et al, 2020; Zhao et al, 2020; Li et al, 2018; Guan et al, 2019; Padula et al, 2018; Seabrook et al, 2019; Stieb et al, 2019; Ranjbaran et al, 2020). This is a particularly rapidly growing area of research: of these 34 studies, 24 were published during 2017-2020, so would not have been captured by the literature search of the most recent meta-analysis for this pollutant-outcome pair evaluated above (search cut-off date in 2016) (Klepac et al, 2018). The remainder (n=10) were published between 20112016. There were 23 retrospective cohort studies using birth records or electronic medical records, 3 prospective cohort studies, one study pooling multiple birth cohorts, 3 case-control studies, 2 cross-sectional studies, and 2 time-series studies. Geographically they were spread as follows: USA (n=16 (of which 7 in California)), Canada (n=3), Europe (n=3), Asia (n=11 (of which 7 in China)), Australia (n=1). It is important to note that it is possible that some of the earlier studies were not included in meta-analyses above due to decisions authors made regarding which study to include where there were several examining the same/similar population.

As stated in section 2.1.4, as we consider here only studies which do not appear anywhere in the core meta-analyses, there may, in addition, be individual studies that would add to the evidence for a specific exposure window which are not considered here.

Seven of these 34 studies would not have met the criteria for the specific meta-analyses we evaluated (that is, did not examine Whole Pregnancy, trimester 1, trimester 2, trimester 3 exposure windows, or were time-series studies examining short-term effects) or were not particularly informative as the exposure window was not clearly reported or defined (Trasande et al, 2013; Li et al, 2018; Guan et al, 2019; Padula et al, 2018; Seabrook et al, 2019; Stieb et al, 2019; Ranjbaran et al, 2020). As they are not comparable with the specific exposure windows considered in meta-analyses above, we do not comment on them further here.

Brief commentary on selected key studies of interest (Appendix 12 Table A 48) which are likely to expand the evidence base (but not a commentary on all):

Most of these studies (n=27) examined exposures in the Whole Pregnancy, trimester 1, trimester 2, or trimester 3 time windows (Giorgis-Allemand et al, 2017; Chen et al, 2018; Hyder et al, 2014; Mendola et al, 2016; Rappazzo et al 2014; Zhu et al, 2019; Smith et al, 2020; Sheridan et al, 2019; Laurent et al, 2016a; Basu et al, 2017; Chang et al, 2015; Huang et al, 2018; Kingsley et al, 2017; Padula et al, 2019; Rudra et al, 2011; Salihu et al, 2012; Chang et al, 2013; Poirier et al, 2015; Siddika et al, 2019; Liang et al, 2019a; Liang et al, 2019b; Liu Y et al, 2019; Guo et al, 2018; Sun et al, 2019; Ye et al, 2018; Yuan et al, 2020; Zhao et al, 2020). We consider these as they have potential to add to the evidence from meta-analyses above, in relation to these specific exposure windows, for example in any future meta-analysis. Geographically they were spread as follows: USA (n=14), Canada (n=1), Europe (n=3), Asia (n=8), Australia (n=1).

Of 14 studies in the USA (Hyder et al, 2014; Mendola et al, 2016; Rappazzo et al 2014; Zhu et al, 2019; Sheridan et al, 2019; Laurent et al, 2016a; Basu et al, 2017; Chang et al, 2015; Huang et al, 2018; Kingsley et al, 2017; Padula et al, 2019; Rudra et al, 2011; Salihu et al, 2012; Chang et al, 2013) we highlight the following:

- Sheridan et al (2019), a retrospective cohort study using routine birth records with a very large sample size (n=2,293,218) across California (expanding beyond the well-studied Los Angeles (LA) and Orange Counties of many previous Californian studies). In particular, out of all the large Californian studies, it was the only one to use survival analysis, which reduces risk of bias when analysing the relationship between Whole Pregnancy and trimester 3 exposures in relation to PTB. However, like most other Californian studies it still used environmental monitoring approaches for exposure assessment and did not adjust for smoking. It found significant positive associations between PTB and $PM_{2.5}$ in Whole Pregnancy, and trimesters 2 and 3, but no association for trimester 1 exposures (HRs (95% CIs) of PTB associated with a 10 $\mu g/m^3$ increase in $PM_{2.5}$: Whole Pregnancy: 1.12 (1.09, 1.14); trimester 1: 0.99 (0.98, 1.00); trimester 2: 1.04 (1.03, 1.06); trimester 3: 1.05 (1.04, 1.07)).

- Laurent et al (2016a) is a state-wide nested case-control study with a large sample size (n=1,326,942) across the whole of California, and uses modelling in exposure assessment, and adjusts for smoking – the latter 2 factors being rare in studies based in California. It analysed $PM_{2.5}$ exposure in Whole Pregnancy only. It found a significant positive association between PTB and $PM_{2.5}$ in Whole

Pregnancy (OR (95% CI) of PTB associated with a 6.45 μg/m^3 increase in $PM_{2.5}$: 1.133 (1.118,1.148)).

One study from Australia (Chen et al, 2018) is highlighted as it was larger in size (n= 173,800) than previous studies from Australia included in the meta-analyses, it adjusted for smoking, it also used survival analysis, which reduces risk of bias when analysing the relationship between Whole Pregnancy and trimester 3 exposures in relation to PTB (see 'Specific considerations for studies looking at Preterm birth (PTB) and Stillbirth' in section 2.1.4), and examined multi-pollutant models. It found a significant positive (adverse) association between PTB and Whole Pregnancy exposure to $PM_{2.5}$ (HR (95% CI) 1.05 (1.02, 1.08) per IQR (1.32 μg/m^3) increase in $PM_{2.5}$). In multi-pollutant models hazard ratios for $PM_{2.5}$ were attenuated to null after adjustment for SO_2 but remained the same after adjustment for NO_2 and O_3 in turn. Trimester-specific exposures were analysed, but numeric results were not presented; the authors simply commented 'for all air pollutants, highest HRs were present at trimester 3 and the lowest at trimester 1.'

Of 3 studies in Europe (Giorgis-Allemand et al; 2017), Smith et al, 2020; Siddika et al 2019), we highlight:

- Giorgis-Allemand et al (2017), a pooled analysis of European birth cohorts (ESCAPE study) (n=71,493) which covered a wide geographical area of Europe, used modelled exposures, and adjusted for smoking and temperature. They examined the association of PTB with annual mean $PM_{2.5}$ by collating information from several prospective birth cohorts that originated in Sweden, Norway, Netherlands, Greece, Lithuania, Hungary, Spain, Italy, and Denmark. A common protocol for LUR was used, exposure was based on maternal address and accounted for moving during pregnancy in some centres. LUR models were validated with specific monitoring campaigns (although these were only for short periods). Whole Pregnancy and trimester specific estimates for trimester 1 and trimester 2 were derived. Logistic regression analyses were adjusted for major confounders including maternal smoking and were used for trimester 1 and trimester 2 exposure windows. Survival analysis was used to assess associations with Whole Pregnancy exposure. There was no association between PTB and exposure to $PM_{2.5}$ in Whole Pregnancy (OR 0.96 95% CI 0.87, 1.04 per 5 μg/m^3 increase in $PM_{2.5}$, n= 46,791), nor in trimester 1 or trimester 2 (trimester 1 OR 0.98 95% CI 0.91, 1.05, n=46,242: trimester 2 OR 0.96 95% CI 0.90, 1.03 n=47,153, respectively).

- Smith et al (2020) examined London, UK, highlighted due to its large sample size, used modelled exposures, and multi-pollutant models. Smith et al (100) used information from large number of routine birth records (n=581,774) to examine the association between PTB and a range of particulate matter and gaseous air pollutant exposures in London, UK. The study used highly spatially resolved modelled exposures (monthly estimates from dispersion model at 20m x 20m grid point resolution), assigned to maternal address. They assessed the association of PTB with $PM_{2.5}$ exposure in trimester 1 and trimester 2 only. Logistic regression analyses were adjusted for a wide range of potential confounders, including area-level adjustment for tobacco sales (as a proxy for

smoking), and road traffic noise. Multi-pollutant models examined $PM_{2.5}$ adjusted for a range of other air pollutants. This study showed a negative association between PTB and $PM_{2.5}$ in trimesters 1 and 2 (OR 0.93 (95% CI 0.91, 0.95) per 3.2 μg/m^3 in trimester 1; OR 0.85 (95% CI: 0.83, 0.87) per 3.6 μg/m^3 in trimester 2). In multi-pollutant models adjusting for O_3, negative associations remained for $PM_{2.5}$ and PTB, with some attenuation.

Of 8 studies from Asia (Liang et al, 2019a; Liang et al, 2019b; Liu Y et al, 2019; Guo et al, 2018; Sun et al, 2019; Ye et al, 2018; Yuan et al, 2020; Zhao et al, 2020) all but one were from China. We highlight:

- A study by Liang et al (2019a) in the Pearl River Delta area of China, due to its large sample size (n=1,455,026), its use of survival analysis, exclusion of caesarean deliveries (which will exclude some if not all medically indicated early deliveries) and analysis of multi-pollutant models. Liang et al (2019a) examined associations of PTB with several air pollutants in a large retrospective birth cohort using routine birth records, across 9 cities in the Pearl River Delta region of China (Guangzhou, Shenzhen, Zhuhai, Dongguan, Foshan, Jiangmen, Zhongshan, Huizhou and Zhaoqing). Exposure assessment was via environmental monitoring. It analysed trimester-specific exposure windows. There were significant positive (adverse) associations between PTB and $PM_{2.5}$ in trimesters 2 and 3 (HRs (95% Cis) of PTB associated with a 10 μg/m^3 increase in $PM_{2.5}$: trimester 1: 1.01 (1.00, 1.02); trimester 2: 1.06 (1.05, 1.07); trimester 3: 1.07 (1.06, 1.08). In multi-pollutant models adjusting for O_3, NO_2 and SO_2, in turn, associations between $PM_{2.5}$ and PTB were very similar, remaining significant and positive for trimesters 2 and 3, but null for trimester 1.

- A study by Guo et al (2018), the next largest (n=426,246), for its broad coverage of China (births from 129 cities from 30 provinces in China), its use of survival analysis and adjustment for smoking. Exposure assessment was via environmental monitoring. Survival analyses were adjusted for major confounders. The study found significant positive associations between PTB and $PM_{2.5}$ during Whole Pregnancy and each trimester (HR (95% CI) per 10 μg/m^3 increase in $PM_{2.5}$: Whole pregnancy 1.06 (1.05, 1.06); trimester 1 1.04 (1.03, 1.04); trimester 2 1.02 (1.02, 1.02); trimester 3 1.06 (1.06, 1.07)).

The results from these highlighted that some studies demonstrated statistically significant increased risks of PTB associated with long-term exposure to $PM_{2.5}$ during pregnancy, (Chen et al, 2018; Sheridan et a, 2019; Laurent et al, 2016a; Liang et al, 2019a; Guo et al, 2018) one found inverse associations (Smith et al, 2020), and one no association (Giorgis-Allemand, 2017).

Conclusion regarding evidence for $PM_{2.5}$ and PTB from key individual studies not included in the core meta-analyses

There were a very large number of studies not captured in/published since the core meta-analyses, and the results from key studies highlighted were inconsistent (section 2.1.4).

OVERALL CONCLUSION for $PM_{2.5}$ and PTB from the epidemiological evidence

Based on the meta-analyses available, we conclude there is weak evidence (section 2.1.3) for an association between PTB and $PM_{2.5}$ exposure. Examination of key highlighted studies points to the need for an up-to-date meta-analysis to synthesise all the available evidence.

PM_{10} and PTB

Commentary on results from meta-analyses

After exclusions, the core meta-analyses that examined associations of PM_{10} exposure with PTB were Sapkota et al (2012) for Whole Pregnancy and Stieb et al (2012) and Klepac et al (2018) for Whole Pregnancy and trimesters 1, 2 and 3. A summary of results are provided in Table 8, with more detailed information provided in Appendix 11, Table A 37.

Table 8. Summary of the results from the core meta-analysis examining preterm birth and PM_{10} exposure

Meta-analyses (Increment)	OR (95% CI)			
	WP	T1	T2	T3
Klepac et al 2018 (Per 10 µg/m³)	1.09 (1.03, 1.16)	1.04 (1.01, 1.08)	1.04 (0.98, 1.09)	1.00 (0.99, 1.00)
Sapkota et al 2012 (Per 10 µg/m³)	1.02 (0.99, 1.04)	Excluded	Not performed	Excluded
Stieb et al 2012 (Per 20 µg/m³)	1.35 (0.97, 1.90)	0.97 (0.87, 1.07)	0.95 (0.91, 0.99)	1.06 (1.03, 1.11)
Individual studies in common between meta-analyses?	Sapkota & Stieb=2	Klepac & Stieb=2	No	No

Klepac et al 2018

Klepac et al (2018) reported summary estimates for PTB per 10 µg/m³ PM_{10} by exposure period as follows: Whole Pregnancy OR 1.09 (95% CI 1.03, 1.16), trimester 1 OR 1.04 (95% CI 1.01, 1.08), trimester 2 OR 1.04 (95% CI 0.98, 1.09) and trimester 3 OR 1.00 (95% CI 0.99, 1.00) (Table 8).

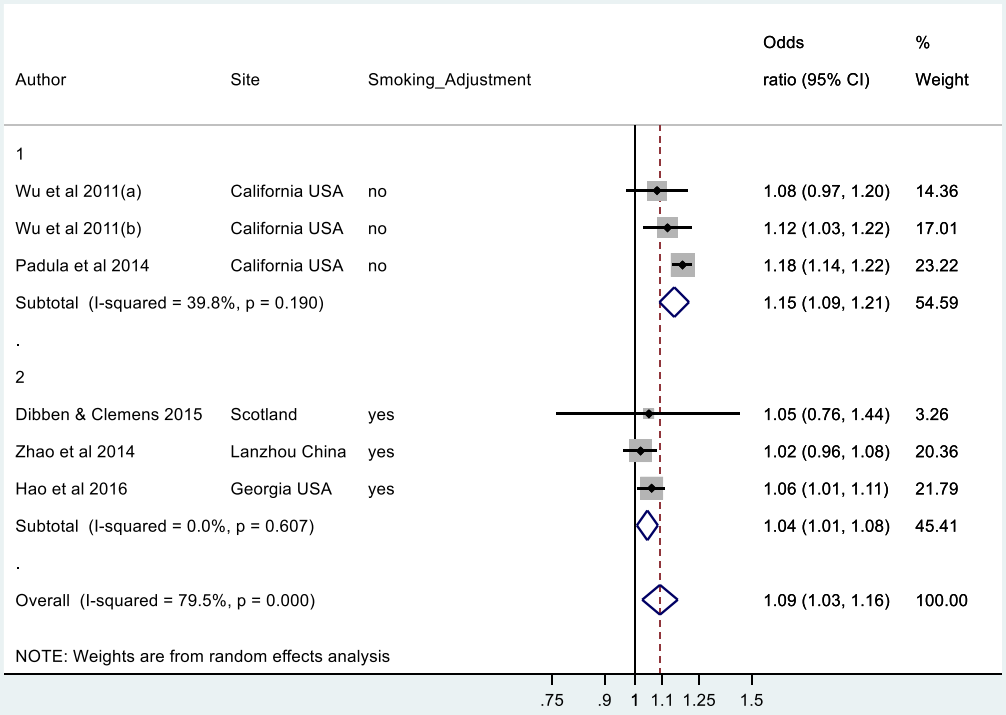
The Whole Pregnancy estimate incorporated 6 estimates from 5 studies, trimester 1 incorporated 11 estimates from 10 studies, trimester 2 incorporated 7 estimates from 6 studies, and trimester 3 incorporated 5 estimates from 5 studies; overall 11 studies were used at least once. Considerable heterogeneity was observed in the Whole

Pregnancy (I^2 79%; p=0.0002) and trimester 1 (I^2 91%; p<0.00001) and 2 (I^2 95%; p<0.00001) estimates. No important degree of heterogeneity (I^2 0%, p=0.90) was observed in the trimester 3 estimate (93% of the weight of this estimate came from one study (Lin et al, 2015). Publication bias was not formally examined.

Of the 11 studies included in the meta-analysis, 6 adjusted for smoking (3/5 for Whole Pregnancy, 5/10 for trimester 1, 3/6 for trimester 2, 3/5 for trimester 3 (Appendix 11, Table A 37). For Whole Pregnancy, and trimesters 1 and 2, just over 50% of the weight of the estimates came from studies with no adjustment for smoking (54.6% for Whole Pregnancy, 53.1% for trimester 1, 52.6% for trimester 2); only 2.1% of the weight of the estimates came from studies with no adjustment for smoking for trimester 3 (93% of the estimate came from Lin et al (2015), which did adjust for smoking). This was not explored further in the meta-analysis through stratification by studies that had adjusted for smoking. We have redrawn the information stratifying by smoking adjustment for Whole Pregnancy and trimester 1 where there was a significant summary estimate and evidence of heterogeneity (Figure 7). Visual inspection and assessment of the effect estimates suggested that, in studies that had not adjusted for smoking, there was greater heterogeneity (39.8% vs 0% for Whole Pregnancy and 94.7% vs 18.3% for trimester 1) and stronger reported positive (adverse) associations (OR 1.10 vs OR 1.04 for Whole Pregnancy and OR 1.10 vs OR 1.01 for trimester 1) for both estimates. We noted weak positive associations for the Whole Pregnancy estimates arising from the 3 studies that had adjusted for smoking (OR 1.04, 95% CI 1.01, 1.08). These effect estimates may still be subject to residual confounding by under-reporting of smoking by expectant mothers.

As in other meta-analyses, there were 2 studies (Lin et al, 2015; Huang et al, 2015) that provided estimates for each trimester but not for the Whole Pregnancy.

In view of the heterogeneity in reported associations, and the weak associations when considering only studies that have adjusted for smoking, we conclude that Klepac et al (2018) provides **no evidence of trimester 1 specific effects,** if we take account of the weaker associations observed when only studies that have adjusted for smoking are considered, and the risk of residual confounding by maternal smoking. There was **no evidence for trimester 2 or trimester 3 specific effects.** There was **weak evidence to support an association of PTB with PM_{10} exposure during Whole Pregnancy**. However, for Whole Pregnancy and trimester 3 this conclusion is based on limited evidence as these meta-analyses contained only 5 studies.



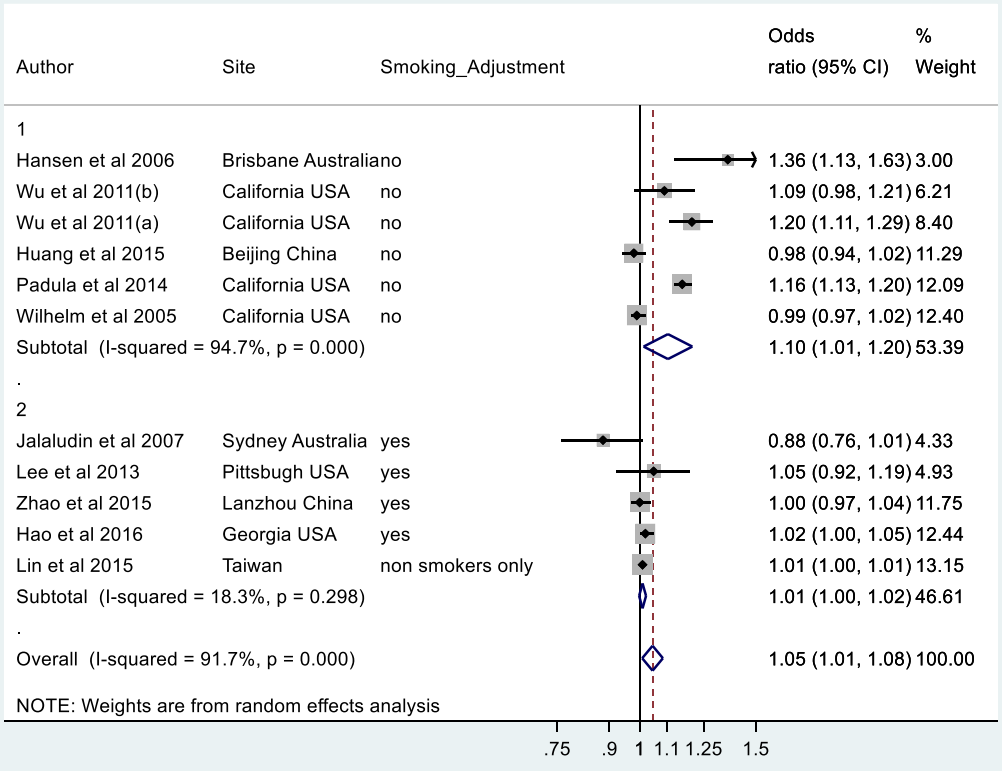


Figure 7. Redrawn Klepac et al (2018) estimates for PM_{10} and PTB, stratified by adjustment for smoking (Top = Whole Pregnancy; Bottom = trimester 1)

This is a forest plot of the meta-analyses used to derive separate summary effect estimates for studies with and without adjustment for confounding, as described in the text.

Sapkota et al 2012

Sapkota et al (2012) reported summary estimates for PTB per 10 μg/m^3 PM_{10} for Whole Pregnancy of OR 1.02 (95% CI 0.99, 1.04) (Table 8).

The Whole Pregnancy estimate incorporated 8 estimates from 8 studies. There was substantial heterogeneity (I^2 73%; p=0.001), (Appendix 11, Table A 37). It was not possible to infer the location and magnitude of the heterogeneity, as weight of estimates were not provided in the study. There was no significant publication bias (p>0.05) according to both the Begg's and Egger's bias tests.

The meta-analysis included 8 individual studies of which 4 adjusted for smoking (Appendix 11, Table A 37). We have redrawn the information stratifying by smoking adjustment, in an attempt to see the influence of smoking adjustment and to obtain overall weights within the analysis (Figure 8). This showed that 53% of the weight of the summary estimates come from studies that have not adjusted for smoking, but there was little difference in the estimates between studies with and without adjustment for smoking (OR 1.01 (95% CI 0.97, 1.04) vs OR 1.03 (95% CI 0.99, 1.06)).

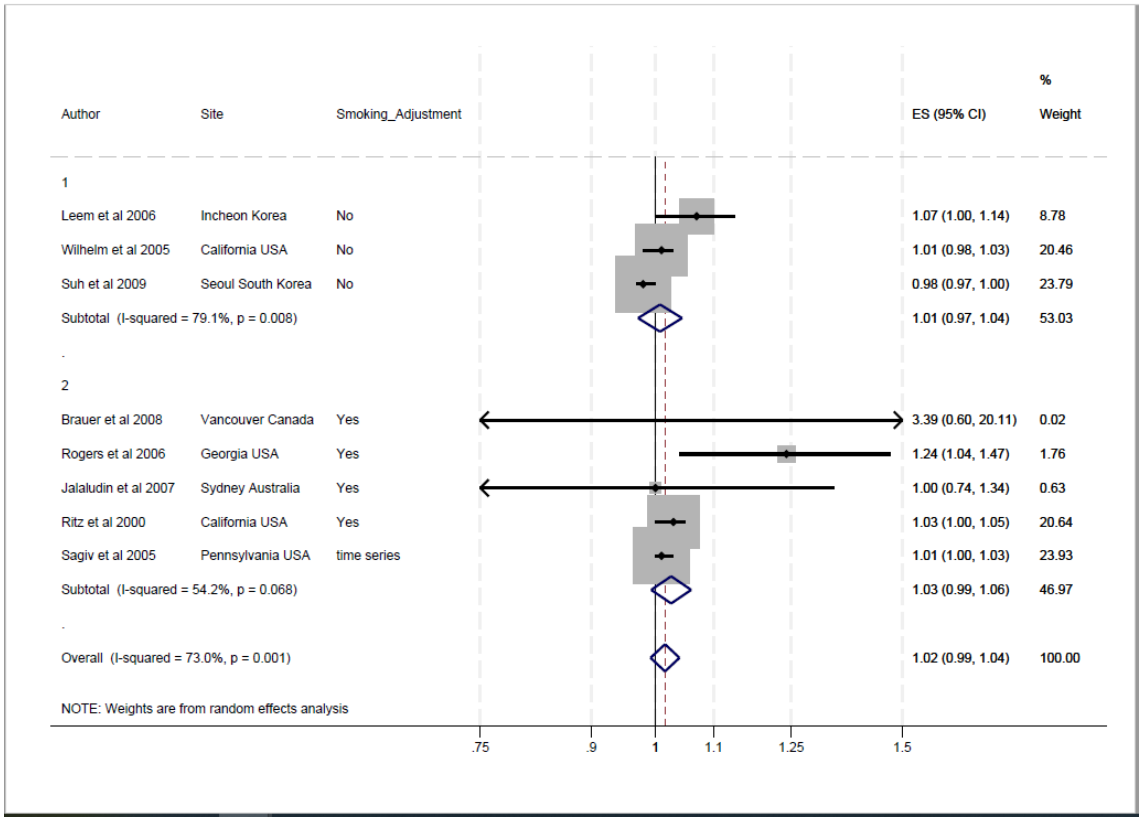


Figure 8. Redrawn Sapkota et al (2012) estimates for PM_{10} and PTB for Whole Pregnancy, stratified by adjustment for smoking. This is a forest plot of the meta-analyses used to derive separate summary effect estimates for studies with and without adjustment for confounding, as described in the text

We conclude Sapkota et al (2012) shows **no evidence to support an association of PTB with PM_{10} exposure during Whole Pregnancy**.

Stieb et al 2012

Stieb et al (2012) reported summary estimates for PTB per 20 μg/m^3 PM_{10} by exposure period as follows: Whole Pregnancy OR 1.35 (95% CI 0.97, 1.90), trimester 1 OR 0.97 (95% CI 0.87, 1.07), trimester 2 OR 0.95 (95% CI 0.91, 0.99) and trimester 3 OR 1.06 (95% CI 1.03, 1.11) (Table 8).

The meta-analysis included 8 individual studies (6 for trimester 1, 3 for trimester 2, 6 for trimester 3 and 3 for the Whole Pregnancy estimate). Examination of the 6 estimates contributing to the trimester 1 summary estimates showed almost 70% of the weight came from 4 studies with point estimates that were less than 1.00. The considerable heterogeneity (I^2 85.3%; p<0.001) is highly likely due to the strong positive (adverse) association in Hansen et al (2006) (5.88% of the weight) that did not adjust for smoking. Only 3 studies contributed to the trimester 2 summary estimate, which shows a weak negative association. There was no important degree of heterogeneity (I^2 0.0%; p=0.461), although this should be interpreted with caution as this estimate included only 3 studies. Estimates contributing to the trimester 3 and Whole Pregnancy summary estimates were broadly consistent (I^2 20.1%; p=0.282 and I^2 16.9%; p=0.300 respectively).

Overall, the methodological quality of this meta-analysis was considered to be good although for the association of PM_{10} and PTB the results of Begg's and/or Egger's bias tests were not presented (they were for other associations examined).

Of these 8 studies, 4 adjusted for smoking (2 out of 3 for Whole Pregnancy, 2 out of 6 for trimester 1, zero out of 3 for trimester 2, 2 out of 6 for trimester 3) (Appendix 11, Table A 37). For all trimester-specific estimates more than half the weight came from studies that did not adjust for smoking – (66.79% for trimester 1, 100% for trimester 2, 60.52% for trimester 3. The influence of adjustment for smoking was not explored within the report.

We conclude Stieb et al (2012) shows **no evidence of an association of exposure to PM_{10} in trimester 1, trimester 2 nor Whole Pregnancy with PTB**. Whilst the estimate for trimester 3 reached conventional statistical significance, given the considerable potential for residual confounding by maternal smoking in this estimate, we conclude it provides **weak evidence for an association of PTB with PM_{10} exposure during trimester 3.** However, for Whole Pregnancy and trimester 2 this conclusion is based on limited evidence as these meta-analyses contained only 3 studies.

Synthesis

Sapkota et al (2012) reviewed publications up to 2009, Stieb et al (2012) reviewed publications up to 2011 and Klepac et al (2018) included more recent literature, collating publications up to 2016. Remarkably few of the studies included in the earlier meta-analyses were included in the later one. None of the studies included in the Sapkota et al (2012) were included in the Klepac et al (2018) Whole Pregnancy estimate. None of the studies in the Stieb et al (2012) Whole Pregnancy estimate, or trimester 2 and 3 estimates were included in the equivalent estimates from Klepac et

al (2018). Furthermore, while Stieb et al (2012) suggest positive (adverse) associations for trimester 3 (but not for trimester 1), Klepac et al (2018) show positive associations for trimester 1 (but not for trimester 3). About half the weight of all summary estimates came from studies that did not adjust for smoking, which is a recognised major risk factor for PTB. Also, of note, Stieb et al (2012) and Sapkota et al (2012) used very different Whole Pregnancy estimates from the Brauer et al (2008) study. In Sapkota et al (2012) there was a positive (adverse) association with very wide interval, whereas in Stieb et al (2012) there was a negative association with a wide interval. This is because Stieb et al (2012) were using estimates for a standard definition for PTB (delivery at <37 weeks gestation), whereas Sapkota et al (2012) used a definition for very PTB (delivery at <30 weeks gestation)[7].

Conclusion regarding evidence for PM_{10} and PTB from core meta-analyses

The 3 meta-analyses provide weak evidence (section 2.1.3) for an association of PTB with PM_{10} and no indication as to a critical gestational window.

Commentary on evidence from key individual studies regarding PM_{10} and Preterm Birth (PTB) (not included in the core meta-analyses)

Following the flowchart process as outlined in Figure 6, Endpoint D was reached (Appendix 10).

Brief description of available studies:

24 studies examining the association between PM_{10} exposure during pregnancy and PTB were captured by our literature search and were not included in any of the core meta-analyses above (Giorgis-Allemand et al, 2017; Mendola et al, 2016; Smith et al, 2020; Hannam et al, 2014; Padula et al, 2019; Salihu et al, 2012; Poirier et al, 2015; Siddika et al, 2019; Liu Y et al, 2019; Sun et al, 2019; Ye et al, 2018; Zhao et al, 2020; Mendola et al, 2019; Weber et al, 2019; Arroyo et al, 2019a; Ghosh et al, 2019; Moreira et al, 2020; Dastoorpoor et al, 2018; Kim et al, 2019; Liang et al, 2018; Liu et al, 2018; Qian et al, 2016; Han et al, 2018; Le et al, 2012). Of these 20 were published between 2016-2020 and would not have been available to the authors of even the most recent meta-analysis on PTB evaluated above (Klepac et al, 2018), where the search date was 2016. The remainder of studies identified by our literature search were published between 2012-2016.
The majority of these additional studies used electronic health records, with 1 study pooling multiple cohorts and another reporting on a small birth cohort study.

The most striking difference in geographical spread compared to previous meta-analyses was the number of European studies (n=7). In particular, there had only been one study from Scotland in the previous meta-analyses. In addition, there were 6 from USA, 1 from Canada and 10 from Asia (7 from China, 1 from Mongolia, 1 from Korea and 1 from Iran).

[7] Stieb et al report that Brauer et al provided further information beyond what was in their published article

As stated in section 2.1.4, as we consider here only studies which do not appear anywhere in the core meta-analyses, there may, in addition, be individual studies that would add to the evidence for a specific exposure window which are not considered here.

Brief commentary on selected key studies of interest (Appendix 12 Table A 49) which are likely to expand the evidence base (but not a commentary on all):

Of the 7 new studies conducted in Europe we highlight:

- Smith et al (2020), which used information from large number of routine birth records (n=581,774) to examine the association between PTB and a range of particulate matter and gaseous air pollutant exposures in London, UK. The study used highly spatially resolved modelled exposures (monthly estimates from dispersion model at 20m x 20m grid point resolution), assigned to maternal address. They assessed the association of PTB with PM_{10} exposure in trimester 1 and trimester 2 only. Logistic regression analyses were adjusted for a wide range of potential confounders, including area-level adjustment for tobacco sales (as a proxy for smoking), and road traffic noise. Multi-pollutant models examining PM_{10} adjusted for a range of other air pollutants. This study showed a negative association of PTB with PM_{10} in trimester 1 or trimester 2 (OR 0.88 95% CI 0.86, 0.90 per 4.0 μg/m^3 in trimester 1: OR 0.82 95% CI 0.81, 0.84 in trimester 2).

- Giorgis-Allemand et al (2017), a pooled analysis of European birth cohorts (ESCAPE study) (n=71,493) which covers a wide geographical area of Europe, use of modelled exposures, and adjustment for smoking and temperature. The study examined the association of PTB with annual mean PM_{10} by collating information from several prospective birth cohorts that originated in Sweden, Norway, Netherlands, Greece, Lithuania, Hungary, Spain, Italy, and Denmark. A common protocol for LUR was used, exposure was based on maternal address and accounted for moving during pregnancy in some centres. LUR models were validated with specific monitoring campaigns (although these were only for short periods). Whole Pregnancy and trimester specific estimates for trimester 1 and trimester 2 were derived. Logistic regression analyses were adjusted for major confounders including maternal smoking and were used for trimester 1 and trimester 2 exposure windows. Survival analysis was used to assess associations with Whole Pregnancy exposure. There was no association of exposure to PM_{10} in Whole Pregnancy 0.97 (95% CI 0.87, 1.07 n= 46,971), nor in trimester 1 or trimester 2 (0.98 95% CI 0.90, 1.07 n=46,242: 0.98 95% CI 0.90, 1.06 n=47,153, respectively). There was no association for Whole Pregnancy exposure, trimester 1, or trimester 2 exposure to coarse particles ($PM_{10\text{-}2.5.}$)

- Hannam et al (2014) used information from hospital records from 2004-2008 (>250,000 births) conducted in North West England. Logistic regression analyses were conducted, adjusting for multiple confounders including maternal smoking to assess associations with modelled exposure (Defra's pollution climate mapping model). The PCM-modelled annual estimates were multiplied by monthly adjustment factors (MAF) enabling investigation of pollution effects during specific pregnancy periods. Due to missing data for smoking and BMI, the

sample size in adjusted models was considerably reduced. Odds ratios (OR) are presented as the fourth quartile (>46.3 μg/m^3) of pollution exposure compared to the first (<35.4 μg/m^3). There was no association of PTB with exposure to PM_{10} (0.90 95% CI 0.74, 1.11) in Whole Pregnancy or in any trimester specific period (trimester 1: 1.06 95% CI 0.92, 1.21; trimester 2: 0.95 95% CI 0.83, 1.09; trimester 3: 0.97 95% CI 0.84, 1.11).

Other studies based in Europe, but which are unlikely to be informative include Arroyo et al (2019a) which used administrative records of births from 2001-2009 in 46 Spanish provinces. Exposure was based on the monitor of the provincial capital city. There was no adjustment for smoking, and it is not clear if the published estimates coming from a meta-analysis are correctly derived (the figure suggests they are based on only the provinces in which statistically significant positive associations were observed, rather than all provinces). Ghosh et al (2019) only considered emissions from waste incinerators. Moreira et al (2020) focused on the effects of dust storms. Siddika et al (2019) only had information on 2568 births (195 PTBs).

Of the 6 studies in the USA, we highlight the following:

- Mendola et al (2016) examined associations of PTB with several air pollutants in over 220,000 deliveries (about 11% were PTB) in a large US obstetric cohort (Consortium on Safe Labor), that collected detailed medical records from 19 hospitals/15 hospital referral regions across the US (in the states of California (LA), Utah, Texas, Illinois, Ohio, Indiana, Florida, Washington DC, Maryland, Delaware, NY State and Maine). The main stated aim of the analysis was to assess whether associations were different in mothers with and without diagnosed asthma. Exposures were estimated using a modified Community Multiscale Air Quality Modelling System (CMAQ) model. Researchers did not have access to maternal address, so exposure assignment was based on the average concentrations in the 15 non-overlapping delivery hospital referral regions of hospital of delivery (area range: 415–312,644 km^2), weighted for population density. Authors presented 14 analyses stratified by asthma (one each for each gestational week – weeks 23 to 36) for each of CO, O_3, oxides of nitrogen (NO_x), $PM_{2.5}$, PM_{10}, and SO_2 and tested interactions. A correction for potential false discovery was made. Estimates were adjusted for maternal age, race, pre-pregnancy BMI, smoking and alcohol use, study site, parity, insurance status, marital status, and comorbidities. After allowing for false discovery there was no evidence that associations were different in women with and without asthma at any particular gestational week (a positive significant interaction had been observed at week 29 for PM_{10} prior to correction for false discovery). Whole pregnancy exposure to PM_{10} for early PTB (before 34 weeks) was 1.09 95% CI 1.00, 1.20 (mothers with asthma) and 1.03 95% CI 0.99, 1.06 for mothers without asthma. For the exposure window ‘last 6 weeks of pregnancy’ there was evidence that associations of PM_{10} with PTB were different in mothers with and without asthma (1.13 95% CI 1.03, 1.24 and 0.98 95% CI 0.85, 1.00 respectively).

- Other studies based in USA but which are unlikely to be informative include Mendola et al (2019) which used modelled exposures (CMAQ), and adjusted for smoking, but its analytical design and exposure contrasts (PM_{10} level high/ medium/ low in second vs first pregnancy) may make it difficult to incorporate estimates into any new meta-analysis. Padula et al (2019) and Weber et al (2019) considered the same population as in Padula et al (2014) (which was included in the Klepac et al (2018) meta-analysis). Salihu et al (2012) had population overlap with Mendola et al (2016) and Le et al (2012) only considered exposure in the first month and the last month.

We highlight one study from Canada:

- Poirier et al (2015) used birth records to examine associations of several pollutants with a range of birth outcomes including PTB (1,043 cases and 13,709 controls). Whole Pregnancy exposure was determined from a LUR model and based on the post-code of mother's address. The LUR model was able to predict 42% of the variation in PM_{10}. Logistic regression models were adjusted for mother's age, parity, smoking, and neighbourhood income. There was no association of PM_{10} exposure during pregnancy with PTB (OR 0.93 95% CI 0.88, 0.98 per IQR of PM_{10})

There were 10 studies from Asia (7 from China, 1 from Korea, 1 from Mongolia, 1 from Iran). It was not easy to identify the studies likely to be most informative as all used environmental monitoring for exposure assignment (not modelling), the exposures to PM_{10} were very high, the vast majority adopted logistic regression approaches rather than time to event/time series approaches and most did not adjust for smoking (although they may report that smoking levels in mothers in China/Asia are very low).

- Sun et al (2019) used Euclidean distance matching to the nearest monitor for 6,275 pregnancies (372 PTB) and adopted a time to event analysis. They did not adjust for smoking. PM_{10} ranged from 34.3 $\mu g/m^3$ to 146.3 $\mu g/m^3$.

- Ye et al (2018) used station-based inverse distance weighting (IDW) to calculate exposure value at maternal address (n~24,000, all reported as non-smoking, with ~1,500 PTBs) from the 4 monitoring stations and conducted logistic regression to generate risks per IQR change in PM_{10}. Whole Pregnancy exposure to PM_{10} was high (Mean 108.6, 10.3 Min, 80.7 Max, 146.0 25th centile, 101.7 75th centile).

- Kim et al (2019) in Korea included over 1.7 million births but it is unclear how PM_{10} measures were derived from the 266 monitoring stations in Korea and results are presented with PM_{10} considered in quartiles or as binary (with cut off at 70 $\mu g/m^3$).

- Dastoorpoor et al (2018) in Iran conducted a time series on PTB but the levels of 24-hour average PM_{10} (average of 4 stations) were very high (mean 162.5 $\mu g/m^3$).

- Liu et al (2018) was a time series reporting excess risks of PM_{10} exposure in the 7 days prior to birth on PTB.

- Liu Y et al (2019) was small case control study (PTB: 687 Controls: 1,766) and exposure was based on average city level. Logistic regression models were not adjusted for smoking.

- Qian et al (2016) used nearest monitor exposure assessment in a sample of about 97,000 births. Logistic regression models were adjusted for exposure to tobacco smoke.

- Zhao et al (2020) examined associations in 70,000 births (5,842 PTB) through nearest monitor methods for exposure assignment. They did not mention or adjust for smoking behaviours in mothers in their logistic regression models.

- Han et al (2018) examined associations in 6,693 pregnancies using logistic regression models with exposure assigned through inverse distance weighting. There was no adjustment for smoking.

- Liang et al (2018) evaluated the effect of a cold spell on PTB.

Overall conclusion regarding evidence for PM_{10} and PTB from key individual studies not included in the core meta-analyses

There were a large number of studies not captured in/published since the core meta-analyses. We note that the key highlighted studies that have been conducted in Europe do not support an association between PM_{10} and PTB, but there are a number of studies with varying design and methodology from Asia that need further consideration.

OVERALL CONCLUSION for PM_{10} and PTB from the epidemiological evidence

Based on the meta-analyses available, we conclude there is weak evidence (section 2.1.3) for an association between PTB and PM_{10} exposure. Examination of key highlighted studies points to the need for an up-to-date meta-analysis to synthesise all the available evidence.

NO_2 and PTB

Commentary on results from meta-analyses

After exclusions, the core meta-analyses that examined associations between NO_2 exposure and PTB were Stieb et al (2012) for trimesters 1, 2 and 3 and Whole Pregnancy and Klepac et al (2018) for trimesters 2 and 3 and Whole Pregnancy. Results are provided in Table 9, with more detailed information provided in Appendix 11, Table A 38.

Table 9. Summary of the results from the core meta-analyses examining preterm birth and NO_2 exposure

Meta-analyses (Increment)	OR (95% CI)			
	WP	T1	T2	T3
Klepac et al 2018 (Per 10 ppb)	1.05 (0.99, 1.11)	Excluded	1.02 (0.97, 1.08)	1.02 (0.96, 1.08)
Stieb et al 2012 (Per 20 ppb)	1.16 (0.83, 1.63)	0.87 (0.64, 1.17)	1.03 (0.77, 1.39)	1.06 (0.96, 1.18)
Individual studies in common between meta-analyses?	Klepac&Stieb=1	n/a	Klepac&Stieb=1	Klepac&Stieb=1

Klepac et al 2018

Klepac et al (2018) report summary estimates for PTB per 10 ppb NO_2, by exposure period, as follows: Whole Pregnancy OR 1.05 (95% CI 0.99, 1.11), trimester 2 OR 1.02 (95% CI 0.97, 1.08), and trimester 3 OR 1.02 (95% CI 0.96, 1.08) (Table 9).

The Whole Pregnancy estimate incorporated 8 estimates from 7 studies, trimesters 2 incorporated 10 estimates from 9 studies, and trimester 3 incorporated 7 estimates from 7 studies; overall 10 studies were used at least once. There was considerable heterogeneity in all of the summary estimates (I^2 95%; p<0.00001 for Whole Pregnancy, I^2 96%; p<0.00001 for trimester 2, and I^2 97%; p<0.00001 for trimester 3). Publication bias was not evaluated.

Of the 10 studies included in the meta-analyses, 3 adjusted for smoking (2 out of 7 for Whole Pregnancy, 3 out of 9 for trimester 2, and 3 out of 7 for trimester 3) (Appendix 11, Table A 38). For Whole Pregnancy, and trimester 3, over 60% of the weight of the estimates came from studies with no adjustment for smoking (68.9% for Whole Pregnancy, 62.6% for trimester 3); 41.3% of the weight of the estimates came from studies with no adjustment for smoking for trimester 2.

The summary estimate for Whole Pregnancy was slightly elevated (5%) but did not reach conventional levels of significance. Summary estimates for trimesters 2 and 3 were around null and provide no evidence for an association. Klepac et al (2018) **provides no evidence to support an association of NO_2 exposures and PTB.**

Stieb et al 2012

Stieb et al (2012) reported summary estimates for PTB per 20 ppb NO_2, by exposure period, as follows: Whole Pregnancy OR 1.16 (95% CI 0.83, 1.63), trimester 1 OR 0.87 (95% CI 0.64, 1.17), trimester 2 OR 1.03 (95% CI 0.77, 1.39) and trimester 3 OR 1.06 (95% CI 0.96, 1.18) (Table 9).

The Whole Pregnancy estimate included 5 studies, trimester 1 included 6 studies, trimester 2 included 2 studies, and trimester 3 included 6 studies; overall 8 studies were used at least once. Whilst heterogeneity might not be important for trimesters 2 (I^2 21.6%; p=0.259) and 3 (I^2 19.5%; p=0.286), there was substantial heterogeneity for Whole Pregnancy (I^2 53.3%; p=0.073) and considerable heterogeneity for trimester 1 (I^2 89.1%; p<0.001) estimates. There was no evidence of publication bias.

Of the 8 studies included in the meta-analyses, 2 adjusted for smoking (Jalaludin et al (2007), which was included in the trimester 1 and 3 estimates and Brauer et al (2008), which was used in the Whole Pregnancy estimate) (Appendix 11, Table A 38). For trimesters 1, 2 and 3, over 80% of the weight of the estimates came from studies with no adjustment for smoking (82.32% for trimester 1, 100.00% for trimester 2, and 95.79% for trimester 3); 62.13% of the weight of the estimates came from studies with no adjustment for smoking for Whole Pregnancy.

Whilst summary estimates for Whole Pregnancy and trimesters 2 and 3 were slightly elevated, none of the summary estimates in Stieb et al (2012) reached conventional levels of significance. Stieb et al (2012) **provides no evidence to support an association of NO_2 exposures and PTB. However, this conclusion is based on limited evidence as the meta-analysis for Whole Pregnancy and for trimester-specific exposures contained small numbers of studies.**

Synthesis

The evidence from the 2 meta-analyses is largely consistent, in that the associations were all around null or slightly elevated and none reached conventional levels of significance. This was the case even though the study findings are reported in each meta-analysis by different exposure increments. It is reassuring to find this consistency despite differences between the meta-analyses in the number of studies incorporated, their geographies and populations, and the differential control for confounding by smoking.

Conclusion regarding evidence for NO_2 and PTB from core meta-analyses

Evidence from the 2 meta-analyses does not support an association (section 2.1.3) of prenatal exposure to NO_2 and PTB in Whole Pregnancy or any of the individual trimesters.

Commentary on evidence from key individual studies regarding NO_2 and Preterm Birth (PTB) (not included in the core meta-analyses)

Following the flowchart process as outlined in Figure 6, Endpoint D was reached (Appendix 10).

Brief description of available studies

31 studies examining the association between NO_2 exposure during pregnancy and PTB were captured by our literature search and were not included in any of the core

meta-analyses above (Giorgis-Allemand et al, 2017; Chen et al, 2018; Trasande et al, 2013; Smith et al, 2020; Hannam et al, 2014; Ritz et al, 2007; Padula et al, 2019; Poirier et al, 2015, Siddika et al, 2019, Liang et al, 2019b; Guo et al, 2018; Ye et al, 2018; Zhao et al, 2020; Stieb et al, 2019; Ranjbaran et al, 2020; Mendola et al, 2019; Weber et al, 2019; Arroyo et al, 2019a; Dastoorpoor et al, 2018; Liu et al,2018; Qian et al, 2016; Le et al, 2012; Liu Y et al, 2019; Wilhelm et al, 2012; Llop et al, 2010; Panasevich et al, 2016; Ji X et al, 2019; Lu et al, 2020; Leem et al, 2006; Saldiva et al, 2018; Deguen et al, 2018). Of these, 21 were published during 2017-2020, so would not have been captured by the literature search of the most recent meta-analysis evaluated above (search cut-off date in 2016) (Klepac et al, 2018). The remainder (n=10) were published between 2006-2016. There were 17 retrospective cohort studies using birth records or electronic medical records, 3 prospective cohort studies, 1 study pooling multiple birth cohorts, 4 case-control studies, 1 cross-sectional study, 1 ecological study and 4 time-series studies. Geographically they were spread as follows: USA (n=7 (of which 4 were in California)), Canada (n=2), South America (n=1), Europe (n=8), Asia (n=12 (of which 8 were in China)), and Australia (n=1). N.B., it is possible that some of the earlier studies were not included in meta-analyses above due to decisions authors made regarding which study to include where there were several examining the same/similar population.

As stated in section 2.1.4 as we consider here only studies that did not appear anywhere in the core meta-analyses, there may, in addition, be individual studies that would add to the evidence for a specific exposure window which are not considered here.

Nine of these 31 studies would not have met the criteria for the specific meta-analyses we evaluated (that is, did not examine Whole Pregnancy, trimester 1, trimester 2, trimester 3 exposure windows/ or were time-series studies examining short-term effects) or were not particularly informative as the exposure window was not clearly reported. As they are not comparable with the specific exposure windows considered in meta-analyses above, we do not comment on them further here (Trasande et al, 2013; Stieb et al, 2019; Ranjbaran et al, 2020; Arroyo et al, 2019a; Dastoorpoor et al, 2018; Liu et al, 2018; Le et al, 2012; Saldiva et al, 2018; Deguen et al, 2018).

Brief commentary on selected key studies of interest (shown in Appendix 12 Table A 50*) which are likely to expand the evidence base (but not a commentary on all):*

22 of the 31 studies examined exposures in the Whole Pregnancy, trimester 1, trimester 2, or trimester 3 time windows – we consider these as they have potential to add to the evidence from meta-analyses above, in relation to these specific exposure windows, for example, in any future meta-analysis (Giorgis-Allemand et al, 2017; Chen et al, 2018; Smith et al, 2020; Hannam et al, 2014; Ritz et al, 2007; Padula et al, 2019; Poirier et al, 2015; Siddika et al, 2019; Liang et al, 2019b; Sun et al, 2019; Ye et al, 2018; Zhao et al, 2020; Mendola et al, 2019; Weber et al, 2019; Qian et al, 2016; Liu Y et al, 2019; Wilhelm et al, 2012; Llop et al, 2010; Panasevich et al, 2016; Ji X et al, 2019; Lu et al, 2020; Leem et al, 2006). Geographically they were spread as follows: USA (n=5), Canada (n=1), Europe (n=6), Asia (n=9, of which 7 were in China), and Australia (n=1).

The one study from Canada is highlighted, as it expanded geographical coverage of Canada beyond the areas covered in the meta-analyses: Poirier et al (2015) used birth records to examine associations of several pollutants with a range of birth outcomes including PTB (1,043 cases and 13,709 controls) in Halifax, Nova Scotia (Canada). Whole Pregnancy exposure was determined from a LUR model and based on the post-code of mother's address. The LUR model was able to predict 90% of the variation in NO_2. There were low concentrations of pollutants, and likely limited exposure contrast. Logistic regression models were adjusted for mother's age, parity, smoking, and neighbourhood income. There was no association of NO_2 exposure during pregnancy with PTB (OR 0.94 95% CI 0.85, 1.04 per IQR of NO_2).

The one study from Australia (Chen et al, 2018) is highlighted as it is larger in size (n= 173,800) than previous studies from Australia included in the meta-analyses. Unlike the previous Brisbane study (75) it adjusts for smoking. It also uses survival analysis, which reduces risk of bias when analysing the relationship between whole pregnancy and trimester 3 exposures in relation to PTB (see 'Specific considerations for studies looking at Preterm birth (PTB) and Stillbirth' in section 2.1.4), and examines multi-pollutant models. It found an elevated risk of PTB associated with Whole Pregnancy exposure to NO_2 (HR 1.07 95% CI 1.01, 1.13) per IQR (1.19 ppb) increase in NO_2). In multi-pollutant models hazard ratios for NO_2 were attenuated to null after adjustment for other pollutants ($PM_{2.5}$, SO_2, NO_2) individually and jointly. Trimester-specific exposures were analysed, but numeric results were not presented, the authors simply commented 'for all air pollutants, highest HRs were present at trimester 3 and the lowest at trimester 1.'

Of 6 studies in Europe (Giorgis-Allemand et al, 2017; Smith et al, 2020; Hannam et al, 2014; Siddika et al, 2019; Llop et al, 2010; Panasevich et al, 2016) we highlight:

- Giorgis-Allemand et al (2017), a pooled analysis of European birth cohorts (ESCAPE study) (n=71,493), which covered a wide geographical area of Europe, used modelled exposures, and adjusted for smoking and temperature. The study examined the association of PTB with annual mean NO_2 by collating information from several prospective birth cohorts that originated in Sweden, Norway, Netherlands, Lithuania, Hungary, Spain, Italy and Denmark, France, and England. A common protocol for LUR was used, exposure was based on maternal address and accounted for moving during pregnancy in some centres. LUR models were validated with specific monitoring campaigns (although these were only for short periods). Whole Pregnancy and trimester specific estimates for trimester 1 and trimester 2 were derived. Logistic regression analyses were adjusted for major confounders including maternal smoking and were used for trimester 1 and trimester 2 exposure windows. Survival analysis was used to assess associations with whole pregnancy exposure. There was no association of exposure to NO_2 in Whole Pregnancy with PTB (OR 0.96 95% CI 0.91, 1.01 n= 56,977), nor in trimester 1 or trimester 2 (trimester 1 OR 0.97 95% CI 0.92,1.02, n=55,811: trimester 2 OR 0.96 95% CI 0.92, 1.01 n=55,892, respectively).

- Smith et al (2020) examining London, UK, due to its large sample size, used modelled exposures, and multi-pollutant models. Smith et al (2020) used information from a large number of routine birth records (n=581,774) to examine

the association between PTB and a range of particulate matter and gaseous air pollutant exposures in London, UK. The study used highly spatially resolved modelled exposures (monthly estimates from dispersion model at 20m x 20m grid point resolution), assigned to maternal address. They assessed the association of PTB with NO_2 exposure in trimester 1 and trimester 2 only. Logistic regression analyses were adjusted for a wide range of potential confounders, including area-level adjustment for tobacco sales (as a proxy for smoking), and road traffic noise. Multi-pollutant models examining NO_2 adjusted for a range of other air pollutants. This study showed a negative association of PTB with NO_2 in trimester 1 or trimester 2 (OR 0.92 95% CI 0.90, 0.94 per 14.0 μg/m^3 in trimester 1: OR 0.88 95% CI 0.86, 0.90 per 14.4 μg/m^3 in trimester 2). In multi-pollutant models adjusting for O_3, the trimester 1 OR for NO_2 was attenuated to null, whilst a negative association remained for trimester 2 NO_2.

- Hannam et al (2014) due to its use of modelled exposures, and adjustment for smoking. Hannam et al (2014) used information from hospital records from 2004-2008 (>250,000 births) conducted in North West England. Logistic regression analyses were conducted, adjusting for multiple confounders including maternal smoking, to assess association of modelled exposure (Defra's pollution climate mapping model). The PCM-modelled annual estimates were multiplied by monthly adjustment factors (MAF) enabling investigation of pollution effects during specific pregnancy periods. Due to missing data for smoking and BMI, the sample size in adjusted models was considerably reduced. Odds ratios (OR) were presented as the fourth quartile (>63.6 μg/m^3) of pollution exposure compared to the first (<42.9 μg/m^3). There was no association of PTB with exposure to NO_2 (1.10 95% CI 0.82, 1.48) in whole pregnancy or in any trimester specific period (trimester 1 1.08 95% CI 0.92, 1.27; trimester 2 1.07 95% CI 0.91, 1.25; trimester 3 1.01 95% CI 0.86, 1.19).

The remaining 3 European studies were from Spain, Finland, and Norway, which further expands European geographic coverage beyond that in the meta-analyses above, and all had adjustment for smoking and modelled exposure estimates but were much smaller than those highlighted above.

Of 9 new studies from Asia (Liang et al, 2019b; Sun et al, 2019; Ye et al, 2018; Zhao et al, 2020; Qian et al, 2016; Liu Y et al, 2019; Ji X et al, 2019; Lu et al, 2020; Leem et al, 2006;) we highlight:

- A study by Liang et al (2019b), due to its large sample size (n=628,439), its use of survival analysis (which reduces risk of bias when analysing PTB), exclusion of caesarean deliveries (which will exclude some if not all medically indicated early deliveries) and analysis of multi-pollutant models. Liang et al (2019b) examined associations of preterm with several air pollutants in a large retrospective birth cohort using birth registration data, across 7 cities in the Pearl River Delta region of China (Guangzhou, Shenzhen, Zhuhai, Dongguan, Foshan, Jiangmen, and Zhongshan). The main stated aim of the analysis was to assess whether associations were different in mothers who were local residents versus migrants, and all analyses were stratified by local resident/migrant status, however no interactions between migrant status and pollutant exposure were tested.

Exposure assessment was via environmental monitoring. It found negative or null associations between PTB and NO_2 and no differences by migration status (HRs (95% CIs) of PTB associated with a 10 μg/m^3 increase in NO_2: Local residents: Whole Pregnancy: 0.96 (0.94, 0.98); trimester 1: 0.97 (0.96, 0.99); trimester 2: 0.98 (0.96, 0.99); trimester 3: 1.01 (0.99, 1.02) and Migrants: Whole Pregnancy: 0.95 (0.94, 0.97); trimester 1: 1.01 (0.99, 1.03); trimester 2: 1.01 (0.99, 1.02); trimester 3: 1.01 (0.99, 1.02)). Multi-pollutant models were analysed for whole pregnancy exposures – associations between NO_2 and PTB were similar (negative) after adjusting for $PM_{2.5}$, NO_2 and SO_2, in turn.

- All the studies from China based exposure assessment on environmental monitoring, except one small retrospective birth cohort (n=25,493) in Shanghai by Ji X et al (2019) which we highlight for its use of modelled estimates (although these were annual concentrations so not temporally highly resolved), multi-pollutant models, and also as it additionally conducted a matched case-control design to account for differing Whole Pregnancy and trimester 3 length between PTBs vs term births which reduces risk of bias. This study showed an association between trimester 3 NO_2 exposure and PTB, but associations were null for Whole Pregnancy and trimester 1 and 2 exposures (OR (95% CI) for 10 μg/m^3 increase in NO_2: Whole Pregnancy: 1.03 (0.96, 1.10); trimester 1: 1.00 (0.95, 1.07); trimester 2: 1.01 (0.96, 1.07); trimester 3**:** 1.07 (1.03, 1.13)). Associations between NO_2 and PTB remained very similar in multi-pollutant models adjusting for $PM_{2.5}$.

All the studies from China were from different regions, so the geographical coverage of China was expanded beyond the one Beijing study (Huang et al, 2015) included in meta-analyses above.

Of 5 new studies in the USA, we did not consider any likely to be particularly informative (Ritz et al, 2007; Padula et al, 2019; Mendola et al, 2019; Weber et al, 2019; Wilhelm et al, 2012). Mendola et al (2019) used modelled exposures (CMAQ), and adjusted for smoking, but its analytical design and exposure contrasts (NO_2 level high/medium/low in second vs first pregnancy) may make it difficult to incorporate estimates into any new meta-analysis. Whilst there were several studies from California, none of these studies adjusted for smoking or expanded geographical/time coverage of the state beyond the meta-analyses above: Padula et al (2019) and Weber et al (2019) considered the same population as in Padula et al (2014) (which was included in the Klepac et al (2018) meta-analysis), but stratified by diabetes and hypertension; and the study populations in Ritz et al (2007) and Wilhelm et al (2012) overlapped with that in Wu et al (2011) (which was included in the Klepac et al (2018) meta-analysis).

The results from these highlighted studies demonstrate some statistically significant increased risks of PTB associated with long-term exposure to NO_2 during pregnancy (Chen et al, 2018; Mendola et al, 2019; Ji X et al, 2019), some significant inverse associations (Smith et al, 2020; Liang et al, 2019b;), whilst some found no significant association (Giorgis-Allemand et al, 2017; Hannam et al, 2014; Poirier et al, 2015).

Conclusion regarding evidence for NO_2 and PTB from key individual studies not included in the core meta-analyses

There were a very large number of studies not captured in or published since the core meta-analyses, and the results from key studies highlighted were inconsistent (section 2.1.4).

OVERALL CONCLUSION for NO_2 and PTB from the epidemiological evidence

We conclude that evidence from the meta-analyses does not support an association (section 2.1.3) between PTB and NO_2 exposure. Examination of key highlighted studies points to the need for an up-to-date meta-analysis to synthesise all the available evidence.

O_3 and PTB

Commentary on results from meta-analyses

After exclusions, the core meta-analyses which examined associations between O_3 exposure and PTB were Stieb et al (2012) for Whole Pregnancy and trimesters 1, 2 and 3, and Klepac et al (2018) for Whole Pregnancy and trimesters 2 and 3. Results are provided in Table 10 and Appendix 11, Table A 39.

Table 10. Summary of the results from the core meta-analyses examining preterm birth and O_3 exposure

Meta-analyses (Increment)	OR (95% CI)			
	WP	T1	T2	T3
Klepac et al 2018 (Per 10 ppb)	1.03 (1.01, 1.04)	Excluded	1.12 (1.05, 1.19)	1.01 (0.99, 1.03)
Stieb et al 2012 (Per 20 ppb)	1.92 (0.38, 9.76)	1.22 (0.91, 1.64)	0.94 (0.88, 1.00)	0.97 (0.86, 1.10)
Individual studies in common between meta-analyses?	No	No	No	No

Klepac et al 2018

Klepac et al (2018) reported summary estimates for PTB per 10 ppb O_3, by exposure period, as follows: Whole Pregnancy OR 1.03 (95% CI 1.01, 1.04), trimester 2 OR 1.12 (95% CI 1.05, 1.19), and trimester 3 OR 1.01 (95% CI 0.99, 1.03) (Table 10). The meta-analyses included 5, 6 and 4 individual studies respectively.

There was considerable heterogeneity in trimester 2 (I^2=98%, p<0.00001) and trimester 3 (I^2=89%, p<0.00001) estimates, but none (I^2=0%) in the Whole Pregnancy estimate. For trimester 2 the odds ratios in the individual studies were all greater than 1 and very similar in magnitude, with the exception of one study which was substantially greater than 1, that is, the heterogeneity was in size of effect. For trimester 3 the odds ratios in the individual studies were both just greater than 1 and just below 1 but were all relatively small in magnitude, that is, the heterogeneity was in direction of the association. Publication bias was not evaluated.

Most of studies included in each meta-analysis adjusted for smoking, meaning that 93%, 76% and 100% of the weight of the Whole Pregnancy, trimester 2 and trimester 3 meta-analysis estimates respectively came from studies which had adjusted for smoking.

Summary estimates for Whole Pregnancy and trimester 2 O_3 exposure reached conventional levels of statistical significance, whereas the summary estimate for trimester 3 did not. Given the consistency in direction of effect for individual studies (for Whole Pregnancy and trimester 2), and that the majority of the weight in summary estimates reflects studies which have adjusted for smoking, and taking into account statistical significance, but balancing this against weaknesses in the methodology or reporting (see section 2.2.1), we conclude that Klepac et al (2018) provides **weak evidence to support an association of PTB with O_3 exposure across Whole Pregnancy (small positive association) and trimester 2 (positive association). The summary estimate for trimester 3 was null and provides no evidence for an association. However, these conclusions are based on limited evidence as the meta-analyses contained small numbers of studies.**

Stieb et al 2012

Stieb et al (2012) reported summary estimates for PTB per 20 ppb O_3, by exposure period, as follows: Whole pregnancy OR 1.92 (95% CI 0.38, 9.76), trimester 1 OR 1.22 (95% CI 0.91, 1.64), and trimester 3 OR 0.97 (95% CI 0.86, 1.10) (Appendix 11, Table A 39). These meta-analyses contained 2, 4 and 4 individual studies respectively. The forest plot for trimester 2 included only a single study and therefore is not a meta-analysis and is not commented upon further.

The 2 studies included in the whole pregnancy meta-analysis showed different results: one null and one positive association. There was considerable heterogeneity in the trimester 1 (89.8%, p<0.0001) meta-analysis, with the odds ratios in the individual studies being both greater than 1 (sometimes substantially) and below 1, that is, the heterogeneity was both in size and in direction of the association. However, with just 4 studies included, this statistic may not be particularly informative. There was no evidence of publication bias (funnel plot asymmetry).

Only one study per meta-analysis adjusted for smoking, meaning that 55%, 75% and 87% of the weight of the whole pregnancy, trimester 1 and trimester 3 meta-analysis estimates respectively came from studies with no adjustment for smoking.

Given that the Whole Pregnancy meta-analysis was based on only 2 studies with differing results, considerable heterogeneity in the trimester 1 meta-analysis, considerable potential for residual confounding by smoking in all these estimates, and that none of the summary estimates reached conventional levels of significance we conclude that Stieb et al (2012) provides **no evidence to support an association of PTB with O_3 exposure. However, this conclusion is based on limited evidence as the meta-analyses contained small numbers of studies.**

Synthesis

The evidence from the 2 meta-analyses is inconsistent. This may reflect a) that none of the studies included in Stieb et al (2012) were included in the later meta-analysis by Klepac et al (2018) and b) differential control for confounding by smoking: the vast majority, if not all, of the weight in each meta-analysis by Klepac et al (2018) came from studies which had adjusted for smoking, whereas the opposite was true for the meta-analyses by Stieb et al (2012) in which over half the weight of all summary estimates came from studies which did not adjust for smoking.

Conclusion regarding evidence for O_3 and PTB from core meta-analyses

One out of the 2 meta-analyses provides weak evidence (section 2.1.3) for an association between PTB and O_3 exposure. Specifically, one meta-analysis (which included more recent studies and much greater control for confounding by smoking) provides some evidence to support an association between PTB and O_3 exposure (Klepac et al, 2018). However, due to weaknesses in the methodology or reporting (see section 2.2.1), it is unclear whether the quantitative estimates presented by Klepac et al (2018) comprehensively reflect the body of evidence available at the time of the meta-analysis. This conclusion is based on limited evidence as both meta-analyses analysed only a small number of studies (ranging from 2 to 6 studies).

Commentary on evidence from key individual studies regarding O_3 and Preterm Birth (PTB) (not included in the core meta-analyses)

Following the flowchart process as outlined in Figure 6, Endpoint D was reached (Appendix 10).

Brief description of available studies:

Twenty-six studies examining the association between O_3 exposure during pregnancy and PTB were captured by our literature search and were not included in any of the core meta-analyses above (Chen et al, 2018; Mendola et al, 2016; Symanski et al, 2016; Trasande et al, 2013; Smith et al, 2020; Ritz et al, 2007; Lee et al, 2013; Huang et al, 2018; Siddika et al, 2019; Liang et al, 2019b; Sun et al, 2019; Li et al, 2018; Seabrook et al, 2019; Stieb et al, 2019; Ranjbaran et al, 2020; Mendola et al, 2019; Arroyo et al, 2019a; Dastoorpoor et al, 2018; Liu et al, 2018; Qian et al, 2016; Han et al, 2018; Le et al, 2012; Liu Y et al, 2019; Saldiva et al, 2018; Gray et al, 2014; Capobussi et al, 2016). Seventeen were published during

2017-2020, so would not have been captured by the literature search of the most recent meta-analysis evaluated above (search cut-off date in 2016) (Klepac et al, 2018). The remainder (n=9) were published between 2007-2016. There were 18 retrospective cohort studies using birth records or electronic medical records, one prospective cohort study, 2 case-control studies, one cross-sectional study, and 4 time-series studies. Geographically they were spread as follows: USA (n=9), Canada (n=2), South America (n=1), Europe (n=4), Asia (n=9), Australia (n=1). N.B., it is possible that some of the earlier studies were not included in meta-analyses summarised above because authors excluded them without reporting reasons (for example, population overlaps).

As stated in section 2.1.4 as we consider here only studies which do not appear anywhere in the core meta-analyses; there may, in addition, be individual studies that would add to the evidence for a specific exposure window which are not considered here.

Nine of these 26 studies (Li et al, 2018; Seabrook et al, 2019; Stieb et al, 2019; Ranjbaran et al, 2020; Arroyo et al, 2019a; Dastoorpoor et al, 2018; Liu et al, 2018; Le et al, 2012; Saldiva et al, 2018) would not have met the criteria for the specific meta-analyses we evaluated (that is, did not examine Whole Pregnancy, trimester 1, trimester 2, trimester 3 exposure windows / or were time-series studies examining short-term effects), or were not particularly informative as the exposure window not clearly reported or defined. As they are not comparable with the specific exposure windows considered in meta-analyses above, we do not comment on them further here.

Brief commentary on selected key studies of interest (Appendix 12 Table A 51) which are likely to expand the evidence base (but not a commentary on all):

Seventeen of these 26 studies specifically examined exposures in Whole Pregnancy, trimester 1, trimester 2, or trimester 3 exposure windows (Chen et al, 2018; Mendola et al, 2016; Symanski et al, 2016; Trasande et al, 2013; Smith et al, 2020; Ritz et al, 2007; Lee et al, 2013; Huang et al, 2018; Siddika et al, 2019; Liang et al, 2019b; Sun et al, 2019; Mendola et al, 2019; Qian et al, 2016; Han et al, 2018; Liu Y et al, 2019; Gray et al, 2014; Capobussi et al, 2016). We consider these as they have potential to add to the evidence from meta-analyses above, in relation to these specific exposure windows, for example in any future meta-analysis. Geographically they were spread as follows: USA (n=8), Europe (n=3), Asia (n=5), Australia (n=1).

Of 8 studies in the USA (Mendola et al, 2016; Symanski et al, 2016; Trasande et al, 2013; Ritz et al, 2007; Lee et al, 2013; Huang et al, 2018; Mendola et al, 2019; Gray et al, 2014), we highlight the following:

- A much larger study (n= 1,448,600) for California by Huang et al (2018) is highlighted, as it expanded beyond the well-studied LA and Orange Counties of the Californian studies (Wilhelm et al, 2005; Wu et al, 2011) included in the meta-analyses above and far exceeded these in sample size, thus offering greater statistical power. It did not advance on these methodologically as it still used environmental monitoring approaches for exposure assessment, did not adjust for smoking, and did not perform multi-pollutant models. It found no association

between Whole Pregnancy exposure to O_3 and PTB (OR (95% CI) per 0.02 ppm increase in O_3: 0.9945 (0.9803, 1.0089). It did not analyse trimester-specific exposure windows.

- Mendola et al (2016) a large study, which expanded geographical coverage of the US, used modelled exposures, and adjusted for smoking. Mendola et al (2016) examined associations of PTB with several air pollutants in over 220,000 deliveries (about 11% were PTB) in a large US obstetric cohort (Consortium on Safe Labor), that collected detailed medical records from 19 hospitals/15 hospital referral regions across the US (in the states of California (LA), Utah, Texas, Illinois, Ohio, Indiana, Florida, Washington DC, Maryland, Delaware, NY State and Maine). The main stated aim of the analysis was to assess whether associations were different in mothers with and without diagnosed asthma. Exposures were estimated using a modified CMAQ model. Researchers did not have access to maternal address, so exposure assignment was based on the average concentrations in the 15 non-overlapping delivery hospital referral regions of hospital of delivery (area range: 415–312,644 km^2), weighted for population density. Authors presented 14 analyses stratified by asthma (one each for each gestational week – weeks 23 to 36) for each of CO, O_3, NO_x, $PM_{2.5}$, PM_{10}, and SO_2 and tested interactions. A correction for potential false discovery was made. Logistic regression estimates were adjusted for maternal age, race, pre-pregnancy BMI, smoking and alcohol use, study site, parity, insurance status, marital status, and comorbidities. The study found some increased risks of preterm (both total preterm <37 weeks, and early PTB <34 weeks) associated with some O_3 exposure windows amongst those without asthma (for example for total PTB, OR 1.08 95% CI 1.03, 1.13 per IQR increase in whole pregnancy O_3 for mothers without asthma vs OR 1.02 95% CI 0.91, 1.15 for mothers with asthma). This was the reverse of the pattern observed for other pollutants for which risk estimates, if elevated, tended to be higher for mothers with asthma vs those without asthma. After allowing for false discovery there was no evidence that associations between O_3 and PTB were different in women with and without asthma at any particular gestational week or chronic exposure window.

- Gray et al (2014) for a large (n=457,642) retrospective cohort study using routine birth records in North Carolina (thus expanding geographical coverage of US), which used of modelled exposures (Bayesian space-time downscaling fusion model at census tract level), and adjusted for smoking. It found no significant association between Whole Pregnancy exposure to O_3 and PTB (OR (95% CI) per IQR increase in O_3: 1.02 (0.99, 1.04). It did not analyse trimester-specific exposure windows.

The one study from Australia (Chen et al, 2018) is highlighted as larger in size (n= 173,800) than those included in the meta-analysis. Unlike the previous Brisbane study (75), it adjusted for smoking. It also used survival analysis, which reduces risk of bias when analysing the relationship between Whole Pregnancy and trimester 3 exposures in relation to PTB (see ‘Specific considerations for studies looking at Preterm birth (PTB) and Stillbirth’ in section 2.1.4), and examined multi-pollutant models. It found an elevated risk of PTB associated with Whole Pregnancy

exposure to O_3 (HR (95% CIs) of PTB per IQR (1.55 ppb) increase in O_3 1.13 (1.10, 1.16)). In multi-pollutant models hazard ratios for O_3 were attenuated after adjustment for other pollutants ($PM_{2.5}$, SO_2, NO_2) individually and jointly but, in all cases, there remained a statistically significant elevated risk of PTB associated with whole pregnancy O_3 exposure. Trimester-specific exposures were analysed, but numeric results were not presented, the authors simply commented 'for all air pollutants, highest HRs were present at trimester 3 and the lowest at trimester 1.'

Of 3 studies in Europe (Smith et al, 2020; Siddika et al, 2019; Capobussi et al, 2016), we highlight that by Smith et al (2020) which examined London, UK, due to its large sample size, its use of modelled exposures, and multi-pollutant models. Smith et al (2020 used information from large number of routine birth records (n=581,774) to examine the association between preterm and a range of particulate matter and gaseous air pollutant exposures in London, UK. The study used highly spatially resolved modelled exposures (monthly estimates from dispersion model at 20m x 20m grid point resolution), assigned to maternal address. They assessed the association of PTB with NO_2 exposure in trimester 1 and trimester 2 only. Logistic regression analyses were adjusted for a wide range of potential confounders, including area-level adjustment for tobacco sales (as a proxy for smoking), and road traffic noise. Multi-pollutant models examining NO_2 adjusted for a range of other air pollutants. It found associations between PTB and O_3 exposure (trimester 1: OR (95% CI) 1.15 (1.11, 1.18) per IQR 17.6 µg/m^3 increase O_3; trimester 2: OR (95% CI) 1.15 (1.11, 1.18) per IQR 19.1 µg/m^3 increase O_3). In multi-pollutant models, trimester 1 and 2 O_3 exposures remained associated with elevated risk of PTB when adjusted for other pollutants (NO_2, NO_x, $PM_{2.5}$ traffic-exhaust, $PM_{2.5}$ traffic-non-exhaust, $PM_{2.5}$, PM_{10}, and road traffic noise in turn), except PM_{10} for trimester 2.

All 5 studies from China (Liang et al, 2019b; Sun et al, 2019; Qian et al, 2016; Han et al, 2018; Liu Y et al, 2019) based exposure assessment on environmental monitoring. We highlight the following 2 studies:

- A study by Liang et al (2019b), due to its large sample size (n=628,439), its use of survival analysis (which reduces risk of bias when analysing PTB), exclusion of caesarean deliveries (which will exclude some if not all medically indicated early deliveries) and analysis of multi-pollutant models. Liang et al (2019b) examined associations of PTB with several air pollutants in a large retrospective birth cohort using birth registration data, across 7 cities in the Pearl River Delta region of China (Guangzhou, Shenzhen, Zhuhai, Dongguan, Foshan, Jiangmen, and Zhongshan). The main stated aim of the analysis was to assess whether associations were different in mothers who were local residents versus migrants, and all analyses were stratified by local resident/migrant status; however, no interactions between migrant status and pollutant exposure were tested. It found elevated risk of PTB associated with Whole Pregnancy and trimester 3 exposure for both local residents and migrants (HRs (95% CIs) of PTB associated with a 10 ppb increase in O_3: Local residents: Whole Pregnancy: 1.18 (1.12, 1.23); trimester 1: 1.01 (0.97, 1.05); trimester 2: 0.92 (0.89, 0.96); trimester 3: 1.08 (1.04, 1.13). Migrants: Whole Pregnancy: 1.32 (1.27, 1.39); trimester 1: 1.01 (0.98, 1.05); trimester 2: 1.03 (0.99, 1.06); trimester 3: 1.08 (1.04, 1.12)). Multi-pollutant models were analysed for Whole Pregnancy exposures – statistically

significant elevated risks between PTB and Whole Pregnancy O_3 remained after adjusting for $PM_{2.5}$, NO_2 and SO_2, in turn.

- We also highlight Qian et al (2016), as it was the next largest study (n=95,911) from China, it adjusted for smoking, and analysed multi-pollutant models. Qian et al (2016) examined associations of preterm birth with several air pollutants in a retrospective birth cohort using routine electronic hospital and birth records, in Wuhan, China. It found increased risk of PTB with each 10 μg/m^3 increase in O_3 in whole pregnancy and trimester 3 (Whole Pregnancy OR = 1.05; 95% CI: 1.02, 1.07); trimester 3: 1.04 (1.02, 1.06)). Trimester 1 and 2 analysis results were not reported for O_3; the authors reported results for Whole Pregnancy and 'the most susceptible exposure windows' only for each pollutant. Multi-pollutant models were analysed for whole pregnancy exposures – associations between PTB and Whole Pregnancy O_3 were very similar and remained statistically significant after adjusting for $PM_{2.5}$, PM_{10}, SO_2, NO_2 and CO, in turn.

The results from most of these highlighted studies demonstrate some statistically significant increased risks of PTB associated with long-term exposure to O_3 during pregnancy, (Chen et al, 2018; Mendola et al, 2016; Smith et al, 2020; Liang et al, 2019b; Qian et al, 2016), but some find no association (Huang et al, 2018; Gray et al, 2014).

Conclusion regarding evidence for O_3 and PTB from key individual studies not included in the core meta-analyses

There were a large number of studies not captured in or published since the core meta-analyses, and the results from key studies highlighted were inconsistent (section 2.1.4).

OVERALL CONCLUSION for O_3 and PTB from the epidemiological evidence

Based on limited evidence from the meta-analyses available, we conclude there is weak evidence (section 2.1.3) for an association between PTB and O_3 exposure. Examination of key highlighted studies points to the need for an up-to-date meta-analysis to synthesise all the available evidence.

Commentary on results from studies examining Term Low Birth Weight (TLBW)

$PM_{2.5}$ and TLBW

Commentary on results from meta-analyses

The core meta-analyses that examined $PM_{2.5}$ exposure with TLBW were Stieb et al (2012) and Zhu et al (2015) for Whole Pregnancy. Li et al (2020) remained for trimester-specific exposure windows. A summary of results is presented in Table 11, with more detailed information provided in Appendix 11, Table A 40.

Table 11. Summary of the results from the core meta-analyses examining term low birth weight and $PM_{2.5}$ exposure

Meta-analyses (Increment)	OR (95% CI)			
	WP	T1	T2	T3
Li et al 2020 (Per 10 µg/m³)	Excluded	1.03 (0.97, 1.09)*	1.03 (0.98, 1.08)*	1.05(1.01, 1.09)*
Zhu et al 2015 (Per 10 µg/m³)	1.05 (1.02–1.07)	Not performed	Not performed	Not performed
Stieb et al 2012 (Per 10 µg/m³)	1.05 (0.99, 1.12)	Not performed	Not performed	Not performed
Individual studies in common between meta-analyses?	Stieb & Zhu = 2	n/a	n/a	n/a

*For Li et al (2020) these values represent RR

Li et al 2020

Li et al (2020) reported summary estimates for TLBW per 10 µg/m^3 $PM_{2.5}$ by exposure period as follows: trimester 1 RR 1.03 (95% CI 0.97, 1.09), trimester 2 RR 1.03 (95% CI 0.98, 1.08), and trimester 3 RR 1.05 (95% CI 1.01, 1.09)

Trimester 1 results were based on estimates from 19 studies, and trimester 2 and trimester 3 results were based on estimates from 20 studies from a search up to June 2020 (no start date specified). Considerable heterogeneity was observed for all 3 trimesters (I^2 95.1%, 91.5% and 92.0% for trimesters 1, 2 and 3 respectively, $p<0.001$ in all cases). The forest plots showed that heterogeneity was in both direction and size for trimesters 1 and 2. For trimester 3 there were a small number of studies showing an OR below 1 but most of the heterogeneity was in size rather than direction.

There were 9 out of 19 studies that controlled for smoking for trimester 1, 10/20 for trimester 2 and 10/20 for trimester 3. The estimates from these studies adjusting for smoking accounted for 35%, 37% and 37% respectively of the weight in the random-effects meta-analysis. There was no stratification by control for smoking but, examining the forest plots by eye, there were no obvious differences in estimate size between those that controlled for smoking and those that did not. There was no evidence of publication bias. Term low birthweight was a selection criterion for the meta-analysis, so the studies should all be for term low birthweight rather than low birthweight.

This study was regarded as providing **no evidence for an association for $PM_{2.5}$ exposure in trimesters 1 and 2** (estimates do not reach conventional statistical significance, considerable heterogeneity) and **weak evidence for a positive association for $PM_{2.5}$ exposure in trimester 3** (reasonable number of studies and confidence intervals above 1, but a fairly small odds ratio with considerable heterogeneity, and including several studies that did not adjust for smoking).

Zhu et al 2015

Zhu et al (2015) reported an OR 1.05 (95% CI 1.02, 1.07) per 10 μg/m^3 between Whole Pregnancy $PM_{2.5}$ exposure and TLBW.

Zhu et al (2015) identified 6 studies from a search conducted up to March 2014. There was no indication of any important degree of heterogeneity (I^2 39.7%; p=0.141). No evidence of publication bias was found.

Five of the 6 studies included adjustment for smoking, Morello-Frosch et al (2010) did not. The size of the odds ratio for this study was not noticeably different compared with the studies that did adjust for smoking, but it should be noted that this study had a weight of 89.17% (fixed effects) in the analysis.

Study selection was not limited to TLBW, but only Gray et al (2014) (weight 3.26%, fixed effect analysis) did not use that definition. Gray et al (2014) did adjust for gestational age, and excluded multiple births and congenital anomalies, 2 general factors that could lead to LBW before term, thus reducing potential for confounding by gestational age.

This study was regarded as providing **weak evidence from a limited evidence base for a positive association with $PM_{2.5}$ exposure across the Whole Pregnancy** (confidence intervals above 1, little heterogeneity but a small odds ratio, high weighting for a study that did not adjust for smoking and a small number of studies).

Stieb et al 2012

Stieb et al (2012) reported an OR 1.05 (95% CI 0.99, 1.12) per 10 μg/m^3 $PM_{2.5}$ between Whole Pregnancy $PM_{2.5}$ exposure and TLBW.

There was considerable heterogeneity (I^2 85.5%; p<0.001) with the odds ratios in the individual studies being both greater than 1 (sometimes substantially) and just below 1 (that is, the heterogeneity was both in size and direction of the association). The $PM_{2.5}$ association was not tested for publication bias.

Five of the 6 studies included adjustment for smoking; again Morello-Frosch et al (2010) did not (this study had a weight of 31.37% in the meta-analysis and was not obviously different to other studies).

Study selection was not restricted to TLBW, but all studies except Bell et al (2007) (weight 17.29%, random effects analysis) used a TLBW definition (Appendix 5). Nonetheless Bell et al (2007) did adjust for gestational age, thus reducing potential for confounding by gestational age. The authors considered the number of studies to be insufficient for an investigation of whether this contributed to the heterogeneity of the association with low birthweight.

On balance, given that there was considerable heterogeneity, and a meta-analysis summary estimate which did not reach conventional levels of significance, we

conclude that Stieb et al (2012) provides **no evidence to support an association of TLBW with $PM_{2.5}$ exposure across the Whole Pregnancy but this is based on a limited amount of evidence.**

Synthesis

With regards to Whole Pregnancy, it is possible to conclude that there may be a small positive association between $PM_{2.5}$ and TLBW. The results of the 2 meta-analyses are consistent in magnitude of summary estimates (if not in achieving statistical significance) despite only minor overlap in the studies included. Thus, the evidence considered can be provisionally represented by an OR=1.05 per 10 $\mu g/m^3$ considering only effect size. The key question though is the strength of the evidence for this. The confidence intervals around this informal summary OR are likely to be narrower than those in Stieb et al (2012) (and therefore possibly reaching statistical significance) but possibly wider than those in Zhu et al (2015) (because while more studies would be added if all studies across the 2 meta-analyses were included, heterogeneity may also increase). However, the above conclusions do not take account of the following:

a. **Overlap between studies** (for example, if Parker et al (2008) (a null study of a sub sample of births across the whole USA) was chosen over Morello-Frosch et al (2010) (a large study in California, with a significant positive association), this could reduce the size of the association. This was regarded as a tolerable overlap but might not be in a sensitivity analysis which eliminates all overlaps).

b. **Excluding studies that did not adjust for smoking** Morello-Frosch et al (2010) did not adjust for smoking and was an influential study in both Zhu et al (2015) and Stieb et al (2012). It might be hypothesised that by not adjusting for smoking, this may give an overestimate of the effect, but the estimate from Morello-Frosch et al (2010) was not markedly different from other studies. A stricter approach might exclude it.

c. **Whether studies examined TLBW or LBW.** It was not always clear whether the studies included in Zhu et al (2015) and Stieb et al (2012) examined TLBW or LBW, although further investigation indicated that the majority examined TLBW. Including LBW studies might lead to an overestimate of the effect, but, if still included after considering overlaps, they would still have relatively little weight in the overall meta-analysis.

Li et al (2020) found a positive association between trimester 3 $PM_{2.5}$ exposure and TLBW with a larger number of studies than the whole pregnancy meta-analyses discussed above, providing some support for the whole pregnancy results. The positive associations between trimester 1 or 2 $PM_{2.5}$ exposure and TLBW did not attain statistical significance but the confidence intervals overlap with the estimate for trimester 3. Therefore, no clear conclusion as to which trimester is most important can be reached.

Conclusion regarding evidence for $PM_{2.5}$ and TLBW from the core meta-analyses

There is weak evidence (section 2.1.3) for a small positive association between Whole Pregnancy exposure to $PM_{2.5}$ and term low birthweight. This conclusion is based on limited evidence. It should be noted that, for Whole Pregnancy, this is based on meta-analyses that each only considered part of the available evidence published up to 2014. There is a need for a more complete meta-analysis of pre-2014 studies supplemented with a check for more recent studies. Li et al, 2020 (2020) provided a more recent meta-analysis for trimester specific exposures that also included more studies. The analyses did not support an association for trimester 1 and 2 and provided weak evidence for trimester 3. Methodological improvements could be made in a new meta-analysis for example, stratification by control for smoking.

Commentary on evidence from key individual studies regarding $PM_{2.5}$ and Term Low Birth Weight (TLBW) (not included in the core meta-analyses)

Following the flowchart process as outlined in Figure 6, Endpoint D was reached (Appendix 10).

Brief description of available studies:

18 studies examining the association between $PM_{2.5}$ exposure during pregnancy and TLBW were captured by our literature search that were not included in any of the core meta-analyses above for any of the exposure periods for this pollutant and outcome (Pedersen et al, 2013; Coker et al, 2016; Laurent et al, 2013; Twum et al, 2016; Laurent et al, 2016a; Poirier et al, 2015; Yuan et al, 2020; Wilhelm et al, 2012; Brown et al, 2015; Ha et al, 2017; Nieuwenhuijsen et al, 2019; Enders et al, 2019; Kumar, 2012; Ng et al, 2017; Laurent et al, 2016b; Ghosh et al, 2012; Coker et al, 2015; Laurent et al, 2014). 18 studies were for the whole pregnancy period: 14 each for trimesters 1 and 3 and 13 for trimester 2. For whole pregnancy, 18 studies represents a considerable increase over the previous selected meta-analyses (6 studies each for Zhu et al (2015) and Stieb et al (2012). For trimester specific exposures, the previously selected meta-analysis was more recent (2020) but 13-14 studies identified in our updated literature search were not all the same as in Li et al (2020). 9 of these studies in our literature search (for any exposure period) were published in 2016-2020, with the remaining 9 studies published from 2012 to 2015. Two studies were based on prospective cohorts (with overlap between them); many of the remaining studies were retrospective cohort studies. The geographical spread was dominated by studies in the US: Europe (n=2); Asia (n=1); North America (n=15, one in Canada, the rest US). Many of the US studies were from similar regions to those studied previously, for example California, but with some new areas of the US and some new US wide studies. There is likely to be considerable population overlap between these studies and between these studies and earlier ones.

As stated in section 2.1.4, we consider here only studies that do not appear in any of the core meta-analyses for this pollutant-outcome pair; there may, in addition, be individual studies that would add to the evidence for a specific exposure window which are not considered here. This is particularly important for this pollutant-

outcome pair because the Whole Pregnancy result in the meta-analysis by Li et al (2020) was excluded, while the trimester specific analyses were included. Thus, for example, Smith et al (2017), is not included in our list of new studies because it had already been included in the Li et al (2020) meta-analysis for trimester-specific exposures, although the whole pregnancy result from Smith et al (2017) was not. In addition to Smith et al (2017), there are another 18 studies (with some population overlaps) of whole pregnancy exposure to $PM_{2.5}$ and TLBW that do not appear in the list because the studies were previously considered in Li et al (2020) for trimester specific exposures.

Brief commentary on selected key studies of interest (Appendix 12 Table A 52) which are likely to expand the evidence base (but not a commentary on all):

Whole pregnancy: The previous meta-analyses included 3 European studies of $PM_{2.5}$ and TLBW but a pooled analysis across European birth cohorts for the ESCAPE analysis was subsequently published (in 2013) (Pedersen et al, 2013). While a more recent study was published in 2019 by Nieuwenhuijsen et al (Nieuwenhuijsen et al, 2019), this included some cohorts for which $PM_{2.5}$ exposure had not been directly validated and was a less traditional analysis with a smaller population size, being designed to look at biomarkers. Therefore, the results from Pedersen et al (2013) are highlighted here. This study was a pooled analysis of European birth cohorts (ESCAPE study) (including n=70,551 term births, of which 1,257 were TLBW (1.8%)). The authors examined the association of TLBW with annual mean $PM_{2.5}$ by collating information from several prospective birth cohorts that originated in Sweden, Norway, Netherlands, Greece, Lithuania, Hungary, Spain, Italy, and Denmark. A common protocol for LUR was used, exposure was based on maternal address and accounted for moving during pregnancy in some centres. LUR models were validated with specific monitoring campaigns (although these were only for short periods). Model validation showed that the LUR models explained 60–88% of $PM_{2.5}$ variability depending on the area. $PM_{2.5}$ concentrations were lower in northern areas than in central and southern areas. The range of $PM_{2.5}$ concentrations included the UK range although it extended to higher levels (annual mean $PM_{2.5}$ exposure 16.5 $\mu g/m^3$, 75th percentile 20 $\mu g/m^3$). Whole pregnancy and trimester specific estimates were derived. Logistic regression analyses were adjusted for major confounders including maternal smoking and individual socio-economic status. Multi-pollutant models were examined where data were available. The study found a positive and statistically significant association (OR 1.18 95% CI 1.06, 1.33) per 5 $\mu g/m^3$) with TLBW, which was robust to a variety of sensitivity analyses, including adjustment for NO_2 and $PM_{2.5}$ absorbance. The main analysis pooled the raw data, but a meta-analytical approach gave similar results with no important heterogeneity between cohorts. The odds ratio remained positive and statistically significant for exposures less than 25, 20 and 15 $\mu g/m^3$.

We highlight studies from the US that (i) cover new areas of the US (ii) are reasonably large (iii) adjusted for smoking and (iv) used modelling for exposure assessment. Previously the US areas covered were, Connecticut (Bell et al, 2007), California (Morello-Frosch et al, 2010), Atlanta (Darrow et al, 2011), and North Carolina (Gray et al, 2014), plus, 2 wider area studies (a subset of births across the

whole US, Parker et al (Parker et al, 2008) and births across the North East and mid-Atlantic US (Ebisu and Bell, 2012).

Our literature search identified studies in further specific US areas: New York State/New York (Brown et al, 2015); Chicago (Kumar, 2012); Georgia (Twum et al, 2016); plus 10 more studies in California and another cross-US study (Ha et al, 2017). Of these we highlight the cross-US study, and the large study in New York.

Despite the large number of studies in California, none (i) controlled for smoking and (ii) used modelling for exposure assessment. Therefore, no studies from California are highlighted in Appendix 12 Table A 52.

The cross-US study by Ha (2017) controlled for smoking and used modelled exposure data. This study investigated the potential association of extreme ambient temperature, 5 criteria air pollutants, and 6 particulate matter constituents with TLBW, in a large US obstetric cohort (Consortium on Safe Labor), that collected detailed medical records from 19 hospitals/15 hospital referral regions across the US (in the states of California (LA), Utah, Texas, Illinois, Ohio, Indiana, Florida, Washington DC, Maryland, Delaware, NY State and Maine). Exposures were estimated using a modified CMAQ model. Researchers did not have access to maternal address, so exposure assignment was based on the average concentrations in the 15 non-overlapping delivery hospital referral regions of hospital of delivery (area range: 415–312,644 km^2), weighted for population density. Over 220,000 births (about 2% of which were TLBW) remained in the regression analyses which was able to adjust for a wide range of confounders including maternal smoking in pregnancy and ethnicity but had few adjustments for socio-economic status (it used insurance status). The study states Poisson regression was used but gives little detail. A positive (adverse) association was found but it lacked statistical significance.

Brown et al (2015) was a large study (n=480,430) in New York State[8] which expanded geographical coverage of the US beyond that of the selected core meta-analyses. Exposures were modelled (USEPA Bayesian hierarchical space time modelling) on a 12x12km grid scale. Exposure was assigned according to the 12km^2 grid cell in which the maternal address was located. A wide range of potential confounders were adjusted for, including maternal smoking during pregnancy. The results for whole pregnancy were null.

There is also a study in Canada, with particularly low concentrations (Poirier et al, 2015), which we highlight. It used birth records to examine associations of several pollutants with a range of birth outcomes including TLBW (226 cases and 13,483 controls) in Halifax, Nova Scotia (Canada). Whole pregnancy exposure was determined from a LUR model and based on the post-code of mother's address. The LUR model was able to predict 51% of the variation in $PM_{2.5}$. It has low concentrations of pollutants, and likely limited exposure contrast. Logistic regression

[8] This study may partially overlap with Ha S et al. Ambient temperature and air quality in relation to small for gestational age and term low birthweight. Environmental research. 2017;155:394-400, 179.(which includes Brooklyn)

models were adjusted for mother's age, parity, smoking, and neighbourhood income. The result was null for $PM_{2.5}$ and TLBW.

It is important to note that there are some large studies with control for smoking and use of modelled concentrations that are neither discussed here nor included in the core meta-analyses for whole pregnancy exposures. This is because of the approach discussed in section 2.1.4 of only including studies that were new for all exposure periods. Because of the pattern and dates of included and excluded meta-analyses the number of studies involved was more extensive for $PM_{2.5}$ and TLBW.

Trimester-specific estimates: The new trimester-specific studies also examined whole pregnancy and the reasoning for highlighting particular studies are the same.

Pedersen et al (2013) found positive and statistically significant associations for all 3 trimesters, in line with their findings for whole pregnancy. The trimester-specific results for Brown et al (2015) were also in line with the whole pregnancy results - that is, null. The results for Ha (2017) were also null, although the RR for trimester 1 was raised and almost statistically significant.

Conclusion regarding evidence for $PM_{2.5}$ and TLBW from key individual studies not included in the core meta-analyses

While the results of the highlighted studies are inconsistent (section 2.1.4), it is difficult to tell how a summary of the whole literature would balance out. A strong quality study such as Pedersen et al (2013) that had a positive and statistically significant result could be outweighed by a multiplicity of other studies.

OVERALL CONCLUSION for $PM_{2.5}$ and TLBW from the epidemiological evidence

Based on limited evidence from the meta-analyses available, we conclude there is weak evidence (section 2.1.3) for an association between TLBW and $PM_{2.5}$ exposure. Examination of key highlighted studies points to the need for an up-to-date meta-analysis to synthesise all the available evidence.

PM_{10} and TLBW

Commentary on results from meta-analyses

After exclusions, the core meta-analyses that examined associations of PM10 exposure with TLBW were Li et al (2020) for each trimester, Sapkota et al (2012) for Whole Pregnancy and trimester 3 and Stieb et al (2012) for Whole Pregnancy and trimesters 1 and 2. A summary of results are provided in Table 12, with more detailed information provided in Appendix 11, Table A 41.

Table 12. Summary of the results from the core meta-analyses examining term low birth weight and PM_{10} exposure

Meta-analyses (Increment)	OR (95% CI)			
	WP	T1	T2	T3
Li et al 2020 (Per 10 µg/m³)	Excluded	1.02 (1.00, 1.05)*	1.01 (1.01, 1.02)*	1.00 (1.00, 1.01)*
Sapkota et al 2012 (Per 10 µg/m³)	1.02 (0.99, 1.05)	Excluded	Not performed	1.00 (0.99, 1.01)
Stieb et al 2012 (Per 20 µg/m³)	1.10 (1.05, 1.15)	1.03 (0.95, 1.11)	1.02 (0.96, 1.09)	Excluded
Individual studies in common between meta-analyses?	Sapkota&Stieb=4	n/a Li & Stieb=5	n/a Li & Stieb=5	Sapkota&Stieb=2 Sapkota & Li =2 Li & Stieb=5

*For Li et al (2020) these values represent RR

Li et al 2020

Li et al (2020) reported summary estimates for TLBW per 10 µg/m^3 PM_{10} by exposure period as follows: trimester 1 RR 1.02 (95% 1.00, 1.05), trimester 2 RR 1.01 (95% 1.01, 1.02), and trimester 3 RR 1.00 (95% 1.00, 1.01).

These trimester-specific results are based on estimates from the same 13 studies from a search up to June 2020 (no start date specified). Considerable heterogeneity was observed for trimester 1 (I^2 71.5%, p<0.001) but there was less variation for the other 2 trimesters (I^2 <30% p>0.1). Effect estimates for trimester 1, trimester 2 and trimester 3 were small and presented to 3 significant figures (NB these have been rounded to 2 significant figures in Table 12), with the estimate for trimester 2 just reaching conventional level of statistical significance.

Only 4 out of 13 studies controlled for smoking, with most of the weight of the meta-analyses (~90% for trimester 1, ~99% for trimester 2, ~98% for trimester 3) coming from studies with no adjustment for smoking and conducted in China and Korea (where smoking rates among pregnant women are rarely recorded but reported to be low by authors). There was no evidence of publication bias. TLBW was a selection criterion for the meta-analysis, so the studies should all be for TLBW rather than LBW.

This study was regarded as providing **weak evidence for an association for PM_{10} exposure in trimester 2** (estimates just reached conventional statistical significance, there was limited heterogeneity but several studies that do not adjust for smoking were included), but **no evidence for an association for PM_{10} exposure in trimester 1 and 3** (neither estimate reached conventional levels of significance and several of the studies included did not adjust for smoking).

Sapkota et al 2012

Sapkota et al (2012) reported summary estimates for TLBW per 10 μg/m^3 PM_{10} by exposure period as follows: Whole Pregnancy OR 1.02 (95% CI 0.99, 1.05) and trimester 3 OR 1.00 (95% CI 0.99, 1.01) (Table 12)

The meta-analyses for Whole Pregnancy and trimester 3 included 11 and 7 individual studies respectively. The Whole Pregnancy estimates showed moderate heterogeneity (I^2 54.5%; p=0.015). The studies included in the trimester 3 estimates can be inferred from Table 1 in Sapkota et al (2012), but the estimates extracted from each study used to generate the summary estimate were not provided. The level of heterogeneity was not reported nor was the weight contribution from each study. In the absence of detailed information on the trimester 3 study estimates and a non-significant summary estimate (p<0.05) the trimester 3 summary estimates are not discussed further. There was no significant publication bias according to both the Begg's and Egger's test.

We have redrawn the information for the Whole Pregnancy stratifying by smoking adjustment (Figure 9). This shows that just over 50% of the weight of the Whole Pregnancy estimate came from studies that did not adjust for smoking, but that there was little difference in the estimates arising from studies with and without adjustment for smoking (OR 1.03 95% CI 0.98, 1.09) vs OR 1.02 95% CI 0.98, 1.06)).

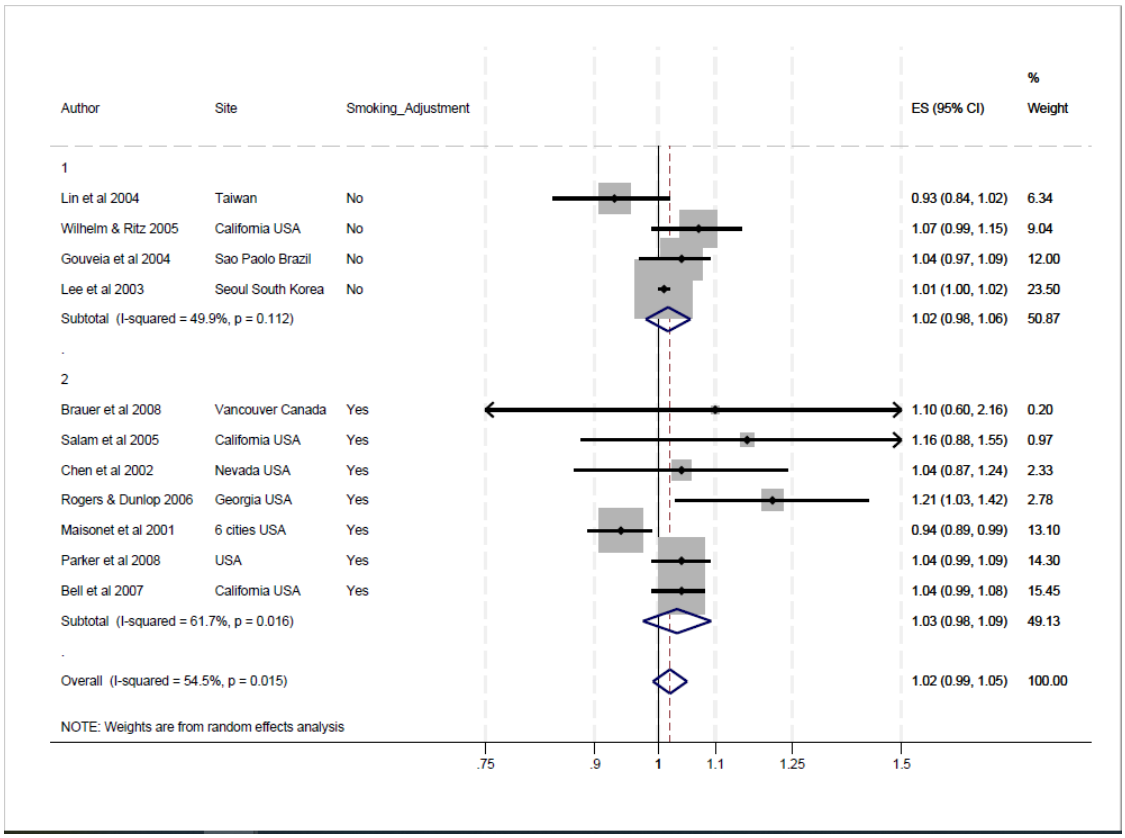


Figure 9. Redrawn Sapkota et al (2012) estimates for PM_{10} and TLBW for Whole Pregnancy, stratified by adjustment for smoking. This is a forest plot of the meta-analyses used to derive separate summary effect estimates for studies with and without adjustment for confounding, as described in the text

Of note, 2 of the 11 studies Bell et al (2007) and Rogers and Dunlop (2006) reported associations with LBW not at term. These 2 studies reported positive (adverse) associations (Bell et al (2007) OR 1.04 (95% CI 0.99, 1.08) and Rogers and Dunlop (2006) OR 1.21 95% CI 1.03, 1.42), both adjusted for smoking, and, combined, they contribute about 20% of the weight (see Figure 9). Removal of these 2 studies would weaken the non-significant positive (adverse) associations reported.

We conclude Sapkota et al (2012) shows **no evidence to support an association of TLBW with PM_{10} exposure during Whole Pregnancy** and **no evidence for an association of TLBW with PM_{10} exposure during trimester 3.**

Stieb et al 2012

Stieb et al (2012) reported summary estimates for TLBW per 20 μg/m^3 PM_{10} by exposure period as follows: Whole Pregnancy OR 1.10 (95% CI 1.05, 1.15), trimester 1 OR 1.03 (95% CI 0.95, 1.11) and trimester 2 OR 1.02 (95% CI 0.96, 1.09) (Table 12).

Seven studies contributed to the trimester-specific estimates and 8 studies to the Whole Pregnancy estimate. One of the studies (Seo et al; 2010) included in the Whole Pregnancy meta-analysis provided 7 separate estimates based on stratification by geographical region within South Korea. There was substantial heterogeneity among studies included in the Whole Pregnancy estimate (I^2 68.4%; p<0.001). There was moderate heterogeneity among studies included in the trimester 1 estimate (I^2 41.6%; p=0.114), and no important degree of heterogeneity among studies included in the trimester 2 (I^2 22.6%; p=0.256) estimates.

There was asymmetry in the published funnel plot and evidence of publication bias by formal test (Begg's test p-value= 0.04), but this appears to be based on a single test/funnel plot of all estimates for Whole Pregnancy and each of trimesters 1, 2 and 3 combined (Figure 5 in Stieb et al (2012)). It is likely that it is strongly influenced by the 7 Whole Pregnancy estimates coming from the Seo et al (2010) paper which are imprecise and strongly positive.

Elsewhere in their report the authors considered birth weight (rather than the binary outcome TLBW) and reported a reduction in the Whole Pregnancy summary estimate if only studies with adjustment for socio-economic status were considered. They did not comment on whether adjustment for smoking may have influenced the Whole Pregnancy estimate for birth weight (or for TLBW). We have redrawn the information for the Whole Pregnancy TLBW estimate stratifying by smoking adjustment (Figure 10). There was little difference in the estimates coming from studies with and without adjustment for smoking (OR 1.11 (95% CI 1.06, 1.16) vs OR 1.09 (95% CI 1.03, 1.14)).

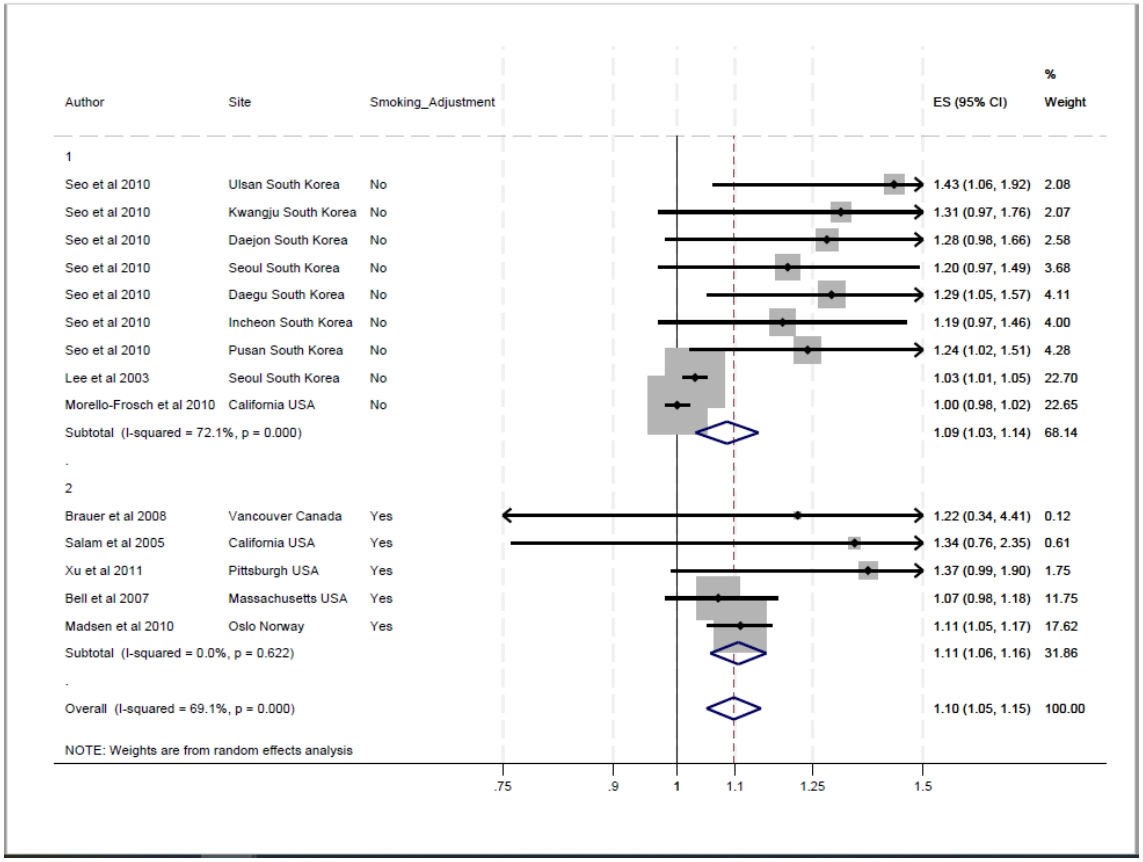


Figure 10. Redrawn Stieb et al (2012) estimates for PM_{10} and TLBW for Whole Pregnancy, stratified by adjustment for smoking. This is a forest plot of the meta-analyses used to derive separate summary effect estimates for studies with and without adjustment for confounding, as described in the text

Of note, 1 of the 8 studies for Whole Pregnancy estimates reported associations with LBW not TLBW (Bell et al (2007), 11.28% of the weight).

The same 7 studies contribute to the trimester 1 and trimester 2 estimates. The results from each study within the trimester 1 and 2-specific meta-analysis are broadly consistent, showing weak non-significant associations that are very similar.

We conclude that Stieb et al (2012) provides **weak evidence to support an association of TLBW with PM_{10} exposure across whole pregnancy** and **no evidence of an association for trimester 1 or 2.**

Synthesis

There was inconsistency in the findings from the meta-analyses. Sapkota et al (2012) provided no evidence at all for an association, whereas Stieb et al (2012) provided weak evidence for an association with whole pregnancy exposure (but not trimesters 1 or 2) and Li et al (2020) provided weak evidence for an association in trimester 2 (but not trimester 1 or 3). Sapkota et al (2012) reviewed publications up to 2009, Stieb et al (2012) reviewed publications up to 2011 and Li et al reviewed

publications up to 2020. Only 4 of the publications in Sapkota et al (2012) were included in Stieb et al (2012). Although Li et al (2020) reviewed all studies up to 2020 it included only 2 of 7 publications in Sapkota et al (2012), and 5 of 7 studies in Stieb et al (2012). Of concern Stieb et al (2012) was considered to have publication bias.

Conclusion regarding evidence for PM_{10} and TLBW from core meta-analyses

The 3 meta-analyses provide weak evidence (section 2.1.3) for an association of TLBW with PM_{10} and no indication as to a critical gestational window.

Commentary on evidence from key individual studies regarding PM_{10} and Term Low Birth Weight (TLBW) (not included in the core meta-analyses)

Following the flowchart process as outlined in Figure 6, Endpoint D was reached (Appendix 10).

Brief description of available studies:

15 studies examining the association between PM_{10} exposure during pregnancy and TLBW were captured by our literature search and were not included in any of the core meta-analyses above (Pedersen et al, 2013; Laurent et al, 2013; Poirier et al, 2015; Kim et al, 2019; Wilhelm et al, 2012; Ha et al, 2017; Nieuwenhuijsen et al, 2019; Enders et al, 2019; Kumar, 2012; Ghosh et al, 2012; Warren et al, 2018; Guo P et al, 2020; Yorifuji et al, 2015; Habermann and Gouveia, 2014; Romão et al, 2013). Of these 15 studies only 1 was published in 2020 and therefore may not have been captured by the search in Li et al (2020) (the most recent meta-analysis). The remainder may have been considered for inclusion by Li et al (2020) – but as Li et al (2020) did not provide a list of excluded studies with reason for exclusion, this is uncertain. Of the 14 studies published before 2020, one (Pedersen et al, 2013) had been included in Li et al. (2020) whole pregnancy analysis but not in trimester specific analyses), one (Laurent et al, 2013) was included in Li et al. (2020) whole pregnancy analysis and one (Kumar, 2012) was the same population as Kumar (2016) (which was included in whole pregnancy and trimester specific meta-analyses in Li et al. (2020)).

Of the 15 studies identified in the literature search, there were 12 retrospective cohort studies using birth records or electronic medical records, 2 studies pooling multiple birth cohorts, and one large survey in which survey data were linked to administrative birth records. Geographically they were spread as follows: USA (n=7) Canada (n=1), Europe (n=2), Asia (n=3) and South America (n=2).

As stated in section 2.1.4 as we consider here only studies which do not appear anywhere in the core meta-analyses, there may, in addition, be individual studies that would add to the evidence for a specific exposure window which are not considered here.

Brief commentary on selected key studies of interest (Appendix 12 Table A 53) that are likely to expand the evidence base (but not a commentary on all):

Of the 15 studies above, we highlight 4 below.

Ha (2017) investigated the potential association of extreme ambient temperature, 5 criteria air pollutants, and 6 particulate matter constituents with TLBW, in a large US obstetric cohort (Consortium on Safe Labor), that collected detailed medical records from 19 hospitals/15 hospital referral regions across the US (in the states of California (LA), Utah, Texas, Illinois, Ohio, Indiana, Florida, Washington DC, Maryland, Delaware, NY State and Maine). Exposures were estimated using a modified CMAQ model. Researchers did not have access to maternal address, so exposure assignment was based on the average concentrations in the 15 non-overlapping delivery hospital referral regions of hospital of delivery (area range: 415–312,644 km^2), weighted for population density. Over 220,000 births (about 2% of which were TLBW) remained in the regression analyses which was able to adjust for a wide range of confounders including maternal smoking in pregnancy and ethnicity but had few adjustments for socio-economic status (used insurance status). The study stated Poisson regression was used but gave little detail. For TLBW, PM_{10} exposure during trimester 2 (RR: 1.06, 95% CI: 1.00, 1.13) and Whole Pregnancy (RR 1.08 95% CI 1.01, 1.16) were positively and significantly associated. However, after adjusting for a false discovery rate using standard methods (multiple exposures had been examined) no statistically significant association remained.

This study provides weak evidence for a positive association of TLBW exposure to PM_{10} in whole pregnancy and in trimester 2. Associations were not present having accounted for false discovery – this adjustment has been rarely applied in studies considered in this report but highlights the risk of false discovery and selective reporting when multiple pollutants-outcomes-exposure periods are tested.

Poirier et al (2015) used birth records to examine associations of several pollutants with a range of birth outcomes including TLBW (226 cases and 13,483 controls) in Halifax, Nova Scotia (Canada). Whole pregnancy exposure was determined from a LUR model and based on the post-code of mother's address. The LUR model was able to predict 42% of the variation in PM_{10}. It has low concentrations of pollutants, and likely limited exposure contrast. Logistic regression models were adjusted for mother's age, parity, smoking, and neighbourhood income. There was a non-significant increased risk per IQR of PM_{10} (N.B. IQR could not be determined from the presented information) of 1.04 (95% CI 0.90, 1.19).

This study provides **no evidence of an association of whole pregnancy PM_{10} exposure with TLBW.**

Pedersen et al (2013) conducted a pooled analysis of European birth cohorts (ESCAPE study) (including n=70,551 term births, of which 1257 were TLBW (1.8%)). The authors examined the association of TLBW with annual mean PM_{10} by collating information from several prospective birth cohorts that originated in Sweden, Norway, Netherlands, Greece, Lithuania, Hungary, Spain, Italy, and Denmark. A common protocol for LUR was used, exposure was based on maternal address and accounted for moving during pregnancy in some centres. LUR models were

validated with specific monitoring campaigns (although these were only for short periods). Model validation showed that the LUR models explained 50–75% of PM_{10} variability depending on the area. $PM_{2.5}$ and PM_{10} concentrations were lower in northern areas than in central and southern areas. Whole Pregnancy and trimester specific estimates were derived. Logistic regression analyses were adjusted for major confounders including maternal smoking and individual socio-economic status. Multi-pollutant models were examined where data were available. There was an association of TLBW with exposure to PM_{10} in Whole Pregnancy 1.16 (95% CI 1.00, 1.35 n= 50,151) which was little altered by adjustment for NO_2 1.19 (95% CI 1.00, 1.42; n=49,285). There was no association for Whole Pregnancy exposure to $PM_{2.5-10}$ - 1.01 (95% CI 0.88, 1.15; n=48,995). The association for trimester-specific exposures with PM_{10} were for trimester 1: 1.13 95% CI 1.00, 1.28, trimester 2: 1.12 95% CI 0.97, 1.28, trimester 3: 1.10 95% CI 0.96, 1.27).

This study provides **some evidence of an association of TLBW with Whole Pregnancy exposure to PM_{10} but does not provide any evidence for trimester specific associations.** There was no evidence of an association of TLBW with $PM_{10\text{-}2.5}$.

Enders et al (2019) explored whether $PM_{10–2.5}$ exposure during pregnancy was associated with increased risk of TLBW in California using data from 2002 to 2013. Mothers were assigned $PM_{10–2.5}$ values from the closest monitor within 20 km of the population-weighted centroid of her residential zip code tabulation area – therefore the sample was biased towards a more urban population. In logistic regression analyses extensive adjustment for a range of appropriate confounders including maternal smoking during pregnancy, individual markers of maternal and paternal educational level and area levels of poverty was included. It is likely that the sample was biased towards married and co-habiting mothers as births were excluded if paternal information was not reported. Over 2.7 million births were included in this study of which about 55000 (~2%) were TLBW. Exposure to $PM_{2.5\text{-}10}$ was considered in quartiles and adjusted for $PM_{2.5}$ levels. Statistically significant associations were observed for fourth quartile exposure to $PM_{10\text{-}2.5}$ for Whole Pregnancy (Quartile 2 OR 1.00 95% CI 0.98-1.03; Quartile 3 OR 1.03 95% CI 1.00-1.06; Quartile 4 OR 1.04 95% CI 1.01-1.07) and third trimester (Quartile 4 OR 1.04 95% CI 1.01-1.07), and for third quartile (but not fourth quartile) exposure for first trimester (Quartile 3 OR 1.04 95% CI 1.01-1.06).

Statistically significant associations were also seen for $PM_{10\text{-}2.5}$ adjusted associations of $PM_{2.5}$ with TLBW for third quartile (first trimester) and fourth quartile (Whole Pregnancy and first and second trimester). Of note, stratified analyses by paternal education showed that the ORs for the associations between $PM_{10–2.5}$ exposure and TLBW were elevated for fathers with less than high school education, high school diploma only, and some college education. On the other hand, fathers with a college degree or beyond showed no association. Stratifying by mother's education elicited similar results (results not shown). While the authors suggest this may identify sub-groups at risk it may also reflect residual confounding by another factor related to educational level or SES (for instance: alcohol consumption).

This study provides **weak evidence for a positive association of TLBW exposure to the coarse fraction of PM in Whole Pregnancy and in each trimester specific period.** This is in contrast to the findings of Pedersen et al (61) (reported above).

Conclusion regarding evidence for PM_{10} and TLBW from key individual studies not included in the core meta-analyses

There were several large studies not captured in/published since the core meta-analyses, which have inconsistent findings (section 2.1.4).

OVERALL CONCLUSION for PM_{10} and TLBW from the epidemiological evidence

Based on the meta-analyses available, we conclude there is weak evidence (section 2.1.3) for an association between TLBW and PM_{10} exposure. Examination of key highlighted studies points to the need for an up-to-date meta-analysis to synthesise all the available evidence.

NO_2 and TLBW

Commentary on results from meta-analyses

After exclusions, the core meta-analyses that examined associations between NO_2 exposures and LBW were Li et al (2020) and Stieb et al (2012) for trimesters 1, 2 and 3 and Stieb et al (2012) for Whole Pregnancy. Results are provided in Table 13, with more detailed information provided in Appendix 11, Table A 42.

Table 13. Summary of the results from the core meta-analyses examining term low birth weight and NO_2 exposure

Meta-analyses (Increment)	OR (95% CI)			
	WP	T1	T2	T3
Li et al 2020 (Per 10 ppb)	Excluded	1.02 (1.01, 1.04)*	1.01 (0.99, 1.04)*	1.01 (0.97, 1.06)*
Stieb et al 2012 (Per 20 ppb)	1.05 (1.00, 1.09)	1.03 (0.99, 1.06)	1.04 (1.01, 1.08)	0.98 (0.87, 1.10)

*For Li et al (2020) these values represent RR

Li et al 2020

Li et al (2020) reported summary estimates for LBW per 10 ppb NO_2, by exposure period, as follows: trimester 1 RR 1.02 (95% CI 1.01, 1.04), trimester 2 RR 1.01 (95% CI 0.99, 1.04) and trimester 3 RR 1.01 (95% CI 0.97, 1.06) (Table 13).

The trimester 1 estimate included 12 studies, trimester 2 includes 13 studies, and trimester 3 included 13 studies; overall the same studies were used for each trimester with the addition of Lavigne et al (2016) for trimesters 2 and 3. Heterogeneity was not important for trimester 1 (I^2 10.6%), with the odds ratios for all the studies with greatest weight at 1 or above and ~60% of the weight coming from one study, Lee et al (2003). However, there was substantial heterogeneity in the estimates for trimesters 2 and 3 respectively (I^2 74.9% and I^2 78.1%), with the odds ratios in the individual studies being both greater than 1 and below 1 that is, the heterogeneity was both in size and in direction of the association.

There was no indication of publication bias from funnel plots or Egger's test (p-values all >0.05) for the 3 trimester-specific analyses.

The trimester 1, 2 and 3 meta-analyses contained only 3, 4, and 4 individual studies that had adjusted for smoking.

TLBW was a selection criterion for the meta-analysis, so the studies should all be for TLBW rather than LBW.

Summary estimates for all the trimesters were very slightly elevated above 1 (1-2%), however, only those for trimester 1 reached conventional levels of significance. A finding of this small magnitude could have been due to residual confounding from smoking or from multiple testing. Overall, Li et al (2020) provides **very weak evidence to support an association of NO_2 exposures and TLBW in trimester 1, but no evidence for other exposure windows.**

Stieb et al 2012

Stieb et al (2012) reported summary estimates for LBW per 20 ppb NO_2, by exposure period, as follows: Whole Pregnancy OR 1.05 (95% CI 1.00, 1.09), trimester 1 OR 1.03 (95% CI 0.99. 1.06), trimester 2 OR 1.04 (95% CI 1.01, 1.08) and trimester 3 OR 0.98 (95% CI 0.87, 1.10) (Table 13).

The Whole Pregnancy estimate includes 10 estimates from 9 studies, trimester 1 includes 5 studies, trimester 2 includes 4 studies, and trimester 3 includes 5 studies; overall 11 studies were used at least once. Whilst heterogeneity might not be important for trimesters 1 (I^2 0%; p=0.905) and 2 (I^2 0%; p=0.863), there was substantial heterogeneity in Whole Pregnancy (I^2 78.4%; p<0.001) and trimester 3 (I^2 69.7%; p=0.010) estimates. There was no evidence of publication bias.

None of the studies in each meta-analysis adjusted for smoking, apart from one study in the meta-analysis for Whole Pregnancy (Brown et al, 2015), which accounts for 1.38% of the weight of the estimate.

Study selection was not restricted to TLBW, but all studies except Jackson et al (2009) and Liu et al (2003) used a TLBW definition.

Summary estimates for Whole Pregnancy and trimesters 1 and 2 were slightly elevated, however, only those for trimester 2 reached conventional levels of

significance. This could have been due to residual confounding or multiple testing. Overall, Stieb et al (2012) provides **very weak evidence to support an association of NO_2 exposures and TLBW in trimester 2, but no evidence for other exposure windows. It is important to note however that conclusions for the 3 trimester specific exposures are based on limited evidence, as each meta-analysis contained a small number of studies.**

Synthesis

The evidence from the 2 meta-analyses is largely consistent, in that all associations are around the null, even though the study findings are reported by different exposure increments. The main difference was that Li et al (2020) found a small elevated risk in trimester 1 that was unlikely to be due to chance whereas Stieb et al (2012) found a small elevated risk that was unlikely to be due to chance in trimester 2. It is difficult to pinpoint the reason for this difference but it may reflect a) a number of more recent studies included in the later meta-analysis by Li et al (2020) or b) differential control for confounding by smoking in the 2 meta analyses or c) a random occurrence due to the large amounts of multiple testing (including sensitivity testing) in both studies.

Conclusion regarding evidence for NO_2 and TLBW from core meta-analyses

Evidence from the 2 meta-analyses does not support an association (section 2.1.3) of NO_2 exposures and TLBW in whole pregnancy or any of the individual trimesters.

Commentary on evidence from key individual studies regarding NO_2 and Term Low Birth Weight (TLBW) (not included in the core meta-analyses)

Following the flowchart process as outlined in Figure 6, Endpoint D was reached (Appendix 10).

Brief description of available studies:

12 studies examining the association between NO_2 exposure during pregnancy and TLBW were captured by our literature search and were not included in any of the core meta-analyses above (Pedersen et al, 2013; Coker et al, 2016; Laurent et al, 2013; Laurent et al, 2016a; Poirier et al, 2015; Wilhelm et al, 2012; Nieuwenhuijsen et al, 2019; Kumar, 2012; Guo P et al, 2020; Yorifuji et al, 2015; Wu et al, 2018; Lin et al, 2004). N.B. If studies only looked at LBW not restricted to term we do not include them here. Five of these studies were published in 2016-2020, the remainder (n=7) between 2004-2015. Of the 12 studies included, there were 8 retrospective cohort studies using birth records or electronic medical records, 2 nested case-control studies, 1 study pooling multiple birth cohorts, and 1 population-based longitudinal survey. Geographically they were spread as follows: USA (n=5 (of which 4 are in California)), Canada (n=1) Europe (n=2), Asia (n=4 (of which 2 are in China)). All 12 studies examined whole pregnancy and/or trimester-specific exposure windows.

As stated in section 2.1.4, as we consider here only studies which do not appear anywhere in the core meta-analyses, there may be additional individual studies that would add to the evidence for a specific exposure window which are not considered

here. An example of such a study is Smith et al (2017), examining 540,365 births in London, UK – this was covered in the remaining Li et al (2020) meta-analysis on NO_2 and TLBW for trimester 1, trimester 2, and trimester 3, but is missing from the meta-analysis commentary for the whole pregnancy exposure window as the Li et al (2020) whole pregnancy meta-analysis was excluded due to population overlaps.

Brief commentary on selected key studies of interest (Appendix 12 Table A 54*) which are likely to expand the evidence base (but not a commentary on all):*

Of 5 new studies in the USA, we highlight Kumar (2012) as it was a large study (n=398,120) in Chicago, thus it expanded geographical coverage of the US beyond that of the selected meta-analyses (California, Connecticut, Massachusetts), and adjusted for smoking. It used environmental monitoring with interpolation for exposure assessment and examined exposure-outcome associations for a range of different distance buffers (3 mile, 6 mile, 9 mile and 12 mile buffers) plus a county-level exposure). It found no association between NO_2 and TLBW at any distance buffer/exposure level, for either whole pregnancy or trimester-specific exposure, all results were null (see Appendix 12 Table A 54 for ORs).

Whilst there were 4 other USA studies (Coker et al, 2016; Laurent et al, 2013; Laurent et al, 2016a; Wilhelm et al, 2012) these are not highlighted, as they were all from California and did not expand geographical coverage of the state beyond the meta-analyses above (populations in Los Angeles and Orange Counties). Moreover, they did not adjust for smoking (although one did a sensitivity analysis with control for smoking), and none increased on the sample size of Morello-Frosch et al (2010) study (~3.3 million births) included in the meta-analysis by Stieb et al (2012), so they would not increase statistical power. However, 3 out of 4 did use modelling for exposure assessment of NO_2.

The one study from Canada by Poirier et al (2015) is highlighted as it expanded geographic coverage of Canada to include Halifax, Nova Scotia (beyond Vancouver (Brauer et al, 2008), (Liu et al, 2003), and Ontario (Lavigne et al, 2016) covered in the meta-analyses above). The study used birth records to examine associations of several pollutants with a range of birth outcomes including TLBW (226 cases and 13,483 controls) in Halifax, Nova Scotia (Canada). Whole Pregnancy exposure was determined from a LUR model and based on the post-code of mother's address. The LUR model was able to predict 90% of the variation in NO_2. There were low concentrations of pollutants, and likely limited exposure contrast. Logistic regression models were adjusted for mother's age, parity, smoking, and neighbourhood income. The study analysed Whole Pregnancy exposure only and found no association between NO_2 and TLBW (OR 1.02 (95% CI: 0.83, 1.26) per IQR increase in NO_2, N.B. IQR could not be determined from the presented information) in the most fully adjusted model.

The previous meta-analyses included 6 European studies of NO_2 and TLBW but a pooled analysis across European birth cohorts for the ESCAPE analysis was subsequently published in 2013 (Pedersen et al, 2013). While a more recent study was published in 2019 by Nieuwenhuijsen et al, this varied in the type of model used for NO_2 estimation between cohorts (some LUR, some dispersion) and was a less

traditional analysis with a smaller population size, being designed to look at biomarkers. Therefore, the study by Pedersen et al (2013) is highlighted here. Pedersen et al (2013) was included in Li et al (2020) meta-analysis of Whole Pregnancy exposure to NO_2 (which has been excluded from further consideration in this report due to population overlaps). It was not included in the Li et al (2020) meta-analysis trimester specific meta-analysis (reason uncertain but may be because estimates are present in the supplemental tables only). Pedersen et al (2013), was a pooled analysis of European birth cohorts (ESCAPE study) (including n=70,551 term births, of which 1257 were TLBW (1.8%)). The authors examined the association of TLBW with annual mean NO_2 by collating information from several prospective birth cohorts that originated in Sweden, Norway, Netherlands, Lithuania, Hungary, Spain, Italy, Denmark, France and England. A common protocol for LUR was used, exposure was based on maternal address and accounted for moving during pregnancy in some centres. LUR models were validated with specific monitoring campaigns (although these were only for short periods). Model validation showed that the LUR models explained 58-90% of NO_2 variability depending on the area. Whole Pregnancy and trimester-specific estimates were derived. Logistic regression analyses were adjusted for major confounders including maternal smoking and individual socio-economic status. Multi-pollutant models were examined where data were available. There was an association of exposure to NO_2 in Whole Pregnancy with TLBW (OR 1.09 (95% CI: 1.00, 1.19) per 10 $\mu g/m^3$ NO_2 n= 61452). The associations for trimester-specific exposures were very similar (trimester 1: 1.08 (95% CI: 1.00, 1.17); trimester 2: 1.08 (95% CI: 1.00, 1.17); trimester 3: 1.07 (95% CI: 0.98, 1.16). However, no statistically significant association was observed when analysing the subset of study areas that also had $PM_{2.5}$ exposure information (Whole Pregnancy OR 1.05 (95% CI: 0.95,1.16) per 10 $\mu g/m^3$ NO_2, n=49,285), and adjustment for $PM_{2.5}$ attenuated the OR to null (Whole Pregnancy OR 1.01 (0.91, 1.11) per 10 $\mu g/m^3$ NO_2, n=49,285). The main analysis pooled the raw data, but a meta-analytical approach gave similar results with no important heterogeneity between cohorts.

Of 4 studies from Asia, we highlight one study by Guo P et al (2020) which covers 21 cities in Guangdong Province in China. It expanded geographical coverage of China beyond cities covered in the selected meta-analyses (cities of Changsha in Hunan province (Lu et al, 2020), Taizhou in Zhejiang province (Ye et al, 2018), and Shenzhen in Guangdong province (Wang et al, 2019), and was without population overlap with the latter study in Guangdong province. It had a very large sample size (over 2.5 million births), adjustment for smoking and analysis of multi-pollutant models. Exposure assessment was via daily monitoring data with inverse distance weighting to residential address. The study found no association between NO_2 and TLBW for any exposure window (OR (95% CI) per IQR increase in NO_2: Whole Pregnancy: 0.97 (0.92, 1.03), trimester 1: 0.98 (0.94, 1.02), trimester 2: 0.98 (0.93, 1.03), trimester 3: 0.97 (0.92, 1.03)). ORs were very similar after adjustment for other pollutants (CO, O_3, PM_{10}, $PM_{2.5}$, SO_2) in 2-pollutant models, for all exposure windows.

The 3 remaining studies from Asia (Yorifuji et al, 2015; Wu et al, 2018; Lin et al, 2004) are not highlighted as 2 lacked adjustment for smoking (Wu et al 2018, Lin et al 2004), one was very small (n=1845) (Wu et al 2018), and all used basic monitoring

approaches for exposure assessment. One (Lin et al, 2004) had previously been captured in the systematic review by Stieb et al (2012) but was not incorporated into the published meta-analysis.

The 4 highlighted studies provide little evidence for an association between TLBW and long-term exposure to NO_2 during pregnancy – 3 found no associations at all (Poirier et al, 2015; Kumar, 2012; Guo P et al, 2020), and the statistically significant association reported by Pedersen et al (2013) was not robust in subset analyses/to adjustment for other air pollutants.

Conclusion regarding evidence for NO_2 and TLBW from key individual studies not included in the core meta-analyses

The key highlighted studies do not support an association (section 2.1.4) of NO_2 exposure and TLBW. With specific reference to the ESCAPE study (Pedersen et al, 2013), our conclusion reflects that the association between NO_2 and TLBW, observed in the ESCAPE study, was attenuated to null after adjustment for $PM_{2.5}$.

OVERALL CONCLUSION for NO_2 and TLBW from the epidemiological evidence

We conclude that evidence from the meta-analyses and key highlighted studies does not support an association (section 2.1.3) between TLBW and NO_2 exposure. We nonetheless recommend an up-to-date meta-analysis to synthesise all the available evidence.

O_3 and TLBW

Commentary on results from meta-analyses

After exclusions, the core meta-analyses that examined associations between O_3 exposure and TLBW were Li et al (2020) for trimesters 1, 2, and 3 and Stieb et al (2012) for whole pregnancy and trimesters 1, 2 and 3. Results are provided in Table 14 and Appendix 11, Table A 43.

Table 14. Summary of the results from the core meta-analysis examining term low birth weight and O_3 exposure

Meta-analyses (Increment)	OR (95% CI)			
	WP	T1	T2	T3
Li et al 2020 (Per 10 ppb)	Excluded	1.00 (0.95, 1.05)*	1.02 (0.95, 1.09)*	1.09 (0.99, 1.20)*
Stieb et al 2012 (Per 20 ppb)	1.01 (0.82, 1.25)	0.99 (0.91, 1.08)	0.95 (0.79, 1.15)	1.03 (0.84, 1.26)

*For Li et al (2020) these values represent RR

Li et al 2020

Li et al (2020) reported summary estimates for TLBW per 10 ppb O_3, by exposure period, as follows: trimester 1 RR 1.00 (95% CI 0.95, 1.05), trimester 2 RR 1.02 (95% CI 0.95, 1.09), and trimester 3 RR 1.09 (95% CI 0.99, 1.20) (Appendix 11, Table A 43). These meta-analyses contained 9, 8, and 9 individual studies respectively.

There was considerable heterogeneity in the meta-analyses for all exposure windows (I^2 78.5 to 95.8%, p<0.001 in all cases), with the odds ratios in the individual studies being both greater than 1 and below 1 that is, the heterogeneity was both in size and in direction of the association, for all trimesters. There was no indication of publication bias from funnel plots or Egger's test (p-values all >0.2) for any trimester-specific analysis.

The majority of the weight of the meta-analysis estimates (68%, 86%, and 72% for trimester 1, 2 and 3 respectively) came from studies which had adjusted for smoking.

All studies included in these meta-analyses analysed LBW restricted to term births, except for Michikawa et al (2017) for which the estimate included in the meta-analysis is for 'adverse birth weight at term' defined as 'LBW in addition to SGA at term'. Michikawa et al (2017) contributed 13-14% of the weight to the meta-analyses but excluding it did not materially alter the meta-analysis summary estimates or their statistical significance.

The summary estimate for trimester 3 was in a positive (adverse) direction and borderline significant. However, given considerable heterogeneity in the trimester-specific analyses, and that none of the summary estimates reached conventional levels of significance we conclude that Li et al (2020) provides **no evidence to support an association of TLBW with O_3 exposure in any trimester.**

Stieb et al 2012

Stieb et al (2012) reported summary estimates for LBW per 20 ppb O_3, by exposure period, as follows: Whole Pregnancy OR 1.01 (95% CI 0.82, 1.25), trimester 1 OR 0.99 (95% CI 0.91, 1.08), trimester 2 OR 0.95 (95% CI 0.79, 1.15), and trimester 3 OR 1.03 (95% CI 0.84, 1.26) (Table 14). These meta-analyses contained 3, 5, 3, and 5 individual studies respectively.

There was considerable heterogeneity in the trimester 3 estimate 75.6 %, p=0.003), with the odds ratios in the individual studies being both greater than 1 and below 1 that is, the heterogeneity was both in size and in direction of the association. However, with such small numbers of studies included this statistic may not be particularly informative.

The authors combined estimates from all exposure windows (Whole Pregnancy, trimesters 1, 2 and 3) in a single funnel plot and report funnel plot asymmetry, suggesting 'greater than expected number of positive than negative effect sizes

among small, imprecise studies with larger standard errors' in relation to O_3 and LBW (p-value for Egger's bias coefficient was 0.03).

Only 2 individual studies included by Stieb et al (2012) adjusted for smoking, meaning 78%, 95%, 82% and 85% of the weight of the whole pregnancy, trimester 1, 2 and 3 meta-analysis estimates respectively came from studies with no adjustment for smoking.

All studies included in these meta-analyses analysed LBW restricted to term births, except one (Liu et al, 2003). Liu et al (2003) contributed 25% and 23% of the weight to the trimester 1 and 3 meta-analyses respectively, but had it been excluded it is unlikely that the meta-analysis summary estimates would have altered materially/reached conventional statistical significance.

Given the considerable heterogeneity in trimester 3 estimate, considerable potential for residual confounding by smoking in all these estimates, and that none of the summary estimates reached conventional levels of significance we conclude that Stieb et al (2012) provides **no evidence to support an association of TLBW with O_3 exposure, but this is on the basis of limited evidence as the meta-analyses contained small numbers of studies (<=5).**

Synthesis

The evidence from the 2 meta-analyses is largely consistent, in that almost all associations are null. The only divergence between these meta-analyses is for trimester 3, where the summary estimate from Li et al (2020) is suggestive of a positive association (although this fails to reach statistical significance), whereas in Stieb et al (2012) the association was very clearly null. This difference may reflect a) a number of more recent studies included in the later meta-analysis by Li et al (2020) and b) differential control for confounding by smoking: a much higher proportion of studies included in Li et al (2020) adjusted from smoking, whereas for Stieb et al (2012) 85% of the weight of the trimester 3 estimate came from studies which did not adjust for smoking. However, the same comments would be applicable if comparing the meta-analysis for trimester 1 and trimester 2, for which the results from both meta-analyses were null.

Conclusion regarding evidence for O_3 and TLBW from core meta-analyses

We conclude the 2 meta-analyses do not support an association (section 2.1.3) between prenatal exposure to O_3 and TLBW. However, this conclusion regarding whole pregnancy O_3 exposure is based on limited evidence, as it rests on a single meta-analysis of 3 individual studies.

Commentary on evidence from key individual studies regarding O_3 and Term Low Birth Weight (TLBW) (not included in the core meta-analyses)

Following the flowchart process as outlined in Figure 6, Eenpoint D was reached (Appendix 10).

Brief description of available studies:

Twelve studies examining the association between O_3 exposure during pregnancy and TLBW were captured by our literature search and were not included in any of the core meta-analyses above (Laurent et al, 2013; Laurent et al, 2016a; Le et al, 2012; Wilhelm et al, 2012; Capobussi et al, 2016; Brown et al, 2015; Ha et al, 2017; Kumar, 2012; Ghosh et al, 2012; Laurent et al, 2014; Guo P et al, 2020; Lin et al, 2004). N.B. If studies only looked at LBW not restricted to term; we do not include them here. Four of these studies were published during 2016-2020, the remainder (n=8) between 2004-2015. There were 11 retrospective cohort studies using birth records or electronic medical records, and 1 nested case-control study within a retrospective cohort. Geographically they were spread as follows: USA (n=9), Europe (n=1), Asia (n=2). N.B., it is possible that some of the earlier studies were not included in the core meta-analyses above due to decisions authors made regarding which study to include where there were several examining the same/similar population or for other reasons, but the basis for exclusion from a meta-analysis is not always transparent. For example, the earliest study from 2004 (Lin et al) was captured in the systematic review by Stieb et al (2012) but was not incorporated into their published meta-analysis.

As stated in section 2.1.4, as we consider here only studies which do not appear anywhere in the core meta-analyses, there may, in addition, be individual studies that would add to the evidence for a specific exposure window which are not considered here. An example of such a study is Smith et al (2017), examining 540,365 births in London, UK – this was covered in the remaining Li et al (2020) meta-analysis on O_3 and TLBW for trimester 1, trimester 2, and trimester 3, but is missing from the meta-analysis commentary for the whole pregnancy exposure window as the Li et al (2020) Whole Pregnancy meta-analysis was excluded due to population overlaps.

Brief commentary on selected key studies of interest (Appendix 12 Table A 55) which are likely to expand the evidence base (but not a commentary on all):

In particular, the geographical coverage of the research/evidence base is expanded beyond that of the meta-analyses by the studies highlighted below:

Of 9 studies in the USA (Laurent et al, 2013; Laurent et al, 2016a; Le et al, 2012; Wilhelm et al, 2012; Brown et al, 2015; Ha et al, 2017; Kumar, 2012; Ghosh et al, 2012; Laurent et al, 2014;) we highlight the following:

- A study by Ha et al (2017) is highlighted as it examined a wide range of sites across the US, used modelled exposures and adjusted for smoking. Ha et al (2017) investigated the potential association of extreme ambient temperature, 5 criteria air pollutants, and 6 particulate matter constituents with TLBW, in a large US obstetric cohort (Consortium on Safe Labor), that collected detailed medical records from 19 hospitals/15 hospital referral regions across the US (in the states of California (LA), Utah, Texas, Illinois, Ohio, Indiana, Florida, Washington DC, Maryland, Delaware, NY State and Maine). Exposures were estimated using a modified CMAQ model. Researchers did not have access to maternal address,

so exposure assignment was based on the average concentrations in the 15 non-overlapping delivery hospital referral regions of hospital of delivery (area range: 415 to 312,644 km^2), weighted for population density. Over 220,000 births (about 2% of which were TLBW) remained in the regression analyses which were able to adjust for a wide range of confounders including maternal smoking in pregnancy and ethnicity but had few adjustments for socio-economic status (used insurance status). The study stated Poisson regression was used but gave little detail. It found no statistically significant associations between O_3 and TLBW for any exposure window (RR (95% CI) per IQR increase in O_3: Whole Pregnancy RR 1.01 (0.91, 1.11), trimester 1 RR 1.00 (0.92, 1.07), trimester 2 RR 1.03 (0.95, 1.11), trimester 3 RR 1.00 (0.93, 1.08).

- Brown et al (2015) is highlighted as it was a large study (n=480,430) in New York State, which expanded geographical coverage of the US beyond that of the selected core meta-analyses and used modelled exposures (USEPA Bayesian hierarchical space-time modelling on a 12x12km grid scale). Exposure was assigned according to the $12km^2$ grid cell in which the maternal address was located. A wide range of potential confounders were adjusted for, including maternal smoking during pregnancy. It analysed exposure as a categorical variable (quartiles) and found no statistically significant associations between O_3 and TLBW for any exposure quartile, although there was a suggestion of a possible exposure-response relationship for trimester 1 exposure only, (OR (95% CI) for O_3 exposure >75th percentile vs ≤25th percentile (reference category): Whole Pregnancy OR 0.98 (0.92, 1.04), trimester 1 OR 1.07 (0.99, 1.14), trimester 2 OR 0.98 (0.92, 1.05), trimester 3 OR 0.95 (0.90, 1.02).

- Kumar (2012) is highlighted as it was a large study (n=398,120) in Chicago, which expanded geographical coverage of the US beyond that of the selected core meta-analyses and adjusted for smoking. It used environmental monitoring with interpolation for exposure assessment and examined exposure-outcome associations for a range of different distance buffers (3 mile, 6 mile, 9 mile and 12 mile buffers) plus a county-level exposure. It found no association between O_3 and TLBW at any distance buffer/exposure level, for either whole pregnancy or trimester-specific exposure, all results were null (see Appendix 12 Table A 55 for ORs).

- Le et al (2012) was a medium sized study (n=139,951) in Detroit, highlighted as it expanded geographical coverage of the US beyond that of the selected meta-analyses and adjusted for smoking. Le et al (162) analysed trimester-specific exposures (no Whole Pregnancy exposure) as categorical variables (quartiles) and found no statistically significant associations between O_3 and TLBW for any exposure quartile, nor any indication of an exposure-response relationship (either when examining all seasons or just high season). (OR (95% CI) for O_3 exposure >75th percentile vs ≤25th percentile (reference category) for all seasons given here: trimester 1: OR 0.92 (0.80, 1.05), trimester 2: OR 0.90 (0.78, 1.03), trimester 3: OR 0.93 (0.83, 1.05).

Whilst there were 5 other studies from the USA (Laurent et al, 2013; Laurent et al, 2016a; Wilhelm et al, 2012; Ghosh et al, 2012; Laurent et al, 2014), these were all

from California (covering populations in Los Angeles and Orange Counties), and none of these are highlighted as they used environmental monitoring approaches for exposure assessment of O_3, did not adjust for smoking, did not perform multi-pollutant models, and therefore do not advance these aspects of the research over the Californian studies included in the Stieb et al 2012 meta-analysis. Also, none increased on the sample size of Morello-Frosch et al (175) study (~3.3 million births) included in the meta-analysis by Stieb et al (2012), so they do not increase statistical power.

The only European study covered the Como Region of Italy (Capobussi et al, 2016). It is not highlighted as small in size, did not use modelling for exposure assessment and lacked adjustment for smoking.

Of 2 studies in Asia (Guo P et al, 2020; Lin et al, 2004), we highlight the study by Guo P et al (2020) which covered 21 cities in Guangdong Province in China. It expanded geographical coverage, as China is not covered by the selected core meta-analyses. The study had a very large sample size (over 2.5 million births), adjustment for smoking and analysis of multi-pollutant models. Exposure assessment was via daily monitoring data with inverse distance weighting to residential address. Associations between O_3 and TLBW did not reach statistical significance for any exposure window (OR (95% CI) per IQR increase in O_3: Whole Pregnancy: OR 1.05 (0.99, 1.10), trimester 1: OR 1.02 (0.98, 1.06), trimester 2: OR 1.04 (0.98, 1.11), trimester 3: OR 0.99 (0.94, 1.05)). ORs were very similar after adjustment for other pollutants (CO, NO_2, PM_{10}, $PM_{2.5}$, SO_2) in 2-pollutant models, for all exposure windows.

The exposure-outcome associations from the 5 highlighted studies (Le et al, 2012; Brown et al, 2015; Ha et al, 2017; Kumar, 2012; Guo P et al, 2020) failed to reach statistical significance and were largely null.

Conclusion regarding evidence for O_3 and TLBW from key individual studies not included in the core meta-analyses

There are several large studies not captured in/published since the core meta-analyses, which expand geographical coverage of the evidence base. The key highlighted studies do not support an association (section 2.1.4) of O_3 exposure and TLBW.

OVERALL CONCLUSION for O_3 and TLBW from the epidemiological evidence

We conclude evidence from the meta-analyses (limited evidence) and key highlighted studies does not support an association between TLBW and O_3 exposure (section 2.1.3). We nonetheless recommend an up-to-date meta-analysis to synthesise all the available evidence.

2.2.4 Commentary on results from studies examining Stillbirth

All Pollutants ($PM_{2.5}$, PM_{10}, NO_2 and O_3) and Stillbirth

Commentary on results from meta-analyses

Siddika et al (2016) was the only meta-analysis that examined associations of $PM_{2.5}$, PM_{10}, NO_2 and O_3 exposure with stillbirth for Whole Pregnancy, and trimesters 1,2 and 3. A summary of results are provided in Table 15, with more detailed information provided in Appendix 11, ($PM_{2.5}$: Table A 44; PM_{10}: Table A 45; NO_2:Table A 46 and O_3: Table A 47).

Table 15. Summary of the results from Siddika et al (2016)

Pollutant	(Increment)	OR (95% CI)			
		WP	T1	T2	T3
$PM_{2.5}$	**(Per 4 µg/m³)**	1.02 (1.00. 1.05)	1.04 (0.92, 1.18)	1.04 (0.94, 1.15)	1.00 (0.98. 1.02)
PM_{10}	**(Per 10 µg/m³)**	1.01 (0.95, 1.09)	1.00 (0.94, 1.06)	1.01 (0.91, 1.12)	1.02 (0.92, 1.13)
NO_2	**(Per 10 ppb)**	1.07 (0.97, 1.18)	1.04 (0.98, 1.09)	1.01 (0.95, 1.07)	1.02 (0.98, 1.05)
O_3	**(Per 10 ppb)**	1.00 (0.97, 1.03)	1.00 (0.98. 1.02)	0.99 (0.94, 1.04)	1.01 (0.97, 1.06)

The summary estimates, based on information from 2 studies, were as follows:

- For $PM_{2.5}$: Whole Pregnancy OR 1.02 (95% CI 1.00, 1.05), trimester 1 OR 1.04 (95% CI 0.92, 1.18), trimester 2 OR 1.04 (95% CI 0.94, 1.15) and trimester 3 OR 1.00 (95% CI 0.98, 1.02) (Table 15). The 2 studies included were Faiz et al (2012) and Green et al (2015) (Appendix 11, Table A 44).

- For PM_{10}: Whole Pregnancy OR 1.01 (95% CI 0.95, 1.09), trimester 1 OR 1.00 (95% 0.94, 1.06), trimester 2 OR 1.01 (95% CI 0.91, 1.12), and trimester 3 OR 1.02 (95% CI 0.92, 1.13) (Table 15). The 2 studies comprise of Hwang et al (2011) and Kim et al (2007) (Appendix 11, Table A 45).

- For O_3: Whole Pregnancy OR 1.00 (95% CI 0.97, 1.03), trimester 1 OR 1.00 (95% CI 0.98, 1.02), trimester 2 OR 0.99 (95% CI 0.94, 1.04), and trimester 3 OR 1.01 (95% CI 0.97, 1.06). These meta-analyses each included 2 individual studies (Table 15). The 2 studies included were Green et al (2015) and Hwang et al (2011) (Appendix 11, Table A 47).

The summary estimates, based on information from 3 studies, were as follows:

- For NO_2: Whole Pregnancy OR 1.07 (95% CI 0.97, 1.18), trimester 1 OR 1.04 (95% CI 0.98, 1.09), trimester 2 OR 1.01 (95% CI 0.95, 1.07), and trimester 3 OR 1.02 (95% CI 0.98, 1.05) (Table 15). The 3 studies included were Faiz et al (2012), Green et al (2015) and Hwang et al (2011) (Appendix 11, Table A 46).

Heterogeneity and publication bias were assessed, but as the authors point out, these are relatively uninformative when only 2 or 3 studies are included in a meta-analysis (as was the case here).

For PM_{10} the effect estimates from each study were opposite for each analysis (reflected in the I^2 estimate, see Appendix 11, Table A 45). For example, for the Whole Pregnancy estimate, the Hwang et al (2011) point estimate suggested negative associations (OR 0.96, CI 0.94, 1.01) whereas Kim et al (2007 showed positive (adverse) associations as (OR 1.05, CI 1.01. 1.09).

Of the 2 studies included for $PM_{2.5}$ in the meta-analysis, only 1 (Faiz et al, 2012), adjusted for smoking; Green et al (2015) did not. Green et al (2015) comprised 98% of the weight for the Whole Pregnancy estimate (70% for trimester 1, 75% for trimester 2 and 99% for trimester 3). For the 2 studies included for PM_{10} and O_3, neither adjusted for smoking. Of the 3 studies included for NO_2, only 1 of the 3 studies (Faiz et al, 2012) had adjusted for smoking and this study had the least weight for Whole Pregnancy and each of the trimesters (<15%).

Conclusion regarding evidence for All Pollutants ($PM_{2.5}$, PM_{10}, NO_2 and O_3) and stillbirth from the core meta-analysis

The meta-analysis (Siddika et al, 2016;) does not support an association (section 2.1.3) between stillbirth with $PM_{2.5}$, PM_{10}, NO_2 or O_3 exposure, but this is on the basis of limited evidence as the meta-analyses contained small numbers of studies (<=3). Our conclusion reflects the considerable potential for residual confounding by smoking in these estimates, and that none of the summary estimates reached conventional levels of significance.

Commentary on evidence from key individual studies regarding all pollutants ($PM_{2.5}$, PM_{10}, NO_2 and O_3) and Stillbirth (not included in the Siddika et al (63) meta-analysis)

Following the flowchart process as outlined in Figure 6, Endpoint D was reached (Appendix 10).

Brief description of available studies:

Thirteen studies examining air pollution and stillbirth were captured by our literature search that were not included in the meta-analysis above.

The following studies were not considered further:

- DeFranco et al (DeFranco et al, 2015) was considered for the Siddika et al (2016) meta-analysis but excluded as thought to be of too low quality.
- Dastoorpoor et al (2018) and Sarovar et al (Sarovar et al, 2020) which considered short-term effects of pollution in Iran and California respectively.
- Butler et al (2021) based in Massachusetts and Rhode Island, which considered pollution exposure as ‘distance to road’ only.

- Rammah et al (Rammah et al, 2019a) based in Texas, USA, which considered levels of 10 metal constituents of **$PM_{2.5}$** (aluminium (Al), chromium (Cr), copper (Cu), iron (Fe), lead (Pb), manganese (Mn), nickel (Ni), strontium (Sr), titanium (Ti) and zinc (Zn)). Of the 10 metal constituents considered there was a single significant association of stillbirth with zinc levels (1.11 95% CI 1.02-1.20 per IQR increase in Zn). Of note, this study furthered the analysis of the health data from another paper by the same authors that was also captured in our literature search and which is discussed in detail below (Rammah et al, 2019b).

- Ghosh et al (2019) based in Great Britain which considered **PM_{10}** emissions only from municipal waste incinerators (MWI), where this was considered as a proxy for MWI emissions. There was no evidence of any associations with stillbirth.

This left 7 studies of interest – all of them published since the Siddika et al (2016) systematic review and meta-analysis. Five were retrospective cohort studies using birth records or electronic medical records, 1 adopted a nested case control design using birth records and one was a multi-national case control study based on self reported stillbirth (controls were a livebirth from a prior or later pregnancy in the same mother).

They were geographically spread as follows:

- For $PM_{2.5}$: USA (n=2), China (n=2) UK (n=1), Africa (n=1) – the Siddika et al (2016) meta-analysis had only 2 USA based studies.

- For PM_{10}: USA (n=1), China (n=2) UK (n=1) – the Siddika et al (2016) meta-analysis had only 2 Asia based studies.

- For NO_2: USA (n=1), China (n=2) UK (n=1) - the Siddika et al (2016) meta-analysis had only 2 USA and 1 Asia based studies.

- For O_3: USA (n=2), China (n=2) UK (n=1), the Siddika et al (2016) meta-analysis had only 1 USA and 1 Asia based study.

Brief commentary on relevant studies identified by the literature search (Appendix 12 Table A 56)

One study (Xue et al, 2019) based in 22 countries in Africa used advanced satellite remote sensing technology linked with participant residential history to assess **$PM_{2.5}$** exposure. Information on history of pregnancy loss (representing miscarriage and stillbirth) was collected by questionnaire from participants in the African Demographic and Health Survey, with participants acting as their own controls (that is, prenatal exposure for a pregnancy loss was compared with prenatal exposure for a prior or later successful term birth). Stillbirth was considered as pregnancy loss after 5 months' gestation[9]. To account for biases that may occur when gestational

[9] Stillbirth is usually defined in the UK as a baby born after 24 completed weeks gestation showing no signs of life. Different timing cut-offs, typically ranging from 20-28 weeks' gestation, are used for defining stillbirth in different countries

length is very different in cases and controls, they matched the time window for effective exposure for both the case and control by calculating the gestational length adjusted $PM_{2.5}$. For a case, the effective exposure was calculated as the average monthly $PM_{2.5}$ during the reported full gestation period. Next, for the corresponding controls, the maternal exposure on the basis of the gestational length of the matched case, regardless of the control's gestational length was calculated. Mean modelled $PM_{2.5}$ exposures were high (mean whole pregnancy average for live births 25.7 $\mu g/m^3$) and while there was evidence that $PM_{2.5}$ exposure was associated with pregnancy loss the association was more robust for miscarriage (1.125 95% CI 1.109, 1.142 per 10 $\mu g/m^3$) than for stillbirth (overall estimate 1.094 95% CI 1.051, 1.138 per 10 $\mu g/m^3$). For stillbirth, there was suggestion of a non-linear effect with a threshold for effect at 80 $\mu g/m^3$ and evidence of negative confounding (that is, effects became stronger with further adjustment). Furthermore, effects were weaker if exposure in last month, last 2 months, last 3 months and last 5 months were considered (the authors consider these as fixed lags). The authors state '...fixed-lag models did not detect a significant effect of $PM_{2.5}$ on stillbirth, possibly because of potential exposure misclassifications,' and '…The function for stillbirth was non-linear, and its pairwise CIs suggested non-significant effects for most mothers, except a minor group of 188 mothers (1.5% of the 12,534 mothers who reported stillbirth) with exposure concentrations greater than 81 $\mu g/m^3$' and 'Both (linear and non-linear models) indicated strong effects of $PM_{2.5}$ on miscarriage but weak effects on stillbirth'. Sub-group analysis suggested that effects were stronger in older women and those less well educated. While this study is of interest the methods of outcome assessment and the populations included limit generalisability to the UK population.

Two very similar studies from China, one in Wuhan (Yang et al, 2018) and one in Yancheng (Zang et al, 2019) used information from maternity-related administrative records (all pregnant women must register in these databases during the first trimester) linked to pollutant level information **($PM_{2.5}$, PM_{10}, NO_2, O_3)** from local air quality monitoring stations adopting the nearest monitor method for exposure assignment (Whole Pregnancy, trimester 1, trimester 2 and trimester 3). Daily pollutant exposures were averaged to generate mean trimester specific exposure and Whole Pregnancy exposure. No attempt was made to account for potential biases in exposure that may arise from the differences in gestational length of stillbirths and live births. The definition of stillbirth in each study was slightly different (28-42 weeks for Zang et al (2019); 20-42 weeks in Yang et al (2018). The appendix to the Yang et al (2018) provides tables with risk estimates based on a definition of stillbirth requiring occurrence between 28- and 42-weeks' gestation (these estimates have been used in Tables 16 to 20). In logistic regression analyses adjustments were made for a range of factors but smoking was not one of these because 'none of the women reported smoking (or alcohol drinking) during the pregnancy' (Zang et al, 2019) and 'because of the extremely low prevalence of smokers among our study population' Yang et al (2018). Both studies provide 2 pollutant models for exposure through the entire pregnancy.

One study in London, UK (Smith et al, 2020), used information from a large number of routine birth records (n=581,774) to examine the association between stillbirth and a range of particulate matter and gaseous air pollutant exposures in London, UK

($PM_{2.5}$, PM_{10}, NO_2, O_3). The study used highly spatially resolved modelled exposures (monthly estimates from dispersion model at 20m x 20m grid point resolution), assigned to maternal address. Stillbirth was defined as occurring from 24 weeks and they assessed the association of stillbirth with pollutant exposure in trimester 1 and trimester 2 only. Analyses were limited to these 2 periods to avoid spurious associations that may occur for the longer exposure periods when there are marked differences in gestational length of stillbirths and live births. Logistic regression analyses were adjusted for a wide range of potential confounders, including area-level adjustment for tobacco sales (as a proxy for smoking), and road traffic noise. Multi-pollutant models were assessed. Two different categories of stillbirth were assessed – asphyxia related, and congenital malformation related.

The remaining 3 studies were based in the USA.

The largest of these was from a multicentre study (Consortium on Safe Labor - CSL) in USA (228,438 deliveries) using electronic health records from participating hospitals linked to pollutant level information generated from a CMAQ model that provided hourly levels (**$PM_{2.5}$, PM_{10}, NO_2, O_3**) then averaged to daily levels (Mendola et al, 2017). Exposure assignment was based on the referral hospital (mothers residential address was not available for researchers). For the Whole Pregnancy exposure analysis, to ensure comparable length of exposure, they stratified stillbirths by gestational age and truncated the exposures among ongoing births as a comparison.

Stillbirth was defined as occurring from 24 weeks and the chronic O_3 exposure periods were for Whole Pregnancy and trimester 1 only. This study also examined acute short-term exposure around the time of birth (delivery day, and then 1, 2, 3, 4, 5, 6 and 7 days before delivery). Adjustments were made for a range of factors including self-reported maternal smoking in pregnancy. Two pollutant models were tested.

Ebisu et al (2018) conducted a matched case-control study using administrative records for live birth and stillbirth in California from 1 January 2002 to 31 December 2009, and the analysis was restricted to **$PM_{2.5}$** (total mass, sources, and chemical constituents). There was considerable overlap with the population studied in Green et al (2015) (included in the Siddika et al (2016) meta-analysis), which used information from California between 1 January 1999 and 31 December 2009. These 2 studies came from the same research group, and exposure assignment methods were broadly similar (information from local air quality network monitors assigning exposure based on zip code centroid because researchers did not have access to full maternal residential address). Weekly averaged pollutant levels were first calculated, and then combined to generate mean gestational exposure levels for mothers with stillbirth. After matching, they assigned gestational age-adjusted exposure levels to controls based on the gestational length of the matched case, regardless of the gestational lengths of the controls. For instance, if the gestational length of a stillbirth case was 33 weeks, the exposures for the controls were also calculated based on 33 weeks of exposure. The study did not adjust for maternal smoking. Ebisu et al (2018) examined associations using conditional logistic regression and also considered different categories of stillbirth based on ICD_10

codes – congenital malformations, fetal disorders, fetal growth, maternal complications, and obstetric complications.

Rammah et al (2019b) used administrative records in Harris County, Texas linked to hourly pollutant levels from local air quality monitoring networks applying an inverse-distance weighting to derive the daily maximum 8-hour average **O_3** concentration or the daily average **$PM_{2.5}$** or **NO_2** concentrations for the 3 monitoring stations nearest to the mother's geocoded residential address. Mean exposure for each trimester was determined. Stillbirth was defined as occurring from 20 weeks, and the analysis was conducted as a time to event analysis with time varying exposure and generation of hazard ratios. This approach overcomes potential bias in exposure assignment that may arise if there are very different gestational lengths between stillbirths and the comparison group of livebirths. This study also examined short term exposures using a case cross-over design. Adjustments were made for a range of factors including self-reported maternal smoking in pregnancy. Two pollutant models were tested (with $PM_{2.5}$ or NO_2).

Due to the limited number of studies for stillbirth it is feasible to summarise and compare the evidence from the meta-analysis and key highlighted studies, in tabular form for each pollutant, as follows.

$PM_{2.5}$ and Stillbirth

Table 16. The Siddika meta-analysis effect estimates and those from the 6 key studies of interest for $PM_{2.5}$ and Stillbirth

Paper (Increment)	OR (95% CI)			
	WP	T1	T2	T3
Siddika et al (2016) (Per 4 µg/m³)	1.02 (0.99, 1.04)	1.04 (0.92, 1.18)	1.04 (0.94, 1.15)	1.00 (0.98, 1.02)
Xue et al 2019 (Per 10 µg/m³)	**1.09 (1.05, 1.14)**			
Yang et al 2018 (Per 10 µg/m³) From Table 1(Appendix)	**1.13 (1.05 1.21)**	**1.06 (1.01, 1.12)**	**1.07 (1.03, 1.11)**	**1.02 (0.98, 1.07)**
Zang et al 2019 (Per 10 µg/m³)	**1.14 (1.08, 1.20)**	**1.14 (1.06, 1.22)**	**1.11 (1.06, 1.16)**	**1.15 (1.09, 1.21)**
Smith et al 2020 (Per IQR)	**-**	**0.88 (0.82, 0.93)⁺**	**0.91 (0.85, 0.99)⁺⁺**	**-**
Mendola et al 2017 (Per IQR)	1.08 (0.75, 1.55)^	0.91 (0.77, 1.07)^^	-	-
Ebisu et al 2018 (Per IQR) NB population overlap with Green et al included Siddika et al	**1.06 (1.01, 1.11)$**		-	-

⁺ IQR = 3.2 µg/m³; ⁺⁺IQR= 3.6 µg/m³
^ IQR =4.7 µg/m³; ^^IQR 5.3 µg/m³
$ IQR = 7.23 µg/m³

Six studies (Smith et al, 2020; Xue et al, 2019; Yang et al, 2018; Zang et al, 2019; Mendola et al, 2017; Ebisu et al, 2018) reported associations of $PM_{2.5}$ with stillbirth (whole pregnancy: 5 studies, trimester 1: 4 studies, trimester 2: 3 studies, trimester 3: 2 studies).

Two studies reported associations for Whole Pregnancy only (Xue et al, 2019; Ebisu et al, 2018). The former was a multicentre study in Africa reporting an overall positive association with some suggestion of a threshold effect above 80 µg/m³) (Table 16). The methodology adopted (self-reported stillbirth, use of earlier or later pregnancy as control), and the location of the study (low- and middle-income countries with high $PM_{2.5}$ levels) may limit generalisability of results to the UK and other settings. One study from California (Ebisu et al, 2018) reported positive (adverse) associations of stillbirth with whole pregnancy in California – there is

population overlap with one of the studies included in the Siddika et al (2016) meta-analysis. Ebisu et al (2018) also examined categories of causes of stillbirth and showed that strongest associations were for stillbirths categorised as 'fetal growth' (OR 1.23 95% CI 1.06, 1.44 per IQR $PM_{2.5}$) and for 'obstetric complications' (OR 1.13 95% CI 1.03, 1.24 per IQR $PM_{2.5}$) with no associations observed for 'maternal complications', 'congenital malformations' or 'fetal disorders' Ebisu et al (2018) did not adjust for smoking and associations of whole pregnancy $PM_{2.5}$ exposure with 'fetal growth related stillbirth' could be confounded by maternal smoking.

Positive (adverse) associations of stillbirth with $PM_{2.5}$ exposure during pregnancy were reported from 2 studies in China (statistically significant associations reported for whole pregnancy and all trimester specific exposures in both studies) (Yang et al, 2018; Zang et al, 2019) and one in Africa (Whole Pregnancy only) (Xue et al, 2019). In both studies $PM_{2.5}$ levels are high compared to UK, and maternal smoking levels are very low. In both studies, the statistical approach does not account for potential biases that may occur due to the shorter gestational length of stillbirths compared to live births (see 'Specific considerations for studies looking at Preterm birth (PTB) and Stillbirth' in section 2.1.4). Of note in Yang et al (2018) analyses were conducted even though almost 20% of days had no $PM_{2.5}$ measurement (measurements could not be made if it was raining). Positive associations were seen for trimester 1 and trimester 2 in both studies (Table 16).

One large well conducted study from London UK (Smith et al, 2020), which included a proxy area level adjustment for tobacco consumption, showed negative associations of stillbirth with trimester 1 and trimester 2 modelled $PM_{2.5}$ exposure that reached conventional levels of statistical significance.

We note that 4 of the 5 reported Whole Pregnancy associations were positive (adverse), 2 of the 4 reported trimester 1 and trimester 2 associations were positive, and both of the 2 reported trimester 3 associations were positive. However, the 2 well conducted large studies in populations most similar to the UK setting (Smith et al, 2020; Mendola et al, 2017) reported negative or no associations.

Conclusion regarding evidence for $PM_{2.5}$ and stillbirth from key individual studies not included in the core meta-analyses

Between 2 large UK-relevant studies, and the other studies identified, the differences are likely related to differences in geographical area, background levels of $PM_{2.5}$, whether the study adjusted for smoking and, critically, whether the overall statistical approach accounted for differences in gestational length of stillbirths and live births. We conclude these new studies provide inconsistent evidence (section 2.1.4) for an association of stillbirth with exposure to $PM_{2.5}$ in whole pregnancy or any specific trimester.

OVERALL CONCLUSION for $PM_{2.5}$ and Stillbirth from the epidemiological evidence

We conclude that the limited evidence from the meta-analysis available does not support an association (section 2.1.3) between stillbirth and $PM_{2.5}$ exposure. Examination of key highlighted studies points to the need for an up-to-date meta-analysis to synthesise all the available evidence.

PM_{10} and Stillbirth

Table 17. The Siddika meta-analysis effect estimates and those from the 4 key studies of interest for PM_{10} and Stillbirth

Paper (Increment)	OR (95% CI)			
	WP	T1	T2	T3
Siddika et al (2016) (Per 10µg/m³)	1.01 (0.94, 1.08)	0.99 (0.94, 1.06)	1.00 (0.90, 1.16)	1.02 (0.91, 1.13)
Yang et al 2018 (Per 20 µg/m³) from Table 1 Appendix A **	1.18 (0.91, 1.51)	0.96 (0.84, 1.10)	1.09 (0.96, 1.23)	**1.20 (1.01, 1.44)**
Zang et al 2019 (Per 10 µg/m³)	1.04 (0.96, 1.12)	**1.09 (1.04, 1.14)**	1.05 (0.94, 1.15)	1.02 (0.97, 1.17)
Smith et al 2020 (Per IQR)	-	**0.84 (0.79, 0.88)⁺**	**0.91 (0.86, 0.97)⁺⁺**	-
Mendola et al 2017 (Per IQR)	1.04 (0.86,1.25)^	0.97 (0.85, 1.12)^^	-	-

** Note that the table states the estimates are per 20 µg/m³ although estimates quoted as per 10 µg/m³ throughout rest of paper

⁺ IQR = 4.0 µg/m³; ⁺⁺IQR= 4.5 µg/m³

^ IQR =4.4 µg/m³; ^^IQR 6.3 µg/m³

We highlight 4 studies (Smith et al, 2020; Yang et al, 2018; Zang et al, 2019; Mendola et al, 2017) reporting associations of PM_{10} with stillbirth (Whole Pregnancy: 3 studies, trimester 1: 4 studies, trimester 2: 3 studies, trimester 3: 2 studies).

All 3 studies reporting associations for Whole Pregnancy observed weak non-significant associations of stillbirth with PM_{10} exposure but 2 of these studies (Yang et al, 2018; Zang et al, 2019) did not use statistical approaches that account for potential biases that may occur due to the shorter gestational length of stillbirths compared to live births. Furthermore, in both studies $PM_{2.5}$ levels were high compared to UK and maternal smoking levels were very low. One large, well conducted study (Mendola et al, 2017) from the USA that adjusted for smoking showed no association (Table 17).

There were heterogeneous results from the 4 studies reporting associations of stillbirth with PM_{10} exposure in trimester 1 with one study from China (Zang et al, 2019) reporting statistically significant positive (adverse) associations and a well conducted UK based study (Smith et al, 2020) reporting statistically significant negative associations. There was also variation in the estimates reported from the 3 studies reporting associations for trimester 2 (Table 17) again with Smith et al (Smith et al, 2020) reporting statistically significant negative associations.

Positive (adverse) associations of stillbirth with PM_{10} exposure during trimester 3 were reported from the 2 studies in China (Yang et al, 2018; Zang et al, 2019) but as stated above both studies did not use a statistical approach that accounted for potential biases that may occur due to the shorter gestational length of stillbirths compared to live births (see 'Specific considerations for studies looking at Preterm birth (PTB) and Stillbirth' in section 2.1.4).

Conclusion regarding evidence for PM_{10} and stillbirth from key individual studies not included in the core meta-analyses

The key highlighted studies have inconsistent results (section 2.1.4) but provide little convincing evidence for an association of stillbirth with exposure to PM_{10} in whole pregnancy or any specific trimester.

OVERALL CONCLUSION for PM_{10} and Stillbirth from the epidemiological evidence

Limited evidence from the one meta-analysis available does not support an association (section 2.1.3) between stillbirth and PM_{10} exposure. Examination of key highlighted studies points to the need for an up-to-date meta-analysis to synthesise all the available evidence more succinctly, but from the evidence base we have considered it is unlikely to alter this conclusion.

NO_2 and Stillbirth

Table 18. The Siddika meta-analysis effect estimates and those from the 4 key studies of interest for NO_2 and Stillbirth

Paper (Increment)	OR (95% CI)			
	WP	T1	T2	T3
Siddika et al (2016) (Per 10ppb NO_2)	1.07 (0.96, 1.17)	1.03 (0.98, 1.09)	1.00 (0.95, 1.07)	1.02 (0.98, 1.05)
Yang et al 2018 (Per 10 µg/m^3 NO_2) from Table 1 Appendix A	1.00 (0.88, 1.13)	0.95 (0.87, 1.02)	1.00 (0.93, 1.08)	1.02 (0.92, 1.13)
Zang et al 2019 (Per 10 µg/m^3 NO_2)	1.01 (0.98, 1.04)	1.03 (0.96, 1.10)	1.01 (0.95, 1.07)	1.03 (0.97, 1.09)
Smith et al 2020 (Per NO_2 IQR)	-	0.93 (0.87, 1.00)$^+$	**0.92 (0.86, 0.99)$^{++}$**	-
Mendola et al 2017 (Per NOx IQR)	0.96 (0.62,1.49)^	0.85 (0.66, 1.09)^^	-	-

$^+$ IQR = 14.0 µg/m^3; $^{++}$IQR= 14.4 µg/m^3

^ IQR =24.2 ppb; ^^IQR 30.2ppb – NB Mendola estimates are for NOx

We highlight 4 studies (Smith et al, 2020; Yang et al, 2018; Zang et al, 2019; Mendola et al, 2017) reporting associations of NO_2 or NO_x with stillbirth (whole pregnancy: 3 studies, trimester 1: 4 studies, trimester 2: 3 studies, trimester 3: 2 studies).

All of the 3 studies (Yang et al, 2018; Zang et al, 2019; Mendola et al, 2017) reporting associations of stillbirth with Whole Pregnancy NO_2 or NO_x exposure suggested no association was present with effect estimates close to null. Similarly, both reported effect estimates for the association of stillbirth with trimester 3 NO_2 exposure were close to null (Table 18).

Three of the 4 studies (Smith et al, 2020; Yang et al, 2018; Mendola et al, 2017) reporting associations of stillbirth with trimester 1 showed negative associations with none of them reaching conventional levels of statistical significance and the remaining study (Zang et al, 2019) reported estimates close to the null.

One large well conducted study from London UK (Smith et al, 2020), which included a proxy area level adjustment for tobacco consumption, showed statistically significant negative associations for trimester 2 (Table 18).

Conclusion regarding evidence for NO_2 and stillbirth from key individual studies not included in the core meta-analyses

The key highlighted studies do not suggest an association (section 2.1.4) of stillbirth with exposure to NO_2 in whole pregnancy or any specific trimester.

OVERALL CONCLUSION for NO_2 and Stillbirth from the epidemiological evidence

We conclude that evidence from the meta-analyses (limited evidence) and key highlighted studies does not support an association (section 2.1.3) between stillbirth and NO_2 exposure. We nonetheless recommend an up-to-date meta-analysis to synthesise all the available evidence more succinctly, but from the evidence base we have considered it is unlikely to alter this conclusion.

O_3 and Stillbirth

Table 19. The Siddika meta-analysis effect estimates and those from the 5 key studies of interest for O_3 and Stillbirth

Paper (Increment)	OR (95% CI)			
	WP	T1	T2	T3
Siddika et al 2016 (Per 10ppb)	1.00 (0.97, 1.03)	1.00 (0.98, 1.02)	0.99 (0.94, 1.04)	1.01 (0.97, 1.06)
Yang et al 2018 (Per 10 µg/m^3) from Table 1 Appendix A	0.99 (0.95, 1.03)	0.99 (0.95, 1.03)	1.01 (0.98, 1.04)	0.96 (0.92, 1.000)
Zang et al 2019* (Per 10 µg/m^3)	1.03 (0.97, 1.09)	**1.05 (1.01, 1.09)**	1.01 (0.96, 1.06)	**1.04 (1.00, 1.08)**
Smith et al 2020* (Per IQR)	-	**1.17 (1.07, 1.27)$^+$**	**1.20 (1.09, 1.32)$^{++}$**	-
Mendola et al 2017* (Per IQR)	**1.39 (1.05,1.84)^**	**1.18 (1.00, 1.39)^^**	-	-
Rammah et al 2019b** (Per IQR)	**1.09 (1.01,1.18)$^\$$**			

+ IQR = 17.6 µg/m^3; ++IQR= 14.4 µg/m^3

^ IQR =7.8ppb; ^^IQR 12.3ppb

$ 3.6ppb – effect estimate based on survival analysis with time varying exposure

*estimates adjusted for season of conception

** authors explored adjustment of season of conception but report no change in estimate and therefore not included in final model

Two studies conducted in China reported significant associations of some trimester specific associations of stillbirth with O_3 but one of these was a negative association (Yang et al, 2018; for trimester 3) and the others were positive (adverse) (Zang et al, 2019; for trimester 1 and trimester 3) (Table 19).

Two studies conducted in USA, both of which adjusted for maternal smoking, reported statistically significant positive (adverse) associations for stillbirth with exposure in Whole Pregnancy, with Mendola et al (2017) also reporting a statistically significant association for trimester 1 (Table 19). Both studies also adjusted for mean temperature. One large, well conducted study from London UK (Smith et al, 2020), which included a proxy area level adjustment for tobacco consumption, although no adjustment for temperature, showed positive (adverse) associations of stillbirth with O_3 exposure in both trimester 1 and trimester 2. In the London study estimates of association were slightly stronger for 'asphyxia-related stillbirth' compared to 'congenital malformation related stillbirth'.

Conclusion regarding evidence for O_3 and stillbirth from key individual studies not included in the core meta-analyses

The key highlighted studies provide inconsistent evidence (section 2.1.4) for an association between stillbirth and O_3 exposure. However, we note that the well conducted studies from the USA and the UK provide some evidence of an association of stillbirth with O_3 exposure in pregnancy, although the small number of studies preclude robust evaluation of whether this is related to a trimester specific exposure or to Whole Pregnancy exposure.

OVERALL CONCLUSION for O_3 and Stillbirth from the epidemiological evidence

We conclude the limited evidence from the meta-analysis available does not support an association (section 2.1.3) between stillbirth and O_3 exposure. Examination of key highlighted studies points to the need for an up-to-date meta-analysis to synthesise all the available evidence, although it may, like our narrative review, be limited by the small number of studies and the limitations this would place on interpretation of a formal meta-analysis.

2.3 Conclusions from epidemiological studies

The box below presents summaries of the overall conclusions from examination of the epidemiological evidence for each pollutant-outcome pair.

Conclusions from epidemiological studies for each pollutant-outcome pair

Overall conclusion for $PM_{2.5}$ and PTB

Based on the meta-analyses available there is weak evidence for an association between PTB and $PM_{2.5}$ exposure. Examination of key highlighted studies points to the need for an up-to-date meta-analysis to synthesise all the available evidence.

Overall conclusion for PM_{10} and PTB

Based on the meta-analyses available there is weak evidence for an association between PTB and PM_{10} exposure. Examination of key highlighted studies points to the need for an up-to-date meta-analysis to synthesise all the available evidence.

Overall conclusion for NO_2 and PTB

Evidence from the meta-analyses available does not support an association between PTB and NO_2 exposure. Examination of key highlighted studies points to the need for an up-to-date meta-analysis to synthesise all the available evidence.

Overall conclusion for O_3 and PTB

Based on limited evidence from the meta-analyses available there is weak evidence for an association between PTB and O_3 exposure. Examination of key highlighted studies points to the need for an up-to-date meta-analysis to synthesise all the available evidence.

Overall conclusion for $PM_{2.5}$ and TLBW

Based on limited evidence from the meta-analyses available there is weak evidence for an association between TLBW and $PM_{2.5}$ exposure. Examination of key highlighted studies points to the need for an up-to-date meta-analysis to synthesise all the available evidence.

Overall conclusion for PM_{10} and TLBW

Based on the meta-analyses available there is weak evidence for an association between TLBW and PM_{10} exposure. Examination of key highlighted studies points to the need for an up-to-date meta-analysis to synthesise all the available evidence.

Overall conclusion for NO_2 and TLBW

Evidence from the meta-analyses available and key highlighted studies does not support an association between TLBW and NO_2 exposure. We nonetheless recommend an up to date meta-analysis to synthesise all the available evidence.

Overall conclusion for O_3 and TLBW

Evidence from the meta-analyses available (based on limited evidence) and key highlighted studies does not support an association between TLBW and O_3 exposure. We nonetheless recommend an up to date meta-analysis to synthesise all the available evidence.

Overall conclusion for $PM_{2.5}$ and Stillbirth

Limited evidence from the meta-analysis available does not support an association between stillbirth and $PM_{2.5}$ exposure. Examination of key highlighted studies points to the need for an up-to-date meta-analysis to synthesise all the available evidence.

Overall conclusion for PM_{10} and Stillbirth

Limited evidence from the meta-analysis available does not support an association between stillbirth and PM_{10} exposure. Examination of key highlighted studies points to the need for an up-to-date meta-analysis to synthesise all the available evidence more succinctly, but from the evidence base we have considered it is unlikely to alter this conclusion.

Overall conclusion for NO_2 and Stillbirth

Evidence from the meta-analysis available (limited evidence) and key highlighted studies does not support an association between stillbirth and NO_2 exposure. We nonetheless recommend an up to date meta-analysis to synthesise all the available evidence more succinctly, but from the evidence base we have considered it is unlikely to alter this conclusion.

Overall conclusion for O_3 and Stillbirth

Limited evidence from the meta-analysis available does not support an association between stillbirth and O_3 exposure. Examination of key highlighted studies points to the need for an up-to-date meta-analysis to synthesise all the available evidence, although it may, like our narrative review, be limited by the small number of studies and the limitations this would place on interpretation of a formal meta-analysis.

For context and interpretation of these conclusions, please also refer to the definition of the strength of evidence used in this report, as defined in section 2.1.3. A conclusion of weak evidence for an association from the meta-analyses, is defined (in section 2.1.3) as where the evidence from different meta-analyses (or different exposure windows within the same meta-analysis) is inconsistent (either in direction or statistical significance), but there is evidence from some portion of the meta-analyses indicating an association and/or associations observed in the meta-

analyses are subject to considerable heterogeneity in direction in forest plots and/or considerable potential for residual confounding. Statements on the strength of the evidence are distinct from any qualifying statements regarding the sufficiency of evidence. Conclusions are qualified as being based on limited evidence, where these are based on meta-analyses analysing a small number of studies (around 5 studies or fewer, as an approximate guide) and/or studies of limited quality or size, and/or where the meta-analyses may not have fully captured the evidence base available at the time.

Chapter 3
Mechanistic evidence

3.1 Approach

Relevant mechanistic evidence was identified using a similar approach to the epidemiology working group (see section 2.1), in that the KCL scoping review (Fussell and Walton, 2018) was used as a basis and was supplemented by an updated literature search, using revised search terms based on the KCL scoping review (Appendix 4).

The KCL scoping review summarised mechanistic evidence from key studies that were published prior to 2018. The updated search identified mechanistic studies published between January 2018 and March 2019. The search was conducted using PubMed, and the search terms are provided in Appendix 4. In addition, alerts from PubMed from this search were set up, and collected evidence continued to be collated from March 2019 to June 2021. As part of the search process, we also checked the references of recently published reviews.

Potentially relevant studies were first categorised into 'animal', 'human' or 'in vitro' mechanistic studies. Information on the data collected (for example smoking status via questionnaires and biological sample type), sample analytical methods, statistical analysis, quality controls, air pollution exposure assessment methods and exposure window were extracted from all potentially relevant studies. For 'human' mechanistic studies, further information including the study setting, population details and exclusions, cohort details, confounder information and exposure assessment (for example, personal monitoring) were also extracted to consider the quality of the studies. For 'animal' mechanistic studies, information on the species and number of animals was also extracted.

The mechanistic evidence was quality assessed using the Texas Commission Environmental Quality guidelines (TCEQ, 2017). Each of the relevant studies was grouped by mechanism and textually summarised, making specific comment on the following: the pollutant(s) studied; the size of the study population or number of samples; the outcomes studied; the biomarkers identified; and notes about any adjustment for potential confounders (Appendix 13).

3.2 Identified literature

Studies that exposed human and animal cells to air pollution in-vitro were excluded as were animal studies using an oral or injection method of exposure rather than

inhalation. Of the identified literature, 53 publications were included in the review, 42 were mechanistic studies from studies in humans and 11 were animal studies.

However, it was decided that the emphasis of the summary of evidence would be on 'human' mechanistic studies since it is recognised that animal studies are less useful for understanding the effects of air pollution on adverse birth outcomes in humans. This is because of differences between humans and most laboratory mammals such as rodents with respect to placentation, gestational sac, anatomy, fetal and maternal circulations, fetal/placental weight ratios, length of embryogenesis and organogenesis phases and multiple gestation pregnancy (litter size 8 to 18 pups), such that they are not comparable to the human uterine environment in terms of diffusion of chemicals/drugs/toxins and effect on feto-placental development.

Most of the human mechanistic studies identified involved measurements of biomarkers in participants in cohort studies. The characteristics of some of the high-profile named birth cohorts (referred to here as "key" cohorts) used in these studies are provided in Table 20. Table 21 provides the characteristics of other birth cohort studies that provide human mechanistic data.

Table 20. The characteristics of key birth cohort studies that provide human mechanistic data

Cohort name	Study setting and period	Approximate total number of women enrolled*	Exposure assessment	Studies in identified literature using data from these cohorts
The ENVIRonmental influence ON early AGEing (ENVIRONAGE) **Study profile provided in Janssen et al (2017a)**	Genk, Belgium. 2010 to present.	1,080 (as of February 2015)	Daily air pollutant concentrations of $PM_{2.5}$, PM_{10}, NO_2, and BC were calculated at the mother's residential address by spatial interpolation, using data from Belgium's air quality monitoring networks, accounting for land-cover data from satellite images. Air pollution estimates were calculated during the entire pregnancy (and specific exposure windows of interest during gestation) and at preconception. Maternal smoking was assessed via a questionnaire during pregnancy.	Bové et al 2019 Clemente et al 2016 Grevendonk et al 2016 Gruzieva et al 2017 Janssen et al 2012 Janssen et al 2013 Janssen et al 2015 Janssen et al 2017b Robinson et al 2018 Saenen et al 2016 Saenen et al 2017 Tsamou et al 2018 Vos et al 2021
Early Autism Risk Longitudinal Investigation (EARLI) **Study profile provided in Newschaffer et al (2012)**	US Network sites: 1) Drexel/Children's Hospital of Philadelphia, 2) Johns Hopkins/Kennedy Krieger Institute, 3) University of California (UC) Davis, 4) Kaiser Permanente Northern California 2009-2012	232	Maternal exposure was modelled at the home address using monthly air quality data: NO_2, O_3, from monitoring stations located within 50 km of each residence, with spatial interpolation based on inverse distance-squared weighting of data from up to 4 closest stations. Maternal smoking was assessed via questionnaire.	Ladd-Acosta et al 2019 Gruzieva et al 2017

Cohort name	Study setting and period	Approximate total number of women enrolled*	Exposure assessment	Studies in identified literature using data from these cohorts
GENERATION R **Study profile provided in Jaddoe et al (2010)**	Rotterdam, The Netherlands 2001-2005	9,978	Exposure of women to PM_{10} and NO_2 for all addresses in the study area was gathered using 3 Dutch national standard methods for air quality modelling. Hourly concentrations of PM_{10} and NO_2 were derived and were adjusted for background concentrations, using hourly air pollution measurements from 3 continuous monitoring stations to give exposure for 1 week, 2 weeks and 4 weeks and total pregnancy. Maternal smoking before and during pregnancy in each trimester was assessed by questionnaires.	Gruzieva et al 2017 van den Hooven et al 2012a van den Hooven et al 2012b van den Hooven et al 2012c
HEALTHY START STUDY **Study profile Provided in Starling et al (2020)**	Colorado, USA 2009-2014	1,125	Average ambient $PM_{2.5}$ and O_3 concentrations were estimated using measurements obtained from stationary monitors, located within 50 km of at least one study participant, in the US Environmental Protection Agency Air Quality System. Participants reported their place of residence via questionnaire completed at the time of enrollment. Maternal smoking during pregnancy was assessed via self-report at 2 study visits during pregnancy and again shortly after delivery	Friedman et al 2021
INMA **Study profile provided in Guxens et al (2012)**	Spain (6 study areas: Ribera d'Ebre, Menorca, Granada, Valencia, Sabadell, Asturias and Gipuzkoa) 1997-2008.	3,944	Ambient concentrations of NO_2 were measured with the aid of passive samplers distributed outside across the study areas according to geographic criteria. Samplers remained exposed during various 7-day sampling periods through pregnancy.	Clemente et al 2016 Gruzieva et al 2017 Robinson et al 2018

Cohort name	Study setting and period	Approximate total number of women enrolled*	Exposure assessment	Studies in identified literature using data from these cohorts
			Temporally adjusted land use regression (LUR) models were used to predict NO_2 levels at women's residential addresses and taking into account residential changes if women lived at least 2 months of pregnancy in the residence. To calculate individual exposure during pregnancy, annual average NO_2 estimates from the LUR were temporally adjusted using serial records from the network of monitoring stations covering each of the 4 INMA study areas. Data on parental smoking habits was collected by questionnaire.	
Rhea **Study profile provided in Chatzi et al (2009)**	Crete, Greece 2007 - 2009	1,610	LUR	Gruzieva et al 2017 Robinson et al 2018
Piccolipiù **Study profile provided in Farchi et al (2014)**	Five Italian cities (Florence, Rome, Trieste, Turin and Viareggio) October 2011 - 2014	3,000 children	LUR	Gruzieva et al 2017 Robinson et al 2018
Southern California Children's Health Study (CHS) **Study profile provided in Urman et al (2014)**	Eight southern Californian communities (Anaheim, Glendora, Long Beach, Mira Loma, Riverside, San Dimas, Santa Barbara and Upland)	5,341	Exposure to regional ozone (O_3), NO_2, particulate matter with aerodynamic diameter less than 10 µm (PM_{10}) and 2.5 µm ($PM_{2.5}$) was measured continuously at a central monitoring location in each community from cohort recruitment (2002) to the recording of lung function tests (2007).	Gruzieva et al 2017

Cohort name	Study setting and period	Approximate total number of women enrolled*	Exposure assessment	Studies in identified literature using data from these cohorts
	1992 - present			
PRogramming of Intergenerational Stress Mechanisms (PRISM) **Study profile provided in Brunst et al (2014)**	Boston, MA, USA 2012 - 2014	592	Hybrid land use regression and satellite- based model	Gruzieva et al 2017
Project Viva **https://www.hms.harvard.edu/viva/project-viva-history.html**	Eastern Massachusetts USA 1999-2003	2,128 (as of 2002)	Hybrid land use regression and satellite-based model	Gruzieva et al 2017

* At the time of report writing.

Table 21. The characteristics of other birth cohort studies that provide human mechanistic data

Study	Location	Cohort Name OR Brief Description	Exposure Assessment
Ambroz et al 2016	Czech Republic: Karvina and Ceske Budejovic	Case control study of non-smoking women and newborns born at term in summer and winter in different regions	$PM_{2.5}$ was collected by an air sampler in 4 localities of the Czech Republic. Sampling was conducted for 24hr each day for 30-35 days in the winter season of 2008/09.
Brunst et al 2018	Boston, Massachusetts	Single cohort from local antenatal clinic at the Beth Israel Deaconess Medical Center (BIDMC)	Moderate Resolution Imaging Spectroradiometer satellite-derived Aerosol Optical Depth measurements based on the Multi-Angle Implementation of Atmospheric Correction algorithm were used at a 1 km × 1 km grid spatial resolution to predict daily $PM_{2.5}$ levels across New England between March 2011 and March 2021.
Cai et al 2017	Wenzhou, China	Local cohort study at the Second Affiliated Hospital of Wenzhou Medical College	Daily air pollution data for PM_{10}, SO_2 and NO_2 were averaged for 4 fixed-site monitoring stations and averaged for each of the 3 trimesters. Air pollution data for 2010 in addition to the years in which eligible births occurred, 2011–2013, were used to estimate exposure for births in 2011.
Carvalho et al 2016	São Paulo, Brazil	Procriar cohort	Passive personal samplers (n=366) were used during each trimester of pregnancy to measure NO_2 and O_3 between April 2011 and December 2013.
Contreras et al 2018	Los Angeles, California	Behaviour in pregnancy cohort	Exposures to traffic-related air pollutants was estimated during pregnancy and were assessed at participant's residential address between 1993 and 1996 using land use regression–derived NO_2 and CALINE4-derived NOx exposure.
He et al 2018	Zhengzhou City, China	Cohort from Houzhai Center Hospital	24-h average concentrations of PM_{10}, SO_2 and NO_2 were obtained from 2 stations and averaged over 4 exposure windows (WP, T1, T2 and T3). The permanent residences of participants were 2.9 km apart from the nearest monitoring station on average.

Study	Location	Cohort Name OR Brief Description	Exposure Assessment
Herbstman et al 2012	New York City, USA	Columbia Center for Children's Environmental Health (CCCEH)	Prenatal personal PAH assessment was carried out by personal monitoring (n=263) during the third trimester of pregnancy. Women were asked to wear a small backpack containing a personal monitor during the daytime hours for 2 consecutive days and to place the monitor near the bed at night.
Kingsley et al 2016	Rhode Island, USA	Women delivering at Women and Infants Hospital of Rhode Island (WIHRI)	Daily $PM_{2.5}$ levels, during the study period (2002-2012), were estimated at each maternal residential address using a hybrid of land-use regression and satellite remote sensing (1km grid). Daily black carbon (2004-2011) was estimated at each address using an extended version of a validated spatial-temporal land-use regression model.
Lavigne et al 2018	Ontario, Canada	Retrospective cohort in 31 cities between 1 January 2006 and 31 March 2012. Birth data were obtained from the Better Outcomes Registry & Network (BORN)	All participants residing within 5 km of a given monitoring site during pregnancy were assigned the daily environmental data corresponding to the different periods of their pregnancy.
Lee et al 2011	Allegheny County, Pennsylvania	Prenatal Exposures and Pre-eclampsia Prevention study	Space-time ordinary Kriging Interpolation Method was used to estimate daily air pollution concentrations ($PM_{2.5}$, PM_{10}, NO_2, SO_2, O_3 and CO), from 1996-2001, at the zip code level by calculating the average estimated zip code concentration for each pollutant based on all grid centroid values that fall within a zip code.
Liu X et al 2019	Guangzhou, China	Birth cohort	The 24-h average concentrations of $PM_{2.5}$, PM_{10}, SO_2, NO_2, CO, and 8-h average concentration of O_3 were collected from 111 monitoring site stations during 1 January 2014 – 31 December 2015.
Liu et al 2021	United Kingdom	Royal London Hospital	Participants' mean PM_{10} and $PM_{2.5}$ exposure 12 months prior to the delivery date was modelled for their home addresses at a resolution of 20 m × 20 m.
Maciel-Ruiz et al 2019	Mexico City	Pregnant women between the ages of 19–35 at the Unidad Médica del Hospital de Alta Especialidad en Ginecología y Obstetricia No. 3, Centro Médico Nacional 'La Raza'	The daily records of atmospheric $PM_{2.5}$, PM_{10}, and O_3 levels were obtained, for the years 2014 to 2015, from which monthly medians were estimated.

Study	Location	Cohort Name OR Brief Description	Exposure Assessment
Maghbolli et al 2018	Tehran, Iran	Nested case-control study of pregnant healthy women attending prenatal care clinics	Regional background daily levels of PM_{10} and $PM_{2.5}$ were calculated, from 21 monitoring stations in 2015, for each mother's home address from the date of the last menstrual period (LMP) to delivery, from 9 stations in 4×4 km grids in 2 selected regions .
Nachman et al 2016	Boston, USA	Birth Cohort	Individual exposures to $PM_{2.5}$ were estimated by assigning each subject, from 1999-2012, to the closest monitor by their residential address, reported at the time of study recruitment. No limits were placed on the distance between subjects and monitors. Only data from monitors that had at least one measurement per week for 75% of the study period were included. Measurements were recorded every 3 days for the monitors in the study except for a few short periods during which measurements were recorded daily.
Perera et al 1999	Krakow and Limanowa	70 mother/newborn pairs and 90 mother/newborn pairs retrospectively	Daily ambient monitoring data for Krakow (1990-1992) were provided from 5 monitoring stations. Ambient particulate data were available for 69/70 subjects from Krakow. Each woman's exposure to ambient particulates, was estimated by averaging PM_{10} measurements at the monitoring station closest to her residence for each of the past 2 years and the month prior to her delivery date.
Rosa et al 2017	Mexico	Programming Research in Obesity, Growth, Environment and Social Stressors (PROGRESS) Study	Daily exposure to $PM_{2.5}$ were estimated for each cohort participant during pregnancy at a 1km × 1 km spatial resolution, for the years 2004-2014.
Rossner et al 2011	Two districts in the Czech Republic (Teplice and Prachatice)	Case–control study conducted between 1994 and 1999	Concentrations of ambient air pollutants ($PM_{2.5}$ and c-PAHs) in both districts were continuously measured using a Versatile Air Pollution Sampler (VAPS)
Suh et al 2008	Seoul, Korea	Case–control study performed at Ewha Womans University Hospital between January 2003 and March 2007	27 monitoring stations measured 24-h averages of PM_{10} between January 2003 and March 2007. The level of PM_{10} exposure was evaluated for each participant in accordance with the level at the monitoring station closest to the participant's home address considering pregnancy period.

Study	Location	Cohort Name OR Brief Description	Exposure Assessment
Suter et al 2019	Harris County, Texas	Peri-Bank, population-based biobank	PAHs were analysed using gas chromatography coupled mass spectrometry a single reaction monitoring strategy with electron impact ionization. PAH levels were estimated using a standard curve.
Tang et al 2006	Tongliang, China	Chinese women who gave birth at the Tongliang County Hospital, Tongliang Traditional Chinese Medicine Hospital, Tongliang Maternal Children Health Hospital, and Bachuan Hospital between 4 March 2002 and 19 June 2002	The estimated duration of in utero exposure to the power plant operation was based on the number of months of pregnancy that overlapped with the period of the plant operation (1 December 2001 to 31 May 2002). All births occurred between 4 March 2002 and 19 June 2002. Thus, the duration of prenatal exposure to power plant emissions ranged from 3.13 to 6 months for individual children.
Tang et al 2014	Tongliang, China	2002 cohort: Women who gave birth between 4 March 2002 and 19 June 2002 at 4 hospitals: the Tongliang County Hospital, the Traditional Chinese Medicine Hospital, the Tongliang Maternal Children's Health Hospital, and the Bachuan Hospital. 2005 cohort: Women who gave birth between 2 March 2005 and 23 May 2005 at 4 hospitals: at the same hospitals as the 2002 cohort.	Integrated 72-h PAH samples were collected at 3 sites in Tongliang between March 2002 and February 2003 and between March 2005 and February 2006. Data on BaP from the months of December to May, the active period of the power plant in 2002, were used to calculate the averages from 3 monitoring stations.
Topinka et al 1997	Czech Republic: Teplice and Prachatice districts	Each woman involved in the study was systematically selected from the whole cohort of mothers giving birth to living baby during the relevant season intervals. Summer: June and September 1994 and Winter: November 1994 and January 1995	Ambient air SO_2 and NO_x concentrations were continuously measured in the stationary monitoring stations in both districts between June and September 1994 and between November 1994 and January 1995.

Study	Location	Cohort Name OR Brief Description	Exposure Assessment
Zhao et al 2021	Taiyuan, China	A birth cohort of pregnant women who came to the First Affiliated Hospital of Shanxi Medical University in Taiyuan, China for delivery between March 2012 and November 2013 with gestational age ≥20 weeks, who had no mental illness, and who were aged 18 years or older	Chinese Air Quality Index (AQI), calculated based on the levels of SO_2, NO_2, PM_{10}, $PM_{2.5}$, CO, and O_3, measured at the monitoring stations throughout each city

3.3 Evidence from human mechanistic studies

The studies, summarised in this section, plus in tabulated form in Appendix 13, have investigated the effects associated with exposure to particulate matter (PM) (various size fractions), nitrogen dioxide (NO_2), nitrogen oxides (NO_x), ozone (O_3), black carbon (BC), polycyclic aromatic hydrocarbons (PAHs) and diesel engine exhaust (DE). Placental tissue, maternal, cord and/or newborn blood samples from participants in epidemiological studies have been analysed to investigate oxidative stress, inflammation, DNA damage, epigenetic modifications, thyroid hormone synthesis alteration and metabolomic changes. In addition, studies have researched placental circulations and haemodynamics. Prior to summarising the data originating from these studies, (sections 3.3.2 to 3.3.8) we consider how pollutants might reach the placenta and discuss the studies providing evidence for the translocation of PM to the placenta (section 3.3.1).

3.3.1 Does inhaled particulate matter cross the placental membrane?

When the utero-placental circulation is fully established and the umbilico-placental circulation fully functional at the end of the first trimester, materno-fetal exchanges increase rapidly and expose the placenta and possibly the fetus to PM and surface adhered constituents such as metals and PAHs. While pollutants such as volatile organic compounds (VOCs) are likely to be transported by the trophoblast via simple diffusion, inhaled PM and adhered metals, if transported to the placenta via the systematic circulation, are likely to enter the villous mesenchyme via pinocytosis (Figure 11). The role of mechanisms described in section 1.4.2.2 such as pinocytosis remains to be determined.

Liu et al (2021) reported light and electron microscopy (EM) analysis of macrophages extracted from 5 placentas after elective caesarean section. Of the 3,500 placental macrophages analysed, 60 cells contained a total of 72 separate black particles under light microscopy. Under EM, macrophages from 2 placentas contained black nanoparticle (NP)-like structures within macrophage phagolysosomes.

Bové et al (2019) examined stored placentas from 20 term placentas from mothers enrolled within the ENVIRONAGE birth cohort (described in Table 20) and 5 placentas from preterm deliveries. Term placentas came from n=10 mothers with high residential BC exposure and n=10 mothers with low residential BC exposure. NP-like structures were not observed directly, but by femtosecond pulsed laser illumination that causes black particles to emit ‘dots’ of white light which are then quantified. Particles were present in the 20 term placentas. Each 0.5 $\mu g/m^3$ increment in modelled exposure to BC at the home address during the whole of pregnancy was associated with an increase of 0.45×10^4 BC particles/mm^3. The 5 placentas from PTBs also emitted white dots compatible with carbon particles under laser light illumination. Bové et al (2019), also provided evidence of NP in placental tissue by EM (see Figure 11). They concluded that their study provides ‘compelling evidence for the presence of BC particles originating from ambient air pollution in human placenta and suggests the direct fetal exposure to those particles during the

most susceptible period of life.' We note that a later study from the same group of researchers (Bongaerts et al 2022) found evidence of the presence of BC particles in cord blood and a strong correlation between residential exposure to BC and the particle loads in maternal blood, term placenta and cord blood. BC particles were also found in first and second trimester fetal tissues: liver, lung and brain.

Valentino et al (2016) provides additional evidence from an animal model that inhaled PM reach placental tissue. This study exposed pregnant New-Zealand white female rabbits to either diluted DE (1 mg/m^3) or clean air for 2 hours per day, 5 days per week, for 20 days over a 31-day gestation period. Black structures compatible with PM from DE and classified by the researchers as 'NP-like' were found in maternal erythrocytes and plasma, and also i) along the microvilli of placental syncytial membranes, ii) scattered in the nucleus of some trophoblastic cells, iii) in the lumen of fetal vessel, and iv) in endocytotic vesicles in fetal erythrocytes. Since no NP-like structures were found in any control animals, the researchers concluded that 'NP from DE are able to cross the placenta and reach the fetal circulation'.

Overall, these studies suggest that inhaled NP do reach the placenta and the fetal tissues. But it remains unclear whether the relatively small amount of NP observed in placental tissue is sufficient to affect placental growth and/or function.

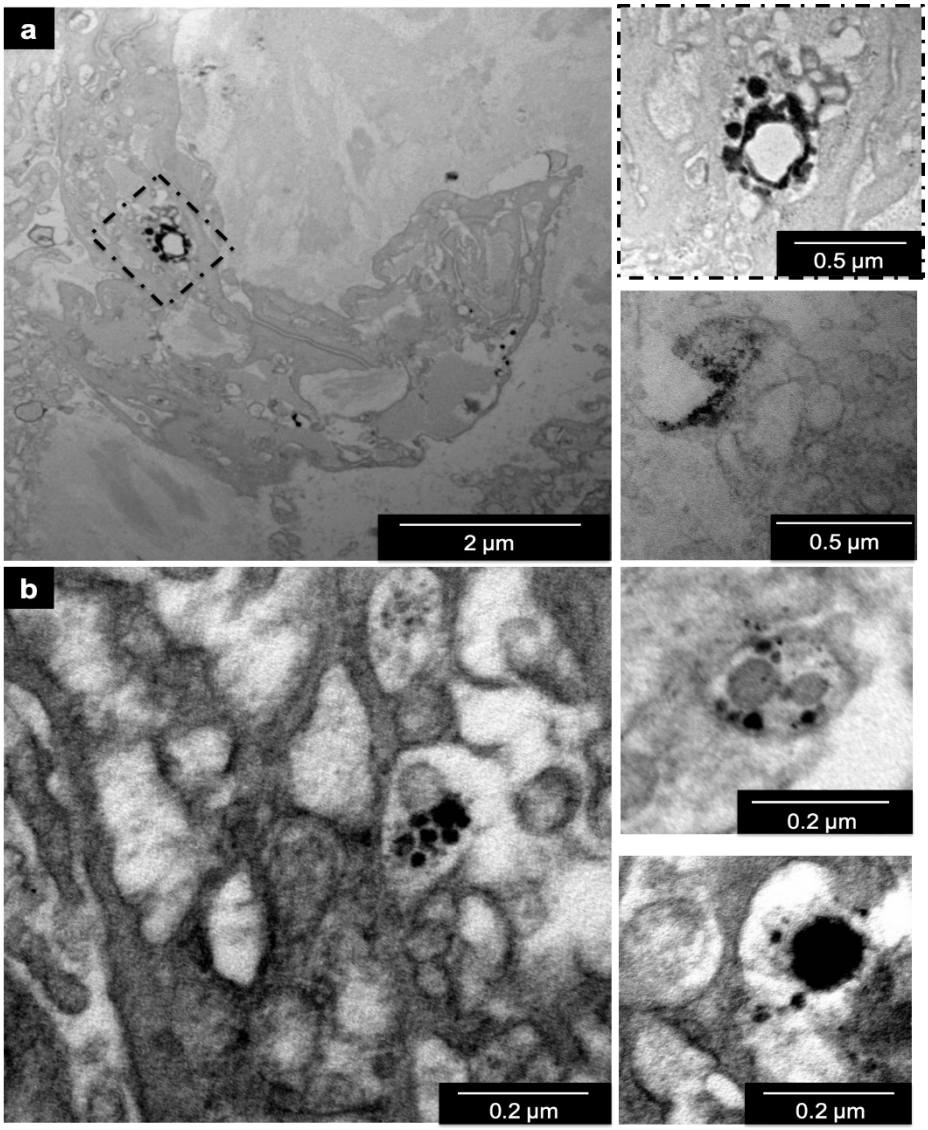


Figure 11. Transmission electron microscopy imaging of BC particles inside placenta thin sections of 2 (a-b) high exposed mothers. Taken from Bové et al (2019)

3.3.2 Oxidative stress

Altered physiological processes and increased energy expenditure during a standard pregnancy are associated with increased levels of oxidative stress (Fialova et al, 2006). In response, several antioxidants increase in concentration to protect the mother and the fetus from undue harm; glutathione (GSH) in particular is thought to have an important role in detoxifying reactive oxygen species before they could impact the developing child (Mistry and Williams, 2011).

Factors that can increase oxidative stress in uterine and placental tissue include maternal active and passive smoking, exposure to other forms of environmental pollution (chemical, radiation), maternal diabetes, hypertension, preeclampsia and the mode and length of delivery (Cindrova-Davies et al, 2007). The persuasive evidence that oxidative stress is a critically important cause and consequence of cardiorespiratory diseases elicited by inhaled air pollutants offers biological plausibility that any association between exposure to air pollution and changes in the

placenta and fetus are mediated, in part, by this key pathway. Several approaches have been undertaken to investigate an exacerbation in oxidative stress in response to ambient air pollution exposure during gestation. These include studies on biomarkers and oxidative potential as well as ones focusing on genetic susceptibility due to polymorphisms in antioxidant genes.

Biomarkers of oxidative stress

Biomarkers of oxidative stress reported in studies include: 3-nitrotyrosine (3-NT), 8-oxo-7,8-dihydro-2'-deoxyguanosine (8-OHdG or 8-oxodG), F2-isoprostanes and mitochondrial DNA (mtDNA):

- 3-NT: the stable product of tyrosine nitration with the ROS peroxynitrite. Nitration of placental proteins is evident in standard pregnancies but, at higher levels, it may elicit pronounced perturbations of placental function, as has been reported in placental vessels in states of preeclampsia and gestational diabetes (Bosco et al, 2012; Lyall et al, 1998). The presence of nitrosative stress has also been linked with diminished vascular reactivity of the fetal placental circulation (Myatt, 2010).

- 8-OHdG: a DNA base modification and the major DNA lesion induced by oxidative stress. It is therefore a marker of oxidative DNA damage.

- F2-isoprostanes (for example, 8-iso-prostaglandin F2α (8-iso-PGF2α) & 15-F2t-isoprostane (15-F2t-IsoP)): produced by the non-enzymatic peroxidation of arachidonic acid in membrane phospholipids (Basu, 2010)

- mtDNA: In regulating energy, mitochondria within the cells of the placenta are essential in the nourishment, growth, and development of the fetus. Mitochondria are also the major intracellular sources and primary targets of ROS. Moreover, compared to nuclear DNA, mtDNA lacks a protective chromatin structure, histones, and introns and possesses fewer efficient repair mechanisms (Lee and Wei, 2000, Payne et al, 2013;). mtDNA are therefore particularly vulnerable to ROS-induced damage and have a high mutation rate (Linnane et al. 1989). Compensatory mechanisms result in a change in mtDNA copy number (a change in mtDNA content) – an established marker of mitochondrial damage and dysfunction (Hou et al, 2010).

3-NTp

Saenen et al (2016) studied 330 mother-newborn pairs from Genk, Belgium to investigate the association between placental 3-NT (3-NTp) and measures of exposure to ambient air pollutants including $PM_{2.5}$, BC, and NO_2 during various time windows of pregnancy. 3-NT levels in the placenta were positively associated with $PM_{2.5}$ levels during gestation and they found that associations were driven by exposures incurred during the first and second trimesters of gestation. For BC, the association with 3-NTp was mainly driven by the first-trimester exposure window. No association was found between placental 3-NT and NO_2 exposure during pregnancy. Furthermore, the results for $PM_{2.5}$ remained the same after additional adjustment for BC or NO_2 exposure. However, no significant associations were found between 3-NTp and birth weight or birth length.

8-OHdG and isoprostanes

Grevendonk et al (2016) explored the association between PM_{10} and $PM_{2.5}$ exposure in each trimester of pregnancy and mitochondrial 8-OHdG in maternal (n=224) and fetal (cord; n=293) blood in subjects from Genk, Belgium. Increasing PM_{10} and $PM_{2.5}$ exposures during the whole pregnancy and in the third trimester were significantly associated with an increase in 8-OHdG in maternal blood. In cord blood, concentrations of 8-OHdG were significantly associated with maternal PM_{10} exposure in the first and second trimester but not during whole pregnancy. The authors concluded that PM exposure increases oxidative stress in the mother and the fetus, although the evidence for an effect in fetal (cord) blood samples was weak. Neither PM exposure nor 8-OHdG levels in cord and maternal blood were associated with birth weight or SGA.

Ambroz et al (2016) studied oxidative DNA damage as well as lipid peroxidation 15-F2t-IsoP in the urine and blood from mothers and newborns from a clean (n=99/100) and polluted (n=70-74) district of the Czech Republic. The measurements were controlled for cotinine levels (a marker of active and passive smoking) in maternal urine. No significant differences in lipid peroxidation in mothers between localities were found and, unexpectedly, levels of 8-oxodG in mothers' samples were elevated in clean versus polluted district. For newborns, significant elevations were observed in 8-oxodG in samples derived from polluted district in the winter season only. In summer lipid peroxidation was significantly higher in the clean air samples whilst in the winter period no significant differences were evident. Results of multivariate regression analysis in newborns from polluted districts, showed $PM_{2.5}$ concentrations to be a significant predictor for 8-oxodG excretion and $PM_{2.5}$ and BaP (benzo[a]pyrene) concentrations to be a significant predictor for 15-F2t-IsoP levels. In mothers $PM_{2.5}$ concentrations were shown to be a significant predictor of 8-oxodG levels.

Rossner (2011) investigated associations between levels of 8-OHdG measured in placental DNA, exposure to BaP or $PM_{2.5}$, averaged per month of the pregnancy and for the whole pregnancy, and pregnancy outcomes (IUGR and TLBW). The study population constituted 891 newborns born in the Czech Republic in 2 districts with different concentrations of air pollution. No difference in 8-OHdG levels between the 2 districts were found. Using univariate analysis, 8-OHdG levels in placental DNA was not associated with BaP, whereas $PM_{2.5}$ exposure during the second month of pregnancy was significantly associated with 8-OHdG levels in placental DNA. Borderline statistically significant elevated 8-OHdG levels were found in the IUGR group compared to the non-IUGR group ($p=0.06$). Slightly elevated levels of 8-OHdG were found in the TLBW group compared to the non-TLBW group ($p<0.05$). Multivariate-adjusted analysis showed above-median levels of 8-OHdG were significantly associated with IUGR ($p=0.02$) and TLBW ($p=0.01$). Exposure to $PM_{2.5}$ was not associated with IUGR but in the first month of pregnancy was associated with the risk of TLBW. BaP exposure was associated with neither IUGR nor LBW.

8-iso-PGF2α

Maciel-Ruiz et al (2019) investigated the influence of $PM_{2.5}$, PM_{10} and ozone concentrations on 8-iso-PGF2α in maternal and cord blood. The study focused on

non-smoking, cotinine-negative pregnant women (n=106) who lived in in Mexico City's Metropolitan Area - a highly polluted area impacted by industry and vehicle emissions. Maternal and cord blood isoprostanes correlated significantly, whilst cord blood levels were significantly higher ($P < 0.0001$). A significant association between maternal 8-iso-PGF2α and ozone (but not $PM_{2.5}$ or PM_{10}) concentrations was observed only during the last month of pregnancy. No correlation was found between concentrations of cord blood 8-iso-PGF2α and ozone concentrations.

mtDNA

Janssen et al (2012) assessed the association between mtDNA content of placental tissue (n=174 samples) and maternal exposure to PM_{10} and distance to a major road. In a multivariate-adjusted analysis, a 10 $\mu g/m^3$ increase in PM_{10} exposure during the last month and third trimester of pregnancy was associated with decreases of 16.1% and 17.4% in placental mtDNA respectively. In addition, each doubling in residential distance to major roads was associated with a 4% increase in placental mtDNA content (p=0.03). No association was found between mtDNA content in cord blood (n=176 samples) and PM_{10} exposure.

Rosa et al (2017) investigated associations between exposure to $PM_{2.5}$ and mtDNA in cord blood from newborns of 456 mothers in Mexico. Significant associations were found between higher $PM_{2.5}$ exposure during late pregnancy (35-40 weeks) and lower mtDNA content in cord blood, after adjustments. When stratifying by sex, a significant association between increased $PM_{2.5}$ and lower cord blood mtDNA content was found among boys (at 37–40 weeks gestation), but there were no significant associations among girls.

Brunst et al (2018) investigated the associations between a joint exposure to prenatal $PM_{2.5}$ and maternal lifetime traumatic stressors, and mtDNA copy number (cn) in cord blood (n=147) and placenta (n=162) in a socio-demographically diverse pregnancy cohort in Boston, USA. They found that increased $PM_{2.5}$ exposure across pregnancy was associated with decreased mtDNAcn in cord blood. While in the same direction, an effect of $PM_{2.5}$ on placental mtDNAcn did not reach significance. $PM_{2.5}$ exposure between 25- and 40-weeks' gestation was significantly associated with increased placental mtDNAcn among boys of mothers reporting high trauma and decreased placental mtDNAcn among girls of mothers reporting low trauma.

Clemente et al (2016) reported a combined analysis from the ENVIRONAGE (n=550) and the INMA (n=376) cohort studies and investigated whether changes in placental mtDNA is a mediator in the association between NO_2 exposure during pregnancy and birth weight. Pooled estimates indicated that a 10 $\mu g/m^3$ increment in average NO_2 exposure during pregnancy was associated with a 4.9% decrease in placental mtDNA content and a 48g decrease in BW. Placental mtDNA content was positively and significantly associated with birth weight in both cohorts and in the pooled analysis. An interquartile range increase in mtDNA content was significantly associated with a 66g increase in birthweight in boys but there was no significant difference in girls.

Oxidative potential of PM

Lavigne et al (2018), in another attempt to mechanistically link exposure to birth outcomes, evaluated whether city-level differences in the oxidative potential (OP; the capacity of particles to directly cause damaging oxidative reactions) of $PM_{2.5}$ modifies the relationship between $PM_{2.5}$ mass concentrations and PTB, TLBW, and term birth weight during specific periods of pregnancy (daily exposures in the last 4 weeks before delivery, weekly exposures throughout pregnancy, each trimester and whole-pregnancy). In this retrospective cohort study of 196,171 singleton births that occurred in 31 cities in the province of Ontario, Canada, OP was measured as the ability of $PM_{2.5}$ filter extracts of given concentration to deplete the antioxidants glutathione (GSH; OP^{GSH}) and ascorbate (AA; OP^{AA}) in the simulated respiratory tract lining fluid. The term 'oxidative burden' described the product of $PM_{2.5}$ mass concentrations and OP.

In general, associations of $PM_{2.5}$ and GSH- and AA-related oxidative burden with TLBW and PTB, were not statistically significant for exposures during any time period. An IQR increase in $PM_{2.5}$ GSH-related oxidative burden was, however, more strongly associated with TLBW than an IQR increase in $PM_{2.5}$ mass concentration alone. When associations between IQR increases in $PM_{2.5}$ and the birth outcomes were stratified by quartiles of OP^{GSH} there was little evidence of effect modification overall. However, IQR increase (2.6 $\mu g/m^3$) in first trimester $PM_{2.5}$ was positively (adversely) associated with TLBW among women in the highest quartile of OP^{GSH} but not among those in the lowest quartile. $PM_{2.5}$ on the day of delivery was also associated with PTB among women in the highest quartile of OP^{GSH} but not in the lowest quartile. Ascorbate-related oxidative potential was not an effect modifier in models for TLBW or PTB.

Genetic susceptibility to oxidative stress

Further evidence for a mechanistic role of oxidative stress has come from studies assessing interactions between air pollution, PTB and genetic polymorphisms that increase vulnerability to oxidative stress and reduce metabolic detoxification. Frequently studied polymorphisms are in the glutathione S-transferase (GST) genes. These genes encode phase II metabolising enzymes, and biotransform and detoxify xenobiotics and endogenous compounds. Two polymorphic genes of this family, GSTM1 and GSTT1, are present as 'null alleles.' Null alleles do not produce the enzyme if the genotype is homozygous and may therefore increase vulnerability to oxidative stress.

Suh et al (2008), in a case-control study in Seoul of 117 women with preterm deliveries and 118 women with term deliveries, investigated the association between the risk of preterm delivery and each metabolic gene of glutathione S-transferases mu 1 (GSTM1), theta 1 (GSTT1) with exposure to PM_{10}. Exposure to high levels of PM_{10} (≥75th percentile) during the third trimester of pregnancy was associated with an increased risk for PTB, only in the presence of the GSTM1 null genotype.

Zhao et al (2021) investigated the interactive effects of maternal air pollution (assessed by the Chinese Air Quality Index (AQI)) and single nucleotide polymorphisms on risk of PTB among 436 women in Taiyuan, China. The study

found higher levels of maternal AQI exposure to be associated with a higher risk of PTB among women who carried polymorphisms in genes for glutathione peroxidase 4 (GPX4), glutaredoxin, endothelial growth factor A and interleukin alpha (IL_{α}). After adjustment for multiple comparisons, only single nucleotide polymorphisms GPX4-rs376102 and AQI interaction remained statistically significant.

Summary of evidence on oxidative stress in adverse birth outcomes following exposure to air pollution

A number of studies have investigated the effect of oxidative stress on placental and fetal tissue with the potential to adversely influence birth outcomes.

Biomarker studies have shown consistent associations between exposure of pregnant women to PM and markers of nitrosative/oxidative stress (3-NTp, 8-OHdG, mtDNA) in the placenta and/or cord blood (Brunst et al, 2018; Ambroz et al, 2016; Rosa et al, 2017; Rossner et al, 2011; Saenen et al, 2016; Grevendonk et al, 2016; Janssen et al, 2012). Fewer studies have looked for associations between these markers and NO_2/ PAHs, and these have yielded more inconsistent results (Ambroz et al, 2016; Rossner et al, 2011; Saenen et al, 2016; Clemente et al, 2016).

One study suggested that 10% of LBW associated with prenatal NO_2 exposure could be explained by reduced placental mtDNA content. Others that examined the combination of exposure, markers of oxidative stress and an adverse birth outcome do not provide evidence that an increased oxidative burden with increased exposure is associated with PTB or negative indices of fetal growth (Maciel-Ruiz et al, 2019; Rossner et al, 2011; Saenen et al, 2016; Grevendonk et al, 2016).

Studies focusing on genetic susceptibility due to polymorphisms in antioxidant genes (Suh et al, 2008; Zhao et al, 2021;), suggest an interaction between exposure to air pollution during pregnancy and genetic polymorphisms in oxidative stress genes for increased risk of adverse pregnancy endpoints (preterm delivery), suggesting evidence for a gene–environment interaction in pregnant women.

Possible effect modifications by $PM_{2.5}$ oxidative potential, examined in only one study, reported that between-city differences in GSH-related oxidative potential may modify the impact of $PM_{2.5}$ on the risk of term LBW and PTB (Lavigne et al, 2018).

3.3.3 Inflammation

Since there is strong evidence that associations between air pollution and cardiopulmonary diseases are mediated by inflammatory pathways, it is plausible that air pollution may influence the fetus via inflammation, either in the placenta, the fetus, or both.

Indeed, PTB is generally regarded as a syndrome resulting from multiple causes including inflammation (Goldenberg et al, 2008) through the early activation of cytokines such as IL-1β and tumour necrosis factor-α (TNF-α), which are otherwise

part of the body's preparatory steps for term deliveries (Engel et al, 2005; Keelan et al, 2003).

Furthermore, placental inflammation may affect its growth, development and function, which in turn can lead to fetal growth restriction (FGR) (Williams et al, 2000), whilst cytokines can cross the placental barrier and interfere with fetal development (Jonakait, 2007).

Nachman et al (2016) assessed the association between maternal exposure to relatively low concentrations of $PM_{2.5}$ (9.5 to 16.3 μg/m^3) and intrauterine inflammation (IUI) in 5,059 mother-infant pairs in the Boston Birth Cohort, predominantly an urban low-income minority population. IUI was assessed based on intrapartum fever > 38°C and placenta pathology consistent with acute or chronic inflammation in any of several locations including the decidua, chorion, amnion, chorionic plate, and umbilical cord. The researchers examined associations of IUI by quartile of maternal exposure to residential ambient $PM_{2.5}$ using multivariable logistic regression.

Comparing the highest with the lowest $PM_{2.5}$ exposure quartiles, the multi-adjusted association with IUI was significant for all pregnancy periods considered, including 3 months before conception, each trimester and whole pregnancy.

Two studies have investigated associations between air pollution exposure during pregnancy and high sensitivity C-reactive protein (hsCRP) in maternal blood during the first trimester of pregnancy and in cord blood at birth: Lee et al (2011) and van den Hooven et al (2012b). CRP is an acute-phase protein of hepatic origin that increases following IL-6 secretion. As a result of the inflammatory response to the pregnancy, maternal CRP levels increase slightly during standard pregnancies, generally peaking during the third trimester (Thornton, 2010; von Versen-Hoeynck et al, 2009). However, a greater increase in CRP levels has been reported in women whose pregnancies are complicated by FGR, preeclampsia and preterm delivery (Guven et al, 2009; Pitiphat et al, 2005; Tjoa et al, 2003). There is an association, albeit weak, between increased mid trimester maternal plasma CRP and spontaneous PTB (267). Furthermore, elevated CRP levels in cord blood have been measured in SGA neonates (Trevisanuto et al, 2007).

van den Hooven et al (2012b) investigated associations between PM_{10} exposure in Rotterdam during pregnancy and CRP in maternal blood (n=5,067) during the first trimester of pregnancy and in cord blood (n=4,450) at birth. Short-term (that is, prior one to 2 weeks) maternal PM_{10} exposure was modestly associated with elevated (>8 ng/mL) maternal CRP levels in early pregnancy but no clear dose–response relationships were observed. PM_{10} exposure for 1, 2, and 4 weeks preceding delivery was not consistently associated with elevated (>1mg/L) fetal CRP levels, but associations were seen between higher long-term maternal PM_{10} and NO_2 exposure (Total Pregnancy) and fetal CRP levels at delivery.

Lee et al (2011) studied 1,696 pregnant women in Allegheny County, PA and reported associations between high (≥8 ng/mL) CRP in maternal blood in early (before week 22 of gestation) pregnancy and $PM_{2.5}$, PM_{10} and ozone concentrations averaged over 0-21 and 0-28 days. A 4.6 μg/m^3 increase in $PM_{2.5}$ and a 9.2 μg/m^3

increase in PM_{10} was significantly associated with high CRP. The odds ratio was 1.49 (95% CI 0.75 - 2.96) per 7.9 ppb increase in ozone during summer. There were no associations for NO_2.

Friedman et al (2021) further characterised the relationship between ambient air pollution concentrations and inflammatory biomarkers (CRP, IL-6, TNFα) among 515 mother-infant pairs in Denver, Colorado, USA. Positive associations were reported between trimester 1, trimester 2 and prior 30-day $PM_{2.5}$ concentrations and IL-6 and TNFα in maternal blood mid-pregnancy. In contrast to the findings of Lee et al (2011) no significant associations were found between $PM_{2.5}$ or O_3 concentrations and maternal CRP levels. An explanation given for this discrepancy was the variation in air pollution exposure levels by geographic location, as average particulate matter concentrations in Colorado are lower than those found in the northeastern US. Negative associations were reported between recent (7 and 30 days prior to sample collection) O_3 concentrations and maternal IL-6, whilst positive associations were found between O_3 and IL-6 and trimester 1, trimester 2 & prior 30-day concentrations and maternal TNFα levels. Patterns were inconsistent for associations between $PM_{2.5}$ and O_3 and cord blood inflammatory biomarkers, and generally 95% confidence intervals included the null. Concentrations of these maternal inflammatory biomarkers were not associated with infant birth weight, whilst unexpectedly, maternal CRP levels were positively associated with infant gestational age at birth (β; 1.30, 95% CI 0.40–2.21).

Summary of evidence on inflammation in adverse birth outcomes following exposure to air pollution

Only 4 studies were identified that looked at a possible inflammatory role in eliciting adverse birth outcomes in response to maternal exposure to air pollution during pregnancy (Lee et al, 2011; Nachman et al, 2016; van den Hooven et al, 2012b; Friedman et al, 2021).

These studies, measuring either IUI or inflammatory biomarkers in maternal blood in early/mid pregnancy and/or cord blood at birth, reported an association between inflammation and exposure to PM (Lee et al, 2011; Nachman et al, 2016; van den Hooven et al, 2012b; Friedman et al, 2021) and ozone (Lee et al, 2011; Friedman et al, 2021).

Whilst exposure to NO_2 was not associated with a maternal inflammatory response (Lee et al, 2011; van den Hooven et al, 2012b), associations between ambient NO_2 concentrations and cord blood CRP concentrations were reported (van den Hooven et al, 2012b).

Only one of the 4 studies extended their analysis to look for associations between the maternal (but not cord blood) inflammatory biomarkers and birth outcomes (Friedman et al, 2021). No associations were found observed for birth weight, whilst elevated levels of maternal CRP were associated with modestly older gestational age at birth.

3.3.4 DNA damage

DNA adducts

DNA adducts are generated when metabolites, originating for example from PAHs, react with cell macromolecules, such as DNA nitrogenous bases. These 'additions' or alterations of DNA structure increase the likelihood of gene mutations and DNA strand breaks. The formation of DNA adducts is a multifactorial process, dictated by numerous genetic and environmental factors, that previous studies have shown to be operative in human placenta (Gottlieb and Manchester, 1986; Manchester et al, 1988; Manchester et al, 1992). Since levels of PAH-DNA adducts measured in a biological sample reflect individual variation in exposure, absorption, metabolic activation, detoxification and DNA repair, adduct levels are considered biomarkers of a biologically effective dose and risk markers for exposure to PAH (Godschalk et al, 2003; Poirier, 1997).

Topinka et al (1997) studied DNA adducts in the placenta of 98 mothers living in Teplice, a polluted industrial area (mines, brown coal power plants) in the Czech Republic, and Prachatice, an agricultural area but where many residential heating systems used predominantly brown coal. While SO_2 and NO_x levels were much higher in Teplice, there was relatively little difference in PAH concentrations between the 2 districts. DNA adduct levels in the placentas of mothers from both districts were found to be very low and a significant difference between districts was found for the winter sampling period only. Higher placental DNA adduct levels were detected in the group of mothers with GSTM1-null genotype, and this was more significant for those living in the polluted district of Teplice than for those in Prachatice, suggesting that the enzyme may play a significant role in the detoxifying metabolism of DNA adduct forming reactive intermediates.

A study conducted by Perera et al (1999) studied 160 mother-newborn pairs who lived in industrial Krakow (n=70), Poland, versus Limanowa (n=90), a small town in an agricultural district. There was no difference in PAH-DNA adduct levels in placenta or maternal/cord blood from the 2 districts, possibly because of higher indoor air concentrations of PAH from coal burning in Limanowa. In Krakow, ambient PM_{10} was significantly associated with the amount of PAH-DNA adducts in both maternal and infant cord white blood cells, but not placenta. Newborns with elevated PAH-DNA adducts had significantly decreased BW, birth length, and head circumference compared to those with lower adducts. An increased susceptibility of the developing fetus to DNA damage was suggested when PAH DNA adduct levels in 112 paired maternal and infant samples were compared. White blood cell adduct levels were higher in infants than mothers, but not significantly so. These were deemed notable results given the estimated 10-fold lower PAH dose to the fetus compared with the mother. Maternal adduct levels were not associated with GSTM1 genotype.

Tang et (2006) examined the relationship between prenatal PAH exposure, as measured by maternal and cord blood BaP adducts, and fetal growth in a cohort of 150 women in Tongliang, China, where a seasonally coal-fired power plant was the main contributor to pollution. Cord adduct level was negatively but not significantly correlated with distance from the power plant. High adduct levels in cord blood were

significantly associated with decreased birth head circumference in females, but not birth length or weight, after controlling for potential confounders. Longer distance from the power plant was significantly associated with greater birth length ($p = 0.03$). Maternal blood adduct levels were not significantly correlated with cord blood adduct level or associated with fetal growth. Further work by the same investigators enrolled a second cohort of 58 non-smoking women following closure of the power plant in 2004 (Tang et al, 2014). There was a significant reduction of PAH-DNA adducts in the 2005 cohort ($p \leq 0.001$) reflecting the reduced emissions and reduced levels of ambient PAHs ($p \leq 0.01$). Head circumference was greater in children of the second cohort, compared with the first, along with significantly reduced levels of cord blood PAH-DNA adducts and reduced levels of ambient PAHs.

Suter et al (2019) investigated whether higher levels of exposure to PAHs and PAH-DNA adducts in placenta of women living near Superfund sites[10] (versus those living in non-Superfund site areas) in Harris County, Texas, USA contribute to the increased rate of PTBs. The study included samples from 104 subjects. Bulky PAH-DNA adducts were more prevalent in placental tissue from individuals residing near Superfund sites. Levels of BaP, benzo[b]fluoranthene (BbF) and dibenz[a,h]anthracene (DBA), were higher in placentae from preterm deliveries compared with term deliveries in women living near Superfund sites, whereas this was not the case for women living in non-Superfund site areas. Levels of DBA in the placenta were significantly higher than those of BaP and BbF and, moreover, were inversely correlated with gestational age at delivery and birth weight.

Maciel-Ruiz et al (2019) investigated the influence of $PM_{2.5}$ and PM_{10} concentrations on the presence of PAH-DNA adducts in maternal and cord blood. The study focused on non-smoking, cotinine-negative pregnant women (n=106) who lived in in Mexico City's Metropolitan Area - a highly polluted area impacted by industry and vehicle emissions. The levels of maternal DNA adduct were positively associated with cord blood levels, whilst, 60% of neonates cord blood DNA samples showed significantly higher levels of adducts than their mothers. The presence of adducts was significantly associated with concentrations of $PM_{2.5}$ and PM_{10}, mainly during the last trimester and these associations were always higher in cord blood. For $PM_{2.5}$, the association with DNA adducts in cord blood showed a linear dose–response trend. No clear associations were observed between PM and/or adduct levels and birth weight or length.

Summary of evidence on DNA adduct formation in adverse birth outcomes following exposure to air pollution

Six studies have investigated the possible role of DNA adduct formation in eliciting adverse birth outcomes by assessing whether increased exposure to air pollution is associated with elevated levels of PAH-DNA adducts in maternal/cord blood samples and/or placental tissue. Five of these studies included an evaluation of birth

[10] Abandoned or uncontrolled hazardous waste sites in the US requiring clean up

outcomes (Maciel-Ruiz et al, 2019; Perera et al, 1999; Suter et al, 2019; Tang et al, 2006; Tang et al, 2014).

It is noteworthy that all of these studies were located in highly polluted areas impacted primarily by industrial emissions and in the main, compared results obtained from mother-infant pairs who resided in these areas characterised by poor air quality, with those living in less polluted areas.

With the exception of Suter et al (2019), reporting a higher prevalence of placental DNA adduct from mothers residing near sites where PAH are contaminants of concern, there was little evidence of increased adducts in samples from polluted versus rural residential districts (possibly owing to confounding by indoor air pollution) (Perera et al, 1999; Tang et al, 2006; Topinka et al, 1997). Significant associations between ambient particulate pollution (PM_{10} and $PM_{2.5}$) and the levels of PAH-DNA adducts in maternal and cord blood have, however, been reported (Maciel-Ruiz et al, 2019; Perera et al, 1999).

Studies investigating whether transplacental PAH exposure to the fetus compromises fetal development have found inconsistent evidence as to whether elevated cord blood adduct levels are associated with birth weight and length, but findings have been consistent with respect to a reduced head circumference (Perera et al, 1999; Tang et al, 2006; Tang et al, 2014). The one study investigating a link to PTB also found an association with levels of both placental PAH and DNA adducts (Suter et al, 2019).

3.3.5 Epigenetic modifications

Epigenetic mechanisms influence gene expression without changing the underlying DNA sequences. Therefore, unlike genetic mutations, which lead to permanent changes of genes, epigenetic modifications are reversible and responsive to different environmental factors including lifestyle, diet and exposure to chemicals (Crews D et al, 2014; Foley et al, 2009).

Epigenetic processes include DNA methylation (the potential of a cytosine base to be methylated at its fifth carbon if followed by a guanine base in the DNA code (called a CpG site)), microRNAs (single-stranded, short (approximately 18-22 nucleotides) non-coding RNA sequences) and histone modifications (Chuang and Jones, 2007; Maccani and Marsit, 2009; McCarthy and Nugent, 2015).

Perturbations in the fetal epigenome may provide a link between early life exposure to environmental factors and both adverse pregnancy outcomes and later life events (Odom and Taylor, 2010; Perera and Herbstman, 2011; Szyf, 2009) because DNA synthesis rates are high and epigenetic remodelling is extensive through critical stages of development (Hochberg et al, 2011). For example, during the prenatal period, the embryo undergoes genome-wide DNA methylation, demethylation and re-methylation (Chaillet et al, 1994). The human placenta also expresses a distinct subset of miRNAs, and although the functions of these placental epigenetic marks

are largely unknown, research has revealed a functional role for miRNAs in placental biology (Gu et al, 2013).

Studies exploring whether prenatal exposure to air pollution is associated with epigenetic modifications and unfavourable birth outcomes have primarily focused on DNA methylation and, to a lesser extent, miRNA expression as potential factors.

DNA methylation

Janssen (2013) investigated the association between global DNA methylation in placental tissue from 240 newborns born in Genk, Belgium and $PM_{2.5}$ exposure over both the whole of pregnancy and for each trimester. Placental global DNA methylation was inversely associated with $PM_{2.5}$ during whole pregnancy and decreased by 2.19% for each 5 $\mu g/m^3$ increase in $PM_{2.5}$. In a multi-lag model in which all 3 trimester exposures were fitted as independent variables in the same regression model, only exposure to $PM_{2.5}$ during trimester 1 was significantly associated with lower global DNA methylation (−2.13% per 5 $\mu g/m^3$ increase, p=0.009).

Janssen et al (2015) studied placental mtDNA methylation by bisulphite-pyrosequencing in 2 regions (the displacement (D)-loop control region and 12S rRNA (MT-RNR1)) in 381 mother-newborn pairs living in Genk, Belgium. The D-loop is a crucial promoter region from which mtDNA replication and transcription are initiated. The MT-RNR1 sequence encodes the mitochondrial ribosomal 12S rRNA and has been posited to affect normal function and integrity of the mitochondrial ribosome. An IQR increment in $PM_{2.5}$ exposure over the entire pregnancy was positively associated with mtDNA methylation (MT-RNR1: +0.91%, P = 0.01 and D-loop: +0.21%, P = 0.05) and inversely associated with mtDNA content in placental tissue. No association however was found between altered placental mtDNA methylation and LBW or SGA.

Methodological literature on mtDNA methylation has raised concerns about the validity of the findings of Janssen et al (2015), due to findings that indicate that mtDNA methylation values derived by bisulphite pyrosequencing have been prone to overestimation (Liu et al, 2016; Mechta et al, 2017; van der Wijst et al, 2017). As a result, the same workers conducted a validity study, investigating placental mtDNA methylation (as well as other mtDNA content and birth weight) while accounting for methodological issues such as nuclear contamination, bisulphite conversion, and polymerase chain reaction bias (Vos et al, 2021). Placental samples used from the ENVIRONAGE cohort include those from mothers with high ($PM_{2.5}$: 16.0 ± 1.4 $\mu g/m^3$, BC: 1.8 ± 0.3 $\mu g/m^3$) and low ($PM_{2.5}$: 10.6 ±1.7 $\mu g/m^3$, BC: 0.9 ±0.1 $\mu g/m^3$) air pollutant exposure during the whole period of their pregnancy. Absolute D-loop methylation levels were higher for mothers that were highly exposed to air pollutants (+0.47%) and furthermore, D-loop methylation levels were negatively correlated with placental mtDNA content (p = 0.002) and inversely associated with birth weight (−106.98 g for an IQR increase in D-loop methylation), showing that the earlier findings of Janssen et al (2015) based on bisulphite pyrosequencing are legitimate and remain robust.

Saenen et al (2017) investigated in 361 placental tissue samples, whether DNA methylation of the leptin (LEP) promoter is associated with $PM_{2.5}$ and whether the oxidative/nitrosative stress biomarker 3-NTp is involved. The study was conducted in Genk, Belgium. During pregnancy, placental LEP plays a functional role in embryo implantation, intrauterine development and fetal growth (Sagawa et al, 2002). When LEP binds to its receptor, it stimulates angiogenic factors such as the vascular endothelial growth factor, thereby activating p38, MAPK, and Akt pathways that induce proliferation, motility, and angiogenesis (Garonna et al, 2011). These processes are critical in placental development, angiogenesis in villi and fetal-derived vascularisation (Demir et al, 1997). The study by Saenen et al (2017) found placental LEP methylation was 1.4% lower in association with an IQR (7.5 $\mu g/m^3$) increase in second trimester $PM_{2.5}$ exposure and 0.43% lower in association with a doubling of placental 3-NTp content. No associations were observed between LEP methylation and an IQR increment in the first or third trimester $PM_{2.5}$ exposure.

Liu X et al (2019) utilised a birth cohort in Guangzhou, China to prospectively estimate the effects of prenatal exposure to PM of different sizes (PM_{10}, $PM_{2.5}$, and PM_1) on the risk of PTB and further investigate the roles of both maternal and umbilical cord blood long interspersed nucleotide element (LINE)-1 methylation. LINE-1 is an evolutionarily conserved superfamily, mainly distributed in the adenine-thymine enrichment region. The decrease in promoter methylation level can induce the activation and transcription of retrotransposons, which may trigger the immune system and induce an inflammatory response (Crow, 2010), an important initiating mechanism of PTB (Vadillo-Ortega et al, 2014). Moreover, studies have found that maternal LINE-1 methylation levels are negatively correlated with PTB (Burris et al, 2012). Liu X et al (2019) reported that the risk of PTB birth was positively associated with $PM_{2.5}$ and PM_1 during the 12th to 20th gestational weeks. Observations included significant negative associations of PM_{10}, $PM_{2.5}$ and PM_1 concentrations with cord blood LINE-1 methylation levels, and a negative association between PM_1 concentration and maternal LINE-1 methylation level. Whilst the risk of PTB was negatively associated with maternal blood LINE-1 methylation levels and positively correlated with cord blood LINE-1 methylation levels, these findings were not significant.

Maghbooli et al (2018) investigated the level of global DNA methylation in newborns exposed in utero to contrasting concentrations of $PM_{2.5}$ and PM_{10}. The nested case-control study enrolled 92 pregnant non-smoking women not living with a smoker who were residents of either polluted (n=48) or unpolluted (n=44) regions in Tehran, Iran, one of the most polluted cities in the world. No significant correlations were found between birth outcomes (gestational age, weight, length, head and chest circumference) and (a) polluted versus non-polluted regions, (b) $PM_{2.5}$ and PM_{10} concentrations or (c) placental global DNA methylation levels. Although no significant difference was observed in placental global DNA methylation levels between mothers who lived in polluted and non-polluted regions, in a univariate model, after adjusting for gestational age, newborn sex, and region, significant positive associations between $PM_{2.5}$ and PM_{10} ($p = 0.01$) in the first trimester and placental global DNA methylation levels were observed.

Cai et al (2017) investigated associations between prenatal PM_{10} exposure during each trimester and DNA methylation of LINE-1, 11β-hydroxysteroid dehydrogenase 2 (HSD11B2) and NR3C1 in human placenta from 101 standard pregnancies and 80 complicated by FGR in Shanghai, China. HSD11B2 and NR3C1 are involved in glucocorticoid metabolism and play a critical role in fetal growth (Benediktsson et al, 1993; Oakley and Cidlowski, 2013). In FGR cases (but not newborns with normal growth), placental LINE-1 methylation decreased by 1.78% for each first trimester 10 $\mu g/m^3$ increase in PM_{10} exposure. Placental HSD11B2 DNA methylation was associated with both first and second trimester PM_{10} exposure, increasing by 1.03% and 2.33% for each 10 $\mu g/m^3$ increase in exposure. These associations were much more evident in FGR newborns than those in newborns with normal growth. No significant associations were observed between NR3C1 promoter methylation and all trimesters of PM_{10} exposure.

Gruzieva et al (2017) investigated meta-analysed associations between NO_2 concentrations during pregnancy and cord blood DNA methylation in 4 European and North American studies (*n* = 1,508). Epigenome-wide significant associations between NO_2 and DNA methylation in newborns were found for 3 CpG sites in mitochondria-related genes: cg12283362 (*LONP1*), cg24172570 (3.8 kbp upstream of *HIBADH*), and cg08973675 (*SLC25A28*). Further analysis of antioxidant and anti-inflammatory genes revealed differentially methylated CpGs in *CAT* (catalase) and *TPO* (thyroid peroxidase).

Kingsley et al (2016) examined associations between living close to a major roadway (defined as living ≤ 150 m from a primary highway or primary road or ≤ 50 m from a secondary road), markers of fetal growth (birth weight, SGA) and DNA methylation of repetitive elements (LINE-1 and AluYb8) in a cohort of 471 mother-infant pairs from the Rhode Island Child Health Study. LINE-1 and AluYb8 methylation levels have been found to be significantly positively associated with birth weight percentile and differ significantly among infants exposed to tobacco smoke and alcohol (Wilhelm-Benartzi et al, 2012). Living close to a major roadway was associated with statistically significant lower birth weight, increased odds of being SGA, lower levels of placental LINE-1 DNA methylation and differential placental DNA methylation at 7 loci, compared to those living further from major roadways. No association between residential proximity to major roadway and mean AluYb8 methylation levels were observed. Additional adjustment for placental methylation did not attenuate the association between roadway proximity and birth weight, suggesting that the association between residential proximity to major roadways and fetal growth is not explained by the observed placental epigenetic changes.

He et al (2018) explored whether exposure to PM_{10} and NO_2 in Zhengzhou, China during pregnancy was associated with H19/DMR (differentially methylated regions) methylation in maternal and cord blood of 527 newborns, as well as looking at effects on birth weight and length. The imprinted genes H19 and insulin like growth factor 2 (Igf2) are regulated by the DMRs or the imprinting control region (Bartolomei et al, 1991; Dechiara et al, 1991) and are associated with growth alterations when deletion of the maternally expressed Igf2 receptor or the adjacent imprinted locus, H19, or over-expression of the Igf2 gene occur (Le et al, 2013; Leighton et al, 1995; Sun et al, 1997). PM_{10} and NO_2 exposure during the whole pregnancy were

associated with lower H19 DMR methylation and higher H19 methylation in cord blood respectively. After stratification by pregnancy trimesters, the association of H19 methylation in cord blood with PM_{10} exposure also was found. There were statistically significant associations between NO_2 and PM_{10} exposures on birth weight and length but again these varied by different trimesters and were sometimes positively associated and sometimes negatively associated. No mediating effects of H19/DMR methylation in maternal or cord blood were found in the relationship between air pollutants exposure during pregnancy and birth sizes.

Ladd-Acosta et al (2019) sought to identify site-specific epigenetic changes, at the genome-scale, related to levels of prenatal exposure to NO_2 and O_3 in both cord blood and placenta. The study was conducted among 175 mother-child pairs at 4 major metropolitan locations across the US (Philadelphia, Baltimore, San Francisco Bay Area, and Sacramento). For increased O_3 exposure, a significant decrease in DNA methylation in cord blood at open sea regions and in placenta at shelf regions of the genome was observed. For NO_2, there were suggestive decreases in placenta DNA methylation levels for both the genome-wide and island genomic regions associated with increased exposure. No differences in global DNA methylation levels were observed in cord blood related to NO_2 exposure. The DMRs identified included those located at genes (eg RNF39 and PTPRH) involved in immune and inflammatory responses. Six DMRs associated with prenatal NO_2 exposure were identified including 3 that were sex specific. An additional 3 sex specific DMRs were associated with prenatal O_3 exposure levels. Comparison of exposure-related differences in methylation across placenta and cord blood revealed tissue-specific and as well as shared DMRs across tissues.

Among 164 participants in a New York city birth cohort, Herbstman et al (2012) reported that prenatal exposure to PAHs, measured using a personal air monitor during the third trimester, was associated with decreased genomic methylation in cord blood ($p=0.01$). In addition, higher global methylation levels were positively associated with the presence of detectable BaP–DNA adducts in cord blood, possibly suggesting that the latter may preferentially target methylated genomic regions.

miRNA expression

Tsamou et al (2018) investigated whether in utero exposure to $PM_{2.5}$ and NO_2 during different periods of gestation is associated with expression of 6 candidate miRNAs (miR-16, -20a, -21, -34a, -146a and -222) in placental tissue from 210 mother-newborn pairs living in Genk, Belgium. These candidate miRNAs were chosen based on being related to important cellular processes (Rudov et al, 2014; Vrijens et al, 2015). miR-16 and miR-21 are involved in cell cycle, proliferation, and apoptosis; miR-146a has been described as a regulator of inflammation; miR-20a, miR-34a, and miR-222 function in angiogenesis. Moreover, Maccani et al (2010) showed that maternal smoking during pregnancy was inversely associated with placental expression of miR-16, miR-21 and miR-146a. Tsamou et al (2018) found that placental expression of miR-21, miR-146a and miR-222 was inversely associated with $PM_{2.5}$ exposure during trimester 2 of pregnancy, while placental expression of miR-20a and miR-21 was positively associated with trimester 1 exposure. Similar findings for NO_2 exposure were observed: miR-20a, miR-21 and miR-146a were

inversely associated with NO_2 exposure during trimester 2, while a positive association was observed for the placental expression of miR-21 with exposure during the trimester 1 of pregnancy. These observations have potential to have important health consequences as down regulation of miR-21 has been shown to be significantly associated with FGR (Maccani et al, 2011).

Summary of evidence on epigenetic modifications in adverse birth outcomes following exposure to air pollution

A number of studies have investigated whether exposure to ambient air pollution during different periods of prenatal life is associated with differences in either global DNA methylation or differential gene-specific methylation in placental tissue or maternal/cord blood at birth.

Those that have examined global methylation or DNA methylation within LINE-1 repetitive elements (frequently used as surrogate markers for global methylation) have observed a lower degree of methylation in association with exposure to particulate air pollution (Liu X et al, 2019; Cai et al, 2017; Janssen et al, 2013). Decreased LINE-1 placental DNA methylation associated with living close to a major roadway has also been observed (Kingsley et al, 2016). The finding of a lower degree of placental global DNA methylation in association with exposure to particulate air pollution in early pregnancy (Janssen et al, 2013), including the critical stages of implantation, is of interest considering that disturbance of maintenance DNA methylation in placental tissue is associated with early pregnancy loss in humans and abnormal embryonic implantation and development in the mouse model (Yin et al, 2012). Decreased DNA methylation has also been observed in the few studies that looked at associations with PAH (Herbstman et al, 2012), NO_2 and O_3 (Ladd-Acosta et al, 2019).

With respect to differential gene-specific methylation (and importantly genes that are relevant in placental/fetal development), studies have found that prenatal exposure to particulate or gaseous pollutants are associated with altered DNA methylation of HSD11B2 (glucocorticoid metabolism) (Cai et al, 2017), LEP promotor (embryo implantation, intrauterine development and fetal growth) (Saenen et al, 2017) and H19 (fetal growth) (He et al, 2018).

Little evidence exists that such epigenetic changes may be an underlying mechanism by which air pollution may adversely affect birth outcomes. The identified studies either did not extend their analysis to look at a mediating effect of disrupted DNA methylation patterns on the relationship between air pollution exposure during pregnancy and negative birth outcomes or, with the exception of 2 studies (Cai et al, 2017; Vos et al, 2021), found no significant associations between an epigenetic modification and an adverse birth outcome. Cai et al (2017) reported that associations of PM_{10} exposure during early pregnancy with DNA methylation were much more evident in FGR newborns compared to newborns with normal

growth, whilst Vos et al (2021) reported a significant association between mtDNA methylation and birth weight.

Only one study (Tsamou et al, 2018) has looked at associations between air pollution exposure during pregnancy and miRNA expression changes. This suggested that exposure can affect the expression of miRNAs involved in regulation of the cell cycle, inflammation and angiogenesis, including miRNAs which have been shown to be associated with FGR.

3.3.6 Feto-placental haemodynamics/placental growth

van den Hooven et al (2012c) investigated, among 7,801 pregnant women living in Rotterdam, associations between exposure to PM_{10} and NO_2 during pregnancy with measures of placental growth and function: maternal and fetal angiogenic factors (placental growth factor (PlGF)) and soluble fms-like tyrosine kinase 1 (sFlt-1)), which are indices of placental resistance and placenta weight. Whilst PlGF is important for placental development and angiogenesis, sFlt-1 binds to this protein and thereby inhibits its activity (Coolman et al, 2012). PM_{10} and NO_2 exposure were associated with lower maternal second trimester sFlt-1 levels, and weakly associated with higher first-trimester PlGF levels and lower second-trimester PlGF levels. PM_{10} and NO_2 exposures averaged over total pregnancy were associated with higher sFlt-1 (35.8% and 8.9% per 10 $\mu g/m^3$ increase respectively) and lower PlGF (–16.3% and –14.6% respectively) levels in fetal cord blood. These results prompted speculation that air pollution may have contributed to an anti-angiogenic profile (that is, increased sFlt-1 and decreased PlGF levels in fetal cord blood, but not in maternal first- and second-trimester blood). PM_{10} and NO_2 exposures were not consistently associated with second (median 20.5 weeks) or third (median 30.3 weeks)-trimester placental resistance indices or placenta to birth weight ratio but were associated with lower placenta weight.

Contreras et al (2018) assessed the impact of prenatal exposure to traffic-related air pollution (NO_2, NO_x) on uterine vascular resistance (resistance index, pulsatility index, systolic/diastolic ratio) using Doppler velocimetry, in 360 women living in Los Angeles who had never smoked. Associations were reported between land use regression–derived NO_2 and CALINE4-derived NO_x concentrations and increased risk of high uterine artery resistance in late pregnancy (between 35 and 37 weeks of gestation).

Carvalho et al (2016) reported data from a prospective cohort study (Procriar; n=366) in São Paulo, which studied the association between personal monitored concentrations of NO_2 and O_3 on fetal weight, birth weight, and Doppler velocity of the umbilical, middle cerebral, and uterine arteries in the third trimester. Pregnant women used a passive personal sampler 3 times during each pregnancy (once each trimester). Second trimester O_3 exposure was associated with a higher umbilical artery pulsatility index (PI) in the adjusted analyses but, in contrast, third trimester exposure was related to lower umbilical PI. NO_2 did not affect umbilical artery PI and neither NO_2 nor O_3 was related to changes in middle cerebral artery PI and uterine artery PIs. There was no association between NO_2 and O_3 and fetal weight and birth weight. The authors concluded that air pollutants, using O_3 as a proxy, may interfere

with placental function, but that in a healthy population this is not resulting in detectable difference in fetal growth.

Summary of evidence on disturbances in feto-placental haemodynamics/placental growth in adverse birth outcomes following exposure to air pollution

Three studies were identified. These employed different methods to investigate an influence of maternal air pollution exposure on markers of fetoplacental haemodynamics and have yielded inconsistent results (Carvalho et al, 2016; Contreras et al, 2018; van den Hooven et al, 2012c).

An effect of maternal PM_{10} and NO_2 exposure on placental growth and development was supported by one study reporting (a) associations with changes in fetal sFlt-1 and PlGF levels at delivery that are consistent with an anti-angiogenic state and (b) that exposure levels in the prior 2 months, but not during other periods, were associated with lower placenta weight (van den Hooven et al, 2012c).

3.3.7 Thyroid hormones

Thyroid hormone is important for normal fetal development, but a steady balance is required as maternal hyper- and hypothyroidism are associated with increased risk of LBW (Blazer et al, 2003; Millar et al, 1994). Studies also suggest that fetal thyroid function is instrumental in regulating fetal growth (Medici et al, 2013; Shields et al, 2011).

Janssen et al (2017b) assessed the association between third-trimester exposure to $PM_{2.5}$ and thyroid hormones (free thyroid hormones FT3 and FT4; thyroid stimulating hormone (TSH)) in cord (n=499) and maternal (n=431) blood collected from mother-child pairs in the ENVIRONAGE birth cohort (Genk, Belgium). They found that an IQR increment (8.2 μg/m^3) in maternal $PM_{2.5}$ exposure was associated with a 11.6% reduction in TSH (11.6%) and FT4/FT3 ratio (62.7%) in cord but not maternal blood. Neither FT3 nor TSH levels in maternal or cord blood was associated with BW, however a 10th to 90th percentile decrease (11%) in cord blood FT4 was associated with a lowering in birth weight of 71 g ($p < 0.0001$).

Summary of evidence on the role of thyroid hormones in adverse birth outcomes following exposure to air pollution

Findings from a single study suggest a potential influence of third trimester exposure to $PM_{2.5}$ on fetal thyroid function and fetal growth.

3.3.8 The metabolomic approach

Metabolomic research is being increasingly used in maternal-fetal medicine to identify biological changes associated with fetal growth (Govarts et al, 2012). Examples include profiles of maternal plasma and urine to identify metabolites predictive of SGA babies, PTB and BW (Horgan et al, 2011; Maitre et al, 2014; Maitre et al, 2016). It is, however, studies investigating metabolic changes in cord blood that are particularly relevant as the sample contains the essential nutrients,

hormones, and immunological factors as well as potentially harmful xenobiotic metabolites to which the developing fetus is directly exposed.

Robinson et al (2018) employed untargeted metabolomic methods using 481 cord blood samples from 4 European cohorts (ENVIRONAGE (Belgium), INMA (Spain), Piccolipiu (Italy), and Rhea (Greece)) to identify metabolic features and biological pathways that may be disturbed by environmental risk factors (air pollution exposure, maternal and paternal education, maternal smoking, maternal weight gain). Exposure to PM_{10}, $PM_{2.5}$ and ultrafine particulates (UFPs, ≤0.1μm) was estimated using land use regression at the participant's home address. In bivariate analyses (without correction for confounders), all of the pollutants were correlated with birthweight. The correlation with PM_{10} had the smallest *p*-value, and a further analysis found that the association was no longer significant following adjustment for gestational age, cohort, maternal height and pre-pregnancy weight, sex of baby and paternal height. When investigating the potential role of the metabolome in mediating the effect of risk factors on birth weight, findings were limited to an association of vitamin A with both maternal smoking and birth weight.

Yan et al (2019) also conducted an untargeted metabolomics study on 160 serum samples collected from pregnant women during the second trimester of pregnancy who lived in the Central Valley of California. The aim of the study was to identify changes in specific metabolic profiles and biological pathways in response to traffic-related air pollution, but the analysis did not extend to associations with birth outcomes. Whilst information on smoking behaviour was not available, cotinine levels among women with high or low traffic-related air pollution exposure were generally extremely low, leading the authors to suppose that the percentage of active or passive smokers was low. Approximately 200 metabolic features were identified that discriminated between the high (n=98) and low (n=62) exposed groups. Pathway enrichment analysis identified alterations in 24 metabolic pathways. These included linoleate, leukotriene and prostaglandin metabolism, indicating that in maternal serum oxidative stress and inflammation related pathways were altered.

Summary of evidence on disturbance of maternal/fetal metabolities and metabolic pathways and adverse birth outcomes following exposure to air pollution

2 studies investigating metabolic pathways were identified. One (Yan et al, 2019) found that pathways related to maternal oxidative stress and inflammation differed between those with low- and high- exposure to traffic-related air pollutants. However, the analysis did not examine associations with birth outcomes. The other (Robinson et al, 2018) was unable to identify any measured metabolites in cord blood as mediating effects of air pollution exposure on birth weight.

3.4 Studies from animal models

Oxidative stress

Structural changes in umbilical cord (UC) vessels of mice exposed during pregestational and gestational periods to ambient PM (close to roadside in São Paulo, Brazil; mean $PM_{2.5}$ concentrations in filtered: 6.5 μg/m^3 versus non-filtered:

27.5 µg/m^3 exposure chambers) were associated with greater volumes of regions immunostained for 15-F2t-IsoP and endothelin A and B receptors, indicative of oxidative stress in the form of lipid peroxidation and imbalances in endogenous regulators of vascular tone (Veras et al, 2012). Furthermore, the placental injury documented in mice exposed to DE (300 µg/m^3 $PM_{2.5}$ from embryonic day 0.5 to 17.5; 6 hours/day, 5 days/week) was accompanied by elevated 3-NT protein modification, predominantly within perivascular regions in the fetal labyrinth layer (Weldy et al, 2014). This finding is in line with the link between placental nitrosative stress in women exposed to ambient BC during gestation (Saenen et al, 2016).

Inflammation

Fujimoto et al (2005) reported increased mRNA levels of inflammatory cytokines IL-2, IL-5, IL-4, IL-6, IL-12 α, IL-12 β, and GM-CSF (granulocyte-macrophage colony-stimulating factor), in placentas of mice exposed to DE (300, 1000 and 3000 µg/m^3 DE particles from day 2 to day 13 postcoitum; 12 hours/day, 7 days/week). The number of absorbed placentas in the 300 and 3000 µg DE particles/m^3 groups was significantly higher than that in the controls ($P < 0.01$ and 0.05, respectively) however, the number of pregnant females with absorbed placentas in the 1000 µg DE particles/m^3 group was significantly lower than that in the control group ($P < 0.05$).

Placental injury documented in mice exposed to DE (300 µg DE particles/m^3 from embryonic day 0.5 to 17.5; 6 hours/day, 5 days/week) was accompanied by infiltration of CD45+ cells within the decidua layer in placentas, suggesting that inflammation arises from signals originating from the mother (Weldy et al, 2014). Interestingly, observations in the same study of increased placental vascular oxidative stress did not overlap with areas of inflammatory cell infiltration and were largely limited to the fetal labyrinth layer.

de Melo et al (2015) exposed Wistar rats to filtered air or concentrated ambient particles (600 µg/m^3 $PM_{2.5}$; for 1 hour, 5 times/week during the 3 weeks before pregnancy and/or once per day, 7 days a week during pregnancy, starting on the sixth day of gestation). The gene expression of IL-1β, IL-4, IL-6, IL-10, interferon-γ (INF-γ), TNF-α and Toll-like receptor 4 in the placenta were measured as well as serum and placental concentrations of IL-1β, IL-4, IL-6, IL-10, INF-γ and TNF-α. Compared to control animals, IL-4 content was elevated in the fetal portion of the placenta in rats exposed to air pollution before and during pregnancy. No differences between the experimental groups were observed for the other placental or serum cytokines.

Morphological changes in placenta/umbilical cord and fetal growth

Veras et al (2008) evaluated the effects of São Paulo PM prior to and/or throughout pregnancy on functional morphology of the placenta of mice. Placentas were collected from near-term pregnancies, volumes of placental compartments were estimated, and the labyrinth was analysed further in terms of its maternal vascular spaces, fetal capillaries, trophoblast, and exchange surface areas. Pre-gestational exposure to non-filtered air led to significant increases in fetal capillary surface area and total and mass-specific conductance of the intervascular barrier. However, the

calibre of maternal blood spaces was reduced. Exposure to non-filtered air was associated with reduced volumes, calibres and surface areas of maternal blood spaces and greater fetal capillary surfaces and diffusive conductances. The gestational exposure period was associated with the more dramatic morphological changes. Fetal weights were significantly reduced by both pre-gestational and gestational periods of exposure.

Valentino et al (2016) evaluated the effects of gestational exposure to diluted DE (1000 µg/m^3 filtered DE for 2 hours/day, 5 days/week for 20 days in a 31-day gestation) on feto-placental development in a rabbit model. Diesel exhaust exposure induced signs of growth retardation at mid-gestation with decreased head length (-9.6%) and umbilical pulse (-3.6%). Near term, fetal head length (-4%) and plasma insulin (-44.4%) and IGF1 (insulin-like Growth Factor-1) (-91%) concentrations were reduced. Disturbances in placental function included a significant decrease (-12.3%) in placental efficiency (fetal to placental weight ratio) and a reduction in placenta blood flow. In addition, whilst the relative volume fractions of trophoblastic cells (−43 %) and fetal capillaries (−42 %) were significantly decreased and the relative surface density of fetal capillaries was significantly reduced in the labyrinthine area (-31.7%), the relative volume of maternal blood space was significantly larger (+116%).

Weldy et al (2014) exposed pregnant mice to DE (~300 µg/m^3 $PM_{2.5}$, 6 hours/day, 5 days/week) or filtered air from embryonic day 0.5 to 17.5. In utero exposure increased embryo resorption. In assessing fetal and placental weight from viable embryos, DE was not found to affect fetal weight but promoted a reduction in placental weight. Placentas from DE-exposed dams showed evidence of increased necrosis, focal congestion, and haemorrhage at the labyrinth/spongiotrophoblast interface. In addition, in utero DE exposure was associated with diffuse to focally extensive compaction and/or increased stromal density of the highly vascularised labyrinth.

Experimental research investigating whether maternal exposure to airborne PM during pregnancy affects umbilical morphology includes that of Veras et al (2012) in which mice were mated and completed pregnancies inside exposure chambers containing filtered ($PM_{2.5}$ 6.4 µg/m^3) or non-filtered ($PM_{2.5}$ 32.8 µg/m^3) São Paulo ambient air. The cords of mice from non-filtered chambers were approximately 22% smaller in volume due to a 28% loss of mucoid connective tissue and a 60% decrease in volume of collagen. Reductions in total vessel volume and cross-sectional areas did not reach statistical significance; however, reductions were observed in wall thicknesses of umbilical arteries and luminal diameter of umbilical veins.

Yue et al (2019) investigated the effect in mice, of gestational exposure to a suspension of $PM_{2.5}$ (3 mg/kg every other day, by oropharyngeal aspiration) collected from Taiyuan on vascularization of the placenta (Yue et al 2019). Mice were sacrificed at embryonic day 13.5 (E13.5), E15.5 and E18.5. Compared to a saline control group, gestational exposure to $PM_{2.5}$ significantly reduced the area of the labyrinth layer (the largest layer of the placenta and site of nutrient and gaseous exchange between maternal and fetal blood circulations) at E18.5 but not at E13.5 or E15.5. Genes (for example achaete-scute family bHLH transcription factor 2 [Ascl2]

and heart and neural crest derivatives expressed 1 [Hand1]) essential for placental development were altered at E13.5 and E18.5 but not E15.5. At E18.5, $PM_{2.5}$ exposure also decreased the maternal and fetal blood vessels within the labyrinth layer and suppressed protein expression of CD31, a marker of vascularization. Partial activation by $PM_{2.5}$ of the mitogen-activated protein kinase (MAPK; a chain of proteins that regulate a wide range of cellular processes by transmitting extracellular signals to intracellular targets) pathway suggested that the observed effects on placental vessel development might have been mediated through this signaling pathway.

A series of studies (Miller et al, 2017; Miller et al, 2019a; Miller et al, 2019b) have investigated the effects of O_3 exposure (0.4 or 0.8 ppm, 4 hours/day) during implantation (gestational days (GD) 5 and 6) on uterine vascular resistance (on GD 15 and 21) and fetal growth (on GD 21) of Long-Evans rats. Both male and female fetuses from dams exposed to 0.8 ppm O_3 (4 hours) had reduced fetal weight (Miller et al, 2017; Miller et al, 2019a). However, when dams were exposed to 0.4 ppm O_3 (4 hours), the impaired fetal weight gain was only observed in male fetuses (Miller et al, 2017; Miller et al, 2019b). In a standard pregnancy, resistance to blood flow decreases in the uterine artery, which facilitates perfusion of the fetoplacental unit. Whilst uterine arterial resistance did appear to decrease between GD 15 and GD 21 among dams in the control and low ozone exposure group (this decrease was greater in controls), it appeared to increase among dams in the high ozone exposure group, suggesting compromised arterial adaption (Miller et al, 2017; Miller et al, 2019a; Miller et al, 2019b).

DNA methylation

Studies have shown that the vascular adaptive alterations in the uterus during human pregnancy are largely mediated by the estrogen receptor (ER)-induced endothelial (e)NOs expression up-regulation (Magness et al, 2001; Rupnow et al, 2001; Vagnoni et al, 1998) and activity enhancement (Liao et al, 2005; Magness et al, 1996; Salhab et al, 2000; Yi et al, 2010), as well as eventual increased NO production (Miller et al, 1999; Rosenfeld et al, 1996). Dang et al (2018) investigated the effects of $PM_{2.5}$ exposure on birth weight and methylation modification in the ER promoter region in uterine tissue of Sprague-Dawley rats. Animals were exposed throughout gestation to averaged daily concentrations of Xi'an $PM_{2.5}$ of 40.74 ± 8.67 $\mu g/m^3$ (Clean group; n=10), 129.07 ± 14.59 $\mu g/m^3$ (Low group; n=12) or 223.79 ± 20.56 $\mu g/m^3$ (High group; n=8). Compared to the Clean group, birth weights and endometria vascular diameter maximum in the Low and High groups were significantly lower. Compared with the Clean group, the methylation levels of the CpG sites in the ERα promoter region of the uterus were significantly elevated in the Low ($p< 0.05$) and High ($p<0.01$) groups. These findings were accompanied by decreased mRNA and protein expression of the ERα (but not ERβ) and eNOs in the uterus.

Summary of evidence from animal studies on adverse birth outcomes following exposure to air pollution

The 11 experimental studies in laboratory animals that we examined primarily investigated effects of maternal exposure to concentrations of DE (≥ 300 µg PM/m^3) (Valentino et al, 2016; Weldy et al, 2014; Fujimoto et al, 2005), or particulate air pollution generated by urban traffic (Veras et al, 2012; Veras et al, 2008; Yue et al, 2019). Whilst acknowledging the inadequacies of animal models - owing to the fundamental differences in gestational sac structure, placentation, circulations, fetal/placental weight ratios, organogenesis phases and gestational length - results of all 5 relevant studies demonstrated disturbances in placental morphology/function, umbilical cord structure and/or fetal growth (Veras et al, 2008; Valentino et al, 2016; Weldy et al, 2014; Veras et al, 2012; Yue et al, 2019). The available studies suggest a role of an inflammatory response (Fujimoto et al, 2005; Weldy et al, 2014; de Melo et al, 2015), oxidative stress (Veras et al, 2012) and epigenetic control of gene expression (Dang et al, 2018).

Findings from studies investigating the effects of ozone exposure (0.4 or 0.8 ppm), exclusively during the period of implantation receptivity (which has a similar (approximately 2 days) duration in rats and humans), suggest a relationship with vascular dysfunction (diminished uterine artery flow) in the second half of gestation and fetal growth retardation in offspring (Miller et al, 2017; Miller et al, 2019a; Miller et al, 2019b).

Chapter 4 Discussion

In this section, we provide a summary of the epidemiological and mechanistic conclusions, discuss the approach taken for the report and the available literature reviewed, and discuss how the epidemiological and mechanistic evidence integrates. Finally, we highlight the implications of this work.

4.1 Overview of evidence

4.1.1 Epidemiological evidence

There was weak evidence for associations between some air pollutants and some adverse birth outcomes from the available meta-analyses. There were often many studies not reflected in the meta-analysis evidence (section 2.1), and for most pollutant-outcome pairs up-to-date meta-analyses are required to synthesise the full evidence base. The conclusions are summarised below according to adverse birth outcome.

Preterm birth (PTB)

Based on the meta-analyses available there is weak evidence for an association between PTB and exposure to $PM_{2.5}$, PM_{10}, and O_3, but the evidence does not support an association with NO_2 exposure. For O_3, this conclusion is based on limited evidence from the meta-analyses. For each pollutant there were a very large number of studies not captured in/published since the core meta-analyses, and findings from key studies examined were inconsistent for each of the pollutant-outcome pairs. Therefore, we recommended that for all pollutants an up-to-date meta-analysis is needed to synthesise all the available evidence in relation to PTB.

Term low birth weight (TLBW)

Based on the meta-analyses available there is weak evidence for an association between TLBW and exposure to $PM_{2.5}$, PM_{10}, but the evidence does not support an association with NO_2 or O_3 exposure. For $PM_{2.5}$ and O_3, this conclusion is based on limited evidence from the meta-analyses. For each pollutant there were numerous studies not captured in/published since the core meta-analyses. Findings from key additional studies examined were inconsistent for $PM_{2.5}$ and PM_{10}, and for these pollutants an up-to-date meta-analysis is needed to synthesise all the available evidence in relation to TLBW. For NO_2 and O_3 the findings of key additional studies examined are in line with the meta-analysis conclusions, that is, they do not support an association. Up-to-date meta-analyses are nonetheless recommended for these pollutants to synthesise all the available evidence in relation to TLBW.

Stillbirth

The single meta-analysis available does not support an association between stillbirth and exposure to $PM_{2.5}$, PM_{10}, NO_2 or O_3. This conclusion is based on limited evidence from the meta-analysis. For each pollutant there were several studies not captured in/published since this core meta-analysis. Findings from key additional studies examined were inconsistent for $PM_{2.5}$, PM_{10} and O_3, and for these pollutants an up-to-date meta-analysis is needed to synthesise all the available evidence in relation to stillbirth, although we consider it unlikely to alter the conclusion for PM_{10}. For NO_2 the findings of key additional studies examined are in line with the meta-analysis evidence conclusion, that is, they do not support an association. An up-to-date meta-analysis is recommended for NO_2 to synthesise all the available evidence in relation to stillbirth, but we consider it unlikely to alter this conclusion.

4.1.2 Mechanistic evidence

Studies investigating possible mechanisms underlying adverse birth effects have measured biomarkers in human samples or used experimental research in laboratory animals. They have focused on various biochemical or physiological changes that have occurred following exposure to air pollution such as oxidative stress, inflammation, epigenetic alterations and PAH-DNA adduct formation. However, whilst these biomarkers may reflect exposure to air pollution, or its effects, they are not specific to air pollution and may not necessarily be associated with an established or possible health effect that leads to an adverse birth outcome.

It should also be borne in mind that only a limited number of studies have investigated any given mechanism. This prevents any firm conclusions from being made relating to weight of evidence between different pollutants and critical time window of exposure during pregnancy.

Human studies

Oxidative stress

Biomarker studies have shown consistent associations between exposure of pregnant women to PM and markers of nitrosative/oxidative stress (3-NTp, 8-OHdG, mtDNA) in the placenta and/or maternal/cord blood (Brunst et al, 2018; Ambroz et al, 2016; Carvalho et al, 2016; Contreras et al, 2018; Saenen et al, 2016; Grevendonk et al, 2016; Janssen et al, 2012). Saenen et al (2016) was the only study that adjusted for other air pollutants and found that the association between $PM_{2.5}$ and placental 3-NTp was robust to adjustment for BC or NO_2. Fewer studies have looked for associations between these markers and NO_2/O_3/PAHs and have yielded more inconsistent effects (Ambroz et al, 2016; Rossner et al, 2011; Grevendonk et al, 2016; Clemente et al, 2016). Apart from one study suggesting that 10% of LBW caused by prenatal NO_2 exposure could be explained by altered placental mtDNA content (Clemente et al, 2016), studies have not found that an increased oxidative burden with increased exposure is associated with PTB or negative indices of fetal growth (Maciel-Ruiz et al, 2019; Rossner et al, 2011; Saenen et al, 2016; Grevendonk et al, 2016)

Inflammation

Studies measuring either IUI or inflammatory biomarkers in maternal blood in early/mid pregnancy and/or cord blood, reported an association between inflammation and exposure to PM (Lee et al, 2011; Nachman et al, 2016; van den Hooven et al, 2012b; Friedman et al, 2021) and ozone (Lee et al, 2011; Friedman et al, 2021). Whilst exposure to NO_2 has not been associated with a maternal inflammatory response (Lee et al, 2011; van den Hooven et al, 2012b), associations between ambient NO_2 concentrations and cord blood CRP concentrations have been reported (van den Hooven et al, 2012b). The one study that extended analysis to look for associations between the maternal inflammatory biomarkers and birth outcomes, found no associations with birth weight.

PAH-DNA adducts

With the exception of one study reporting higher placental DNA adducts in mothers residing near sites where PAH are contaminants of concern, there is little evidence of increased adducts in samples from polluted versus rural residential districts. Significant associations between ambient particulate pollution (PM_{10} and $PM_{2.5}$) and the levels of PAH-DNA adducts in maternal and cord blood have however been reported. A trend exists for adduct levels to be higher in cord blood when compared to maternal blood, suggesting an increased susceptibility of the developing fetus to DNA damage (Maciel-Ruiz et al, 2019; Perera et al, 1999; Tang et al, 2006). Studies investigating whether PAH-adducts may mediate adverse birth outcomes have found evidence that elevated (a) cord blood adduct levels is associated with decreased head circumference (Perera et al, 1999; Tang et al, 2006; Tang et al, 2014) and (b) placental PAH-DNA adducts are associated with PTB (Suter et al, 2019).

Epigenetics

A number of studies looking for differences in global DNA methylation in placental tissue or maternal/cord blood at birth have reported a lower degree in association with exposure to particulate air pollution (Liu X et al, 2019; Cai et al, 2017; Janssen et al, 2013). Decreased DNA methylation in association with living close to a major roadway (Kingsley et al, 2016) and ambient PAH (Herbstman et al, 2012), NO_2 and O_3 (Blazer et al, 2003) has also been observed.

Studies focusing on differential methylation of relevant genes have shown that prenatal exposure to particulate or gaseous pollutants are associated with altered methylation of HSD11B2 (glucocorticoid metabolism) (Cai et al, 2017), LEP promotor (embryo implantation, intrauterine development and fetal growth) (Saenen et al, 2017) and H19 (fetal growth) (He et al, 2018).

Studies have not generally extended analyses to look for an association between disrupted DNA methylation patterns and negative birth outcomes or have failed to identify such associations (Kingsley et al, 2016; Liu X et al, 2019; He et al, 2018; Maghbooli et al, 2018; Janssen et al, 2015). One study reported that associations between PM_{10} exposure during early pregnancy and DNA methylation were much more evident in FGR newborns compared to newborns with normal growth (Cai et al,

2017) and in another, a significant association between mtDNA methylation and birth weight was reported (Vos et al, 2021).

Feto-placental haemodynamics and placenta/UC growth

An effect of maternal PM_{10} and NO_2 exposure on placental growth and development is supported by one study reporting (a) associations between exposures and changes in fetal sFlt-1 and PlGF levels at delivery that are consistent with an anti-angiogenic state and (b) that exposure concentrations in the prior 2 months were associated with lower placenta weight (van den Hooven et al, 2012c).

Thyroid hormones

Currently, evidence (stemming from a single study (Janssen et al, 2017b)) that air pollution affects hormones with relevance to fetal development is too limited to draw any conclusions.

The metabolomic approach

The use of metabolomics to identify altered biological pathways is a recent area of research. Among the metabolomic features and pathways identified to date, are oxidative stress and inflammation (Yan et al, 2019), mechanisms already implicated using more traditional approaches.

Animal studies

Experimental studies exposing pregnant rodents to DE (≥ 300 µg PM/m^3), (Valentino et al, 2016; Weldy et al, 2014; Fujimoto et al, 2005) or ambient air pollutants generated by urban traffic (Veras et al, 2012; Veras et al, 2008) have demonstrated disturbances in placental morphology/function and umbilical cord structure, accompanied by an inflammatory response and nitrosative/oxidative stress.

Effects of ozone exposure (0.4 or 0.8 ppm) in rats, exclusively during the period of implantation receptivity, suggest a relationship between diminished uterine artery flow in the second half of gestation and fetal growth retardation (Miller et al, 2017; Miller et al, 2019b).

4.2 Strengths and limitations of the report

4.2.1 Epidemiological evidence

We have conducted a rigorous and very detailed evaluation of available meta-analyses. The quality of some published meta-analyses, and of some papers which are included within them, was poor. We wrote to the authors and editors of several published meta-analyses regarding errors, leading to the retraction of 2: Guo et al, 2019 (Guo LQ et al, 2020) and Ji Yet al, 2019 (Ji et al, 2024). When reviewing the available meta-analyses, we encountered some common issues, including: a lack of consistency in the methods used and studies included in the meta-analyses which, in turn, leads to later meta-analyses not building on previously published meta-analyses. Understanding these limitations can help improve meta-analysis research in future. For this report, it means that our conclusions have been drawn on the basis of selecting the most appropriate meta-analyses from those available, using a

methodical justified approach. This involved modifying and applying a tool (AMSTAR 2) to assess the quality of the meta-analyses, and assessing population overlaps within meta-analyses (as detailed in section 2.2.2). These methods have highlighted important considerations for researchers conducting meta-analyses in the future.

We have considered studies that have examined TLBW (a binary outcome) and not considered literature in which term birth weight is considered as a continuous outcome, or other related binary outcomes such as SGA. We note that not all meta-analyses examining LBW applied a criterion to restrict studies specifically examining LBW at term (see 2.1.2); we decided not to exclude these meta-analyses, as we did not want to lose valuable information regarding TLBW, but instead we considered this in the interpretation of the results.

We have considered studies that have examined chronic long term exposures during pregnancy only and have not attempted to consider acute effects of pollution on the outcomes of interest. We have also not considered multi-pollutant models (for neither the epidemiological nor the mechanistic evidence), as this was beyond the scope of the report; there may be additive, or even multiplicative, interactions between pollutants acting on different pathways that increase the risk of harm.

As well as identifying, describing and critiquing available meta-analyses, we have identified and summarised key studies that were not included in the selected meta-analyses, or were published after the selected meta-analyses were conducted, that provide additional information of interest. We did not conduct our own systematic review and meta-analyses as this was beyond the scope of the report.

4.2.2 Mechanistic evidence

We considered evidence from human and animal mechanistic studies and quality assessed the evidence using established guidelines (Texas Commission Environmental Quality guidelines). The emphasis of the summary of evidence was on human mechanistic studies. Whilst we made comment on studies in animal models, which are useful for understanding the health effects of controlled exposures, we did not focus on these, as we recognise that the animals used in such studies are not comparable to a human pregnancy/uterine environment, and therefore are less useful for understanding the effects of air pollution on adverse birth outcomes than human mechanistic studies. Data generated from in vitro studies have not been included.

Within this report, we have begun the process of assigning biological plausibility to epidemiological findings, which has been aided by putting the experimental data in the context of wider evidence (irrespective of air pollution exposure) on mechanisms related to adverse birth outcomes. We have not considered data generated from studies investigating mechanisms that underlie pregnancy complications such as preeclampsia and gestational diabetes that may be on the causal pathway (and thus may represent mediating mechanisms) to a growth deficit/premature birth/stillbirth.

4.3 Limitations of the evidence base

4.3.1 Epidemiological evidence

This report has comprehensively reviewed multiple meta-analyses and summarised major studies that have not been included in them. It is unlikely that our thorough systematic approach has missed any major high-quality studies published before August 2020 (meta-analyses) and March 2019 (individual epidemiological studies) that would alter the conclusions presented in this report.

However, we are aware of limitations within the available evidence and briefly summarise them here.

We consider that adjustment for active maternal smoking behaviour is important for the provision of robust evidence of an association, as there is good supporting evidence from epidemiological studies that active and even passive smoking during pregnancy increases the risks of stillbirth, preterm birth (and other adverse birth outcomes such as miscarriage, oral clefts and neural tube defects) and has a limited and dose-dependent effect on fetal growth (example of passive smoking section 1.7.1). Many studies do not have individual level data on this, although some provide adequate area level smoking levels as a proxy measure. In particular, most studies from California, USA (which are numerous and have dominated this research field) do not adjust for smoking due to lack of individual-level smoking information. No studies that did not adjust for smoking were rejected, although more emphasis was given to those studies that considered active maternal smoking. Studies from Asian countries often do not adjust for maternal smoking and state that the prevalence of maternal smoking in the population is low. Exposure to others' tobacco smoke in the home, work or in other social settings may be an important confounder. Across the world there are now more countries with smoking bans in the workplace and in public areas - so this may be of less concern in more recent studies. However, exposure to second-hand tobacco smoke in the home in some domestic situations may be considerable and is rarely included as an adjustment.

Deprivation has also been shown to be associated with poor birth outcomes in epidemiological studies but disentangling the factors that might mediate this effect is complex (Kim and Saada, 2013; Janevic et al, 2010). Increased frequency of smoking in deprived areas might be one of these mediators. There are some studies that do not adjust for potential confounding by ethnicity or by socio-economic status. PTB, TLBW and stillbirth are more common in women in some ethnic minority groups (Moser et al, 2008) and in those of low socio-economic status – possibly related to factors such as smoking, diet, alcohol consumption, obesity and access to antenatal and good quality health care. There is also evidence that these population sub-groups have higher exposure to traffic related pollutants (Fecht et al, 2015; Williams et al, 2018). However, whether individual socioeconomic status or neighbourhood socioeconomic level is more important, and why, is a matter of debate (Weisskopf, 2017). Deprivation needs to be considered carefully. Some of the studies which investigate associations between air pollution exposure and adverse birth outcomes have not adjusted for potential confounding by ethnicity or by socioeconomic status. This means that some residual confounding may remain, and that the associations reported may over-estimate the effects of pollution. Conversely,

where there are appreciable associations between deprivation and higher air pollution concentrations, adjustment for deprivation may reduce the observed associations with air pollution.

There has been concern that some subgroups may be at greater risk of the effects of air pollution than others, and therefore some of the factors mentioned above may actually be effect modifiers rather than confounders. This is rarely considered, although Thayamballi (2020) conducted a systematic review of all studies in the US in which associations of birth outcomes (including PTB and TLBW) with exposure to PM were presented, stratified by race/ethnicity and/or an individual marker of socioeconomic status (maternal educational level). In a qualitative synthesis of 19 studies (all in the US and likely with population overlaps that are not considered) they conclude that exposure to PM may have a greater effect on birth weight (considered as a continuous outcome with not all studies limiting their analysis to term births) of children of black mothers compared to white mothers No consistent differences in association by ethnicity or by socioeconomic status were observed for the other birth outcomes considered in this study (PTB, stillbirth). Meta-analysis was only possible for ethnicity modification of the relationship between $PM_{2.5}$ and continuous birth weight, for which differences failed to reach conventional levels of statistical significance.

For nearly all pollutant outcome pairs, published evidence is dominated by reports from a limited number of geographical locations in particular from the USA - by comparison the number of reports from Europe and from the UK are limited – and there is an increasing number of reports from China where air pollution levels are particularly high and are not comparable to levels seen in the UK

As with many studies examining health effects of air pollution there are few that include multiple pollutant models

There are specific biases and errors that may influence outcomes from studies that investigate the association of air pollution on birth outcomes – these are rarely considered explicitly. For instance:

- Measurement error in pollution exposure assignment:

  Air pollution exposure assessment methods differ between studies in terms of sophistication, spatial resolution, and temporal resolution - air pollution exposure assessment methodology has evolved over time in this field, from relatively crude metrics, such as proximity to nearest monitor (likely with more error), to highly sophisticated modelling, such as dispersion modelling (likely with less error).

  Exposure assessment is often based on residential address (or another geographic area-level such as postal/zip code) and does not account for migration in and out of these areas. Moreover, a proportion of pregnant women will move home as they prepare for their new child. Relatively few studies have been able to account for mobility during pregnancy and assume that the mother's address at time of birth is her location throughout pregnancy.

- Bias in study design:

  Biases may potentially arise when logistic regression is used to analyse seasonally varying exposures for shorter gestation cases compared to term controls (particularly relevant for studies of PTB and stillbirth).

  Fixed cohort biases may occur in retrospective cohorts when births are identified by their date of delivery rather than date of conception. Although the conception date may be the same, shorter pregnancies are missed at the start of the study (because they have not 'survived' to the start of the study), and longer pregnancies are missed at the end (because they have not delivered before the study end date) - this leads to differences in pregnancy lengths between those births included at the start and those included at the end and can lead to spurious associations with seasonally varying exposures (Strand et al, 2011).

Limitations of meta-analyses available

The quality of some published meta-analyses, and of some papers which are included within them, was poor. We wrote to the authors and editors of several published meta-analyses regarding errors, leading to the retraction of 2: Guo et al, 2019 (Guo LQ et al, 2020) and Ji Y et al, 2019 (Ji et al, 2024).. When reviewing the available meta-analyses, we encountered some common issues, as follows:

- Some systematic reviews, which provide the basis for meta-analyses, have collated literature published over selected short time periods, without justifying this approach (that is they were not updates to previous systematic reviews). Therefore they are not systematic reviews of all of the available evidence

- Detailed description and transparency about why papers were selected for inclusion for meta-analysis once they had been identified in the systematic review (and why some papers were excluded) was often lacking, making evaluation of those selections difficult.

- Some meta-analyses did not account for population overlaps (that is some meta-analyses included multiple studies on the same population over the same period) or did not provide a full description of why one study was preferentially included in the meta-analysis over another conducted on the same population over the same period

- Some meta-analyses did not transform effect estimates to a common exposure metric before meta-analysing

- Some systematic reviews and meta-analyses did not adequately quality assess contributing studies, so that the influence of the inclusion of studies with a less complete exposure assessment and/or lack of adjustment for major confounders could not be easily evaluated

The limitations mentioned above likely explain why there is such limited overlap between the studies included in each of the selected meta-analyses, even though some were based on systematic reviews of literature published over similar periods.

In addition, repeated lack of transparency in methods meant that it was difficult for later meta-analyses to build on earlier meta-analyses.

4.3.2 Mechanistic evidence

When assessing the human mechanistic studies, we noted common limitations, as follows:

- Not all studies have linked concentrations of air pollutants and mechanistic endpoints with an adverse birth outcome, such that the relevance of experimental findings to the epidemiological evidence can be questioned
- Few studies have adjusted for co-pollutants
- Whilst a number of different biological pathways have been investigated, only a limited number of studies have focused on any given mechanism (with the exception of oxidative stress and DNA methylation)
- Throughout the mechanistic literature there has been an emphasis on ambient PM, whilst pollutant gases have generally been neglected
- The majority of studies have used measurements of air pollution exposure generated from regional monitoring stations rather than personal monitors – given the sometimes small sizes of the mechanistic studies, the potential for misclassification of exposure is high, weakening the ability of the studies to show associations
- Many studies did not account for indoor pollution exposure. Some pregnant women may be exposed to higher levels of pollutants indoors compared to outdoors, depending on their location and lifestyle factors.

4.4 Bringing the epidemiological and mechanistic evidence together

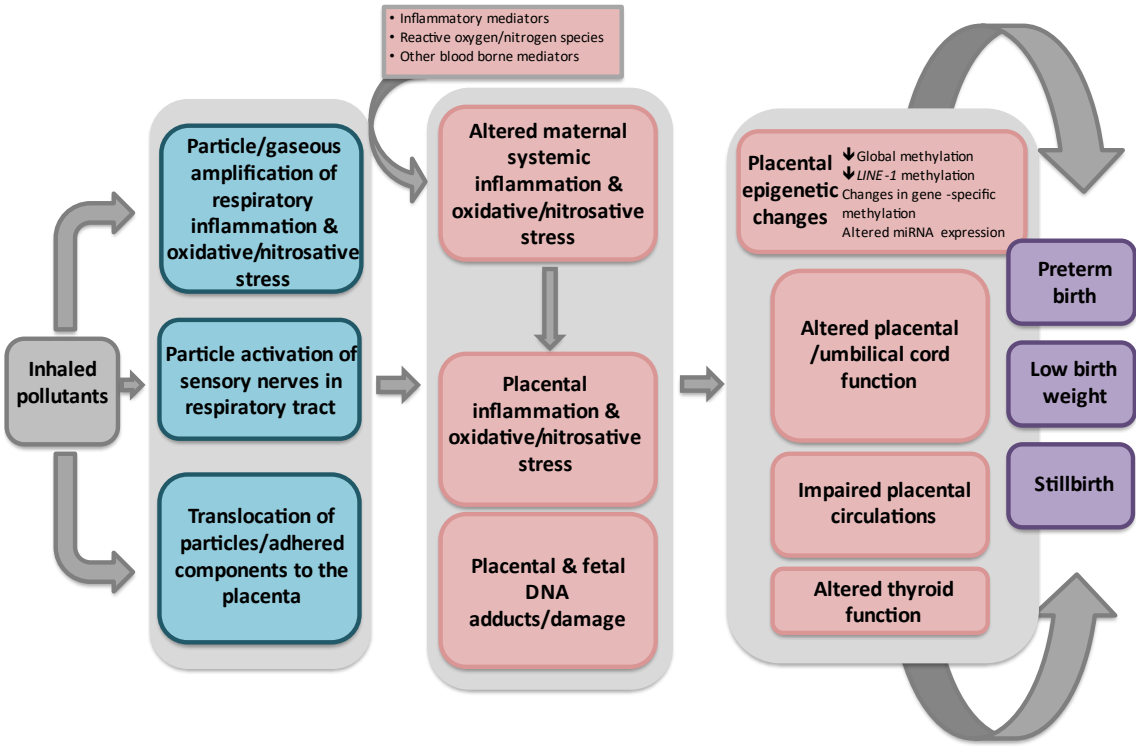
A key issue in assessing the evidence is whether (a) the associations in epidemiological studies can be regarded as causal and (b) potential mechanisms can be identified. Often, an issue in assessing the mechanistic evidence is whether the mechanisms described apply in real exposure situations. However, in this report, much of the mechanistic data did originate from nested studies within epidemiological cohorts.

Some of the issues relating to causality are discussed within the epidemiology sections and in the sections below – these include whether exposure precedes the adverse outcomes, whether confounding has been taken into account sufficiently, whether the results are consistent and whether the associations are of a reasonable size. These aspects were all considered in the conclusions on the strength and availability of evidence in the epidemiology chapter as summarised in section 4.1.1. There was at least some support for causality from consideration of these aspects. These aspects will also have to be considered in evaluating the new studies that have not yet been considered in detail in this report.

Another key aspect of considering causality is biological plausibility. As discussed below, in general, the experimental evidence provides some biological plausibility. At the level of specific pollutant-outcome pairs, this is less certain. Biological plausibility is expanded upon in the text that follows by tying in (a) the mechanistic data generated from studies described earlier in this report with (b) the epidemiological evidence on air pollution and birth outcomes and (c) putting the air pollution experimental data in the context of wider evidence (irrespective of air pollution) on mechanisms related to adverse birth outcomes.

Figure 12 describes the potential biological mechanisms (discussed in this report) linking inhaled air pollutants to adverse birth outcomes.

Figure12. Schematic diagram of possible mechanisms, based on the experimental evidence discussed in this report, underlying associations between ambient air pollution exposure and adverse birth outcomes. Progression of effects (from left to right) is colour coded (grey: exposure; blue: initial effect; pink: intermediate effects; purple: clinical effect).



Text version of Figure 12

This is a schematic diagram of possible mechanisms, based on the experimental evidence discussed in this report, underlying associations between ambient air pollution exposure and adverse birth outcomes. It shows that, following exposure to inhaled pollutants, initial effects might include:

- Amplification of respiratory inflammation and oxidative/nitrosative stress following inhalation of particles and gases
- Activation of sensory nerves in the respiratory tract by particles
- Translocation of particles and/or adhered components to the placenta

These effects might lead, either directly or via altered maternal systemic inflammation and oxidative/nitrosative stress, to:

- Placental inflammation and oxidative/nitrosative stress and
- Generation of placental and fetal DNA adducts and/or DNA damage

Inflammatory mediators, reactive oxygen/nitrogen species or other blood-borne mediators might be involved in some of these responses

Downstream events might include

- Epigenetic changes within the placenta
- Altered placental and/or umbilical cord function
- Impaired placental circulation
- Altered thyroid function

These could increase the risk of the following clinical effects:

- Preterm birth
- Low birth weight
- Stillbirth

End of text version of Figure 12.

Oxidative stress

Pregnancy is a state of increased physiological oxidative stress (Myatt and Cui, 2004). If this is exacerbated by maternal exposure to air pollution, so as to overcome the fetal mechanisms that minimise the deleterious effects of ROS production, it could constitute one of the mechanisms responsible for adverse birth outcomes.

Support for an increased oxidative burden mediated by poor air quality contributing to adverse birth outcomes includes the role that placental oxidative stress is believed to play in the pathophysiology of preeclampsia (Hubel, 1999). Increased oxidative stress, accompanied by reduced translation of proteins, including key signalling molecules, is also present in the placentas of IUGR pregnancies (Takagi et al, 2004; Yung et al, 2008). In such cases, oxidative stress is believed to damage chorionic villi, reducing the functional mass of the syncytiotrophoblast and limiting the capacity of the villi to mediate nutrient transport (Scifres and Nelson, 2009).

The mechanistic studies described in this report have observed associations between exposure of air pollution during pregnancy, and particularly PM, and markers of nitrosative (3-NTp) and oxidative stress (8-OHdG, isoprostanes, mtDNA) in the placenta and/or maternal/cord blood. As outlined below, these findings lend support to an effect of maternal air pollution exposure on adverse birth outcomes, especially as LBW is a consequence of compromised fetal growth.

For example, as described in section 3.3.2, Saenen et al (2016) observed a positive (adverse) association between placental 3-NTp levels in women and concentrations of $PM_{2.5}$ (but not NO_2) during pregnancy, which remained after adjustment for NO_2. In an experimental study in which mice were exposed to diesel exhaust particles (300 µg/m^3 $PM_{2.5}$ 6 hours per day, 5 days per week), Weldy et al (2014) observed elevated 3-NTp, predominantly in perivascular regions in the fetal labyrinth layer. 3-NTp is the stable product of protein nitration with peroxynitrite and whilst nitration of placental proteins is evident in standard pregnancies, higher levels may disturb placental function, as reported in placental vessels in conditions of preeclampsia and gestational diabetes (Bosco et al, 2012; Lyall et al, 1998). Moreover, targets for protein nitration are involved in trophoblast invasion and regulation of placental vascular reactivity, which in turn could compromise placental function (Myatt, 2010) and thus fetal development and growth.

A number of studies have observed significant associations between increased prenatal concentrations of $PM_{2.5}$/PM_{10} (Brunst et al, 2018; Rosa et al, 2017; Janssen et al, 2012) and NO_2 (Clemente et al, 2016) during pregnancy and decreased placental/cord blood mtDNA content. This is consistent with a study reporting a decrease in placental mtDNA content in relation to increased number of cigarettes smoked per day (Bouhours-Nouet et al, ,2005). Mitochondrial DNA (mtDNA) contains the genes essential for the crucial proteins for electron transport, oxidative phosphorylation and mitochondrial homeostasis (Anderson et al, 1981). Abnormalities in the quantity and quality of mtDNA are therefore closely associated with mitochondrial dysfunction, oxidative stress, and apoptosis (Duchen, 2004; Enns, 2003). In regulating energy, placental mitochondria are essential in the proper formation and functioning of the placenta. Oxidative stress therefore has the potential to adversely affect placental mitochondria and, in doing so, impair the ability of the placenta to support the growing fetus.

Inflammation

Pregnancy is primarily characterised by an anti-inflammatory immunological tolerance; however, inflammatory events do take place during different phases, including implantation (Mor et al, 2017), placentation (Brown et al, 2014) and in preparation for delivery (Dubicke et al, 2010).

An increased inflammatory response is also observed during complicated pregnancies (Pitiphat et al, 2005; Amarilyo et al, 2011; Catov et al, 2007). For example, elevated maternal CRP levels have been associated with preterm delivery and FGR (Pitiphat et al, 2005; Ernst et al, 2011), and elevated CRP levels in umbilical cord blood have been reported in infants being born SGA (Trevisanuto et al, 2007; Amarilyo et al, 2011).

These findings support a possible link between exposure to poor air quality and an increased risk of TLBW and PTB, mediated by an enhanced inflammatory response, as suggested by the limited number of mechanistic studies described in section 3.3.3 above, that have investigated such a hypothesis (Lee et al, 2011; Nachman et al, 2016; van den Hooven et al, 2012b; Friedman et al, 2021). Two studies reported associations between elevated (>8 ng/mL) hs-CRP concentrations in maternal blood in early pregnancy and exposure to PM (Lee et al, 2011;van den Hooven et al,

2012b) and ozone (Lee et al, 2011) but not NO_2. In contrast, Friedman et al (2021) did not find associations for maternal CRP levels, but did for maternal IL-6 and TNFα.

A plausible pathway also exists linking an increased inflammatory state to compromised fetal growth through the expression of molecules involved in the formation of a dysfunctional placenta, resulting in excessive trophoblast apoptosis, shallow trophoblast invasion, and impaired spiral artery remodelling (Li and Huang, 2009).

PAH exposure

Some studies (discussed above) have found associations of PAH exposure, indicated by DNA-adduct levels, with adverse birth outcomes.

Maternal exposure to PAH may elicit adverse birth outcomes through a number of mechanisms. For example, it has been suggested that BaP may interfere with growth of uterine tissue during pregnancy due to anti-estrogenic effects (Bui et al, 1986), although the USEPA (2017) noted that findings have been inconsistent and the role of the estrogen receptor in the reproductive toxicity of BaP remains unclear.

Toxicity may also stem from binding of constituents to the AhR to induce P450 enzymes. The latter play a critical role in the oxidative metabolism of xenobiotics, and hence excretion from the body, but in some cases, metabolism leads to bioactivation and enhanced toxicity, as is the case for PAHs (Nebert et al, 2004). It is of interest therefore that the human placenta has high levels of AhR (Manchester et al, 1987). Support for this pathway contributing to compromised fetal growth comes from findings that chronic exposure to PAHs before gestation results in compromised maternal and fetal placental vascularisation and altered labyrinthine architecture in mice, contributing to IUGR, effects that are inhibited by AhR deficiency (Detmar et al, 2008).

Other potential mechanisms include the activation of apoptotic pathways (Ye et al, 2020) and binding of constituents to receptors for placental growth (epidermal growth factor (EGF) and insulin-like growth factors I and II) resulting in decreased feto-placental exchange and, consequently, impaired fetal nourishment and growth (Guyda, 1991; Zhang et al,1995). Using human trophoblastic choriocarcinoma cell lines of placental origin, Zhang et al (1995) showed that BaP mediates a loss of EGF receptors and in doing so, alters trophoblast proliferation as well as its endocrine function.

Products of metabolic activation of PAHs can also be transformed into PAH o-quinones by dihydrodiol dehydrogenase (Penning et al, 1999). That these molecules enter redox cycles and produce ROS (Park et al, 2006) implicates oxidative stress as another mechanism of PAH toxicity.

The high molecular weight PAHs (composed of 4 or more aromatic rings) such as BaP are not only those that are reported to be particularly toxic to human health because of their carcinogenic and mutagenic characteristics but are also one of the primary organic classes known to adsorb onto the surface of $PM_{2.5}$ (Allen et al,

1996). The mechanistic evidence on PAHs outlined above therefore supports a causality argument for $PM_{2.5}$. Epidemiological studies of adverse birth outcomes using PAHs as the pollutant metric were not considered within the scope of the meta-analyses and studies reviewed within this report. This was because PAHs are not monitored as widely as PM_{10} and $PM_{2.5}$ and tend to have less policy emphasis than the pollutants that were included. Nonetheless, given the mechanistic evidence discussed here, it would be useful to examine meta-analyses and additional epidemiological studies for PAHs, if indeed sufficient exist.

We have reviewed the European Environment Agency's mapping of BaP concentrations in the UK and Europe (EEA, 2020) which is presented in Appendix 14. Recommendations for future work into the effects of PAHs on birth outcomes is presented in section 5.1.3.

Epigenetic alterations

The fetal epigenome may provide a biologically plausible link between early life exposure to environmental factors and adverse pregnancy outcomes because DNA synthesis rates are high and epigenetic remodelling is extensive through critical stages of development (Hochberg et al, 2011). The embryo undergoes genome-wide DNA methylation, demethylation and re-methylation (Chaillet et al, 1994), whilst the placenta expresses a distinct subset of miRNAs that are believed to have a functional role in placental biology (Gu et al, 2013). Modification of epigenetic patterns (for instance hypo- or hypermethylation of genes) that control the expression of genes involved in key placental cellular processes has the potential to contribute to abnormal placental or fetal development.

Both global hypomethylation and hypermethylation of specific genes have been associated with cancer and other diseases in humans, while disturbance of maintenance DNA methylation in placental tissue is associated with early pregnancy loss in humans and abnormal embryonic implantation and development in the mouse model (Yin et al, 2012). Of further relevance, changes in placental DNA methylation patterns and MiRNA expression have been reported in association with fetal growth and adverse maternal exposures such as alcohol and tobacco smoke (Maccani and Marsit, 2009; Wilhelm-Benartzi et al, 2012).

Specific findings from studies described in section 3.3.5 above, of potential relevance to biological plausibility are summarised below:

- Studies have sought to explore the functional relevance of higher levels placental D-loop methylation (a crucial promoter region from which mtDNA replication and transcription are initiated) in mothers exposed to higher concentrations of $PM_{2.5}$ (Janssen et al, 2015; Vos et al, 2021) - both studies revealed an inverse association between placental mtDNA methylation and mtDNA content of placental tissue

- Saenen et al (2017) focused on placental methylation of the LEP promoter that plays a functional role in embryo implantation, intrauterine development and fetal growth (Sagawa et al, 2002) - LEP methylation status in the placenta was negatively associated with $PM_{2.5}$ exposure during the second trimester, and with

placental NTp, suggesting that oxidative/nitrosative stress might contribute to associations between $PM_{2.5}$ and placental epigenetic events

- Prenatal PAH exposure (assessed with personal monitoring) was associated with lower global methylation in umbilical cord white blood cells (Herbstman et al, 2012)

- Studies have found that maternal LINE-1 methylation levels are negatively correlated with PTB (Burris et al, 2012). Liu X et al (2019) also observed negative associations of PM_{10}, $PM_{2.5}$ and PM_1 concentrations with cord blood LINE-1 methylation levels, and a negative association between PM_1 concentration and maternal LINE-1 methylation level.

Overall, there is some support, from the rather limited mechanistic evidence reviewed in this report, that it is biologically plausible that exposure to ambient levels of air pollution, particularly particulate matter, could cause adverse birth outcomes. Many of the mechanisms identified appear relevant to TLBW, and stillbirth. Nitrosative/oxidative stress (for example through the nitration of placental proteins involved in trophoblast invasion and regulation of placental vascular reactivity), inflammation (for example through the expression of molecules involved in the formation of a dysfunctional placenta), PAH exposure (for example, through a possible anti-estrogenic effect) and epigenetic changes (for example by altering the expression of genes involved in key placental cellular processes) all have potential to disturb placental function and, as a consequence, appear relevant to compromised fetal growth and development as well as increasing the risk of stillbirth. In addition, given that inflammatory events take place in preparation for delivery and there is an increased inflammatory response during preterm delivery, inflammation is a plausible mechanism mediating the effects of air pollution on PTB.

4.5 Use of single pollutant models

As explained in the Introduction, this review has focused on associations reported from single pollutant models. Apart from the regulatory framework for air quality being on a specific air pollutant basis and from providing consistency with other quantification recommendations that use single-pollutant model recommendations, consideration of multi-pollutant models is complex. This is explained in more detail here.

The Quantification of Air Pollution Risks in the UK subgroup (QUARK) has discussed the use of multi-pollutant models to inform approaches to quantification of effects associated with air pollutants. They noted that “because concentrations of pollutants are often correlated, associations reported in epidemiological studies between pollutants and health outcomes may reflect the effects of a mixture of pollutants” (COMEAP, 2023). Therefore, a coefficient which has not been adjusted for effects associated with other pollutants (that is, a coefficient from a single-pollutant model) likely also reflects, to some extent, effects associated with other correlated pollutants. If effects estimated using a single-pollutant coefficient are added to estimates of the same effect associated with other pollutants, this will likely give an overestimate of the effects of the pollution mix.

Application of 2- or multi-pollutant models allows coefficients to be adjusted for effects associated with other pollutants to some extent. However, information on the concentrations of some potentially causal pollutants are not routinely available, and so cannot be adjusted for. In addition, there are difficulties in interpreting the results when pollutants are highly correlated and subject to measurement error. The effect can be transferred from the pollutant measured with more error to the pollutant measured with less. There is extensive discussion of this issue in the COMEAP (2018b) report on associations of long-term average concentrations of nitrogen dioxide with mortality. QUARK continues to consider methodological developments in this area and it may be possible to make more use of results from multi-pollutant models in future. In the meantime, coefficients for individual pollutants should be understood to partly include effects of other pollutants as well.

4.6 Is there a key exposure window?

This section discusses whether there is mechanistic or epidemiological evidence for a key exposure window for an effect of air pollution in pregnancy. Although it has not yet been possible to come to firm conclusions for effects of air pollution (section 4.7), it is still useful to collate such evidence as there is on this aspect at this stage.

Particulate pollutants have to cross at least 2 anatomical barriers to reach the fetal circulation and developing organs: lung and placenta for air pollutants, gut barrier and placenta for ingested pollutants and skin barrier and placenta following dermal exposure. The maternal blood entering the intervillous chamber of the placenta is limited by the presence of trophoblastic plugs during most of the first trimester of pregnancy. These plugs block the tip of the spiral arteries that are the terminal branches of uterine arteries and limit the passage to non-cellular components of the maternal blood (maternal plasma). Thus, theoretically there is less exposure of the fetal tissue to inhaled particles during the first trimester than in the second and third trimester, although this may depend on the particles' size, molecular weight, and mode of transport in maternal blood and through the villous trophoblast. However, maternally inhaled black carbon air pollution particles have been found to cross the placenta, enter the fetal circulation system and then translocate into human fetal organs (liver, lung and brain) during the first and second trimester of gestation (Bongaerts et al, 2022). These findings are especially concerning because this window of exposure is key to organ formation (embryogenesis) and development (organogenesis) and suggest that black carbon particles can enter the gestational sac from early in pregnancy.

For some exposures that affect fetal development, the timing results in different outcomes. Exposure to environmental hazards and other teratogens during gastrulation or primary morphogenesis (3-8 weeks post fertilization) is associated with the highest risk of early pregnancy loss and major anatomical anomalies (USEPA, 2024). Exposure during the late embryonic and fetal periods may result in abnormal organ growth and function (USEPA, 2024) and it is possible that chronic exposures may impair both the fetal and placental growth. Chemical exposures can also affect reproductive cells and have a transgenerational effect (Cheng et al, 2018).

For air pollution, the epidemiological evidence available from the evaluated meta-analyses does not support the identification of a specific trimester of pregnancy as a critical exposure window, for any of the birth outcomes examined. For PTB there was heterogeneity in findings for specific trimester exposures between different meta-analysis publications. For TLBW and stillbirth, however, there was little to distinguish meta-analysis findings for different specific trimester exposures, within and/or between meta-analysis publications. Similarly, the limited number of mechanistic epidemiological studies that have investigated perturbations (for instance changes in biomarker levels or DNA methylation status) during the entire pregnancy versus specific trimesters, also prevents any firm conclusions from being made relating to a critical time window of exposure to air pollution.

In summary, the current evidence does not provide information on a critical exposure window for effects of air pollution during pregnancy.

4.7 Is there a key air pollutant?

The epidemiological evidence from meta-analyses suggests the possibility of associations between particulate matter and both TLBW and PTB; and O_3 and PTB, but as set out at the start of this chapter, in all these cases the evidence from the meta-analyses is considered weak. The current epidemiological evidence in the meta-analyses provides no evidence for associations between NO_2 and adverse birth outcomes.

There is more evidence from the mechanistic literature of a role for $PM_{2.5}$ in eliciting adverse birth outcomes compared with other air pollutants but this is because the majority of studies have focused on $PM_{2.5}$. For example, a number (n=7) of biomarker studies have shown particularly consistent associations between exposure of pregnant women to PM and markers of oxidative stress in the placenta and/or cord blood, whilst far fewer studies have looked for associations between these markers and NO_2 (n=2) or O_3 (n=1). It should also be borne in mind that only one of the studies looking at $PM_{2.5}$, adjusted for NO_2 exposure. The findings were robust to this adjustment. A role for ambient particulate pollution (which can act as a carrier for PAHs) is also supported by (a) observations of associations between PM and levels of PAH-DNA adducts in maternal and cord blood and (b) evidence that elevated cord blood adduct levels is associated with a compromised head circumference.

Based on the current epidemiological and mechanistic evidence, we are unable to make a conclusion as to whether there is a key air pollutant driving any effects.

4.8 Implications

4.8.1 Contextualising with reports from authoritative bodies

U.S. Environmental Protection Agency Integrated Science Assessments

The US Environmental Protection Agency (USEPA) Integrated Science Assessments (ISAs) form the scientific foundation for the review of the US National Ambient Air Quality Standards. They are the most detailed and comprehensive reviews of epidemiological evidence and toxicological and mechanistic literature

base available (USEPA, 2015). For this reason, we focus on putting our report findings in the context of these reports.

In considering the conclusions from the USEPA ISAs, in the context of the conclusions reached in this report, it is important to note the following:

1) The USEPA and COMEAP have used different approaches to assess and make statements about the evidence.

   The USEPA ISAs evaluate the weight of evidence from both epidemiological and toxicological evidence jointly and apply a framework to classify the health effects evidence according to a 5-level hierarchy (1) Causal relationship; 2) Likely to be causal relationship; 3) Suggestive of, but not sufficient to infer, a causal relationship; 4) Inadequate to infer the presence or absence of a causal relationship; or 5) Not likely to be a causal relationship). The USEPA ISAs for particulate matter (USEPA, 2019) and ozone (USEPA, 2020) present overall statements/conclusions on the evidence for a causal relationship between a given air pollutant and "pregnancy and birth outcomes" as a group of outcomes (comprising maternal health during pregnancy (primarily gestational hypertension disorders and gestational diabetes), fetal growth, birth weight and body length at birth, PTB, birth defects and fetal/infant mortality). The USEPA ISA for oxides of nitrogen (USEPA, 2016) presents statements and conclusions on the evidence for a causal relationship between nitrogen dioxide and 'birth outcomes' as a group of outcomes (comprising fetal growth, PTB, birth weight, birth defects, and early life mortality) Unlike our report, there are no overall conclusions for each pollutant and specific birth outcomes.

   In our report, epidemiological evidence has been considered for each pollutant for each specific birth outcome, whereas toxicological/ mechanistic evidence has been considered according to specific mechanistic pathways. Concluding statements in this report are thus separate for the epidemiological and toxicological or mechanistic evidence, and not directly comparable with those from the USEPA ISAs. Nevertheless, we have attempted to tie in the mechanistic and epidemiological evidence (section 4.4) in considering biological plausibility.

   This report acknowledges that animal studies are less useful for understanding the underlying effects of air pollution on adverse birth outcomes in humans and in doing so, focuses on mechanistic evidence stemming from nested studies within epidemiological cohorts. Reservations about the relevance of animal models do not appear in the USEPA ISAs.

   The USEPA ISAs have focused on evaluating individual epidemiological studies. This report, with a view to answering question iii as stated in the Terms of Reference (Appendix 2) (iii. Is the evidence sufficient to propose quantification of specific adverse birth outcomes associated with air pollution in the UK) has focused on a detailed evaluation of existing meta-analyses of the epidemiological studies. The methodological approach therefore differs

between our report and the USEPA reports; however, the possibility that associations observed are due to chance, confounding or bias has been considered by both.

2) With respect to evidence relating to birth weight, our report examined a single outcome (term LBW) (again with a view to potential quantification), whereas the USEPA ISAs consider multiple birth weight outcomes (birth weight, LBW, term LBW, SGA, birth weight for gestational age (BWGA) z-scores).

3) USEPA ISAs for oxides of nitrogen, PM and O_3 were published in 2016, 2019 and 2020 respectively. The cut-off dates for included literature in these ISAs were August 2014, December 2017 and March 2018 respectively. The timescale for this COMEAP report differs and includes literature published up until March 2019 (with further epidemiological studies of relevance collated until March 2020, and relevant systematic reviews with meta-analyses considered for inclusion until August 2020, and further mechanistic evidence of relevance collated until June 2021.) Since publication of the USEPA ISAs, the evidence base has expanded greatly. Therefore, this report has considered some more recent evidence, compared to that of the USEPA ISAs.

Below, we provide a broad comparison of the main findings of assessments undertaken by the USEPA and COMEAP. It is important to note that if some conclusions of these reports appear divergent, this may simply reflect the differences in approach and timescale noted above.

$PM_{2.5}$ and birth outcomes

The USEPA (2019) concluded that the evidence was suggestive of, but not sufficient to infer, a causal relationship between exposure to $PM_{2.5}$ and pregnancy and birth outcomes. In summing up the epidemiological evidence, the USEPA (2019) noted "There are several well-designed, well-conducted studies that indicate an association between $PM_{2.5}$ and poorer birth outcomes, particularly low birth weight and preterm birth. Albeit, the collective evidence for many of the pregnancy and birth outcomes studies examined is not entirely consistent" (USEPA, 2019, pg. 9-43). We found weak evidence for an association between $PM_{2.5}$ and PTB and $PM_{2.5}$ and TLBW and no evidence to support an association with $PM_{2.5}$ and stillbirth from the core meta-analyses. After examination of key studies not captured by the meta-analyses, we concluded that up-to-date meta-analyses are needed to synthesise all the available evidence. The USEPA (2019) placed greater emphasis on animal toxicological data, assessing 4 additional inhalation studies. Plausible mechanisms by which exposure to $PM_{2.5}$ could alter birth weight or elicit a PTB were summarised as translocation of particles/soluble components, vascular endothelial dysfunction, placental or umbilical cord vascularity changes, altered umbilical cord structure (data from studies using laboratory rodents) and systemic inflammation or oxidative stress, epigenetic changes to the placenta, and impaired fetal thyroid function (data from human studies) (2019). Therefore, both our epidemiological and mechanistic conclusions align with that of the USEPA (2019) report.

PM_{10} and birth outcomes

The USEPA (2019) concluded that the evidence is inadequate to infer the presence or absence of a causal relationship between $PM_{10-2.5}$ exposure and pregnancy and birth outcomes. No toxicological or mechanistic data on $PM_{10-2.5}$ (in relation to pregnancy and birth outcomes) was included and it was concluded that a paucity of information exists in support of potential biological plausibility for $PM_{10-2.5}$ exposure and pregnancy and birth outcomes (2019).

We considered PM_{10} (which includes the $PM_{10-2.5}$ fraction), rather than $PM_{10-2.5}$, and concluded that there was weak evidence for an association between PM_{10} and PTB and PM_{10} and TLBW, and no evidence to support an association with PM_{10} and stillbirth, from the core meta-analyses. However, due to the inconsistent results of key new studies not captured by the meta-analyses, we conclude that up-to-date meta-analyses are needed to synthesise all the available evidence.

Our conclusions broadly align with that of the USEPA (2019) report with respect to coarser particulate matter. It is worth noting that the body of literature examining PM_{10} is much larger than that examining $PM_{10-2.5}$, so there is very limited overlap in terms of the epidemiological studies examined by the USEPA (2019) and this report. Further, the USEPA (2019) did not include any epidemiological studies examining stillbirth in relation to $PM_{10-2.5}$.

NO_2 and birth outcomes

The USEPA (2016) concluded that the evidence is suggestive of, but not sufficient to infer, a causal relationship between exposure to NO_2 and birth outcomes. In summing up the epidemiological evidence, the USEPA (2016) noted "studies that examined PTB, birth weight, birth defects, and infant mortality had inconsistent results (that is, positive and null associations). For PTB, many associations were very close to the null value." (USEPA, 2016, pg. 6-136). The USEPA's causality determination is driven by studies examining FGR (which were considered distinct from the evidence base for birth weight) as they conclude "collectively, the epidemiologic evidence for fetal growth restriction but uncertainty regarding an independent effect of NO_2 exposure is suggestive of, but not sufficient to infer, a causal relationship between NO_2 exposure and effects on birth outcomes" (USEPA, 2016, pg. 6-136). We did not consider studies examining FGR metrics.

In terms of epidemiological evidence, we concluded that there was no evidence from the core meta-analyses to support an association of prenatal exposure to NO_2 and PTB, NO_2 and TLBW and NO_2 and stillbirth. Due to the very large number of studies not captured in/published since the core meta-analyses, and the inconsistent results from key studies highlighted, we conclude that an up-to-date meta-analysis is needed to synthesise all the available evidence for NO_2 and PTB. For NO_2 and TLBW the consistency of results from the selected key new studies point to there being no association. For stillbirth and NO_2, the available evidence (meta-analyses and new studies) does not support an association. Updated meta-analyses are nonetheless recommended by COMEAP in each case.

The toxicological data assessed in the USEPA (2016) was limited to an animal study (not included in this report) dating back to 1971, that found that exposing rats to very high concentrations (1,300 ppb) of NO_2 for 3 months prior to breeding (rather than during gestation) produced pups with statistically significant decreased birth weights. The USEPA report concluded that although the biological mechanisms are not fully understood, several hypotheses have been proposed for the effects of NO_2 on reproductive and developmental outcomes. These include oxidative stress, systemic inflammation, vascular dysfunction, and impaired immune function, thus primarily aligning with findings from human mechanistic studies that were assessed in this report.

Our conclusion may appear divergent with the broad conclusion of the USEPA (2016) report regarding birth outcomes, as we do not view the epidemiological evidence as being suggestive of associations for the specific birth outcomes evaluated. However, our conclusion is aligned with the USEPA's summary regarding the inconsistency of epidemiological evidence relating to PTB, birth weight, and infant mortality (in which category stillbirth was grouped), even with the consideration of an increased evidence base for PTB, TLBW and stillbirth (we additionally considered evidence from 2015 to 2020).

O_3 and birth outcomes

The USEPA ISA for ozone and related photochemical agents (USEPA, 2020) concluded that the evidence is suggestive of, but not sufficient to infer, a causal relationship between O_3 exposure and pregnancy and birth outcomes.

Regarding the epidemiological evidence, the USEPA ISA noted inconsistency in the findings for PTB, fetal growth/birth weight measures, and fetal/infant mortality (into which category stillbirth was grouped), which is compatible with the views of COMEAP. The USEPA ISA summed up that "overall, there is consistent evidence for an association between ozone exposures during early to mid pregnancy with preterm birth in epidemiologic studies" (USEPA, 2020, pg. 7-14). "There are no toxicological studies examining preterm birth in animals or endpoints that are analogous to preterm birth," (USEPA, 2020, pg. 7-14) and that "overall, there is some epidemiologic evidence for the effects of ozone on fetal growth, especially for continuous-term birth weight, a conclusion supported by toxicological evidence in rodents" (USEPA, 2020, pg. 7-13).

From the epidemiological evidence stemming from the core meta-analyses, we concluded that there was weak evidence for an association between O_3 and PTB and no evidence to support an association of prenatal exposure between O_3 and TLBW and O_3 and stillbirth. Due to the inconsistent results of key new studies not captured by the meta-analyses, we conclude that up-to-date meta-analyses are necessary to synthesise all the available evidence for O_3 and PTB, and for O_3 and stillbirth. For O_3 and TLBW, new studies do not suggest an association, but we recommend an up-to-date meta-analysis, nonetheless.

The assessed toxicological data of an effect of O_3 on birth outcomes in the USEPA ISA (2020) was limited to animal inhalation studies that showed a decreased birth weight in pups whose pregnant dams were exposed to O_3 during pregnancy. The

latter included an additional 2 studies that were not included in this report. The biological pathways identified as potentially underlying effects were respiratory/systemic oxidative stress and inflammation, altered uterine artery vascularity of increased resistance and impaired thyroid status.

Therefore, for O_3, our views are broadly compatible with those of the USEPA (2020) report and the identified possible mechanistic pathways align with those discussed in this report. Any apparent divergence with respect to birth weight related outcomes may reflect that we examined only TLBW, whereas the USEPA statements are particularly driven by studies examining continuous birth weight.

Other reports from authoritative bodies

In addition to the USEPA ISAs, we highlight the Health Effects Institute State of Global Air report (HEI, 2020), 3 publications by the WHO (2005; 2013; 2018) and one from the Royal College of Physicians (RCP) and the Royal College of Paediatrics and Child Health (RCPCH) (RCP, 2016). Whilst these reports provide commentary on the effects of ambient pollutants on birth outcomes, they did not systematically review the epidemiological evidence or thoroughly examine the toxicological and mechanistic evidence base.

A working party of the RCP/RCPCH produced a report in 2016 that recognised that reduced fetal growth and birth weight had been linked to particulate air pollution, but that evidence was less consistent for gaseous pollutants (RCP, 2016).

The Health Effects Institute (HEI, 2020) provided a high-level overview of the topic published in 2020 and concluded that there was a growing body of evidence that showed that air pollution increases the risks of LBW and PTB, although stating that the biological reasons for this linkage are not fully known. The report surmises that air pollution may affect a pregnant woman, her developing fetus, or both through pathways similar to those of tobacco smoking, which is a well-known risk factor for LBW and PTB. Plausible mechanisms highlighted in the document were the movement of pollution particles or their components across the membranes of the lungs to other parts of the body to affect placental function and the fetus, systemic inflammation or oxidative stress (HEI, 2020). These high-level conclusions align with our detailed analysis of the literature and overall conclusions.

The 3 reports identified from the WHO were “Effects of air pollution on children’s health and development” (WHO, 2005), “Review of evidence on health aspects of air pollution – REVIHAAP Project” (WHO, 2013) and “Air pollution and child health: prescribing clean air” (WHO, 2018). These reports concluded that there is a growing body of epidemiological evidence that shows air pollution is associated with adverse birth outcomes. It was highlighted that there is evidence implicating air pollution in adverse effects on birth outcomes, but the strength of the evidence differs between outcomes (WHO, 2005). Whilst the overall evidence was deemed persuasive, the available data did not allow precise identification of specific pollutants and timing of exposure that can result in LBW, PTB, IUGR and birth defects (WHO, 2005. In terms of mechanistic data, the earlier (2005) report concluded that molecular epidemiological studies of DNA adduct levels in maternal blood and placentas suggest possible biological mechanisms for the effect on birth weight, premature

birth and intrauterine growth retardation, and support the view that the relationship between pollution and these pregnancy outcomes is genuine (WHO, 2005). Mechanistic data was not included in the REVIHAAP report but the potent oxidising and inflammatory properties of O_3 was discussed as a possible underlying mechanism (WHO, 2013). “Air pollution and child health: prescribing clean air again” did not review individual studies but included a high-level commentary (WHO, 2018). This highlighted the central role of the placenta to the health of the fetus, and that airborne pollutants reaching this crucial organ may cause significant damage. Plausible mechanisms explaining the association between air pollution and adverse birth outcomes were highlighted and included oxidative stress, pulmonary and placental inflammation, changes in fetal blood coagulation or endothelial function, and altered maternal haemodynamic response (WHO, 2018). The report also highlighted the studies that have identified epigenetic changes. All of these points align with our conclusions. In addition, this latest WHO report discussed the possibility of reduced maternal immunity and an associated increased risk of infection that may cause intrauterine growth restriction or preterm labour and an effect of air pollutants on the mother’s respiratory system in compromising the transport and delivery of oxygen and nutrients to the fetus (WHO, 2018).

Although the evidence base was not systematically reviewed in the Royal College of Physicians (RCP, 2016), Health Effects Institute (HEI, 2020) and World Health Organization 2005 and 2018 (WHO, 2005; WHO, 2018) reports, the broad overviews provided align with the detailed analysis of the literature and overall conclusions of our report.

Subsequent to the literature survey conducted for this report, a review by the Health Effects Institute (HEI, 2022) has examined the evidence linking traffic-related air pollution with various health effects, including adverse birth outcomes (HEI, 2022). The HEI review and meta-analyses only consider studies that used exposure assessments which were likely to specifically assess traffic related air pollution (TRAP), as opposed to air pollution exposure more generally, and the associations with single pollutants are regarded as representing the associations of the TRAP mixture. Thus it remains our view that new meta-analyses, representing the overall evidence base for the various pollutants, are needed in order to draw firm conclusions and inform recommendations for quantification.

4.8.2 Highlighting other literature

In this section, we briefly highlight some selected literature that represents both epidemiological and/or mechanistic evidence not included in our evaluation process in this report. Reasons that these studies were not included in our main evaluation are: some are systematic reviews published after our literature cutoff date (for example Simoncic et al, 2020, Zhang et al 2021, Ghosh et al, 2021); other studies did not meet our inclusion criteria because they looked at other pollutants such as traffic-related metrics (for example Wang et al, 2020); one study looked at multi-pollutant effects (for example Laine et al, 2020); others looked at short-term exposures/acute effects (for example Li et al, 2018; Stieb et al, 2019; Ranjbaran et al, 2020).

Other epidemiological literature

More recent systematic reviews

Wang et al. (2020)

Wang et al (2020) conducted a systematic review and meta-analysis of epidemiological studies of maternal exposures to traffic-related metrics (distance to roads, and traffic density) and adverse birth outcomes (PTB, TLBW and SGA). Whilst Wang et al (2020) does not examine pollutants specifically, one would expect that if there was an association with pollutants, there would be a relationship with distance to road. Wang et al (2020) report positive statistically significant associations between decreasing distance to road and SGA, and in sub-group analyses (higher quality score; cohort studies; North American studies) between increasing traffic density and LBW at term. No statistically significant associations were observed for PTB, but considerable heterogeneity was noted.

Simoncic et al. (2020)

Another paper published after the cut-off date was Simoncic et al (2020), that undertook a systematic review of European epidemiological studies examining the relationship between PM and NO_2 and LBW and PTB. Simoncic et al (2020) also conducted a limited number of meta-analyses, of which one is directly relevant to the meta-analyses considered in this report (PTB with whole pregnancy exposure to NO_2). The meta-analysis found a positive (adverse) association between PTB and NO_2 and birth weight decreases with NO_2 increase, but noted that the results were not statistically significant. We note that the number of studies included in each of these meta-analyses was small (n=4-5), so by the criteria used in this report, this would be classed as being based on limited evidence. Although Simoncic et al searched for relevant literature up to April 2020 there were relevant studies published prior to that date that they did not consider (Smith et al, 2020; Hannam et al, 2014). The systematic review captured one small study of PTB and NO_2 (Bertin et al, 2015) not captured by our search strategy, but no studies for TLBW not already captured by the search strategy used in this report, nor exposure-response estimates suitable for quantification, and does not alter our conclusions.

Zhang et al. (2021)

Zhang et al (2021) undertook a systematic review and meta-analysis of ambient air pollution ($PM_{2.5}$, PM_{10}, NO_2, and O_3) and stillbirth. The review searched for studies published in PubMed, Web of Science, Embase, and Cochrane Library until December 2020. The meta-analysis component included studies included in the Siddika et al (2016) meta-analysis (Faiz et al, 2012; Green et al, 2015; Hwang et al, 2011; Kim et al, 2007), but also included DeFranco et al (DeFranco et al, 2015), which Siddika et al (2016) excluded as they thought it to be of too low quality. In addition, most of the studies identified as key selected literature in our report were included in the meta-analyses (Smith et al, 2020; Rammah et al, 2019b; Xue et al, 2019; Yang et al, 2018; Zang et al, 2019; Mendola et al, 2017) with the exception of Ebisu et al (2018). We have not evaluated the quality of this meta-analysis, as it was published after our cut-off date. The meta-analyses of long-term exposures showed

statistically significant positive (adverse) associations with the risk of stillbirth for entire pregnancy $PM_{2.5}$, third trimester $PM_{2.5}$, and first trimester O_3. The number of studies in each meta-analysis ranged from 4 to 8.

The Zhang et al (2021) meta-analysis gives no reason to alter our earlier conclusions of no consistent evidence of associations of stillbirths with PM_{10} exposure and no evidence of associations of stillbirths with NO_2 exposure.

There were some differences from our findings for $PM_{2.5}$ and Ozone, which we discuss in more detail below.

a) In our earlier work, we concluded that the limited evidence from the meta-analysis available did not support association of stillbirth with $PM_{2.5}$., but that the evidence from key additional studies was inconsistent, and that an up-to-date meta-analysis was needed. The Zhang et al meta-analysis reports associations of stillbirth with $PM_{2.5}$ exposure in Whole Pregnancy (OR 1.10 95% CI 1.07-1.13 per 10 $\mu g/m^3$ based on 7 studies all reporting positive associations) and trimester 3 (1.09 95% CI 1.01-1.18 per 10 $\mu g/m^3$ based on 5 studies all reporting positive associations). We do not alter our earlier conclusion recommending a new meta-analysis because Zhang et al (2021) includes:

- one study excluded from an earlier meta-analyses conducted by Siddika et al (2016), as they considered it to be of low quality
- estimates from studies that have not adjusted for differences in gestational length between stillbirths and other pregnancies
- one strongly positive estimate for Whole Pregnancy from a study where stillbirth was defined as occurring from 20 to 42 weeks (even though the study also reported a much weaker association when stillbirth was defined as occurring after 28 weeks)
- a large multinational study from Africa where stillbirth is self-reported, air pollution levels are very high and where authors felt there was likely a threshold effect (this study providing more than 40% of the weight of the final Zhang et al estimate for whole pregnancy), although it should be noted that the meta-analysis result is robust even when this study is excluded

b) Additionally, we identified a couple of errors in the Whole Pregnancy $PM_{2.5}$ meta-analysis by Zhang et al. The OR (95% CI) included in the forest plot for Mendola et al (2017) is incorrect; the original estimate has a wide confidence interval spanning 1 (OR 1.08 (0.75-1.55) per 4.7 $\mu g/m^3$, which converts to OR 1.178 (0.542, 2.541) per 10 $\mu g/m^3$, not 1.084 (1.014, 1.155) per 10 $\mu g/m^3$ as used in the forest plot. Mendola et al (2017) was one of 2 key highlighted studies in our report (section 2.2.4) driving our conclusion of inconsistent evidence. Mendola et al (2017) contributes 16% weight to the $PM_{2.5}$ Whole Pregnancy meta-analysis. Given that the correct OR would be higher, but the confidence interval much wider, it is possible that if the correct estimate was included the pooled results may not have reached statistical significance. We also note that ORs and CIs for DeFranco et al (2015) are

incorrectly interpreted as representing continuous analyses (variably with Zhang et al (2021) as a per 1 μg/m^3 $PM_{2.5}$ increase, but converted as if they represent a per IQR increase), when in fact they represent OR (CI) for High exposure vs reference category (for example for whole pregnancy High exposure >=15.67 μg/m^3 vs reference <15.67 μg/m^3). This latter error reduces our confidence in the meta-analysis overall but is unlikely to influence the meta-analysis results much, as DeFranco et al (2015) contributes a relatively low weight to the meta-analyses.

c) In our earlier work, we concluded that whilst the meta-analysis evidence did not support an association between stillbirth and ozone, there was some inconsistent evidence of associations of stillbirth with ozone in key highlighted studies, although the small number of studies preclude robust evaluation of whether this is related to a trimester specific exposure or to whole pregnancy exposure. Zhang et al (2021) reports associations of stillbirth with ozone exposure in trimester 1 (OR 1.03 95% CI 1.00-1.05 per 10 μg/m^3; I^2 73.5%; estimates from 6 studies) but only weak positive associations below conventional levels of significance for trimester 2 and for Whole Pregnancy, and a null result for trimester 3. Estimates include those from 2 studies that have not adjusted for differences in gestational length between stillbirths and other pregnancies and studies with no adjustment for maternal smoking. As we concluded earlier, a systematic review may be helpful although may be limited by the small number of studies and the limitations this would place on interpretation of a formal meta-analysis.

Ghosh et al. (2021)

Ghosh et al (2021) undertook a systematic review, meta-analysis, and attributable global burden assessment of $PM_{2.5}$ and 4 birth outcomes (birth weight, gestational age, LBW, and PTB). The review searched for articles published anytime until April 4, 2021 in PubMed, Embase, Web of Science, Scopus, Current Contents, Global Health, Cochrane, Toxline, Canadian Research Index, and reference lists of all relevant articles.

Ghosh et al (2021) quantified the relationship between average ambient $PM_{2.5}$ exposure during entire pregnancy and 3 birth outcomes (birth weight, LBW, and PTB) using standard meta-analysis. The authors then generated non-linear exposure response curves for each outcome (using meta-regression-Bayesian regularized trimmed (MRBRT) method) covering the global range of exposures, including those from both ambient and household air pollution. Lastly, Ghosh et al (2021) estimated the global proportion of LBW and PTB attributable to exposures to both ambient and household $PM_{2.5}$ and estimated the attributable reduction in birth weight and gestational age at birth. The burden estimation used the risks from the MRBRT exposure-response curves.

We have not evaluated the quality of this meta-analysis, as it was published after our cut-off date. However, we would make the following observations:

In our report we identified studies (up to March 2020) which did not appear in our core meta-analyses. We note that 9 such relevant papers for $PM_{2.5}$ and PTB (Chang et al, 2015; Padula et al, 2019; Mendola et al, 2016; Rappazzo et al, 2014; Zhu et al,

2019; Poirier et al 2015, Siddika et al 2019, Liu Y et al 2019, Zhao 2020) and 6 such papers for $PM_{2.5}$ and TLBW (Poirier et al, 2015; Wilhelm et al, 2012; Nieuwenhuijsen et al, 2019; Enders et al, 2019; Coker et al, 2016; Brown et al, 2015) were not mentioned/included by Ghosh et al. Ghosh also identified several papers, both for PTB and TLBW, which we did not capture/consider. This may reflect different search strategies/selection criteria between our approaches. Ghosh et al's literature search cut-off date was April 2021, so identified some studies published after the March 2020 cut-off date we applied. Of the studies we identified as not appearing in the core meta-analyses, we specifically highlighted key studies for each pollutant-outcome pair that we considered would be likely to influence, or be important to note, over and above the evidence already provided in the core meta-analyses. For $PM_{2.5}$ and PTB, we highlighted 7 key studies (Laurent et al, 2016a; Giorgis-Allemand et al, 2017; Chen et al, 2018, Liang et al, 2019b; Sheridan et al, 2019; Smith et al, 2020; Guo et al, 2018). Five out of these 7 were included in the Ghosh et al 2021 meta-analysis. Ghosh et al provided reasons for not including the remaining 2 (Smith et al, 2020 did not examine entire pregnancy exposure, and Liang et al, 2019b was not included as it reported stratified results). For $PM_{2.5}$ and TLBW, we highlighted 4 key studies (Ha et al, 2017; Pedersen et al, 2013; Poirier et al, 2015 and Brown et al, 2015). Two out of these 4 were included in the Ghosh et al 2021 meta-analysis, but 2 (Poirier et al, 2015; Brown et al, 2015) were not captured in their review/meta-analysis.

Ghosh et al (2021) include a larger number of studies in their meta-analysis (n=40 for PTB, n=40 for LBW) than any of the core meta-analyses we reviewed for Whole Pregnancy $PM_{2.5}$ and PTB or TLBW (the largest included 8 and 6 estimates, respectively (see Appendix 11, Table A 36 and Table A 40). However, for LBW, Ghosh et al (2021) did not perform a meta-analysis specifically focused on term LBW, instead pooling studies with and without restriction to term births (for LBW and birth weight meta-analyses). Nonetheless, 31 out of 40 LBW studies included in their meta-analysis were restricted to term.

Ghosh et al (2021) report meta-analysis summary estimates of 11% greater risk (1.11 95% CI 1.07, 1.16) of LBW, and 12% greater risk (1.12 95% CI 1.06, 1.19) of PTB, each associated with a 10 $\mu g/m^3$ increase in the entire pregnancy average $PM_{2.5}$ exposure. Heterogeneity was considerable in all cases (>94%).

In sensitivity analyses limiting LBW analyses to studies which did not restrict to term births (9 out of 40 for LBW), the estimates moved away from the null (LBW: 1.25 (95% CI: 1.06, 1.48)). Ghosh et al (2021) suggest that this demonstrates collider bias due to adjustment for gestational age (a concept noted by Wilcox et al (2011). Whilst this could be one possible explanation, we note that 6 out of 9 LBW studies adjusted for gestational age as a continuous covariate (even though they are not restricted to term births), and thus these sensitivity analyses do not give estimates for studies with no adjustment for gestational age. We are not therefore convinced that these results demonstrate collider bias. If part of the association between $PM_{2.5}$ and LBW is mediated through preterm delivery, removing the effect of preterm delivery from LBW would reduce the association, and this may well explain the observed results. Nonetheless, we do acknowledge there is a small risk of collider stratification bias when limiting the examination of LBW to term births, but this is

balanced against the utility of clarifying the effect of an exposure on fetal growth, independent of PTB. It is difficult to interpret birth weight outside the context of gestational age. Future studies on LBW could present results both with and without adjustment for gestational age/restriction to term, to allow greater consideration of this issue.

The authors have gone beyond previous systematic reviews and meta-analyses on this topic, by incorporating non-linear meta-regression and global burden of disease assessment. We have recommended that an up-to-date systematic review and meta-analysis is required to synthesise all the available evidence. We emphasise that this recommendation still stands, due to the following factors:

a) Ghosh et al (2021) only considered entire pregnancy exposure whereas our report considered and commented on the meta-analyses and selected key studies for whole pregnancy, trimester 1, 2 and 3

b) For TLBW, 2 studies we highlighted as likely to be influential were not captured by the Ghosh et al (2021) review/meta-analysis

c) Ghosh et al 2021 did not perform a meta-analysis specifically focused on term LBW, instead pooling studies with and without restriction to term births (for LBW and birth weight meta-analyses), and then doing sensitivity analyses just on the studies which did not restrict to term birth, whereas our report only considers term LBW (Appendix 5)

d) Whilst Ghosh et al (2021) included much larger numbers of studies in meta-analyses compared to the core meta-analyses we were able to examine, we note that Ghosh et al's meta-analyses for both PTB and LBW contain some population overlaps that we considered intolerable when we encountered these in meta-analyses we evaluated (and subsequently excluded as a result) (see Appendix 9, Table A 4, and Table A 20 respectively). Therefore, the assumption of independence may not be met, leading to underestimation of standard errors on the summary effect, giving the appearance of greater precision than is justified. Nonetheless we do note that, given the large number of studies included in Ghosh et al (2021), no particular study contributes more than 4% of the weight to either meta-analysis.

The Ghosh et al (2021) findings are supportive of associations between exposure to $PM_{2.5}$ during pregnancy and both LBW and PTB, and compatible with our conclusions based on the core meta-analyses (which found weak evidence for associations). We would, nonetheless, still recommend an up-to-date meta-analysis for both $PM_{2.5}$ and PTB, and $PM_{2.5}$ and TLBW particularly, to synthesise all the available evidence for all exposure windows, to examine population overlaps/ independence of effect estimates in more detail, and to provide estimates for term LBW specifically.

Epidemiological studies examining acute exposure to air pollution

A small number of epidemiological studies examining short-term exposure/acute effects were identified. For example, 2 of the largest studies examining acute effects of $PM_{2.5}$ on risk of PTB were Stieb et al (2019) and Li et al (2018). Stieb et al (2019) covered 24 cities in Canada and over 1 million births, analysing short-term exposures in the week prior to birth using survival analysis. Li et al (2018) was a time-series study in Hubei Province, China, with over 1.2 million births. Neither of these 2 large studies found any statistically significant associations between PTB with short term exposure to $PM_{2.5}$ in the week prior to delivery (Li et al, 2018; Stieb et al, 2019). In addition, Stieb et al (2019) examined the acute effects of NO_2 on risk of PTB and found a statistically significant inverse association between PTB and exposure to NO_2 in the week prior to delivery. Ranjbaran et al (2020) also examined the acute effects of NO_2 on risk of PTB in a study conducted in Tehran, Iran with over 500,000 births but no effect was identified. A meta-analysis mentioned above, Zhang et al (2021) identified after the cut-off date, also examined short-term exposures. The study reported one statistically significant pooled estimate (O_3 on lag day 4 associated with stillbirth), but the number of studies in each meta-analysis was very small (n=2 to 4), which limits interpretation of these meta-analyses.

Other mechanistic literature

Laine et al (2020) took what they referred to as an 'exposome' approach to determine the effects of prenatal exposure to multiple air pollutants ($PM_{2.5}$, PM_{10}, NO_2, NO_x, ultrafine particles and OP of $PM_{2.5}$) on infant birth weight in a cohort (n=498) pooled from 4 individual European birth cohorts (ENVIRONAGE Belgium, INMA Spain, Rhea Greece, Piccolipiu Italy). This was achieved by incorporating multi-pollutants into models and by assessing mechanistic interactions between metabolomic signatures and targeted inflammatory proteins in cord blood. Exposure to air pollution mixtures, from 3 different mixture pollutant models, was found to be negatively associated with infant birth weight. A 3-pollutant mixture model of $PM_{2.5}$, PM_{10}, and NO_2 estimated that comparing the 75 percentiles to the median level of exposure to the air pollutant mixture could lead to an approximate 96 g decrease in birth weight. This relationship between prenatal exposure to air pollution and birth weight was found to be mediated by inflammatory proteins (IL-17 and epidermal growth factor) and 665 metabolomic features. Moreover, the latter included several features involved in fetal development and growth, including arginine (n=56) and tyrosine metabolism (n=58).

4.8.3 Policy implications

The questions that we aimed to address, as stated in our terms of reference (Appendix 2) are highlighted below:

i. What is the strength and consistency of epidemiological evidence linking ambient air pollutants with adverse birth outcomes?

ii. Are there biologically plausible mechanisms by which air pollutants could affect adverse birth outcomes?

iii. Is the evidence sufficient to propose quantification of specific adverse birth outcomes associated with air pollution in the UK?

iv. Can recommendations be identified from current meta-analyses or should new meta-analyse(s) be undertaken?

The report's description of the evidence in response to (i) and (ii) above should assist in describing the latest thinking in this area. The report also shows that a clear answer for (iii) will only be possible if further work on new meta-analyses is carried out, as recommended as a result of the consideration of question (iv). This means that quantification of effects of air pollution on adverse birth outcomes is not recommended at this stage. However, this is not the result of a clear indication that air pollution does not have an effect on adverse birth outcomes, and this should be reflected in any discussion of omissions from the calculations of benefits of reductions in air pollution. New meta-analyses would provide a clearer position in this respect.

The policy response of reducing air pollution in general would be unaffected should there be a finding of a trimester-specific response.

Information on trimester-specific risks might affect a policy of providing advice to women to avoid higher levels of pollutant exposure, but as this is often difficult to do on an individual basis, the main policy would likely remain that of general reduction in population exposure to air pollution. Therefore, whether any effect of air pollution is or is not trimester-specific is not an essential distinction and has only minor policy implications compared with whether air pollution has effects on adverse birth outcomes or not.

Chapter 5
Conclusions

The Committee set out to respond to the following questions as outlined in Appendix 2.

i. What is the strength and consistency of epidemiological evidence linking ambient air pollutants with adverse birth outcomes?

The epidemiological evidence was summarised in 2 steps:

1. From the core meta-analyses (Figure 5), a commentary on the evidence was prepared for each pollutant, outcome, and exposure window (section 2.2).

2. Key studies that had either been recently published or were not included in the core meta-analyses were identified (section 2.1.4), and commentaries made on the evidence from these (section 2.2).

Preterm birth (PTB)

Based on the meta-analyses available there is weak evidence for an association between PTB and exposure to $PM_{2.5}$, PM_{10}, and O_3, but the evidence does not support an association with NO_2 exposure. For O_3, this conclusion is based on limited evidence from the meta-analyses. For each pollutant there were a very large number of studies not captured in/published since the core meta-analyses and findings from key studies examined were inconsistent for each pollutant-outcome pair.

Term low birth weight (TLBW)

Based on the meta-analyses available there is weak evidence for an association between TLBW and exposure to $PM_{2.5}$, PM_{10}, but the evidence does not support an association with NO_2 or O_3 exposure. This conclusion is based on limited evidence from the meta-analyses for $PM_{2.5}$ and O_3. For each pollutant there were numerous studies not captured in/published since the core meta-analyses. Findings from key additional studies examined were inconsistent for $PM_{2.5}$ and PM_{10}. For NO_2 and O_3 the findings of key additional studies examined are in line with the meta-analysis conclusions, that is, they do not support an association.

Stillbirth

The single meta-analysis available does not support an association between stillbirth and exposure to $PM_{2.5}$, PM_{10}, NO_2 and O_3. This conclusion is based on limited evidence from the meta-analysis. For each pollutant there were several studies not captured in/published since this core meta-analysis. Findings from key additional studies examined were inconsistent for $PM_{2.5}$, PM_{10} and O_3. For NO_2 the findings of key additional studies examined are in line with the meta-analysis evidence conclusion, that is, they do not support an association.

Overall

Overall, there was weak evidence for associations between some air pollutants and adverse birth outcomes from the available meta-analyses. There were often many studies not reflected in the meta-analysis evidence, and for most pollutant-outcome pairs up-to-date meta-analyses are required to synthesise the full evidence base.

ii. Is there evidence of biological mechanisms by which air pollutants have the potential to adversely affect birth outcomes?

The presence of carbon and metal-bearing nanoparticles has been detected in human placenta of mothers. However, it remains unclear whether, what appears to be, a small number of particles compared with those in airway tissue, have the capacity to alter placental development and function and adversely affect the developing fetus.

Overall, there is some support, from the rather limited mechanistic evidence available and reviewed in this report, that it is biologically plausible that exposure to ambient levels of air pollution, particularly particulate matter, could cause adverse birth outcomes. Many of the mechanisms identified appear relevant to TLBW, and stillbirth. Nitrosative/oxidative stress (for example through the nitration of placental proteins involved in trophoblast invasion and regulation of placental vascular reactivity), inflammation (for example through the expression of molecules involved in the formation of a dysfunctional placenta), PAH exposure (for example, through a possible anti-estrogenic effect) and epigenetic changes (for example by altering the expression of genes involved in key placental cellular processes) all have potential to disturb placental function and, as a consequence, appear relevant to compromised fetal growth and development as well increasing the risk of stillbirth. In addition, given that inflammatory events take place in preparation for delivery and there is an increased inflammatory response during preterm delivery, inflammation is a plausible mechanism mediating the effects of air pollution on PTB.

iii. Is the evidence sufficient to propose quantification of specific adverse birth outcomes associated with air pollution in the UK?

In general, the mechanistic evidence provides some biological plausibility for a causal role of air pollution on adverse birth outcomes. At the level of specific pollutant-outcome pairs, this is less certain.

The current evidence does not provide information on a critical exposure window for effects of air pollution during pregnancy, should these be identified.

COMEAP does not consider it appropriate to make recommendations for quantification given the need for an up-to-date meta-analysis. There is some evidence for an association between air pollution and birth outcomes. However, it is not sufficient to undertake quantification at this stage.

iv. Can recommendations be identified from current meta-analyses or should new meta-analyse(s) be undertaken?

We advise that new meta-analyses should be undertaken.

Chapter 6 Recommendations for future work

In this chapter, we make some recommendations for future work. We also make a number of suggestions regarding methods, reporting and publication of epidemiological studies, systematic reviews and meta-analyses that we think would help improve the evidence base in the future.

Recommendations for future work

- New meta-analyses are required to synthesise all of the available evidence and enable recommendations for quantification to be made, if appropriate.
- A new systematic review of the epidemiological evidence for PAHs.
- More studies investigating the mechanisms by which air pollutants impact on birth outcomes are required

Considerations for future meta-analyses

- Greater transparency in reporting of methodology, and decisions made regarding study selection, would help ensure that meta-analyses are replicable and would allow future meta-analyses to build on those undertaken previously.
- We suggest the use of improved and tailored quality assessment that is appropriate to environmental epidemiology studies. Commonly used quality assessment scales, such as the Newcastle-Ottawa scale, are not designed for environmental epidemiology studies, and do not adequately assess exposure assessment (which can vary greatly in method and spatial and temporal resolution) and the risk of specific biases which may arise in this field.
- Detailed assessment of geographical/population overlaps between individual studies should be undertaken before pooling studies, to ensure independence of estimates entered into meta-analyses. Transparency of reporting the decisions made to include/exclude studies based on population overlaps is needed.
- Consideration of the impact of varying exposure assessment over time would be valuable, for example by using sensitivity analyses.

- Focusing only on a limited part of the available evidence base (for example, a limited time window or geographical bounds) may lead to the omission of relevant information and introduce bias into evidence synthesis. We consider that it would be preferable to explore the impacts of different selection criteria (such as time window or geography), using, for example, sensitivity analyses and meta-regression.

- A list of all studies included in a systematic review should be provided (for example as appendices or supplementary material). Reasons for the exclusion of any of the studies from subsequent meta-analysis should be given. Data extracted from original papers, and used in meta-analysis, including any transformations made before pooling, should be provided.

- Several of the systematic reviews identified had quality issues that meant they were not suitable for policy formulation. The quality of the published evidence base could be improved by critical review of meta-analyses by editors/peer-reviewers before publication, reflecting the issues we have highlighted above and in section 4.3.1 of this report. Thorough reviewing of systematic reviews and meta-analyses is an extremely time-consuming process and more generous and realistic journal timescales may be needed if peer reviewers are to be able to complete this process to the appropriate standard.

Considerations for future individual epidemiological studies

- Use of multi-pollutant models would be valuable.

- Studies investigating PAHs in relation to adverse birth outcomes would be valuable. These would also be informative regarding possible associations with PM, as PAHs adsorb to PM.

- We suggest a focus on adverse birth outcomes (for example stillbirth) where the evidence base is currently very limited or insufficient to perform robust meta-analyses.

- Reporting standard numerical information for all analyses and results referred to in a paper (such as numerical values for effect estimates, confidence intervals, exposure increments used for analysis, exposure category boundaries) would be valuable as it would allow the study to be incorporated into future meta-analysis. If some or all results are presented only graphically, or simply alluded to in text, then relevant information from the study may be omitted from future meta-analyses, unless these can be obtained direct from the study authors, which is not always possible.

Considerations for future mechanistic studies

Epidemiological mechanistic studies

- Future studies should consider ways of overcoming the limitation of low study power, for example by greater collaboration between birth cohorts. This might allow links to be made between mechanistic variables of interest (for example an increase in a placental marker of oxidative stress) and adverse birth outcomes

- For smaller scale studies, where assessment of mechanisms requires expensive analysis of placental and cord blood samples, increased precision in exposure estimates is recommended. This could perhaps be achieved by reliable, low-cost personal monitoring of PM and NO_2.

- The use of multi-pollutant models within mechanistic research should be encouraged to investigate potential additive/synergistic effects of the overall pollutant mix, and to distinguish the independent effects of different pollutants.

- Studies should routinely adjust for all relevant potential confounders.

- Research into indoor air pollutants, facilitated by the use of personal monitors.

- Studies should focus on a priori defined variables from published studies (to date it is often unclear whether associations are reported selectively - reporting only significant associations between exposure and mechanistic variables).

Experimental mechanistic studies

- Since PM has been observed in human placental cells, the effect of realistic concentrations of translocated particles on these cells over long periods of time could be examined in sophisticated in vitro systems (for instance relevant multicellular co-culture models) and in animal models.

- More mechanistic studies that assess the effects of ambient gaseous pollutants on adverse birth outcomes are recommended. The majority of mechanistic studies to date have focused on potential pathways underlying effects of airborne particulate matter on adverse birth outcomes.

- Use of cross-omic techniques should be encouraged, to help identify potential interacting genes and pathways that may mediate molecular mechanisms.

References

Allen JO et al. 'Measurement of Polycyclic Aromatic Hydrocarbons Associated with Size-Segregated Atmospheric Aerosols in Massachusetts'. Environmental Science & Technology 1996: volume 30(3), pages 1023-1031

Amarilyo G et al. 'Increased cord serum inflammatory markers in small-for-gestational-age neonates'. Journal of Perinatology 2011: volume 31(1), pages 30-32

Ambroz A et al. 'Impact of air pollution on oxidative DNA damage and lipid peroxidation in mothers and their newborns'. International Journal of Hygiene and Environmental Health 2016: volume 219(6), pages 545-556

Anderson S et al. 'Sequence and organization of the human mitochondrial genome'. Nature 1981: volume 290(5806), pages 457-465

Araban M et al. 'Air pollution and low birth weight: a historical cohort study from Tehran'. Eastern Mediterranean Health Journal 2012: volume 18(6), pages 556-560

Arroyo V et al. 'Premature births in Spain: Measuring the impact of air pollution using time series analyses'. Science of the Total Environment 2019(a): volume 660, pages 105-114

Arroyo V et al. 'Impact of air pollution on low birth weight in Spain: An approach to a National Level Study'. Environmental Research 2019(b): volume 171, pages 69-79

Balakrishnan K et al. 'Exposures to fine particulate matter $PM_{2.5}$ and birthweight in a rural-urban, mother-child cohort in Tamil Nadu, India'. Environmental Research 2018: volume 161, pages 524-531

Bartolomei MS et al. 'Parental imprinting of the mouse H19 gene'. Nature 1991: volume 351(6322), pages 153-155

Basu S. 'Fatty acid oxidation and isoprostanes: oxidative strain and oxidative stress'. Prostaglandins, Leukotrienes, and Essential Fatty Acids 2010: volume 82(4-6), pages 219-225

Basu R et al. 'Effects of fine particulate matter and its constituents on low birth weight among full-term infants in California'. Environmental Research 2014: volume 128, pages 42-51

Basu R et al. 'Association between $PM_{2.5}$ and $PM_{2.5}$ Constituents and Preterm Delivery in California, 2000-2006'. Paediatric and Perinatal Epidemiology 2017: volume 31(5), pages 424-434

Bekkar B et al. 'Association of Air Pollution and Heat Exposure With Preterm Birth, Low Birth Weight, and Stillbirth in the US: A Systematic Review'. JAMA Network Open 2020: volume 3(6), article e208243

Bell ML et al. 'Ambient air pollution and low birth weight in Connecticut and Massachusetts'. Environmental Health Perspectives 2007: volume 115(7), pages 1118-1124

Bell ML et al. 'Prenatal exposure to fine particulate matter and birth weight: variations by particulate constituents and sources'. Epidemiology 2010: volume 21(6), pages 884-891

Benediktsson R et al. 'Glucocorticoid exposure in utero: new model for adult hypertension'. The Lancet 1993: volume 341(8841), pages 339-341

Bertin M et al. 'Association between prenatal exposure to traffic-related air pollution and preterm birth in the PELAGIE mother-child cohort, Brittany, France. Does the urban-rural context matter?'. Environmental Research 2015: volume 142, pages 17-24

Blazer S et al. 'Maternal hypothyroidism may affect fetal growth and neonatal thyroid function'. Obstetrics and Gynecology 2003: volume 102(2), pages 232-241

Bongaerts E et al. 'Maternal exposure to ambient black carbon particles and their presence in maternal and fetal circulation and organs: an analysis of two independent population-based observational studies'. The Lancet Planetary Health 2022: volume 6(10), pages e804-e811

Borm PJ et al. 'Oxidant generation by particulate matter: from biologically effective dose to a promising, novel metric'. Occupational and Environmental Medicine 2007: volume 64(2), pages 73-74

Bosco C et al. 'Oxidative damage to pre-eclamptic placenta: immunohistochemical expression of VEGF, nitrotyrosine residues and von Willebrand factor'. Journal of Maternal-Fetal & Neonatal Medicine 2012: volume 25(11), pages 2339-2345

Bouhours-Nouet N et al. 'Maternal smoking is associated with mitochondrial DNA depletion and respiratory chain complex III deficiency in placenta'. American Journal of Physiology Endocrinology and Metabolism 2005: volume 288(1), pages E171-177

Bové H et al. 'Ambient black carbon particles reach the fetal side of human placenta'. Nature Communications 2019: volume 10(1), article 3866

Brauer M et al. 'A cohort study of traffic-related air pollution impacts on birth outcomes'. Environmental Health Perspectives 2008: volume 116(5), pages 680-686

Brown MB et al. 'M1/M2 macrophage polarity in normal and complicated pregnancy'. Frontiers in Immunology 2014: volume 5, article 606

Brown JM et al. 'Linking air pollution data and adverse birth outcomes: environmental public health tracking in New York State'. Journal of Public Health Management and Practice 2015: volume 21 Suppl 2, pages S68-74

Brunst KJ et al. 'Racial/ethnic and sociodemographic factors associated with micronutrient intakes and inadequacies among pregnant women in an urban US population'. Public Health Nutrition 2014: volume 17(9), pages 1960-1970

Brunst KJ et al. 'Prenatal particulate matter exposure and mitochondrial dysfunction at the maternal-fetal interface: Effect modification by maternal lifetime trauma and child sex'. Environment International 2018: volume 112, pages 49-58

Bui QQ et al. 'A comparative study of the reproductive effects of methadone and benzo[a]pyrene in the pregnant and pseudopregnant rat'. Toxicology 1986: volume 42(2-3), pages 195-204

Burris HH et al. 'Associations of LINE-1 DNA Methylation with Preterm Birth in a Prospective Cohort Study'. Journal of Developmental Origins of Health and Disease 2012: volume 3(3), pages 173-181

Burton GJ et al. 'Nutrition of the human fetus during the first trimester--a review'. Placenta 2001: volume 22 Suppl A, pages S70-77

Burton GJ et al. 'Human early placental development: potential roles of the endometrial glands'. Placenta 2007: volume 28 Suppl A, pages S64-69

Butler L et al. 'Residential proximity to roadways and placental-associated stillbirth: a case-control study'. International Journal of Environmental Health Research 2021: volume 31(4), pages 465-474

Cai J et al. 'Exposure to particulate air pollution during early pregnancy is associated with placental DNA methylation'. Science of the Total Environment 2017: volume 607-608, pages 1103-1108

Campbell HE et al. 'Healthcare and wider societal implications of stillbirth: a population-based cost-of-illness study'. British Journal of Obstetrics and Gynaecology 2018: volume 125(2), pages 108-117

Capobussi M et al. 'Air Pollution Impact on Pregnancy Outcomes in Como, Italy'. Journal of Occupational and Environmental Medicine 2016: volume 58(1), pages 47-52

Carvalho MA et al. 'Associations of maternal personal exposure to air pollution on fetal weight and fetoplacental Doppler: A prospective cohort study'. Reproductive Toxicology 2016: volume 62, pages 9-17

Catov JM et al. 'Inflammation and dyslipidemia related to risk of spontaneous preterm birth'. American Journal of Epidemiology 2007: volume 166(11), pages 1312-1319

Chaillet JR. 'Genomic imprinting: lessons from mouse transgenes'. Mutation Research 1994: volume 307(2), pages 441-449

Chang HH et al. 'Time-to-event analysis of fine particle air pollution and preterm birth: results from North Carolina, 2001-2005'. American Journal of Epidemiology 2012: volume 175(2), pages 91-98

Chang HH et al. 'A spatial time-to-event approach for estimating associations between air pollution and preterm birth'. Journal of the Royal Statistical Society Series C, Applied Statistics 2013: volume 62(2), pages 167-179

Chang HH et al. 'Assessment of critical exposure and outcome windows in time-to-event analysis with application to air pollution and preterm birth study'. Biostatistics 2015: volume 16(3), pages 509-521

Chatzi L et al. 'Metabolic syndrome in early pregnancy and risk of preterm birth'. American Journal of Epidemiology 2009: volume 170(7), pages 829-836

Chen L et al. 'Air pollution and birth weight in northern Nevada, 1991-1999'. Inhalation Toxicology 2002: volume 14(2), pages 141-157

Chen G et al. 'Exposure to low concentrations of air pollutants and adverse birth outcomes in Brisbane, Australia, 2003-2013'. Science of the Total Environment 2018: volume 622-623, pages 721-726

Cheng SF et al. 'Nicotine exposure impairs germ cell development in human fetal ovaries cultured in vitro'. Aging (Albany NY) 2018: volume 10(7), pages 1556-1574

Choe SA et al. 'Association between ambient particulate matter concentration and fetal growth restriction stratified by maternal employment'. BMC Pregnancy and Childbirth 2019: volume 19(1), article 246

Chuang JC et al. 'Epigenetics and microRNAs'. Pediatric Research 2007: volume 61(5 Pt 2), pages 24r-29r

Cindrova-Davies T et al. 'Oxidative stress, gene expression, and protein changes induced in the human placenta during labor'. American Journal of Pathology 2007: volume 171(4), pages 1168-1179

Clemente DB et al. 'Prenatal Ambient Air Pollution, Placental Mitochondrial DNA Content, and Birth Weight in the INMA (Spain) and ENVIRONAGE (Belgium) Birth Cohorts'. Environmental Health Perspectives 2016: volume 124(5), pages 659-665

Coker E et al. 'Modeling spatial effects of $PM_{2.5}$ on term low birth weight in Los Angeles County'. Environmental Research 2015: volume 142, pages 354-364

Coker E et al. 'Multi-pollutant exposure profiles associated with term low birth weight in Los Angeles County'. Environment International 2016: volume 91, pages 1-13

COMEAP. 'The effects of long-term exposure to ambient air pollution on cardiovascular morbidity: mechanistic evidence'. A report by the Committee on the Medical Effects of Air Pollutants 2018a

COMEAP. 'Associations of long-term average concentrations of nitrogen dioxide with mortality'. A report by the Committee on the Medical Effects of Air Pollutants 2018b

COMEAP. 'Summary of COMEAP recommendations for the quantification of health effects associated with air pollutants'. A statement by the Committee on the Medical Effects of Air Pollutants 2023

Contreras ZA et al. 'Prenatal air pollution exposure, smoking, and uterine vascular resistance'. Environmental Epidemiology 2018: volume 2(3), article e017

Coolman M et al. 'Angiogenic and fibrinolytic factors in blood during the first half of pregnancy and adverse pregnancy outcomes'. Obstetrics and Gynecology 2012: volume 119(6), pages 1190-1200

Crews D et al. 'Nature, nurture and epigenetics'. Molecular and Cellular Endocrinology 2014: volume 398(1-2), pages 42-52

Crow MK. 'Long interspersed nuclear elements (LINE-1): potential triggers of systemic autoimmune disease'. Autoimmunity 2010: volume 43(1), pages 7-16

Cutland CL et al. 'Low birth weight: Case definition & guidelines for data collection, analysis, and presentation of maternal immunization safety data'. Vaccine 2017: volume 35(48 Pt A), pages 6492-6500

Dadvand P et al. 'Maternal exposure to particulate air pollution and term birth weight: a multi-country evaluation of effect and heterogeneity'. Environmental Health Perspectives 2013: volume 121(3), pages 267-373

Dadvand P et al. 'Residential proximity to major roads and term low birth weight: the roles of air pollution, heat, noise, and road-adjacent trees'. Epidemiology 2014: volume 25(4), pages 518-525

Dang S et al. '$PM_{2.5}$ exposure during pregnancy induces hypermethylation of estrogen receptor promoter region in rat uterus and declines offspring birth weights'. Environmental Pollution 2018: volume 243(Pt B), pages 851-861

Darrow LA et al. 'Ambient air pollution and birth weight in full-term infants in Atlanta, 1994-2004'. Environmental Health Perspectives 2011: volume 119(5), pages 731-737

Dastoorpoor M et al. 'Acute effects of air pollution on spontaneous abortion, premature delivery, and stillbirth in Ahvaz, Iran: a time-series study'. Environmental Science and Pollution Research International 2018: volume 25(6), pages 5447-5458

de Melo JO et al. 'Inhalation of fine particulate matter during pregnancy increased IL-4 cytokine levels in the fetal portion of the placenta'. Toxicology Letters 2015: volume 232(2), pages 475-480

DeChiara TM et al. 'Parental imprinting of the mouse insulin-like growth factor II gene'. Cell 1991: volume 64(4), pages 849-859

Dedele A et al. 'Individual exposure to nitrogen dioxide and adverse pregnancy outcomes in Kaunas study'. International Journal of Environmental Health Research 2017: volume 27(3), pages 230-240

Defra. 'UK Report on measures for 2018 exceedance of the Target Value for Benzo[a]Pyrene'. Department for Environment Food and Rural Affairs 2020

DeFranco E et al. 'Air pollution and stillbirth risk: exposure to airborne particulate matter during pregnancy is associated with fetal death'. Public Library of Science One 2015: volume 10(3), article e0120594

DeFranco E et al. 'Exposure to airborne particulate matter during pregnancy is associated with preterm birth: a population-based cohort study'. Environmental Health 2016: volume 15, article 6

Deguen S et al. 'Using a Clustering Approach to Investigate Socio-Environmental Inequality in Preterm Birth-A Study Conducted at Fine Spatial Scale in Paris (France)'. International Journal of Environmental Research and Public Health 2018: volume 15(9), article 1895

Demir R et al. 'Classification of human placental stem villi: review of structural and functional aspects'. Microscopy Research and Technique 1997: volume 38(1-2), pages 29-41

Detmar J et al. 'Fetal growth restriction triggered by polycyclic aromatic hydrocarbons is associated with altered placental vasculature and AhR-dependent changes in cell death'. American Journal of Physiology Endocrinology and Metabolism 2008: volume 295(2), pages E519-530

Dibben C et al. 'Place of work and residential exposure to ambient air pollution and birth outcomes in Scotland, using geographically fine pollution climate mapping estimates'. Environmental Research 2015: volume 140, pages 535-541

Dimitriadis E et al. 'Pre-eclampsia'. Nature Reviews Disease Primers 2023: volume 9(1), article 8

Dubicke A et al. 'Pro-inflammatory and anti-inflammatory cytokines in human preterm and term cervical ripening'. Journal of Reproductive Immunology 2010: volume 84(2), pages 176-185

Duchen MR. 'Mitochondria in health and disease: perspectives on a new mitochondrial biology'. Molecular Aspects of Medicine 2004: volume 25(4), pages 365-451

Dugandzic R et al. 'The association between low level exposures to ambient air pollution and term low birth weight: a retrospective cohort study'. Environmental Health 2006: volume 5, article 3

Ebisu K et al. 'Airborne $PM_{2.5}$ chemical components and low birth weight in the northeastern and mid-Atlantic regions of the United States'. Environmental Health Perspectives 2012: volume 120(12), pages 1746-1752

Ebisu K et al. 'Cause-specific stillbirth and exposure to chemical constituents and sources of fine particulate matter'. Environmental Research 2018: volume 160, pages 358-364

EEA. 'Air quality in Europe – 2020 report'. EEA Report No 9/2020. European Environment Agency, Copenhagen, Denmark 2020

Enders C et al. 'Exposure to coarse particulate matter during gestation and term low birthweight in California: Variation in exposure and risk across region and socioeconomic subgroup'. Science of the Total Environment 2019: volume 653, pages 1435-1444

Engel SA et al. 'Risk of spontaneous preterm birth is associated with common proinflammatory cytokine polymorphisms'. Epidemiology 2005: volume 16(4), pages 469-477

Enns GM. 'The contribution of mitochondria to common disorders'. Molecular Genetics and Metabolism 2003: volume 80(1-2), pages 11-26

Ernst GD et al. 'C-reactive protein levels in early pregnancy, fetal growth patterns, and the risk for neonatal complications: the Generation R Study'. American Journal of Obstetrics and Gynecology 2011: volume 205(2), pages 132.e1-132.e12

Estarlich M et al. 'Residential exposure to outdoor air pollution during pregnancy and anthropometric measures at birth in a multicenter cohort in Spain'. Environmental Health Perspectives 2011: volume 119(9), pages 1333-1338

Estarlich M et al. 'Exposure to ambient air pollution during pregnancy and preterm birth: A Spanish multicenter birth cohort study'. Environmental Research 2016: volume 147, pages 50-58

Fairburn C. 'Registration of stillbirth'. House of Commons Library Briefing Paper Number 05595, 1 February 2018 (viewed on 16 July 2024)

Faiz AS et al. 'Ambient air pollution and the risk of stillbirth'. American Journal of Epidemiology 2012: volume 176(4), pages 308-316

Fantuzzi G et al. 'Preterm delivery and exposure to active and passive smoking during pregnancy: a case-control study from Italy'. Paediatric and Perinatal Epidemiology 2007: volume 21(3), pages 194-200

Fantuzzi G et al. 'Exposure to active and passive smoking during pregnancy and severe small for gestational age at term'. Journal of Maternal-Fetal and Neonatal Medicine 2008: volume 21(9), pages 643-647

Farchi S et al. 'Piccolipiù, a multicenter birth cohort in Italy: protocol of the study'. BioMed Central Pediatrics 2014: volume 14, article 36

Fecht D et al. 'Associations between air pollution and socioeconomic characteristics, ethnicity and age profile of neighbourhoods in England and the Netherlands'. Environmental Pollution 2015: volume 198, pages 201-210

Fialová L et al. 'Oxidative stress and inflammation in pregnancy'. Scandinavian Journal of Clinical and Laboratory Investigation 2006: volume 66(2), pages 121-127

Fleischer NL et al. 'Outdoor air pollution, preterm birth, and low birth weight: analysis of the world health organization global survey on maternal and perinatal health'. Environmental Health Perspectives 2014: volume 122(4), pages 425-430

Foley DL et al. 'Prospects for epigenetic epidemiology'. American Journal of Epidemiology 2009: volume 169(4), pages 389-400

Friedman C et al. 'Exposure to ambient air pollution during pregnancy and inflammatory biomarkers in maternal and umbilical cord blood: The Healthy Start study'. Environmental Research 2021: volume 197, article 111165

Fujimoto A et al. 'Diesel exhaust affects immunological action in the placentas of mice'. Environmental Toxicology 2005: volume 20(4), pages 431-440

Fussell JC, Walton HA. ‘The effects of maternal exposure to ambient air pollution on adverse birth outcomes: A scoping review’ prepared for the Department of Health and Social Care by King’s College London 2018.

Garonna E et al. 'Vascular endothelial growth factor receptor-2 couples cyclo-oxygenase-2 with pro-angiogenic actions of leptin on human endothelial cells'. Public Library of Science One 2011: volume 6(4), article e18823

Gehring U et al. 'Traffic-related air pollution, preterm birth and term birth weight in the PIAMA birth cohort study'. Environmental Research 2011: volume 111(1), pages 125-135

Ghosh JK et al. 'Assessing the influence of traffic-related air pollution on risk of term low birth weight on the basis of land-use-based regression models and measures of air toxics'. American Journal of Epidemiology 2012: volume 175(12), pages 1262-1274

Ghosh RE et al. 'Fetal growth, stillbirth, infant mortality and other birth outcomes near UK municipal waste incinerators; retrospective population based cohort and case-control study'. Environment International 2019: volume 122, pages 151-158

Ghosh R et al. 'Ambient and household $PM_{2.5}$ pollution and adverse perinatal outcomes: A meta-regression and analysis of attributable global burden for 204

countries and territories'. Public Library of Science Medicine 2021: volume 18(9), article e1003718

Giorgis-Allemand L et al. 'The Influence of Meteorological Factors and Atmospheric Pollutants on the Risk of Preterm Birth'. American Journal of Epidemiology 2017: volume 185(4), pages 247-258

Godschalk RW et al. 'A critical evaluation of DNA adducts as biological markers for human exposure to polycyclic aromatic compounds'. Journal of Biochemistry and Molecular Biology 2003: volume 36(1), pages 1-11

Goldenberg RL et al. 'Epidemiology and causes of preterm birth'. The Lancet 2008: volume 371(9606), pages 75-84

Gottlieb K et al. 'Twin study methodology and variability in xenobiotic placental metabolism'. Teratogenesis, Carcinogenesis, and Mutagenesis 1986: volume 6(4), pages 253-263

Gouveia N et al. 'Association between ambient air pollution and birth weight in São Paulo, Brazil'. Journal of Epidemiology and Community Health 2004: volume 58(1), pages 11-17

Govarts E et al. 'Birth weight and prenatal exposure to polychlorinated biphenyls (PCBs) and dichlorodiphenyldichloroethylene (DDE): a meta-analysis within 12 European Birth Cohorts'. Environmental Health Perspectives 2012: volume 120(2), pages 162-170

Gray SC et al. 'Assessing the impact of race, social factors and air pollution on birth outcomes: a population-based study'. Environmental Health 2014: volume 13(1), article 4

Green R et al. 'Association of stillbirth with ambient air pollution in a California cohort study'. American Journal of Epidemiology 2015: volume 181(11), pages 874-882

Grevendonk L et al. 'Mitochondrial oxidative DNA damage and exposure to particulate air pollution in mother-newborn pairs'. Environmental Health 2016: volume 15, article 10

Gruzieva O et al. 'Epigenome-Wide Meta-Analysis of Methylation in Children Related to Prenatal NO_2 Air Pollution Exposure'. Environmental Health Perspectives 2017: volume 125(1), pages 104-110

Gu Y et al. 'Differential miRNA expression profiles between the first and third trimester human placentas'. American Journal of Physiology Endocrinology and Metabolism 2013: volume 304(8), pages E836-843

Guan T et al. 'Acute and chronic effects of ambient fine particulate matter on preterm births in Beijing, China: A time-series model'. Science of the Total Environment 2019: volume 650(Pt 2), pages 1671-1677

Guo T et al. 'The association between ambient $PM_{2.5}$ exposure and the risk of preterm birth in China: A retrospective cohort study'. Science of the Total Environment 2018: volume 633, pages 1453-1459

Guo LQ et al. 'Ambient air pollution and adverse birth outcomes: a systematic review and meta-analysis'. Journal of Zhejiang University Science B 2019: volume 20(3), pages 238-252

Guo LQ et al. 'Retraction Note to: Ambient air pollution and adverse birth outcomes: a systematic review and meta-analysis'. Journal of Zhejiang University Science B 2020: volume 21(9), page 756

Guo P et al. 'Ambient air pollution and markers of fetal growth: A retrospective population-based cohort study of 2.57 million term singleton births in China'. Environment International 2020 volume 135, article 105410

Guven MA et al. 'Association of maternal serum CRP, IL-6, TNF-alpha, homocysteine, folic acid and vitamin B12 levels with the severity of preeclampsia and fetal birth weight'. Hypertension in Pregnancy 2009: volume 28(2), pages 190-200

Guxens M et al. 'Cohort Profile: the INMA--INfancia y Medio Ambiente--(Environment and Childhood) Project'. International Journal of Epidemiology 2012: volume 41(4), pages 930-940

Guyda HJ. 'Metabolic effects of growth factors and polycyclic aromatic hydrocarbons on cultured human placental cells of early and late gestation'. Journal of Clinical Endocrinology and Metabolism 1991: volume 72(3), pages 718-723

Ha EH et al. 'Is air pollution a risk factor for low birth weight in Seoul?'. Epidemiology 2001: volume 12(6), pages 643-648

Ha S et al. 'The effects of air pollution on adverse birth outcomes'. Environmental Research 2014: volume 134, pages 198-204

Ha S et al. 'Ambient temperature and air quality in relation to small for gestational age and term low birthweight'. Environmental Research 2017: volume 155, pages 394-400

Habermann M et al. 'Socioeconomic position and low birth weight among mothers exposed to traffic-related air pollution'. Public Library of Science One 2014: volume 9(11), article e113900

Hack M et al. 'Outcomes of children of extremely low birthweight and gestational age in the 1990's'. Early Human Development 1999: volume 53(3), pages 193-218

Han Y et al. 'Maternal air pollution exposure and preterm birth in Wuxi, China: Effect modification by maternal age'. Ecotoxicology and Environmental Safety 2018: volume 157, pages 457-462

Hannam K et al. 'Air pollution exposure and adverse pregnancy outcomes in a large UK birth cohort: use of a novel spatio-temporal modelling technique'. Scandinavian Journal of Work, Environment & Health 2014: volume 40(5), pages 518-530

Hansen C et al. 'Maternal exposure to low levels of ambient air pollution and preterm birth in Brisbane, Australia'. British Journal of Obstetrics and Gynaecology 2006: volume 113(8), pages 935-941

Hao H et al. 'Air Pollution and Preterm Birth in the U.S. State of Georgia (2002-2006): Associations with Concentrations of 11 Ambient Air Pollutants Estimated by Combining Community Multiscale Air Quality Model (CMAQ) Simulations with Stationary Monitor Measurements'. Environmental Health Perspectives 2016: volume 124(6), pages 875-880

Harris G et al. 'The association of $PM_{2.5}$ with full term low birth weight at different spatial scales'. Environmental Research 2014: volume 134, pages 427-434

Hay WW, Jr. 'Placental transport of nutrients to the fetus'. Hormone Research 1994: volume 42(4-5), pages 215-222

He T et al. 'Ambient air pollution, H19/DMR methylation in cord blood and newborn size: A pilot study in Zhengzhou City, China'. Chemosphere 2018: volume 212, pages 863-871

HEI. 'State of Global Air 2020'. Special Report. Health Effects Institute, Boston, MA; 2020

HEI. 'Systematic Review and Meta-analysis of Selected Health Effects of Long-Term Exposure to Traffic-Related Air Pollution'. Special Report 23. HEI Panel on the Health Effects of Long-Term Exposure to Traffic-Related Air Pollution. Health Effects Institute, Boston, MA; 2022

Herbstman JB et al. 'Prenatal exposure to polycyclic aromatic hydrocarbons, benzo[a]pyrene-DNA adducts, and genomic DNA methylation in cord blood'. Environmental Health Perspectives 2012: volume 120(5), pages 733-738

Hochberg Z et al. 'Child health, developmental plasticity, and epigenetic programming'. Endocrine Reviews 2011: volume 32(2), pages 159-224

Horgan RP et al. 'Metabolic profiling uncovers a phenotypic signature of small for gestational age in early pregnancy'. Journal of Proteome Research 2011: volume 10(8), pages 3660-3673

Hou L et al. 'Airborne particulate matter and mitochondrial damage: a cross-sectional study'. Environmental Health 2010: volume 9, article 48

Hoyt AT et al. 'Does Maternal Exposure to Secondhand Tobacco Smoke During Pregnancy Increase the Risk for Preterm or Small-for-Gestational Age Birth?'. Maternal and Child Health Journal 2018: volume 22(10), pages 1418-1429

Huang C et al. 'Ambient air pollution and adverse birth outcomes: a natural experiment study'. Population Health Metrics 2015: volume 13, article 17

Huang H et al. 'Investigation of association between environmental and socioeconomic factors and preterm birth in California'. Environment International 2018: volume 121(Pt 2), pages 1066-1078

Hubel CA. 'Oxidative stress in the pathogenesis of preeclampsia'. Proceedings of the Society for Experimental Biology and Medicine 1999: volume 222(3), pages 222-235

Husby A et al. 'Gestational age at birth and cognitive outcomes in adolescence: population based full sibling cohort study'. British Medical Journal 2023: volume 380, article e072779

Huynh M et al. 'Relationships between air pollution and preterm birth in California'. Paediatric and Perinatal Epidemiology 2006: volume 20(6), pages 454-461

Hwang BF et al. 'Air pollution and stillbirth: a population-based case-control study in Taiwan'. Environmental Health Perspectives 2011: volume 119(9), pages 1345-1349

Hyder A et al. '$PM_{2.5}$ exposure and birth outcomes: use of satellite- and monitor-based data'. Epidemiology 2014: volume 25(1), pages 58-67

Illsley NP. 'Current topic: Glucose transporters in the human placenta'. Placenta 2000: volume 21(1), pages 14-22

Ion RC et al. 'Environmental Tobacco Smoke Exposure in Pregnancy is Associated With Earlier Delivery and Reduced Birth Weight'. Reproductive Sciences 2015: volume 22(12), pages 1603-1611

Jackson CH et al. 'Bayesian graphical models for regression on multiple data sets with different variables'. Biostatistics 2009: volume 10(2), pages 335-351

Jaddoe VW et al. 'The Generation R Study: design and cohort update 2010'. European Journal of Epidemiology 2010: volume 25(11), pages 823-841

Jalaludin B et al. 'Impact of ambient air pollution on gestational age is modified by season in Sydney, Australia'. Environmental Health 2007: volume 6, article 16

Janevic T et al. 'Neighborhood deprivation and adverse birth outcomes among diverse ethnic groups'. Annals of Epidemiology 2010: volume 20(6), pages 445-451

Janssen BG et al. 'Placental mitochondrial DNA content and particulate air pollution during in utero life'. Environmental Health Perspectives 2012: volume 120(9), pages 1346-1352

Janssen BG et al. 'Placental DNA hypomethylation in association with particulate air pollution in early life'. Particle and Fibre Toxicology 2013: volume 10, article 22

Janssen BG et al. 'Placental mitochondrial methylation and exposure to airborne particulate matter in the early life environment: An ENVIRONAGE birth cohort study'. Epigenetics 2015: volume 10(6), pages 536-544

Janssen BG et al. 'Cohort Profile: The ENVIRonmental influence ON early AGEing (ENVIRONAGE): a birth cohort study'. International Journal of Epidemiology 2017a: volume 46(5), pages 1386-1387m

Janssen BG et al. 'Fetal Thyroid Function, Birth Weight, and in Utero Exposure to Fine Particle Air Pollution: A Birth Cohort Study'. Environmental Health Perspectives 2017b: volume 125(4), pages 699-705

Jauniaux E et al. 'The human first trimester gestational sac limits rather than facilitates oxygen transfer to the foetus--a review'. Placenta 2003: volume 24 Suppl A, pages S86-93

Jauniaux E et al. 'Distribution and transfer pathways of antioxidant molecules inside the first trimester human gestational sac'. Journal of Clinical Endocrinology and Metabolism 2004: volume 89(3), pages 1452-1458

Jauniaux E, et al. ‘The placenta’ in ‘Oxford Textbook of Obstetrics and Gynaecology’ edited by Arulkumaran S et al. Oxford University Press 2020: Chapter 9, pages 120-132.

Ji X et al. 'Nitrogen dioxide air pollution and preterm birth in Shanghai, China'. Environmental Research 2019: volume 169, pages 79-85

Ji Y et al. 'Association between exposure to particulate matter during pregnancy and birthweight: a systematic review and a meta-analysis of birth cohort studies'. Journal of Biomedical Research 2019: volume 33(1), pages 56-68

Ji Y et al. 'Retraction note: Association between exposure to particulate matter during pregnancy and birthweight: a systematic review and a meta-analysis of birth cohort studies'. Journal of Biomedical Research 2024: online publication, DOI: 10.7555/JBR.38.20240314

Johnson S et al. 'Ambient Fine Particulate Matter, Nitrogen Dioxide, and Preterm Birth in New York City'. Environmental Health Perspectives 2016: volume 124(8), pages 1283-1290

Jonakait GM. 'The effects of maternal inflammation on neuronal development: possible mechanisms'. International Journal of Developmental Neuroscience 2007: volume 25(7), pages 415-425

Keelan JA et al. 'Cytokines, prostaglandins and parturition--a review'. Placenta 2003: volume 24 Suppl A, pages S33-46

Kelly FJ et al. 'Nitrogen dioxide depletes uric acid and ascorbic acid but not glutathione from lung lining fluid'. Biochemical Journal 1997: volume 325 (Pt 1)(Pt 1), pages 95-99

Kelly FJ. 'Oxidative stress: its role in air pollution and adverse health effects'. Occupational and Environmental Medicine 2003: volume 60(8), pages 612-616

Kelly FJ et al. ‘Particle mediated extracellular oxidative stress in the lung’ in ‘Particle Toxicology’, edited by Donaldson K et al. CRC Press 2007: Chapter 5, pages 89-117.

Kim OJ et al. 'PM_{10} and pregnancy outcomes: a hospital-based cohort study of pregnant women in Seoul'. Journal of Occupational and Environmental Medicine 2007: volume 49(12), pages 1394-1402

Kim D et al. 'The social determinants of infant mortality and birth outcomes in Western developed nations: a cross-country systematic review'. International Journal of Environmental Research and Public Health 2013: volume 10(6), pages 2296-2335

Kim YJ et al. 'Maternal Exposure to Particulate Matter during Pregnancy and Adverse Birth Outcomes in the Republic of Korea'. International Journal of Environmental Research and Public Health 2019: volume 16(4), article 633

Kingsley SL et al. 'Maternal residential proximity to major roadways, birth weight, and placental DNA methylation'. Environment International 2016: volume 92-93, pages 43-49

Kingsley SL et al. 'Maternal ambient air pollution, preterm birth and markers of fetal growth in Rhode Island: results of a hospital-based linkage study'. Journal of Epidemiology and Community Health 2017: volume 71(12), pages 1131-1136

Kirwa K et al. 'Low birth weight and $PM_{2.5}$ in Puerto Rico'. Environmental Epidemiology 2019: volume 3(4), article e058

Klepac P et al. 'Ambient air pollution and pregnancy outcomes: A comprehensive review and identification of environmental public health challenges'. Environmental Research 2018: volume 167, pages 144-159

Kloog I et al. 'Using new satellite based exposure methods to study the association between pregnancy $PM_{2.5}$ exposure, premature birth and birth weight in Massachusetts'. Environmental Health 2012: volume 11, article 40

Kobayashi S et al. 'Association between maternal passive smoking and increased risk of delivering small-for-gestational-age infants at full-term using plasma cotinine levels from The Hokkaido Study: a prospective birth cohort'. British Medical Journal Open 2019: volume 9(2), article e023200

Kumar N. 'Uncertainty in the Relationship between Criteria Pollutants and Low Birth Weight in Chicago'. Atmospheric Environment 2012: volume 49, pages 171-179

Kumar N. 'The Exposure Uncertainty Analysis: The Association between Birth Weight and Trimester Specific Exposure to Particulate Matter ($PM_{2.5}$ vs. PM_{10})'. International Journal of Environmental Research and Public Health 2016: volume 13(9), article 906

Ladd-Acosta C et al. 'Epigenetic marks of prenatal air pollution exposure found in multiple tissues relevant for child health'. Environment International 2019: volume 126, pages 363-376

Laine JE et al. 'Prenatal Exposure to Multiple Air Pollutants, Mediating Molecular Mechanisms, and Shifts in Birthweight'. Environmental Science & Technology 2020: volume 54(22), pages 14502-14513

Lamichhane DK et al. 'A meta-analysis of exposure to particulate matter and adverse birth outcomes'. Environmental Health and Toxicology 2015: volume 30, article e2015011

Langford SD et al. 'Ozone-reactive absorption by pulmonary epithelial lining fluid constituents'. Toxicology and Applied Pharmacology 1995: volume 132(1), pages 122-130

Laurent O et al. 'Investigating the association between birth weight and complementary air pollution metrics: a cohort study'. Environmental Health 2013: volume 12, article 18

Laurent O et al. 'Sources and contents of air pollution affecting term low birth weight in Los Angeles County, California, 2001-2008'. Environmental Research 2014: volume 134, pages 488-495

Laurent O et al. 'A Statewide Nested Case-Control Study of Preterm Birth and Air Pollution by Source and Composition: California, 2001-2008'. Environmental Health Perspectives 2016a: volume 124(9), pages 1479-1486

Laurent O et al. 'Low birth weight and air pollution in California: Which sources and components drive the risk?'. Environment International 2016b: volume 92-93, pages 471-477

Lavigne E et al. 'Ambient air pollution and adverse birth outcomes: Differences by maternal comorbidities'. Environmental Research 2016: volume 148, pages 457-466

Lavigne É et al. 'Fine Particulate Air Pollution and Adverse Birth Outcomes: Effect Modification by Regional Nonvolatile Oxidative Potential'. Environmental Health Perspectives 2018: volume 126(7), citation identifier 077012

Le HQ et al. 'Air pollutant exposure and preterm and term small-for-gestational-age births in Detroit, Michigan: long-term trends and associations'. Environment International 2012: volume 44, pages 7-17

Le F et al. 'In vitro fertilization alters growth and expression of Igf2/H19 and their epigenetic mechanisms in the liver and skeletal muscle of newborn and elder mice'. Biology of Reproduction 2013: volume 88(3), article 75

Lee HC et al. 'Mitochondrial role in life and death of the cell'. Journal of Biomedical Science 2000: volume 7(1), pages 2-15

Lee BE et al. 'Exposure to air pollution during different gestational phases contributes to risks of low birth weight'. Human Reproduction 2003: volume 18(3), pages 638-643

Lee PC et al. 'Particulate air pollution exposure and C-reactive protein during early pregnancy'. Epidemiology 2011: volume 22(4), pages 524-531

Lee PC et al. 'First trimester exposure to ambient air pollution, pregnancy complications and adverse birth outcomes in Allegheny County, PA'. Maternal and Child Health Journal 2013: volume 17(3), pages 545-555

Leem JH et al. 'Exposures to air pollutants during pregnancy and preterm delivery'. Environmental Health Perspectives 2006: volume 114(6), pages 905-910

Leighton PA et al. 'Disruption of imprinting caused by deletion of the H19 gene region in mice'. Nature 1995: volume 375(6526), pages 34-39

Li J et al. ‘Cell-based in vitro models for placental barrier permeability studies’, Chapter 3.10 of ‘Concepts and Models for Drug Permeability Studies (Second Edition), Cell and Tissue Based in Vitro Culture Models’ editors Sarmento B et al. Woodhead Publishing Series in Biomedicine 2024, pages 251-272

Li M et al. 'Innate immunity, coagulation and placenta-related adverse pregnancy outcomes'. Thrombosis Research 2009: volume 124(6), pages 656-662

Li X et al. 'Association between ambient fine particulate matter and preterm birth or term low birth weight: An updated systematic review and meta-analysis'. Environmental Pollution 2017: volume 227, pages 596-605

Li X et al. 'Analysis of short-term and sub-chronic effects of ambient air pollution on preterm birth in central China'. Environmental Science and Pollution Research International 2018: volume 25(19), pages 19028-19039

Li C et al. 'Maternal exposure to air pollution and the risk of low birth weight: A meta-analysis of cohort studies'. Environmental Research 2020: volume 190, article 109970

Li J et al ‘Cell-based in vitro models for placental barrier permeability studies’ in ‘Concepts and Models for Drug Permeability Studies (Second Edition) Cell and Tissue Based in Vitro Culture Models’. Woodhead Publishing Series in Biomedicine 2024: Chapter 3.10, pages 251-272

Liang Z et al. 'Effect of the 2008 cold spell on preterm births in two subtropical cities of Guangdong Province, Southern China'. Science of the Total Environment 2018: volume 642, pages 307-313

Liang Z et al. 'Ambient $PM_{2.5}$ and birth outcomes: Estimating the association and attributable risk using a birth cohort study in nine Chinese cities'. Environment International 2019a: volume 126, pages 329-335

Liang Z et al. 'Migrant population is more vulnerable to the effect of air pollution on preterm birth: Results from a birth cohort study in seven Chinese cities'. International Journal of Hygiene and Environmental Health 2019b: volume 222(7), pages 1047-1053

Liao WX et al. 'Expression of estrogen receptors-alpha and -beta in the pregnant ovine uterine artery endothelial cells in vivo and in vitro'. Biology of Reproduction 2005: volume 72(3), pages 530-537

Lin CM et al. 'Association between maternal exposure to elevated ambient sulfur dioxide during pregnancy and term low birth weight'. Environmental Research 2004: volume 96(1), pages 41-50

Lin YT et al. 'Associations between ozone and preterm birth in women who develop gestational diabetes'. American Journal of Epidemiology 2015: volume 181(4), pages 280-287

Lin L et al. 'The associations of particulate matters with fetal growth in utero and birth weight: A birth cohort study in Beijing, China'. Science of the Total Environment 2020: volume 709, pages 136246

Lin S et al. 'Maternal Passive Smoking, Vitamin D Deficiency and Risk of Spontaneous Abortion'. Nutrients 2022: volume 14(18), article 3674

Linnane AW et al. 'Mitochondrial DNA mutations as an important contributor to ageing and degenerative diseases'. The Lancet 1989: volume 333(8639), pages 642-645

Liu S et al. 'Association between gaseous ambient air pollutants and adverse pregnancy outcomes in Vancouver, Canada'. Environmental Health Perspectives 2003: volume 111(14), pages 1773-1778

Liu B et al. 'CpG methylation patterns of human mitochondrial DNA'. Scientific Reports 2016: volume 6, article 23421

Liu C et al. 'Different exposure levels of fine particulate matter and preterm birth: a meta-analysis based on cohort studies'. Environmental Science and Pollution Research International 2017: volume 24(22), pages 17976-17984

Liu WY et al. 'Association between ambient air pollutants and preterm birth in Ningbo, China: a time-series study'. BioMed Central Pediatrics 2018: volume 18(1), article 305

Liu X et al. 'Effects of prenatal exposure to air particulate matter on the risk of preterm birth and roles of maternal and cord blood LINE-1 methylation: A birth cohort study in Guangzhou, China'. Environment International 2019: volume 133(Pt A), article 105177

Liu Y et al. 'The association between air pollution and preterm birth and low birth weight in Guangdong, China'. BioMed Central Public Health 2019: volume 19(1), article 3

Liu NM et al. 'Evidence for the presence of air pollution nanoparticles in placental tissue cells'. Science of the Total Environment 2021: volume 751, article 142235

Llop S et al. 'Preterm birth and exposure to air pollutants during pregnancy'. Environmental Research 2010: volume 110(8), pages 778-785

Lu C et al. 'Combined effects of ambient air pollution and home environmental factors on low birth weight'. Chemosphere 2020: volume 240, article 124836

Luciano A et al. 'The influence of maternal passive and light active smoking on intrauterine growth and body composition of the newborn'. European Journal of Clinical Nutrition 1998: volume 52(10), pages 760-763

Luo YJ et al. 'Interaction between maternal passive smoking during pregnancy and CYP1A1 and GSTs polymorphisms on spontaneous preterm delivery'. Public Library of Science One 2012: volume 7(11), article e49155

Lyall F et al. 'Increased nitrotyrosine in the diabetic placenta: evidence for oxidative stress'. Diabetes Care 1998: volume 21(10), pages 1753-1758

Maccani MA et al. 'Epigenetics in the placenta'. American Journal of Reproductive Immunology 2009: volume 62(2), pages 78-89

Maccani MA et al. 'Maternal cigarette smoking during pregnancy is associated with downregulation of miR-16, miR-21, and miR-146a in the placenta'. Epigenetics 2010: volume 5(7), pages 583-589

Maccani MA et al. 'miR-16 and miR-21 expression in the placenta is associated with fetal growth'. Public Library of Science One 2011: volume 6(6), article e21210

Maciel-Ruiz JA et al. 'Prenatal exposure to particulate matter and ozone: Bulky DNA adducts, plasma isoprostanes, allele risk variants, and neonate susceptibility in the Mexico City Metropolitan Area'. Environmental and Molecular Mutagenesis 2019: volume 60(5), pages 428-442

Madsen C et al. 'Ambient air pollution exposure, residential mobility and term birth weight in Oslo, Norway'. Environmental Research 2010: volume 110(4), pages 363-371

Maghbooli Z et al. 'Air pollution during pregnancy and placental adaptation in the levels of global DNA methylation'. Public Library of Science One 2018: volume 13(7), article e0199772

Magness RR et al. 'Endothelial vasodilator production by uterine and systemic arteries. I. Effects of ANG II on PGI2 and NO in pregnancy'. American Journal of Physiology Heart and Circulatory Physiology 1996: volume 270(6 Pt 2), pages H1914-1923

Magness RR et al. 'Endothelial vasodilator production by uterine and systemic arteries. VI. Ovarian and pregnancy effects on eNOS and NO(x)'. American Journal of Physiology Heart and Circulatory Physiology 2001: volume 280(4), pages H1692-1698

Maisonet M et al. 'Relation between ambient air pollution and low birth weight in the Northeastern United States'. Environmental Health Perspectives 2001: volume 109 (Suppl 3), pages 351-356

Maitre L et al. 'Urinary metabolic profiles in early pregnancy are associated with preterm birth and fetal growth restriction in the Rhea mother-child cohort study'. BMC Medicine 2014: volume 12, article 110

Maitre L et al. 'Maternal urinary metabolic signatures of fetal growth and associated clinical and environmental factors in the INMA study'. BMC Medicine 2016: volume 14(1), article 177

Manchester DK et al. 'Ah receptor in human placenta: stabilization by molybdate and characterization of binding of 2,3,7,8-tetrachlorodibenzo-p-dioxin, 3-methylcholanthrene, and benzo(a)pyrene'. Cancer Research 1987: volume 47(18), pages 4861-4868

Manchester DK et al. 'Detection of benzo[a]pyrene diol epoxide-DNA adducts in human placenta'. Proceedings of the National Academy of Sciences of the United States of America 1988: volume 85(23), pages 9243-9247

Manchester DK et al. 'Determinants of polycyclic aromatic hydrocarbon-DNA adducts in human placenta'. Cancer Research 1992: volume 52(6), pages 1499-1503

Mangham LJ et al. 'The cost of preterm birth throughout childhood in England and Wales'. Pediatrics 2009: volume 123(2), pages e312-327

Maroziene L et al. 'Maternal exposure to low-level air pollution and pregnancy outcomes: a population-based study'. Environmental Health 2002: volume 1(1), article 6

McCarthy MM et al. 'At the frontier of epigenetics of brain sex differences'. Frontiers in Behavioral Neuroscience 2015: volume 9, article 221

Mechta M et al. 'Evidence Suggesting Absence of Mitochondrial DNA Methylation'. Frontiers in Genetics 2017: volume 8, article 166

Medici M et al. 'Maternal thyroid hormone parameters during early pregnancy and birth weight: the Generation R Study'. Journal of Clinical Endocrinology and Metabolism 2013: volume 98(1), pages 59-66

Mendola P et al. 'Preterm birth and air pollution: Critical windows of exposure for women with asthma'. Journal of Allergy and Clinical Immunology 2016: volume 138(2), pages 432-440.e5

Mendola P et al. 'Chronic and Acute Ozone Exposure in the Week Prior to Delivery Is Associated with the Risk of Stillbirth'. International Journal of Environmental Research and Public Health 2017: volume 14(7), article 731

Mendola P et al. 'Air Pollution and Preterm Birth: Do Air Pollution Changes over Time Influence Risk in Consecutive Pregnancies among Low-Risk Women?'. International Journal of Environmental Research and Public Health 2019: volume 16(18), article 3365

Michikawa T et al. 'Maternal exposure to air pollutants during the first trimester and foetal growth in Japanese term infants'. Environmental Pollution 2017: volume 230, pages 387-393

Millar LK et al. 'Low birth weight and preeclampsia in pregnancies complicated by hyperthyroidism'. Obstetrics and Gynecology 1994: volume 84(6), pages 946-949

Miller SL et al. 'Effect of nitric oxide synthase inhibition on the uterine vasculature of the late-pregnant ewe'. American Journal of Obstetrics and Gynecology 1999: volume 180(5), pages 1138-1145

Miller CN et al. 'Uterine Artery Flow and Offspring Growth in Long-Evans Rats following Maternal Exposure to Ozone during Implantation'. Environmental Health Perspectives 2017: volume 125(12), article 127005

Miller CN et al. 'Aspirin pre-treatment modulates ozone-induced fetal growth restriction and alterations in uterine blood flow in rats'. Reproductive Toxicology 2019a: volume 83, pages 63-72

Miller CN et al. 'Fetal growth outcomes following peri-implantation exposure of Long-Evans rats to noise and ozone differ by sex'. Biology of Sex Differences 2019b: volume 10(1), pages 54

Mistry HD et al. 'The importance of antioxidant micronutrients in pregnancy'. Oxidative Medicine and Cellular Longevity 2011: volume 2011, article 841749

Mor G et al. 'The unique immunological and microbial aspects of pregnancy'. Nature reviews Immunology 2017: volume 17(8), pages 469-482

Moreira I et al. 'Short-term effects of Saharan dust intrusions and biomass combustion on birth outcomes in Spain'. Science of the Total Environment 2020: volume 701, article 134755

Morello-Frosch R et al. 'Ambient air pollution exposure and full-term birth weight in California'. Environmental Health 2010: volume 9, article 44

Moser K et al. 'Birthweight and gestational age by ethnic group, England and Wales 2005: introducing new data on births'. Health Statistics Quarterly / Office for National Statistics 2008: volume (39), pages 22-31, 34-55

Myatt L et al. 'Oxidative stress in the placenta'. Histochemistry and Cell Biology 2004: volume 122(4), pages 369-382

Myatt L. 'Review: Reactive oxygen and nitrogen species and functional adaptation of the placenta'. Placenta 2010: volume 31 (Suppl), pages S66-69

Nachman RM et al. 'Intrauterine Inflammation and Maternal Exposure to Ambient $PM_{2.5}$ during Preconception and Specific Periods of Pregnancy: The Boston Birth Cohort'. Environmental Health Perspectives 2016: volume 124(10), pages 1608-1615

Nebert DW et al. 'Role of aryl hydrocarbon receptor-mediated induction of the CYP1 enzymes in environmental toxicity and cancer'. Journal of Biological Chemistry 2004: volume 279(23), pages 23847-23850

Newschaffer CJ et al. 'Infant siblings and the investigation of autism risk factors'. Journal of Neurodevelopmental Disorders 2012: volume 4(1), article 7

Ng C et al. 'Source apportionment of fine particulate matter and risk of term low birth weight in California: Exploring modification by region and maternal characteristics'. Science of the Total Environment 2017: volume 605-606, pages 647-654

NHS. 'Pre-eclampsia, Symptoms' National Health Service (2021)

NICE. 'Pregnancy and neonates: low birth weight' NICE indicator IND22 National Institute for Health and Care Excellence, 2014

NICE. 'Preterm labour and birth'. Quality standard QS135. National Institute for Health and Care Excellence 2019

Nieuwenhuijsen MJ et al. 'Influence of the Urban Exposome on Birth Weight'. Environmental Health Perspectives 2019: volume 127(4), article 47007

Niu Z et al. 'Placenta mediates the association between maternal second-hand smoke exposure during pregnancy and small for gestational age'. Placenta 2015: volume 36(8), pages 876-880

Oakley RH et al. 'The biology of the glucocorticoid receptor: new signaling mechanisms in health and disease'. Journal of Allergy and Clinical Immunology 2013: volume 132(5), pages 1033-1044

Odom LN et al. 'Environmental induction of the fetal epigenome'. Expert Review of Obstetrics & Gynecology 2010: volume 5(6), pages 657-664

Oken E et al. 'A nearly continuous measure of birth weight for gestational age using a United States national reference'. BioMed Central Pediatrics 2003: volume 3, article 6

Padula AM et al. 'Traffic-related air pollution and risk of preterm birth in the San Joaquin Valley of California'. Annals of Epidemiology 2014: volume 24(12), pages 888-895.e4

Padula AM et al. 'Environmental pollution and social factors as contributors to preterm birth in Fresno County'. Environmental Health 2018: volume 17(1), article 70

Padula AM et al. 'Prenatal exposure to air pollution, maternal diabetes and preterm birth'. Environmental Research 2019: volume 170, pages 160-167

Panasevich S et al. 'Association between pregnancy exposure to air pollution and birth weight in selected areas of Norway'. Archives of Public Health 2016: volume 74, article 26

Park JH et al. 'Polycyclic aromatic hydrocarbon (PAH) o-quinones produced by the aldo-keto-reductases (AKRs) generate abasic sites, oxidized pyrimidines, and 8-oxo-dGuo via reactive oxygen species'. Chemical Research in Toxicology 2006: volume 19(5), pages 719-728

Parker JD et al. 'Influences of study design and location on the relationship between particulate matter air pollution and birthweight'. Paediatric and Perinatal Epidemiology 2008: volume 22(3), pages 214-227

Payne BA et al. 'Universal heteroplasmy of human mitochondrial DNA'. Human Molecular Genetics 2013: volume 22(2), pages 384-390

Pedersen M et al. 'Ambient air pollution and low birthweight: a European cohort study (ESCAPE)'. The Lancet Respiratory Medicine 2013: volume 1(9), pages 695-704

Peleg D et al. 'Intrauterine growth restriction: identification and management'. American Family Physician 1998: volume 58(2), pages 453-460

Penning TM et al. 'Dihydrodiol dehydrogenases and polycyclic aromatic hydrocarbon activation: generation of reactive and redox active o-quinones'. Chemical Research in Toxicology 1999: volume 12(1), pages 1-18

Pereira G et al. 'Fine particulate matter and risk of preterm birth in Connecticut in 2000-2006: a longitudinal study'. American Journal of Epidemiology 2014a: volume 179(1), pages 67-74

Pereira G et al. 'Fine particulate matter and risk of preterm birth and pre-labor rupture of membranes in Perth, Western Australia 1997-2007: a longitudinal study'. Environment International 2014b: volume 73, pages 143-149

Pereira G et al. 'Sources of fine particulate matter and risk of preterm birth in Connecticut, 2000-2006: a longitudinal study'. Environmental Health Perspectives 2014c: volume 122(10), pages 1117-1122

Pereira G et al. 'Fine Particulates, Preterm Birth, and Membrane Rupture in Rochester, NY'. Epidemiology 2016: volume 27(1), pages 66-73

Perera FP et al. 'Molecular epidemiologic research on the effects of environmental pollutants on the fetus'. Environmental Health Perspectives 1999: volume 107(Suppl 3), pages 451-460

Perera F et al. 'Prenatal environmental exposures, epigenetics, and disease'. Reproductive Toxicology 2011: volume 31(3), pages 363-373

Pineles BL et al. 'Systematic review and meta-analysis of miscarriage and maternal exposure to tobacco smoke during pregnancy'. American Journal of Epidemiology 2014: volume 179(7), pages 807-823

Pitiphat W et al. 'Plasma C-reactive protein in early pregnancy and preterm delivery'. American Journal of Epidemiology 2005: volume 162(11), pages 1108-1113

Poirier MC. 'DNA adducts as exposure biomarkers and indicators of cancer risk'. Environmental Health Perspectives 1997: volume 105(Suppl 4), pages 907-912

Poirier A et al. 'Maternal Exposure to Air Pollution and Adverse Birth Outcomes in Halifax, Nova Scotia'. Journal of Occupational and Environmental Medicine 2015: volume 57(12), pages 1291-1298

Qian Z et al. 'Ambient air pollution and preterm birth: A prospective birth cohort study in Wuhan, China'. International Journal of Hygiene and Environmental Health 2016: volume 219(2), pages 195-203

Qiu J et al. 'Passive smoking and preterm birth in urban China'. American Journal of Epidemiology 2014: volume 180(1), pages 94-102

Rammah A et al. '$PM_{2.5}$ metal constituent exposure and stillbirth risk in Harris County, Texas'. Environmental Research 2019a: volume 176, article 108516

Rammah A et al. 'Time-Varying Exposure to Ozone and Risk of Stillbirth in a Nonattainment Urban Region'. American Journal of Epidemiology 2019b: volume 188(7), pages 1288-1295

Ranjbaran M et al. 'Ambient temperature and air pollution, and the risk of preterm birth in Tehran, Iran: a time series study'. Journal of Maternal-Fetal and Neonatal Medicine 2022: volume 35(4), pages 726-737

Rappazzo KM et al. 'Exposure to fine particulate matter during pregnancy and risk of preterm birth among women in New Jersey, Ohio, and Pennsylvania, 2000-2005'. Environmental Health Perspectives 2014: volume 122(9), pages 992-997

RCP 'Every breath we take: the lifelong impact of air pollution. Report of a working party'. Royal College of Physicians and Royal College of Paediatrics and Child Health 2016

Regnault TR et al. 'Transport and metabolism of amino acids in placenta'. Endocrine 2002: volume 19(1), pages 23-41

Ritz B et al. 'Ambient air pollution and preterm birth in the environment and pregnancy outcomes study at the University of California, Los Angeles'. American Journal of Epidemiology 2007: volume 166(9), pages 1045-1052

Ritz B et al. 'Effect of air pollution on preterm birth among children born in Southern California between 1989 and 1993'. Epidemiology 2000: volume 11(5), pages 502-511

Robinson O et al. 'Cord Blood Metabolic Signatures of Birth Weight: A Population-Based Study'. Journal of Proteome Research 2018: volume 17(3), pages 1235-1247

Rogers JF et al. 'Air pollution and very low birth weight infants: a target population?'. Pediatrics 2006: volume 118(1), pages 156-164

Romão R et al. 'The relationship between low birth weight and exposure to inhalable particulate matter'. Cadernos de Saúde Pública 2013: volume 29(6), pages 1101-1108

Rosa MJ et al. 'Identifying sensitive windows for prenatal particulate air pollution exposure and mitochondrial DNA content in cord blood'. Environment International 2017: volume 98, pages 198-203

Rosenfeld CR et al. 'Nitric oxide contributes to estrogen-induced vasodilation of the ovine uterine circulation'. Journal of Clinical Investigation 1996: volume 98(9), pages 2158-2166

Rossner P, Jr. et al. 'Genetic, biochemical, and environmental factors associated with pregnancy outcomes in newborns from the Czech Republic'. Environmental Health Perspectives 2011: volume 119(2), pages 265-271

Rudov A et al. 'Involvement of miRNAs in placental alterations mediated by oxidative stress'. Oxidative Medicine and Cellular Longevity 2014: volume 2014, article 103068

Rudra CB et al. 'Ambient carbon monoxide and fine particulate matter in relation to preeclampsia and preterm delivery in western Washington State'. Environmental Health Perspectives 2011: volume 119(6), pages 886-892

Rupnow HL et al. 'Endothelial vasodilator production by uterine and systemic arteries. VII. Estrogen and progesterone effects on eNOS'. American Journal of Physiology Heart and Circulatory Physiology 2001: volume 280(4), pages H1699-1705

Saenen ND et al. 'Placental Nitrosative Stress and Exposure to Ambient Air Pollution During Gestation: A Population Study'. American Journal of Epidemiology 2016: volume 184(6), pages 442-449

Saenen ND et al. 'Lower Placental Leptin Promoter Methylation in Association with Fine Particulate Matter Air Pollution during Pregnancy and Placental Nitrosative Stress at Birth in the ENVIRONAGE Cohort'. Environmental Health Perspectives 2017: volume 125(2), pages 262-268

Sagawa N et al. 'Possible role of placental leptin in pregnancy: a review'. Endocrine 2002: volume 19(1), pages 65-71

Sagiv SK et al. 'A time-series analysis of air pollution and preterm birth in Pennsylvania, 1997-2001'. Environmental Health Perspectives 2005: volume 113(5), pages 602-606

Salam MT et al. 'Birth outcomes and prenatal exposure to ozone, carbon monoxide, and particulate matter: results from the Children's Health Study'. Environmental Health Perspectives 2005: volume 113(11), pages 1638-1644

Saldiva S et al. 'Small-Scale Variations in Urban Air Pollution Levels Are Significantly Associated with Premature Births: A Case Study in São Paulo, Brazil'. International Journal of Environmental Research and Public Health 2018: volume 15(10), article 2236

Salhab WA et al. 'Regulation of types I and III NOS in ovine uterine arteries by daily and acute estrogen exposure'. American Journal of Physiology Heart and Circulatory Physiology 2000: volume 278(6), pages H2134-2142

Salihu HM et al. 'Particulate pollutants and racial/ethnic disparity in feto-infant morbidity outcomes'. Maternal and Child Health Journal 2012: volume 16(8), pages 1679-1687

Sapkota A et al. 'Exposure to particulate matter and adverse birth outcomes: a comprehensive review and meta-analysis'. Air Quality, Atmosphere & Health 2012: volume 5(4), pages 369-381

Sarovar V et al. 'A case-crossover study of short-term air pollution exposure and the risk of stillbirth in California, 1999-2009'. Environmental Research 2020: volume 191, article 110103

Scifres CM et al. 'Intrauterine growth restriction, human placental development and trophoblast cell death'. Journal of Physiology 2009: volume 587(Pt 14), pages 3453-3458

Seabrook JA et al. 'Geospatial analyses of adverse birth outcomes in Southwestern Ontario: Examining the impact of environmental factors'. Environmental Research 2019: volume 172, pages 18-26

Seo JH et al. 'Population-attributable risk of low birthweight related to PM_{10} pollution in seven Korean cities'. Paediatric and Perinatal Epidemiology 2010: volume 24(2), pages 140-148

Shea BJ et al. 'AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both'. British Medical Journal 2017: volume 358, article j4008

Sheridan P et al. 'Ambient Fine Particulate Matter and Preterm Birth in California: Identification of Critical Exposure Windows'. American Journal of Epidemiology 2019: volume 188(9), pages 1608-1615

Shields BM et al. 'Fetal thyroid hormone level at birth is associated with fetal growth'. Journal of Clinical Endocrinology and Metabolism 2011: volume 96(6), pages E934-938

Siddika N et al. 'Prenatal ambient air pollution exposure and the risk of stillbirth: systematic review and meta-analysis of the empirical evidence'. Occupational and Environmental Medicine 2016: volume 73(9), pages 573-581

Siddika N et al. 'Synergistic effects of prenatal exposure to fine particulate matter ($PM_{2.5}$) and ozone (O_3) on the risk of preterm birth: A population-based cohort study'. Environmental Research 2019: volume 176, article 108549

Simoncic V et al. 'Adverse Birth Outcomes Related to NO_2 and PM Exposure: European Systematic Review and Meta-Analysis'. International Journal of Environmental Research and Public Health 2020: volume 17(21), article 8116

Slama R et al. 'Traffic-related atmospheric pollutants levels during pregnancy and offspring's term birth weight: a study relying on a land-use regression exposure model'. Environmental Health Perspectives 2007: volume 115(9), pages 1283-1292

Slator PJ et al. 'Placenta microstructure and microcirculation imaging with diffusion MRI'. Magnetic Resonance in Medicine 2018: volume 80(2), pages 756-766

Smith RB et al. 'Impact of London's road traffic air and noise pollution on birth weight: retrospective population based cohort study'. British Medical Journal 2017: volume 359, article j5299

Smith RB et al. 'Impacts of air pollution and noise on risk of preterm birth and stillbirth in London'. Environment International 2020: volume 134, article 105290

Starling AP et al. 'Prenatal exposure to traffic and ambient air pollution and infant weight and adiposity: The Healthy Start study'. Environmental Research 2020: volume 182, article 109130

Stieb DM et al. 'Ambient air pollution, birth weight and preterm birth: a systematic review and meta-analysis'. Environmental Research 2012: volume 117, pages 100-111

Stieb DM et al. 'A national study of the association between traffic-related air pollution and adverse pregnancy outcomes in Canada, 1999-2008'. Environmental Research 2016a: volume 148, pages 513-526

Stieb DM et al. 'Associations of Pregnancy Outcomes and $PM_{2.5}$ in a National Canadian Study'. Environmental Health Perspectives 2016b: volume 124(2), pages 243-249

Stieb DM et al. 'Air pollution in the week prior to delivery and preterm birth in 24 Canadian cities: a time to event analysis'. Environmental Health 2019: volume 18(1), article 1

Strand LB et al. 'Methodological challenges when estimating the effects of season and seasonal exposures on birth outcomes'. BioMed Central Medical Research Methodology 2011: volume 11, article 49

Stulc J. 'Placental transfer of inorganic ions and water'. Physiological Reviews 1997: volume 77(3), pages 805-836

Suh YJ et al. 'GSTM1 polymorphism along with PM_{10} exposure contributes to the risk of preterm delivery'. Mutation Research 2008: volume 656(1-2), pages 62-67

Suh YJ et al. 'Different effects of PM_{10} exposure on preterm birth by gestational period estimated from time-dependent survival analyses'. International Archives of Occupational and Environmental Health 2009: volume 82(5), pages 613-621

Sun FL et al. 'Transactivation of Igf2 in a mouse model of Beckwith-Wiedemann syndrome'. Nature 1997: volume 389(6653), pages 809-815

Sun X et al. 'The association between fine particulate matter exposure during pregnancy and preterm birth: a meta-analysis'. BioMed Central Pregnancy and Childbirth 2015: volume 15, article 300

Sun X et al. 'The associations between birth weight and exposure to fine particulate matter ($PM_{2.5}$) and its chemical constituents during pregnancy: A meta-analysis'. Environmental Pollution 2016: volume 211, pages 38-47

Sun Z et al. 'Maternal ambient air pollution exposure with spatial-temporal variations and preterm birth risk assessment during 2013-2017 in Zhejiang Province, China'. Environment International 2019: volume 133(Pt B), article 105242

Suter MA et al. 'Association between elevated placental polycyclic aromatic hydrocarbons (PAHs) and PAH-DNA adducts from Superfund sites in Harris County, and increased risk of preterm birth (PTB)'. Biochemical and Biophysical Research Communications 2019: volume 516(2), pages 344-349

Symanski E et al. 'Evaluating narrow windows of maternal exposure to ozone and preterm birth in a large urban area in Southeast Texas'. Journal of Exposure Science & Environmental Epidemiology 2016: volume 26(2), pages 167-172

Szyf M. 'Early life, the epigenome and human health'. Acta Paediatrica 2009: volume 98(7), pages 1082-1084

Takagi Y et al. 'Levels of oxidative stress and redox-related molecules in the placenta in preeclampsia and fetal growth restriction'. Virchows Archiv 2004: volume 444(1), pages 49-55

Tang D et al. 'PAH-DNA adducts in cord blood and fetal and child development in a Chinese cohort'. Environmental Health Perspectives 2006: volume 114(8), pages 1297-1300

Tang D et al. 'Air pollution effects on fetal and child development: a cohort comparison in China'. Environmental Pollution 2014: volume 185, pages 90-96

Tapia VL et al. 'Association between maternal exposure to particulate matter ($PM_{2.5}$) and adverse pregnancy outcomes in Lima, Peru'. Journal of Exposure Science & Environmental Epidemiology 2020: volume 30(4), pages 689-697

TCEQ. 'TCEQ Guidelines for Systematic Review and Evidence Integration' Texas Commission on Environmental Quality 2017

Thayamballi N et al. 'Impact of Maternal Demographic and Socioeconomic Factors on the Association Between Particulate Matter and Adverse Birth Outcomes: a Systematic Review and Meta-analysis'. Journal of Racial and Ethnic Health Disparities 2021: volume 8(3), pages 743-755

Thornton CA. 'Immunology of pregnancy'. Proceedings of the Nutrition Society 2010: volume 69(3), pages 357-365

Tjoa ML et al. 'Elevated C-reactive protein levels during first trimester of pregnancy are indicative of preeclampsia and intrauterine growth restriction'. Journal of Reproductive Immunology 2003: volume 59(1), pages 29-37

Topinka J et al. 'DNA adducts in human placenta as related to air pollution and to GSTM1 genotype'. Mutation Research/Genetic Toxicology and Environmental Mutagenesis 1997: volume 390(1-2), pages 59-68

Trasande L et al. 'Exploring prenatal outdoor air pollution, birth outcomes and neonatal health care utilization in a nationally representative sample'. Journal of Exposure Science & Environmental Epidemiology 2013: volume 23(3), pages 315-321

Trevisanuto D et al. 'High-sensitivity C-reactive protein in umbilical cord of small-for-gestational-age neonates'. Neonatology 2007: volume 91(3), pages 186-189

Tsamou M et al. 'Air pollution-induced placental epigenetic alterations in early life: a candidate miRNA approach'. Epigenetics 2018: volume 13(2), pages 135-146

Twum C et al. 'Maternal exposure to ambient $PM_{2.5}$ and term low birthweight in the State of Georgia'. International Journal of Environmental Health Research 2016: volume 26(1), pages 92-100

Urman R et al. 'Associations of children's lung function with ambient air pollution: joint effects of regional and near-roadway pollutants'. Thorax 2014: volume 69(6), pages 540-547

USEPA. 'Preamble To The Integrated Science Assessments (ISA)' EPA/600/R-15/06. US Environmental Protection Agency, Washington DC, 2015

USEPA. 'Integrated Science Assessment (ISA) for Oxides of Nitrogen – Health Criteria (Final Report, Jan 2016)'. US Environmental Protection Agency, Washington DC, 2016. Contract No.: EPA/600/R-15/068

USEPA. 'Toxicological review of benzo[a]pyrene [CASRN 50-32-8]' EPA/635/R-17/003FA Integrated Risk Information System, National Center for Environmental Assessment, Office of Research and Development, US. Environmental Protection Agency, Washington DC, 2017

USEPA. 'Integrated Science Assessment (ISA) for Particulate Matter (Final Report, Dec 2019)'. Washington, DC; 2019. Contract No. EPA/600/R-19/188

USEPA. 'Integrated Science Assessment (ISA) for Ozone and Related Photochemical Oxidants (Final Report, Apr 2020)'. Washington, DC, 2020. Contract No.: EPA/600/R-20/012

USEPA 'Children Are Not Little Adults!' USEPA US. Environmental Protection Agency, Washington DC, 2024 (viewed 6 May 2024)

Vadillo-Ortega F et al. 'Air pollution, inflammation and preterm birth: a potential mechanistic link'. Medical Hypotheses 2014: volume 82(2), pages 219-224

Vagnoni KE et al. 'Endothelial vasodilator production by uterine and systemic arteries. III. Ovarian and estrogen effects on NO synthase'. American Journal of Physiology Heart and Circulatory Physiology 1998: volume 275(5), pages H1845-1856

Valentino SA et al. 'Maternal exposure to diluted diesel engine exhaust alters placental function and induces intergenerational effects in rabbits'. Particle and Fibre Toxicology 2016: volume 13(1), article 39

van den Hooven EH et al. 'Air pollution exposure during pregnancy, ultrasound measures of fetal growth, and adverse birth outcomes: a prospective cohort study'. Environmental Health Perspectives 2012a: volume 120(1), pages 150-156

van den Hooven EH et al. 'Chronic air pollution exposure during pregnancy and maternal and fetal C-reactive protein levels: the Generation R Study'. Environmental Health Perspectives 2012b: volume 120(5), pages 746-751

van den Hooven EH et al. 'Air pollution exposure and markers of placental growth and function: the generation R study'. Environmental Health Perspectives 2012c: volume 120(12), pages 1753-1759

van der Wijst MG et al. 'Experimental mitochondria-targeted DNA methylation identifies GpC methylation, not CpG methylation, as potential regulator of mitochondrial gene expression'. Scientific Reports 2017: volume 7(1), article 177

Veras MM et al. 'Particulate urban air pollution affects the functional morphology of mouse placenta'. Biology of Reproduction 2008: volume 79(3), pages 578-584

Veras MM et al. 'The effects of particulate ambient air pollution on the murine umbilical cord and its vessels: a quantitative morphological and immunohistochemical study'. Reproductive Toxicology 2012: volume 34(4), pages 598-606

Vinikoor-Imler LC et al. 'Associations between prenatal exposure to air pollution, small for gestational age, and term low birthweight in a state-wide birth cohort'. Environmental Research 2014: volume 132, pages 132-139

Vinturache A et al. 'Maternal Physiological Changes in Pregnancy' in 'Fetal development and maternal adaptation' edited by Khalil A (Volume 4 of The continuous textbook of women's medicine series – Obstetrics module). The Alliance for Global Women's Medicine, 2021

von Versen-Hoeynck FM et al. 'Plasma levels of inflammatory markers neopterin, sialic acid, and C-reactive protein in pregnancy and preeclampsia'. American Journal of Hypertension 2009: volume 22(6), pages 687-692

Vos S et al. 'Mitochondrial DNA methylation in placental tissue: a proof of concept study by means of prenatal environmental stressors'. Epigenetics 2021: volume 16(2), pages 121-131

Vrijens K et al. 'MicroRNAs as potential signatures of environmental exposure or effect: a systematic review'. Environmental Health Perspectives 2015: volume 123(5), pages 399-411

Wadi MA et al. 'Relationship between birth weight and domestic maternal passive smoking exposure'. Eastern Mediterranean Health Journal 2011: volume 17(4), pages 290-296

Wang Q et al. 'Interaction of Air Pollutants and Meteorological Factors on Birth Weight in Shenzhen, China'. Epidemiology 2019: volume 30 Suppl 1, pages S57-s66

Wang L et al. 'Traffic-related metrics and adverse birth outcomes: A systematic review and meta-analysis'. Environmental Research 2020: volume 188, article 109752

Warren JL et al. 'Investigating the Impact of Maternal Residential Mobility on Identifying Critical Windows of Susceptibility to Ambient Air Pollution During Pregnancy'. American Journal of Epidemiology 2018: volume 187(5), pages 992-1000

Weber KA et al. 'Air Pollution, Maternal Hypertensive Disorders, and Preterm Birth'. Environmental Epidemiology 2019: volume 3(5), article e062

Wei SQ et al. 'Inflammatory cytokines and spontaneous preterm birth in asymptomatic women: a systematic review'. Obstetrics and Gynecology 2010: volume 116(2 Pt 1), pages 393-401

Weisskopf MG et al. 'Trade-offs of Personal Versus More Proxy Exposure Measures in Environmental Epidemiology'. Epidemiology 2017: volume 28(5), pages 635-643

Weldy CS et al. 'In utero exposure to diesel exhaust air pollution promotes adverse intrauterine conditions, resulting in weight gain, altered blood pressure, and increased susceptibility to heart failure in adult mice'. Public Library of Science One 2014: volume 9(2), article e88582

WHO. 'Effects of air pollution on children's health and development: a review of the evidence' World Health Organization Regional Office for Europe, Copenhagen, 2005. Document number: EUR/05/5046027

WHO. 'Review of evidence on health aspects of air pollution – REVIHAAP Project: Technical Report' World Health Organization Regional Office for Europe Copenhagen, 2013

WHO. 'Air pollution and child health: prescribing clean air' World Health Organization, Geneva, 2018. Reference number: WHO/CED/PHE/18.01

WHO. 'Fact sheet on Preterm Birth' World Health Organization, Geneva 2023

Wilcox AJ et al. 'On the pitfalls of adjusting for gestational age at birth'. American Journal of Epidemiology 2011: volume 174(9), pages 1062-1068

Wilhelm M et al. 'Local variations in CO and particulate air pollution and adverse birth outcomes in Los Angeles County, California, USA'. Environmental Health Perspectives 2005: volume 113(9), pages 1212-1221

Wilhelm M et al. 'Traffic-related air toxics and term low birth weight in Los Angeles County, California'. Environmental Health Perspectives 2012: volume 120(1), pages 132-138

Wilhelm-Benartzi CS et al. 'In utero exposures, infant growth, and DNA methylation of repetitive elements and developmentally related genes in human placenta'. Environmental Health Perspectives 2012: volume 120(2), pages 296-302

Williams ML et al. 'The Lancet Countdown on health benefits from the UK Climate Change Act: a modelling study for Great Britain'. The Lancet Planetary Health 2018: volume 2(5), pages e202-e213

Williams MC et al. 'Histologic chorioamnionitis is associated with fetal growth restriction in term and preterm infants'. American Journal of Obstetrics and Gynecology 2000: volume 183(5), pages 1094-1099

Wu J et al. 'Association between local traffic-generated air pollution and preeclampsia and preterm delivery in the south coast air basin of California'. Environmental Health Perspectives 2009: volume 117(11), pages 1773-1779

Wu J et al. 'Comparing exposure assessment methods for traffic-related air pollution in an adverse pregnancy outcome study'. Environmental Research 2011: volume 111(5), pages 685-692

Wu H et al. 'Exposure to fine particulate matter during pregnancy and risk of term low birth weight in Jinan, China, 2014-2016'. International Journal of Hygiene and Environmental Health 2018: volume 221(2), pages 183-190

Xiao Q et al. 'Associations between birth outcomes and maternal $PM_{2.5}$ exposure in Shanghai: A comparison of three exposure assessment approaches'. Environment International 2018: volume 117, pages 226-236

Xu X et al. 'PM_{10} air pollution exposure during pregnancy and term low birth weight in Allegheny County, PA, 1994-2000'. International Archives of Occupational and Environmental Health 2011: volume 84(3), pages 251-257

Xue T et al. 'Association between pregnancy loss and ambient $PM_{2.5}$ using survey data in Africa: a longitudinal case-control study, 1998-2016'. The Lancet Planetary Health 2019: volume 3(5), pages e219-e225

Yan Q et al. 'Maternal serum metabolome and traffic-related air pollution exposure in pregnancy'. Environment International 2019: volume 130, article 104872

Yang S et al. 'Ambient air pollution the risk of stillbirth: A prospective birth cohort study in Wuhan, China'. International Journal of Hygiene and Environmental Health 2018: volume 221(3), pages 502-509

Yaya S et al. 'Exposure to second-hand smoking as a predictor of fetal loss: Egypt Demographic and Health Survey 2014'. International Health 2019: volume 11(6), pages 561-567

Yi FX et al. 'Pregnancy enhances sustained Ca2+ bursts and endothelial nitric oxide synthase activation in ovine uterine artery endothelial cells through increased connexin 43 function'. Biology of Reproduction 2010: volume 82(1), pages 66-75

Ye L et al. 'Associations between maternal exposure to air pollution and birth outcomes: a retrospective cohort study in Taizhou, China'. Environmental Science and Pollution Research International 2018: volume 25(22), pages 21927-21936

Ye Y et al. 'Environmental Pollutant Benzo[a]pyrene Induces Recurrent Pregnancy Loss through Promoting Apoptosis and Suppressing Migration of Extravillous Trophoblast'. BioMed Research International 2020: volume 2020, article 8983494

Yin LJ et al. 'Insufficient maintenance DNA methylation is associated with abnormal embryonic development'. BMC Medicine 2012: volume 10, article 26

Yorifuji T et al. 'Outdoor air pollution and term low birth weight in Japan'. Environment International 2015: volume 74, pages 106-111

Yuan L et al. 'Critical windows for maternal fine particulate matter exposure and adverse birth outcomes: The Shanghai birth cohort study'. Chemosphere 2020: volume 240, article 124904

Yue H et al. 'Gestational exposure to $PM_{2.5}$ impairs vascularization of the placenta'. Science of the Total Environment 2019: volume 665, pages 153-161

Yung HW et al. 'Evidence of placental translation inhibition and endoplasmic reticulum stress in the etiology of human intrauterine growth restriction'. American Journal of Pathology 2008: volume 173(2), pages 451-462

Zang H et al. 'Ambient air pollution and the risk of stillbirth: a population-based prospective birth cohort study in the coastal area of China'. Environmental Science and Pollution Research International 2019: volume 26(7), pages 6717-6724

Zhang L et al. 'Modulation by benzo[a]pyrene of epidermal growth factor receptors, cell proliferation, and secretion of human chorionic gonadotropin in human placental cell lines'. Biochemical Pharmacology 1995: volume 50(8), pages 1171-1180

Zhang H et al. 'Ambient air pollution and stillbirth: An updated systematic review and meta-analysis of epidemiological studies'. Environmental Pollution 2021: volume 278, article 116752

Zhao N et al. 'Ambient air pollutant PM_{10} and risk of preterm birth in Lanzhou, China'. Environment International 2015: volume 76, pages 71-77

Zhao B et al. 'Seasonal response of the synergism of maternal comorbidities and long-term air pollution exposure on birth outcomes'. Ecotoxicology and Environmental Safety 2020: volume 191, article 110232

Zhao N et al. 'Polymorphisms in oxidative stress, metabolic detoxification, and immune function genes, maternal exposure to ambient air pollution, and risk of preterm birth in Taiyuan, China'. Environmental Research 2021: volume 194, article 110659

Zhu X et al. 'Maternal exposure to fine particulate matter ($PM_{2.5}$) and pregnancy outcomes: a meta-analysis'. Environmental Science and Pollution Research International 2015: volume 22(5), pages 3383-3396

Zhu J et al. 'Exposure to ambient $PM_{2.5}$ during pregnancy and preterm birth in metropolitan areas of the state of Georgia'. Environmental Science and Pollution Research International 2019: volume 26(3), pages 2492-2500

Appendices

Appendix 1. Full list of Sub-group and Working Group Members

	Member	Affiliation
Chair	Professor Jonathan Grigg	(Queen Mary University of London)
Members	Professor Debbie Jarvis	(Imperial College London)
	Dr Heather Walton	(Imperial College London)
Co-Opted Members:	Dr Rachel Smith	(Imperial College London)
	Dr Julia Fussell	(Imperial College London)
	Professor Mireille Toledano	(Imperial College London)
	Professor Peter Brocklehurst	(University of Birmingham)
	Professor Eric Jauniaux	(University College London)
Scientific Secretariat:	Dr Philippa Douglas	(UK Health Security Agency)
	Dr Karen Exley	(UK Health Security Agency)
	Eleanor Sykes	(UK Health Security Agency)
Scientific Secretariat (preparation for publication):	Eve Draper	(UK Health Security Agency)
	Dr Naomi Earl	(UK Health Security Agency)
	Alison Gowers	(UK Health Security Agency)

1. Members of the Epidemiological Working Group

	Member	Affiliation
Subgroup Chair	Dr Rachel Smith	(Imperial College London)
Members	Professor Debbie Jarvis	(Imperial College London)
	Dr Heather Walton	(Imperial College London)
	Professor Mireille Toledano	(Imperial College London)
Scientific Secretariat:	Dr Philippa Douglas	(UK Health Security Agency)
	Dr Karen Exley	(UK Health Security Agency)
	Ms Eleanor Sykes	(UK Health Security Agency)

2. Members of the Mechanistic Working Group

	Member	Affiliation
Subgroup Chair	Professor Jonathan Grigg	(NIHR Senior Investigator)
Co-Opted Members:	Dr Julia Fussell	(Imperial College London)
	Professor Peter Brocklehurst	(University of Birmingham)
	Professor Eric Jauniaux	(University College London)
Scientific Secretariat:	Dr Philippa Douglas	(UK Health Security Agency)
	Dr Karen Exley	(UK Health Security Agency)
	Ms Eleanor Sykes	(UK Health Security Agency)

Appendix 2. Terms of reference

Committee on the Medical Effects of Air Pollutants (COMEAP)

Terms of Reference for a COMEAP Sub-group on the effects of maternal exposure to ambient air pollution on adverse birth outcomes

1. Background

In 2016, COMEAP identified early life effects and adverse birth outcomes as a topic for consideration. Since then there have been a number of research studies and the Royal College of Physicians and Royal College of Paediatrics and Child Health published its report 'Every breath we take: the lifelong impact of air pollution (RCP, 2016). As there has been growing interest from the public and the government about the effects of air pollution on early life, the Department of Health and Social Care (DHSC) requested that COMEAP begin work on this topic.

To support COMEAP's work, DHSC approached the National Institute of Health Research (NIHR) Health Protection Research Unit (HPRU) on Health Impact of Environmental Hazards at King's College London to undertake an evidence-based scoping review on air pollution and adverse birth outcomes. The scope of the review was discussed and agreed with Public Health England[11].

2. Scope of the work

It is proposed that the COMEAP Sub-group (see below) consider the scoping review in more detail and carry out further work needed for preparing a statement on the current evidence linking maternal exposure to ambient air pollutants with adverse birth outcomes.

It has been agreed that the main health outcomes focus will be on PTB and LBW at term as this is where most of the evidence lies, and also stillbirth. However, other outcomes such as congenital anomalies, head circumference and miscarriage, may still be commented on.

It has been agreed that for the epidemiological studies, the air pollutants of focus would be particulate matter (including $PM_{2.5}$, PM_{10}), nitrogen dioxide (NO_2), and ozone (O_3). For the mechanistic studies, the air pollutants considered would be particulate matter (including $PM_{2.5}$, PM_{10}, black carbon), NO_2, polycyclic aromatic hydrocarbons (including BaP), diesel exhaust, and O_3. It has been agreed that the focus will be on single pollutants and not mixtures.

[11] UK Health Security Agency from the 1 October 2021 (Public Health England was replaced by UK Health Security Agency and Office for Health Improvement and Disparities).

3. Questions to be addressed

Members' views will be invited to address the following questions:

1. What is the strength and consistency of epidemiological evidence linking ambient air pollutants with adverse birth outcomes?
2. Are there biologically plausible mechanisms by which air pollutants could affect adverse birth outcomes?
3. Is the evidence sufficient to propose quantification of specific adverse birth outcomes associated with air pollution in the UK?
4. Can recommendations be identified from current meta-analyses or should new meta-analyse(s) be undertaken?

4. Resources

It is proposed that a COMEAP Sub-group be formed. This will largely comprise current COMEAP Members, but additional Members will be co-opted for particular expertise as required.

COMEAP Secretariat support will be led by Dr Pippa Douglas and Ms Eleanor Sykes, with support from Dr Karen Exley.

5. Timescales

The COMEAP Sub-group will be established during February 2019. Regular follow-up meetings of the group will be organised (the frequency of meetings will be decided at the first Sub-group meeting).

The Sub-group should aim to bring an initial view on the evidence and proposed approach for the work for discussion at the COMEAP meeting in June 2019. It is expected that a statement will be finalised during 2020.

Proposed output

It has been agreed that the first output will be a COMEAP statement on the evidence for the effects of maternal exposure to ambient air pollution on adverse birth outcomes.

The statement may be accompanied by a working paper updating the previous scoping review, incorporating studies that have not yet been included in prior reviews.

If appropriate, quantification of these effects in the UK or the recommendation of concentration-response functions for use in quantification projects will be considered as part of a separate work item later.

Appendix 3. UK Burden of stillbirth, preterm birth, and low birthweight

Table 22. Live births, stillbirths and the stillbirth rate by Index of Multiple Deprivation (IMD) (see notes[1,2]) and country of usual residence, England and Wales, 2019

	Mothers usually resident in England[3]			Mothers usually resident in Wales[3]		
IMD Decile	**Live births**	**Stillbirths**	**Stillbirth rate (stillbirths per 1,000 live births and stillbirths)**	**Live births**	**Stillbirths**	**Stillbirth rate (stillbirths per 1,000 live births and stillbirths)**
1	82,979	443	5.31	4,081	18	4.39
2	74,770	357	4.75	3,658	24	6.52
3	71,411	308	4.29	3,009	11	3.64
4	65,028	270	4.13	3,261	13	3.97
5	61,047	228	3.72	2,889	23	7.90
6	58,068	173	2.97	2,850	10	3.50
7	53,213	183	3.43	2,682	14	5.19
8	51,794	157	3.02	2,436	11	4.50
9	49,050	123	2.50	2,542	10	3.92
10	43,145	104	2.40	2,296	4	1.74
Total	610,505	2,346	3.83	29,704	138	4.62

[1] Index of Multiple Deprivation used is based on The English Indices of Deprivation 2019 and The Welsh Indices of Multiple Deprivation 2019.

[2] Index of Multiple Deprivation deciles range from 1 to 10, with 1 being the most deprived and 10 being the least deprived.

[3] Figures are based on geography boundaries as of May 2020

(Source: Office for National Statistics, released: 16 November 2020)

Table 23. Exemplar of potential costs around stillbirth and preterm birth

Stillbirth	Cost	Reference
Cost per case		
Health and social care costs	£4,191	
Funeral costs	£559	Campbell et al (2018) (37)
Workplace absence (parents and healthcare professionals)	£3,829	
Workplace absence – parents	£2,476	Costs assessed for 2 years after stillbirth.
Productivity loss (child, lifetime)	£213,304	
UK costs		See note (a) for breakdown of 'health and social care costs'
Health and social care cost	£13.6 M/y	
Funerals	£1.8 M/y	
Total productivity losses (note (b))	£345-706 M/y	
Productivity losses (parents)	£8.1 M/y	Costs in 2013/14 UK £724 M/y
Productivity loss (child, lifetime)	£333-694 M/y	

Clinical negligence payments	£2.5 M/y	
Preterm birth		
Cost per case	£22,679	Campbell et al (2018) (37) Breakdown of costs not available
UK costs	£3 billion/y	Mangham et al (2009) (36) Breakdown of costs not available

(a) Health and social care costs from Campbell et al (2018) includes:

(1) Immediate antepartum, intrapartum and postpartum care

(2) Parental anxiety and depression

(3) Experiences of healthcare professionals

(4) Subsequent pregnancy care covering: antenatal care during pregnancies occurring within 12 months of a prior stillbirth; outcome of subsequent pregnancies occurring within 12 months of a prior stillbirth, excess preterm live births

Appendix 4. Search terms used in the updated literature search

Searches were conducted in PubMed. Results were limited to literature published in the English language.

1. Epidemiological evidence search string

The following search string is a replicate of the search string used in the KCL scoping review (Fussell and Walton, 2018):

((air pollution OR particulate matter OR carbon monoxide OR ozone OR nitrogen dioxide OR polyaromatic hydrocarbons OR traffic OR noise)) AND (birth outcomes OR birth weight OR low birth weight OR small for gestational age OR preterm birth OR stillbirth OR spontaneous abortion OR foetal loss OR miscarriage OR missed abortion OR intrauterine death)

2. Mechanistic evidence search string

The following search string is a replicate of the search string used in the KCL scoping review (Fussell and Walton, 2018):

((air pollution OR particulate matter OR carbon monoxide OR ozone OR nitrogen dioxide OR sulphur dioxide OR polyaromatic hydrocarbons OR traffic)) AND (placental weight OR placental structure OR placental morphology OR placental function OR placental inflammation OR placental oxidative stress OR umbilical cord structure OR umbilical cord morphology OR umbilical cord function OR embryonic resorption OR intrauterine growth OR developing fetus OR DNA adducts OR hypoxia OR foetal outcomes OR foetal weight OR foetal oxidative stress OR foetal inflammation)

The following search string was added, in light of the results of the mechanistic evidence summarised in the KCL scoping review (Fussell and Walton, 2018):

((air pollution OR particulate matter OR carbon monoxide OR ozone OR nitrogen dioxide OR sulphur dioxide OR polyaromatic hydrocarbons OR traffic)) AND (hypoxia OR adducts OR epigenetics OR foetal thyroid function OR embryonic growth retardation OR umbilical artery OR hormone synthesis OR progesterone synthesis OR aromatase activity OR endocrine disruption OR gene expression profiling OR gene ontology OR low placenta weight OR foetal thrombotic vasculopathy OR metabolic detoxification OR maternal hyperthyroidism OR maternal hypothyroidism)

Appendix 5. Details of whether included meta-analyses concerned term low birth weight (TLBW) or low birth weight (LBW)

Please note: Papers for each pollutant/exposure have been combined for all exposure windows (WP, T1, T2, T3)				
Meta-analysis	**LBW definition**	**Pollutant/Outcome**	**Included studies**	**TLBW or LBW**
Li et al 2020 (71)	Inclusion criteria to define LBW as birth weight less than 2500 g and gestational age of 37 weeks or more.	**$PM_{2.5}$ TLBW**	Chen et al 2018	TLBW
			Choe et al 2019	TLBW
			Dadvand et al 2014	TLBW
			Ha et al 2014	TLBW
			Hao et al 2016	TLBW
			Harris et al 2014	TLBW
			Hyder et al 2014	TLBW
			Kingsley et al 2017	TLBW
			Kirwa et al 2019	TLBW
			Kumar et al 2016	TLBW
			Lavigne et al 2016	TLBW
			Liang et al 2019a	TLBW
			Smith et al 2017	TLBW
			Stieb et al 2016a	TLBW
			Tapia et al 2020	TLBW
			Vinikoor-Imler et al 2014	TLBW
			Wu et al 2018	TLBW
			Xiao et al 2018	TLBW

Please note: Papers for each pollutant/exposure have been combined for all exposure windows (WP, T1, T2, T3)				
Meta-analysis	**LBW definition**	**Pollutant/Outcome**	**Included studies**	**TLBW or LBW**
			Ye et al 2018	TLBW
			Yuan et al 2020	LBW
		PM_{10}, TLBW	Araban et al 2012	TLBW
			Choe et al 2019)	TLBW
			Dadvand et al 2014	TLBW
			Dugandzic et al 2006	TLBW
			Kim et al 2007	TLBW
			Kumar et al 2016	TLBW
			Lee et al 2003	TLBW
			Lu et al 2020	TLBW
			Maisonet et al 2001	TLBW
			Smith et al 2017	TLBW
			Wang et al 2019	TLBW
			Xu et al 2011	TLBW
			Ye et al 2018	TLBW
		NO_2 and TLBW	Araban et al 2012	TLBW
			Dadvand et al 2014	TLBW
			Ghosh et al 2012	TLBW
			Laurent et al 2014	TLBW
			Lavigne et al 2016	TLBW
			Lee et al 2003	TLBW
			Lu et al 2020	TLBW

Please note: Papers for each pollutant/exposure have been combined for all exposure windows (WP, T1, T2, T3)				
Meta-analysis	**LBW definition**	**Pollutant/Outcome**	**Included studies**	**TLBW or LBW**
			Maroziene and Grazuleviciene 2002	TLBW
			Michikawa et al 2017	TLBW
			Salam et al 2005	TLBW
			Smith et al 2017	TLBW
			Wang et al 2019	TLBW
			Ye et al 2018	TLBW
		O_3 and TLBW	Araban et al 2012	TLBW
			Chen et al 2018	TLBW
			Dugandzic et al 2006	TLBW
			Ha et al 2001	TLBW
			Ha et al 2014	TLBW
			Michikawa et al 2017	TLBW
			Salam et al 2005	TLBW
			Smith et al 2017	TLBW
			Vinikoor-Imler et al 2014	TLBW
Zhu et al 2015 (67)	Inclusion and exclusion criteria section: 'LBW was defined as live births weighing less than 2,500 g'	**$PM_{2.5}$, LBW**	Brauer et al 2008	TLBW
			Darrow et al 2011	TLBW
			Ebisu & Bell 2012	TLBW
			Gehring et al 2011	TLBW

Please note: Papers for each pollutant/exposure have been combined for all exposure windows (WP, T1, T2, T3)				
Meta-analysis	**LBW definition**	**Pollutant/Outcome**	**Included studies**	**TLBW or LBW**
			Gray et al 2014	LBW rather than TLBW but the following was stated: 'We excluded multiple births (3.4%), and infants with diagnosed congenital anomalies at the time of birth (0.8%). These exclusions were chosen as we sought to focus on those pregnancies that could reasonably be expected to go to term and deliver at a normal BWT.
			Morello-Frosch et al 2010	TLBW
Sapkota et al 2012 (62)	Outcome definitions section: 'To allow us to summarize results for the 2 most independent measures of growth and preterm birth, here we only included those studies that examined preterm birth (<37 weeks completed gestation) as the major endpoint, regardless of birth weight and LBW with an emphasis on term LBW.'	**$PM_{2.5}$ LBW**	Bell et al 2007	LBW not restricted to TLBW
			Brauer et al 2008	TLBW
			Chen et al 2002	TLBW
			Gouveia et al 2004	TLBW
			Lee et al 2003	TLBW
			Lin et al 2004	TLBW
			Maisonet et al 2001	TLBW
			Parker and Woodruff 2008	TLBW
			Rogers and Dunlop 2006	LBW
			Salam et al 2005	TLBW
			Wilhelm & Ritz 2005	TLBW
		PM_{10}, LBW and WP	Bell et al 2007	LBW not restricted to TLBW
			Brauer et al 2008	TLBW
			Chen et al 2002	TLBW
			Gouveia et al 2004	TLBW

Please note: Papers for each pollutant/exposure have been combined for all exposure windows (WP, T1, T2, T3)				
Meta-analysis	**LBW definition**	**Pollutant/Outcome**	**Included studies**	**TLBW or LBW**
			Lee et al 2003)	TLBW
			Lin et al 2004	TLBW
			Maisonet et al 2001	TLBW
			Parker & Woodruff 2008	TLBW
			Rogers et al 2006	VLBW (< 1500g), all of whom were preterm infants
			Salam et al 2005	TLBW
			Wilhelm & Ritz 2005	TLBW
Stieb et al 2012	LBW (<2500 g)	**$PM_{2.5}$ LBW**	Bell et al 2007	LBW not restricted to TLBW
			Brauer et al 2008	LBW at term were those with at least 37 weeks' gestation and birth weight < 2,500 g.
			Madsen et al 2010	TLBW
			Morello-Frosch et al 2010	TLBW
			Parker & Woodruff 2008	TLBW
			Slama et al 2007	TLBW
		PM_{10} LBW	Bell et al 2007	LBW not restricted to TLBW
			Brauer et al 2008	TLBW
			Dugandzic et al 2006	TLBW
			Kim et al 2007	TLBW
			Lee et al 2003	TLBW
			Madsen et al 2010	TLBW
			Maisonet et al 2001	TLBW

Please note: Papers for each pollutant/exposure have been combined for all exposure windows (WP, T1, T2, T3)				
Meta-analysis	**LBW definition**	**Pollutant/Outcome**	**Included studies**	**TLBW or LBW**
			Morello-Frosch et al 2010	TLBW
			Salam et al 2005	TLBW
			Seo et al 2010	TLBW
			Wilhelm & Ritz 2005	TLBW
			Xu et al 2011	TLBW
		NO_2 and TLBW	Bell et al 2007	LBW not restricted to TLBW
			Brauer et al 2008	TLBW
			Jackson et al 2009	LBW at all gestational ages
			Lee et al 2003	TLBW
			Liu et al 2003	LBW – excluded births that were <22 weeks.
			Madsen et al 2010	TLBW
			Maroziene and Grazuleviciene 2002	TLBW
			Morello-Frosch et al 2010	TLBW
			Salam et al 2005	TLBW
			Slama et al 2007	TLBW
			Wilhelm & Ritz 2005	TLBW
		O_3, TLBW	Brauer et al 2008	TLBW
			Dugandzic et al 2006	TLBW
			Ha et al 2001	TLBW
			Liu et al 2003	LBW

Please note: Papers for each pollutant/exposure have been combined for all exposure windows (WP, T1, T2, T3)				
Meta-analysis	**LBW definition**	**Pollutant/Outcome**	**Included studies**	**TLBW or LBW**
			Morello-Frosch et al 2010	TLBW
			Salam et al 2005	TLBW
			Wilhelm & Ritz 2005	TLBW

Appendix 6. AMSTAR 2 scoring tool

The original AMSTAR 2 questions and guidance, adapted from Shea et al (2017) are presented in Table A 1.

Table A 1. The original AMSTAR 2 questions and guidance

Question	Guidance
1. Did the research questions and inclusion criteria for the review include the components of PICO?	It is common practice to use PICO description (population, intervention, control group and outcome) as an organising framework for a study question. Sometimes timeframe should be added if this is critical in determining the likelihood of a study capturing relevant clinical outcomes (for example an effect of the intervention is only expected after several years). PICO identifies the elements that should be described in detail in the report of the systematic review and should enable the appraiser to judge selection of studies, and their combinability, and enable the user of the review to determine applicability of the results. Authors of systematic reviews do not always make the elements of PICO explicit, but they should be discernible through a careful reading of the abstract, introduction and methods sections. To score 'Yes' appraisers should be confident that the 4 elements of PICO are described somewhere in the report.
2. Did the report of the review contain an explicit statement that the review methods were established prior to conduct of the review and did the report justify any significant deviations from the protocol?	Systematic reviews are a form of observational research and the methods for the review should be agreed on before the review commences. Adherence to a well-developed protocol reduces the risk of bias in the review. Authors should demonstrate that they worked with a written protocol with independent verification. This can take the form of registration (for example at PROSPERO - https://www.crd.york.ac.uk/PROSPERO/), an open publication journal (for example BMJ Open) or a dated submission to a research office or research ethics board. The research questions and the review study methods should have been planned ahead of conducting the review. At a minimum, this should be stated in the report (scores 'Partial Yes'). To score 'Yes' authors should demonstrate that they worked with a written protocol with independent verification (by a registry or another independent body, for example research ethics board or research office) before the review was undertaken. Appraisers should compare the published report of the review with the registered protocol, when the latter is available. If there are deviations from the protocol appraisers should determine whether these are reported and justified by review authors. Obvious unexplained discrepancies should result in downgrading of the rating.

Question	Guidance
3. Did the review authors explain their selection of the study designs for inclusion in the review?	The selection of study types for inclusion in systematic reviews should not be arbitrary. The authors should indicate that they followed a strategy. The general rule (this may have to be inferred from what the authors actually wrote) is that they asked first whether a review restricted to RCTs would have given an incomplete summary of the effects of a treatment. This might be because there were no relevant RCTs or because of missing outcomes in available RCTs (usually harms), inadequate statistical power, restrictive populations, or unrepresentative control/intervention treatments. If the answer to this general question is yes, the inclusion of nonrandomized studies of the intervention(s) is justified. Conversely, to justify restriction of the review to RCTs the authors should argue that they can provide a complete picture of the effects they are interested in. Restriction of a review to only NRSI is justified when RCTs cannot provide the necessary outcome data, or in the case where reviews of RCTs have been completed and the review of NRSI will complement what is already known. Inclusion of both RCTs and NRSI may be justified to get a complete picture of the effectiveness and harms associated with an intervention. In this situation, we recommend (see below) that these 2 types of studies are assessed and combined independently (if meta-analysis is appropriate). This is a somewhat neglected area and even with guidance it can be difficult to judge the extent to which a review meets the rating criteria. The justification for selection of study designs may have to be inferred from a careful reading of the complete study report.
4. Did the review authors use a comprehensive literature search strategy?	At least 2 bibliographic databases should be searched. The report should include years and databases examined (for example Central, EMBASE, and MEDLINE). Key words and/or MESH terms should be reported and the full search strategy available on request. Searches should be supplemented by checking published reviews, specialized registers, or experts in the particular field of study, and by reviewing the reference list from the studies found. Sometimes it is necessary to approach authors of original studies to clarify results or obtain updates or corrections. Publications in all relevant languages should be sought and a justification provided when there are language restrictions. We have highlighted the need for searching the grey literature in some cases. Grey literature is sometimes important with reports of policy and program evaluations that are only available from web sites (for example government, non-government, or health technology agencies). These may or may not have been subject to peer review and such appraisals should be looked for. Where the grey literature is considered important, authors should have searched appropriate sources, such as trial registries, conference abstracts, dissertations, and unpublished reports on personal websites (for example universities, ResearchGate). In addition, trials of medical interventions may not have been published in peer-reviewed journals but can be obtained directly from company sponsors or directly from investigators. To score 'Yes' appraisers should be satisfied that all relevant aspects of the search have been addressed by review authors.

Question	Guidance
5. Did the review authors perform study selection in duplicate?	Best practice requires 2 review authors to determine eligibility of studies for inclusion in systematic reviews. This involves checking the characteristics of a study (from title, abstract and full text) against the elements of the research question In the response options, we point to the desirability of review authors describing inter-rater agreement across a sample of studies being considered for inclusion in the review. A consensus process should have been used when disagreements arose in study selection. If one individual carried out selection of all studies, with a second reviewer checking agreement on a sample of studies, we recommend that a Kappa score indicating 'strong' agreement (0.80 or greater) should have been achieved. There should have been at least 2 independent assessors for study selection. A consensus process should have been used when disagreements arose in study selection. In the event that one individual carried out selection of studies a second reviewer should have checked agreement on a sample of representative studies, and they should have achieved a kappa score of 0.80 or greater.
6. Did the review authors perform data extraction in duplicate?	As in Item 5, there should have been at least 2 independent assessors performing data extraction. A consensus process should have been used when disagreements arose. In the event that one individual carried out data extraction a second reviewer should have checked agreement on a sample of studies, and they should have achieved a kappa score of 0.80 or greater.
7. Did the review authors provide a list of excluded studies and justify the exclusions?	This item requires review authors to provide a complete list of potentially relevant studies with justification for the exclusion of each. Non-inclusion of studies may be necessary for a range of reasons, based on inappropriate/ irrelevant populations, interventions, and controls. Exclusion should not be based on risk of bias, which is dealt with separately and later in the review process. Unjustified exclusion may bias the review findings and we encourage an inclusive approach in the early stages of a review. This item requires review authors to provide a complete list of potentially relevant studies with justification for the exclusion of each one.
8. Did the review authors describe the included studies in adequate detail?	The description of subjects, interventions, controls, outcomes, design, analysis, and settings of the studies should be provided. The detail should be sufficient for an appraiser, or user, to make judgments about the extent to which the studies were appropriately chosen (in relation to the PICO structure) and whether the study populations and interventions were relevant their own practice or policy. The descriptors also provide a framework for studying heterogeneity in intervention effects (for example, by dose, age range, clinical setting and so on)

9.	Did the review authors use a satisfactory technique for assessing the risk of bias (RoB) in individual studies that were included in the review?	This is a crucial part of the appraisal of any systematic review, particularly those that include non-randomised studies of interventions (NRSI). The key appraisal question is whether review authors have taken account of the risk of bias when summarising and interpreting the results. When the review is confined to randomized controlled trials (RCTs) we recommend that you consult the Cochrane Handbook to determine whether the review authors have made an adequate assessment of ROB with individual RCTs. This section is concerned with the challenge posed by RoB in non-randomised studies. Review authors should have used a systematic approach to ROB assessment, preferably with a properly developed rating instrument. If they have used a non-standard instrument you should be satisfied that it was capable of detecting serious methodological flaws. Several ROB instruments (for individual studies) are in common use, including the Newcastle Ottawa Scale, SIGN, and the Mixed Methods Appraisal Tool (MMAT). The most comprehensive assessment instrument is the recently introduced Cochrane instrument, ROBINS-I. It is appreciated that this instrument may not have been available at the time a review was performed. In developing AMSTAR 2 we drew on the Cochrane RoB instruments for RCTs: (http://handbook.cochrane.org/chapter_8/8_5_the_cochrane_collaborations_tool_for_assessing_r isk_of_bias.htm) and NRSI: www.riskofbias.info. In both cases the domain appraisal items are drawn from these instruments. Whatever instrument was used by the review authors, appraisers should be satisfied that it addresses the items listed in item 9 of the instrument. Please note that the guidance given here is not comprehensive – AMSTAR 2 addresses only the most commonly recognised domains of bias. A deeper assessment of risk of bias requires specialist input. In assessing how RoB has been assessed by review authors you should seek methods and content expert advice, if that is not included in your team. Advisors should be asked to provide specific advice on which confounders are important, how to identify selection and measurement biases that are likely, to be relevant to the review under consideration. In addition, you should seek guidance on what adjustment techniques for confounding would be appropriate. The following list of domains of bias has been selected from the ROBINS-I/ACROBAT-NRSI instrument as being the most relevant to systematic reviews that include NRSI: Confounding. Confounding occurs when the effects of 2 associated interventions or exposures (for example smoking and alcohol consumption) have not been separated during analysis. This can result in an effect being attributed to one variable when it is due to the other. In the study of interventions confounding may also be related to the indication for treatment, for instance when one drug is given preferentially to patients with higher rates of co-morbidities than the comparator drug and where these co-morbid conditions are associated with the outcome of interest. These assessments are typically quantified in the baseline data reported in the individual study. Potential baseline confounding can be addressed in several ways, including design (for example matching by propensity score), adjustment (for example logistic regression) and other techniques such as instrumental variable analysis and the inclusion of 'tracer' exposures. It is common to assume that even sophisticated techniques will not adjust completely for all confounders, meaning weak associations, even if statistically significant, should be interpreted cautiously. Sample selection bias. This occurs when subjects are sampled in a biased way that directly distorts the true relationship between exposure and outcome. It requires no third factor, as is the case with confounding. For instance, you should not study the association between smoking and heart disease by recruiting subjects referred to a smoking cessation clinic. The selection of subjects with the exposure of interest should be unrelated to their outcome. Likewise, selection of subjects with the outcome should be unrelated to their exposure status. The timing of selection can be important. If subjects have been using a drug for some time before enrolment (prevalent users) they will be a tolerant group with a lower risk of adverse outcomes. For this reason, contemporary pharmaco-epidemiological studies recruit 'new users' of medications (analogous to starting treatment in an RCT). Other temporal sampling biases

Question	Guidance
	(immortal time bias and inception bias) are sometimes important. It is recommended that users refer to ROBINS-I guidance document for more information. Bias in measurement of exposures and outcomes: measurement of an exposure or treatment may be misclassified if there is no accurate recording made in real time. Typically, modern pharmaco-epidemiological studies use prescriptions or dispensing records as a surrogate for consumption. But adherence to dispensed treatment will not be 100% so actual consumption will be mis-classified by this method. If this error is nondifferential it will be a bias to the null. However, in some fields of research investigators rely on recall (for example ultraviolet exposure and melanoma). This may lead to differential misclassification. For instance, parents of a child who has died of SIDS may have heightened recall of any medications they administered to the baby prior to the event. Measurement of outcomes can also be affected by misclassification and if this is nondifferential it will usually be a bias to the null. However, non-differential misclassification can introduce bias. For instance, if leg ultrasound is performed frequently in women with swollen painful legs who are taking an oral contraceptive, selection of individuals from an ultrasound clinic may bias studies of the association between DVT and oral contraceptives. Selective reporting of outcomes and analyses: large observational studies may analyse population databases that record many outcomes occurring in a defined population. If outcomes are not pre-specified (preferably in a registered protocol) investigators may be tempted to analyse multiple outcomes and selectively report those that appear to be different between exposed and non-exposed individuals. In addition, there are usually several potential methods for analysing a non-randomised data set (including, for example, different ways of categorising the intervention, or different multi-variable adjustment models). If the analytical protocol is not specified in advance of the study, it may be possible to select one set of analyses that appears to show a significant statistical difference that is not apparent in the other analyses. Reviewers should determine whether study authors pre-specified outcomes and analyses. This will become easier as more studies are registered before being conducted.
10. Did the review authors report on the sources of funding for the studies included in the review?	Several investigations have shown that commercially sponsored studies are more likely to have findings that favour a sponsor's product than independently funded studies. It is valuable for review authors to document the funding sources for each study included in the review or to record that the information was not provided in the study reports. Depending on this information it may be possible to analyse separately the results from commercially funded and independently funded studies.

Question	Guidance
11. If meta-analysis was justified did the review authors use appropriate methods for statistical combination of results?	(Only complete this item if meta-analysis of other data synthesis techniques were reported) Review authors should have stated explicitly in the review protocol the principles on which they based their decision to perform meta-analysis of data from the included studies. These include the desire to obtain a single pooled effect (for instance from a number of compatible but underpowered studies) and the extent to which the studies are compatible (in terms of population controls and interventions) and therefore capable of being combined. Where meta-analysis was considered appropriate authors should have explained their decisions to use fixed or random effects models in the case of RCTs and set out the methods they intended to use to investigate heterogeneity. With NRSI study populations vary greatly in size from small cohorts (of tens or hundreds of participants) to studies of hundreds of thousands of individuals and thousands of events. If these results, are going to be combined with those from smaller RCTs the pooled estimates of effect will be dominated by the data from the non-randomised studies. In addition, the results from NRSI may be affected by a range of biases (see above), meaning that the overall pooled estimates may be precise but biased. Review authors should report pooled estimates separately for the different study types. In the case of NRSI, pooling may result in a very precise and 'statistically significant', but biased, estimate of effect. However, the confidence interval is calculated on the assumption that there is no bias (that is, the estimates are as accurate as if obtained from a high quality RCT with the same number of participants). It is rare for a NRSI to have as low risk of bias as a high quality RCT of the same research question and confidence intervals for NRSI (and pooled estimates based on NRSI) should be viewed with caution. This issue is important when considering the varying risk of bias, and uncertainty about the risk of bias across NRSI. Heterogeneity is an important issue in any meta-analysis. It is particularly important in a review of NRSI because of the more diverse methods that are likely to have been used across different studies. In addition to the usual sources of heterogeneity (different comparators, variations in baseline risk of outcomes or other characteristics of the study population, differing interventions (for example dose effects, context/setting, practitioner experience) and different definitions of outcomes), it is important to consider heterogeneity in source of participants, completeness of data, methods of data management and analysis. Statistical adjustment of intervention effects for confounders may result in estimates that are quite different from the unadjusted estimate derived from the raw data. Generally, when combining the results of NRSI review authors should pool the fully adjusted estimates of effect, not the raw data. If they do the latter, there should be a clear justification. However, different studies are very likely to report treatment effects that have been adjusted for different sets of covariates (or covariates measured or fitted in different ways); this diversity represents another source of potential heterogeneity.

Question	Guidance
12. If meta-analysis was performed did the review authors assess the potential impact of RoB in individual studies on the results of the meta-analysis or other evidence synthesis?	In cases where review authors have chosen to include only high quality RCTs there may be little discussion of the potential impact of bias on the results. But where they have included RCTs of variable quality they should assess the impact of this by regression analysis, or by estimating pooled effect sizes with only studies at low ROB. In the case of NRSI they should estimate pooled effect sizes while including only studies at low or moderate risk of bias, and/or only those at low ROB (if there are any). If meta-analyses (or other data synthesis techniques such as regression analysis) were not performed the authors should still provide some commentary on the likely impact of ROB on individual study results.
13. Did the review authors account for RoB in individual studies when interpreting/ discussing the results of the review?	Even if meta-analyses were not conducted review authors should include discussion of the impact of ROB in the interpretation of the results of the review. This is always important, but especially when reviews include RCTs with variable ROB, and with any review that includes NRSI. This discussion should not be limited to the impact of ROB on the pooled estimates (see above) but should also consider whether it may account for differences between the results of individual studies. The authors should make an explicit consideration of ROB if they make any recommendations that are likely to have an impact on clinical care or policy.
14. Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review?	There are many potential causes of heterogeneity in the results of NRSI than in RCTs. Many factors considered in this instrument, including different study designs, different methods of analysis, different populations, and differing intensities of the intervention(s) – dosages in the case of drugs. Both the PICO elements and the domains of bias listed in Item 9 should also be considered as important potential sources of heterogeneity in the results. Review authors should explore these possibilities and discuss the impact of heterogeneity on the results conclusions and any recommendations.

Question	Guidance
15. If they performed quantitative synthesis did the review authors carry out an adequate investigation of publication bias (small study bias) and discuss its likely impact on the results of the review?	This is a very important issue but can be difficult for review authors and appraisers to resolve completely. Typically, statistical tests or graphical displays are used and if they are positive then it indicates the presence of PB. However, negative tests are not a guarantee of the absence of PB as the tests are insensitive. To some extent the importance of PB depends on context and setting. For instance, a series of apparently methodologically sound industry-sponsored studies (for example drugs, devices, putative toxins) might be more likely to be affected by PB than similar studies conducted independently of industry. The key issues are whether the authors have done their best to identify PB through deeper and intensive literature searches (as needed and according to the setting), shown an awareness of the likely impact of PB in their interpretation and discussion of the results and performed a sensitivity analyses to determine how many missing 'null' studies would be needed to invalidate the results they obtained.
16. Did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review?	As noted above (under ROB), individual studies funded by vested interests may generate results that are more likely to favour the intervention than do independent studies. The same assumption applies to systematic reviews and authors should report their direct funding sources. Journals generally will require this. But assessment of the reviewers' conflicts of interest doesn't stop there. They should report their other ties. The review may be independently funded, but the authors have ties to companies that manufacture products included in the systematic review. Professional conflicts of interest are powerful, but harder to discern as they are seldom reported. When investigators have a career-long investment in a field of research, a review that conflicts with their long-held beliefs can be confronting. Potential conflicts of interest of this type will be hard to assess but may be inferred from the fact that the reviewers have published extensively in the field being reviewed and their studies are included in the systematic review. While it can be argued that the effects of competing interests might manifest as flaws in the other domains of bias, we believe that this item should always be rated separately.

Appendix 7. AMSTAR 2 modifications

The modified guidance and scoring for the affected questions are provided in Table A 2. If a question does not appear Table A 2, the original scoring and guidance as described in Table A 1 were used. Please note that meta-analyses were quality assessed 'blind', that is, without considering the effect estimates reported in each meta-analysis, so as not to bias the score that each meta-analysis was assigned.

Table A 2. Modified guidance and scoring for selected AMSTAR 2 questions

	Question	Revised guidance/score
1.	Did the research questions and inclusion criteria for the review include the components of PICO?	To score Yes, the study must clearly specify the exposures and outcomes that are the focus of the review. Population – not relevant Intervention – relevant, as this equates to pollutant exposure Control group – not relevant Outcomes - relevant
2.	Did the report of the review contain an explicit statement that the review methods were established prior to conduct of the review and did the report justify any significant deviations from the protocol?	To score Yes, the study must have registered the study protocol prior to conduct of review.
3.	Did the review authors explain their selection of the study designs for inclusion in the review?	To score Yes, the paper must demonstrate consideration of study designs in the process of selecting studies for the review.
8.	Did the review authors describe the included studies in adequate detail?	The paper would score Yes, if it describes individually for each included study: the subjects, location/setting, study time period, exposures (pollutants/exposure period/exposure assessment/exposure levels), outcomes, design, analysis/covariates. Typically provided in a table format.
9.a. 9.b.	Did the review authors use a satisfactory technique for assessing the risk of bias (RoB) in individual studies that were included in the review?	Answer separately for the systematic review element and the meta-analysis element To score Yes, the study authors should have made a reasonable attempt to assess risk of bias in individual studies included in each element. Use of a grading scale (for example Newcastle Ottawa Scale) is not essential.

	Question	Revised guidance/score
12.	If meta-analysis was performed did the review authors assess the potential impact of RoB in individual studies on the results of the meta-analysis or other evidence synthesis?	Possible answers: Yes/No/Partial To score Yes, the authors should have conducted sensitivity analyses and presented forest plots or stratified summary estimates in text for, for example: • analysis stratified according to Risk of Bias (that is, high-quality vs low quality) AND/OR • analysis excluding low quality studies one by one Score as Partial if the authors have conducted sensitivity analyses as above and comment upon the results, but do not present forest plots or quantitative results in text.
13.	Did the review authors account for RoB in individual studies when interpreting/ discussing the results of the review?	Possible answers: Yes/No/Partial Use original AMSTAR 2 guidance, but score as follows: • To score Yes, the discussion should not be limited to the impact of ROB on the pooled estimates (see 12) but should also consider whether it may account for differences between the results of individual studies. • Score as Partial - if discussion is limited to the impact of ROB on the pooled estimates.
14.	Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review?	Possible answers: Yes/No/Partial To score Yes, the authors must drill down into the detail of their meta-analysis results, and attempt to explain and discuss heterogeneity in specific results/forest plots. Score as Partial if the authors have attempted to provide a thoughtful and detailed explanation for heterogeneity in their results generally (for example discussing differences between specific individual included studies), but without reference to specific results in their meta-analysis. Score as No if the authors do less than above, for example discussion is limited to general reasons for heterogeneity without any attempt to apply this to individual included studies. That is not satisfactory.
16.	Did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review?	Possible answers: 1. Yes, reported and no conflict of interest 2. Yes, reported and there is a potential conflict of interest 3. Not reported If scoring as 2, also state if the potential conflict is of concern.

	Question	Revised guidance/score
17	Have the authors addressed potential confounding by smoking in the meta-analysis, and if yes how?	Additional question, not part of AMSTAR 2. Possible answers: Yes or No Provide detail of what the authors did (for example stratified analysis by adjustment for smoking, or excluding studies which did not adjust for smoking and so on)

Appendix 8. Modified AMSTAR 2 weighting

The weightings applied to each of the modified AMSTAR 2 questions are summarised in Table A 3. Please note that meta-analyses were quality assessed 'blind', that is, without considering the effect estimates reported in each meta-analysis, as not to bias the score that each meta-analysis was assigned.

Table A 3. Weighting applied to the modified AMSTAR 2 questions

	Question	Weighting
1.	Did the research questions and inclusion criteria for the review include the components of PICO?	2
2.	Did the report of the review contain an explicit statement that the review methods were established prior to conduct of the review and did the report justify any significant deviations from the protocol?	1
3.	Did the review authors explain their selection of the study designs for inclusion in the review?	1
4.	Did the review authors use a comprehensive literature search strategy?	3
5.	Did the review authors perform study selection in duplicate?	1
6.	Did the review authors perform data extraction in duplicate?	1
7.	Did the review authors provide a list of excluded studies and justify the exclusions?	2
8.	Did the review authors describe the included studies in adequate detail?	2
9.a.	Did the review authors use a satisfactory technique for assessing the risk of bias (RoB) in individual studies that were included in the review? (Systematic review element)	3

	Question	Weighting
9.b.	Did the review authors use a satisfactory technique for assessing the risk of bias (RoB) in individual studies that were included in the review? (Meta-analysis element)	3
10.	Did the review authors report on the sources of funding for the studies included in the review?	1
11.	If meta-analysis was justified did the review authors use appropriate methods for statistical combination of results?	3
12.	If meta-analysis was performed did the review authors assess the potential impact of RoB in individual studies on the results of the meta-analysis or other evidence synthesis?	3
13.	Did the review authors account for RoB in individual studies when interpreting/ discussing the results of the review?	2
14.	Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review?	3
15.	If they performed quantitative synthesis did the review authors carry out an adequate investigation of publication bias (small study bias) and discuss its likely impact on the results of the review?	2
16.	Did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review?	1
17.	Have the authors addressed potential confounding by smoking in the meta-analysis, and if yes how?	3

Appendix 9. Population overlaps in the meta-analyses

1. Preterm birth (PTB)

a. $PM_{2.5}$, PTB and WP

Table A 4. Population overlap details meta-analyses concerning $PM_{2.5}$, preterm birth, and whole pregnancy

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [12]	Is the population overlap 'tolerable' or 'intolerable'?
Klepac et al 2018	Brauer et al 2008	4.3	**1. Chang et al 2012 and Chang et al 2013** Location: both North Carolina, USA Study period: both 2001-2005 **2. Brauer et al 2008 and Stieb et al 2016b** Location: Brauer et al 2008: Vancouver, Canada, Stieb et al 2016b: Canada Study period: Brauer et al 2008: 1999-2002, Stieb et al 2016b: 1999-2008 **3. Lavigne et al 2016 and Stieb et al 2016b** Location: Lavigne et al 2016: Ontario, Canada, Stieb et al 2016b: Canada Study period: Lavigne et al 2016: 2005-2012, Stieb et al 2016b: 1999-2008	1. Intolerable overlap. Chang et al 2013 is a subset of Chang et al 2012. 2. Intolerable overlap. All of the subjects in Brauer et al 2008 are included in Stieb et al 2016b. 3. Tolerable overlap. Some crossover, but Lavigne et al 2016 adds an additional 4 years of data for Ontario, and Stieb et al 2016b adds an additional 6 years of data for the whole of Canada.
	Chang et al 2012	5.6		
	Chang et al 2013)	6.1		
	Fleischer et al 2014)	9.0		
	Ha et al 2014	9.0		
	Hao et al 2016	8.9		
	Kloog et al 2012	9.0		
	Lavigne et al 2016	8.9		

[12] Population overlaps were assessed 'blind', that is, without considering the effect estimates reported in each meta-analysis

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [12]	Is the population overlap 'tolerable' or 'intolerable'?
	Padula et al 2014	9.0		
	Pereira et al 2014a	3.8		
	Stieb et al 2016b	9.1		
	Wu et al 2011[13]	8.7		
	Wu et al 2011[13]	8.7		
Lamichhane et al 2015	Gehring et al 2011	0.18	1. **Pereira et al 2014a and Hyder et al 2014** Location: Pereira et al 2014a: Connecticut, USA Hyder et al 2014: Connecticut and Massachusetts, USA Study Period: Pereira et al 2014a: 2000-2006, Hyder et al 2014: 2000-2006 2. **Hyder et al 2014 and Kloog et al 2012** Location: Hyder et al 2014: Connecticut and Massachusetts, USA, Kloog et al 2012: Massachusetts. USA Study Period: Hyder et al 2014: 2000-2006, Kloog et al 2012: 2000-2008	1. Intolerable overlap. All of the subjects in Pereira et al 2014a are included in Hyder et al 2014 2. Intolerable overlap. Kloog et al 2012 is almost a complete subset of Hyder et al 2014, adding only 2 additional years of data for Massachusetts. 3. Tolerable overlap. Some crossover, but Huynh et al 2006 includes participants from the entire state of California, and Wu et al 2009 adds an additional 8 years of data for Los Angeles and Orange counties.
	Ha et al 2014	6.72		
	Huynh et al 2006	7.19		
	Hyder et al 2014	6.96		
	Kloog et al 2012	6.85		
	Pereira et al 2014a	0.50		

[13] Klepac et al 2018 included two separate estimates for Wu et al 2011

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [12]	Is the population overlap 'tolerable' or 'intolerable'?
	Wu et al 2009	4.44	3. **Huynh et al 2006 and Wu et al 2009** Location: Huynh et al 2006: California, USA Wu et al 2009: Los Angeles and Orange Counties, California, USA Study Period: Huynh et al 2006: 1999-2000, Wu et al 2009: 1997-2006	
	Gehring et al 2011	(out of 32.84 not 100.00)		
Li et al 2017 (per 10 µg/m³)	Fleischer et al 2014	16.8	Did not contain population overlaps.	N/A
	Hannam et al 2014	12.6		
	Huynh et al 2006	18.		
	Kloog et al 2012	17.4		
	Ritz et al 2007	16.9		
	Stieb et al 2016b	17.9		
Li et al 2017 (per IQR)	Gehring et al 2011	0.2	1. **Pereira et al 2014a and Hyder et al 2014** Location: both Connecticut, USA Study Period: both 2000-2006 2. **Pereira et al 2014a and Pereira et al 2014c** Location: both Connecticut, USA Study Period: 2000-2006 3. **Hyder et al 2014 and Pereira et al 2014c** Location: both Connecticut, USA Study Period: both 2000-2006	1. Intolerable overlap. Full geographical and time period overlap 2. Intolerable overlap. Full geographical and time period overlap 3. Intolerable overlap. Full geographical and time period overlap
	Gray et al 2014	19.4		
	Ha et al 2014	23.6		
	Hyder et al 2014	10.2		
	Lavigne et al 2016	21.9		
	Pereira et al 2014a	1.7		

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [12]	Is the population overlap 'tolerable' or 'intolerable'?
	Pereira et al 2014c	3.5		
	Wu et al 2009	19.6		
Liu et al 2017	Brauer et al 2008	2.0	Did not contain population overlaps	N/A
	Gehring et al 2011	1.3		
	Ha et al 2014	15.4		
	Hannam et al 2014	4.1		
	Pereira et al 2014a	2.3		
	Pereira et al 2014b	1.7		
	Qian et al 2016	18.0		
		(out of 44.9, not 100.0)		
Sapkota et al 2012	Brauer et al 2008	Data not provided in the study.	1. **Huynh et al 2006 and Wilhelm & Ritz 2005**	1. Tolerable overlap. Some crossover, but Huynh et al 2006 includes participants from the entire state of California, and Wilhelm & Ritz 2005 adds an additional 5 years of data for Los Angeles county.
	Huynh et al 2006			

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [12]	Is the population overlap 'tolerable' or 'intolerable'?
	Jalaludin et al 2007		Location: Huynh et al 2006: California, USA: Wilhelm and Ritz 2005: Los Angeles County, California, USA Study Period: Huynh et al 2006: 1999-2000, Wilhelm and Ritz 2005: 1994-2000 2. **Huynh et al 2006 and Wu et al 2009** Location: Huynh et al 2006: California, USA, Wu et al 2009: Los Angeles and Orange counties, California, USA Study Period: Huynh et al 2006: 1999-2000, Wu et al 2009: 1997-2006 3. **Ritz et al 2007 and Wu et al 2009** Location: Ritz et al 2007: Los Angeles county, California, USA Wu et al 2009: Los Angeles and Orange counties, California, USA Study Period: Ritz et al 2007: 2003, Wu et al 2009: 1997-2006 4. **Wilhelm & Ritz 2005 and Wu et al 2009** Location: Wilhelm and Ritz 2005: Los Angeles county, California, USA, Wu et al 2009: Los Angeles and Orange counties, California, USA Study Period: Wilhelm and Ritz 2005: 1994-2000, Wu et al 2009: 1997-2006	2. Tolerable overlap. Some crossover, but Huynh et al 2006 includes participants from the entire state of California, and Wu et al 2009 adds an additional 8 years of data for Los Angeles and Orange counties. 3. Intolerable overlap. All of the subjects in Ritz et al 2007 are included in Wu et al 2009. 4. Tolerable overlap. Some crossover, but Wilhelm & Ritz 2005 add an additional 3 years for Los Angeles county, and Wu et al 2009 adds an additional 6 years of data for Los Angeles and Orange counties.
	Ritz et al 2007			
	Wilhelm & Ritz 2005			
	Wu et al 2009			
Stieb et al 2012	Brauer et al 2008	1.25	1. **Huynh et al 2006 and Wu et al 2009** Location: Huynh et al 2006: California, USA Wu et al 2009: Los Angeles and Orange counties, California, USA Study Period: Huynh et al 2006: 1999-2000, Wu et al 2009: 1997-2006	1. Tolerable overlap. Some crossover, but Huynh et al 2006 includes participants from the entire state of California, and Wu et al 2009 adds an additional 8 years of data for Los Angeles and Orange counties.
	Gehring et al 2011	2.11		
	Huynh et al 2006	79.8		

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [12]	Is the population overlap 'tolerable' or 'intolerable'?
	Wu et al 2009	16.85		
Sun et al 2015	Brauer et al 2008 Chang et al 2015 Fleischer et al 2014 Gehring et al 2011 Gray et al 2014 Ha et al 2014 Hannam et al 2014 Huynh et al 2006 Hyder et al 2014 Kloog et al 2012 Pereira et al 2014a Pereira et al 2014c Wu et al 2009	Data not provided in the study.	1. **Pereira et al 2014a and Hyder et al 2014** Location: both Connecticut, USA Study Period: both 2000-2006 2. **Pereira et al 2014c and Hyder et al 2014** Location: Pereira et al 2014c: Connecticut, USA, Hyder et al 2014: Connecticut and Massachusetts, USA Study Period: Pereira et al 2014c: 2000-2006, Hyder et al 2014: 2000-2006 3. **Hyder et al 2014 and Kloog et al 2012** Location: Hyder et al 2014: Connecticut and Massachusetts, USA, Kloog et al 2012: Massachusetts, USA Study Period: Hyder et al 2014: 2000-2006, Kloog et al 2012: 2000-2008 4. **Huynh et al 2006 and Wu et al 2009** Location: Huynh et al 2006: California, USA, Wu et al 2009: Los Angeles and Orange counties, California, USA Study Period: Huynh et al 2006: 1999-2000, Wu et al 2009: 1997-2006	1. Intolerable overlap. Full geographical and time period overlap 2. Intolerable overlap. Full time period overlap, and Hyder et al 2014 only adds subjects from Massachusetts 3. Intolerable overlap. Kloog et al 2012 is almost a complete subset of Hyder et al 2014, adding only 2 additional years of data for Massachusetts. 4. Tolerable overlap. Some crossover, but Huynh et al 2006 includes participants from the entire state of California, and Wu et al 2009 adds an additional 8 years of data for Los Angeles and Orange counties.
Zhu et al 2015	Brauer et al 2008	0.70		

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [12]	Is the population overlap 'tolerable' or 'intolerable'?
	Gehring et al 2011	1.48	1. **Huynh et al 2006 and Wu et al 2009** Location: Huynh et al 2006: California, USA, Wu et al 2009: Los Angeles and Orange counties, California, USA Study Period: Huynh et al 2006: 1999-2000, Wu et al 2009: 1997-2006	1. Tolerable overlap. Some crossover, but Huynh et al 2006 includes participants from the entire state of California, and Wu et al 2009 adds an additional 8 years of data for Los Angeles and Orange counties.
	Gray et al 2014	27.38		
	Hannam et al 2014	4.78		
	Huynh et al 2006	24.07		
	Kloog et al 2012	8.75		
	Pereira et al 2014a	1.84		
	Wu et al 2009	11.00		

b. $PM_{2.5}$, PTB and trimester 1

Table A 5. Population overlap details meta-analyses concerning $PM_{2.5}$, preterm birth and trimester 1

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [14]	Is the population overlap 'tolerable' or 'intolerable'?
Klepac et al 2018	Chang et al 2012	8.4		
	Ha et al 2014	11.2		
	Hao et al 2016	11.4		

[14] Population overlaps were assessed 'blind', that is, without considering the effect estimates reported in each meta-analysis

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [14]	Is the population overlap 'tolerable' or 'intolerable'?
	Jalaludin et al 2007	4.6	1. **Wilhelm & Ritz 2005 and Wu et al 2011** Location: Wilhelm & Ritz 2005: Los Angeles County, California, USA, Wu et al 2011: Los Angeles County & Orange County, California, USA (but each county was included separately in the meta-analysis, and it is impossible to discern which analysis relates to which county) Study Period: Wilhelm & Ritz 2005: 1994-2000, Wu et al 2011: 1997-2006 2. **Ritz et al 2007 and Wu et al 2011** Location: Ritz et al 2007: 111 Los Angeles County zip codes, California, USA, Wu et al 2011: Los Angeles County & Orange County, California, USA (but each county was included separately in the meta-analysis, and it is impossible to discern which analysis relates to which county) Study Period: Ritz et al 2007: 2003, Wu et al 2011: 1997-2006	1. Tolerable overlap. Wilhelm & Ritz 2005 add an additional 3 years of data for Los Angeles County, and Wu et al 2011 adds an additional 6 years of data for Los Angeles and Orange counties. 2. Intolerable overlap. Ritz et al 2007 is a complete subset of Wu et al 2011
	Johnson et al 2016	11.0		
	Lee et al 2013	6.3		
	Padula et al 2014	11.2		
	Pereira et al 2014a	4.9		
	Ritz et al 2007	4.1		
	Wilhelm & Ritz 2005	10.2		
	Wu et al 2011[15]	8.5		
	Wu et al 2011[15]	8.2		
Lamichhane et al 2015	Gehring et al 2011	2.26	1. **Pereira et al 2014a and Hyder et al 2014** Location: Pereira et al 2014a: Connecticut, USA, Hyder et al 2014: Connecticut and Massachusetts, USA Study Period: both 2000-2006	1. Intolerable overlap. Full time period overlap, and Hyder et al 2014 only adds subjects from Massachusetts
	Wilhelm & Ritz 2005	4.73		
	Pereira et al 2014a	1.83		
	Ha et al 2014	6.96		
	Hyder et al 2014	6.96		

[15] Klepac et al 2018 included two separate estimates for Wu et al 2011

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [14]	Is the population overlap 'tolerable' or 'intolerable'?
		Out of 22.75 (not 100)		
Li et al 2017 (per IQR)	Hyder et al 2014	37.0	1. **Pereira et al 2014a and Hyder et al 2014** Location: both Connecticut, USA Study Period: both 2000-2006 2. **Pereira et al 2014a and Pereira et al 2014c** Location: both Connecticut, USA Study Period: both 2000-2006 3. **Hyder et al 2014 and Pereira et al 2014c** Location: both Connecticut, USA Study Period: both 2000-2006	1. Intolerable overlap. Full geographical and time period overlap 2. Intolerable overlap. Full geographical and time period overlap 3. Intolerable overlap. Full time period overlap, and Hyder et al 2014 only adds subjects from Massachusetts
	Pereira et al 2014a	12.6		
	Pereira et al 2014c	17.6		
	Ha et al 2014	31.7		
	Gehring 2011	1.1		
	Hyder et al 2014	37.0		
Liu et al 2017	DeFranco et al 2015	6.7	Did not contain population overlaps.	N/A
	Gehring et al 2011	3.1		
	Ha et al 2014	17.2		
	Hannam et al 2014	5.9		
	Lee et al 2013	4.6		
	Pereira et al 2014a	3.9		
	Pereira et al 2014b	2.4		

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [14]	Is the population overlap 'tolerable' or 'intolerable'?
	Pereira et al 2016)	6.3		
	Ritz et al 2007	4.9		
		(out of 55.1, not 100.0)		
Sapkota et al 2012	(Not named (in the paper) but Inferred from Table 2)	Data not provided in the study.	1. **Huynh et al 2006 and Wilhelm & Ritz 2005** Location: Huynh et al 2006: California, USA: Wilhelm and Ritz 2005: Los Angeles County, California, USA Study Period: Huynh et al 2006: 1999-2000, Wilhelm and Ritz 2005: 1994-2000	1. Tolerable overlap. Some crossover, but Huynh et al 2006 includes participants from the entire state of California, and Wilhelm & Ritz 2005 adds an additional 5 years of data for Los Angeles county.
	Huynh et al 2006			
	Jalaludin et al 2007			
	Ritz et al 2007			
	Wilhelm & Ritz 2005			
Stieb et al 2012	Gehring et al 2011	22.36	1. **Huynh et al 2006 and Wilhelm & Ritz 2005** Location: both California, USA Study Period: Huynh et al 2006: 1999-2000, Wilhelm & Ritz 2005: 1994-2000	
	Huynh et al 2006	27.95		
	Jalaludin et al 2007	23.38		

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [14]	Is the population overlap 'tolerable' or 'intolerable'?
	Wilhelm & Ritz 2005	26.31		1. Tolerable overlap. Some crossover, but Huynh et al 2006 includes participants from the entire state of California, and Wilhelm & Ritz 2005 adds an additional 5 years of data for Los Angeles county
Sun et al 2015	Chang et al 2015 Gehring et al 2011 Ha et al 2014 Hyder et al 2014 Jalaludin et al 2007 Lee et al 2013 Pereira et al 2014a Pereira et al 2014c Ritz et al 2007 Wilhelm & Ritz 2005	Data not provided in the study.	1. **Pereira et al 2014a and Hyder et al 2014** Location: both Connecticut, USA Study Period: both 2000-2006 2. **Hyder et al 2014 and Pereira et al 2014c** Location: Hyder et al 2014: Connecticut and Massachusetts, USA, Pereira et al 2014c: Connecticut, USA Study Period: both 2000-2006	1. Intolerable overlap. Full geographical and time period overlap 2. Intolerable overlap. Full time period overlap, and Hyder et al 2014 only adds subjects from Massachusetts.
Zhu et al 2015	Gehring et al 2011	14.17	Did not contain population overlaps.	N/A

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [14]	Is the population overlap 'tolerable' or 'intolerable'?
	Hannam et al 2014	16.58		
	Jalaludin et al 2007	15.22		
	Lee et al 2013	17.09		
	Pereira et al 2014a	17.44		
	Wilhelm & Ritz 2005	19.50		

c. $PM_{2.5}$, PTB and trimester 2

Table A 6. Population overlap details meta-analyses concerning $PM_{2.5}$, preterm birth and trimester 2

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [16]	Is the population overlap 'tolerable' or 'intolerable'?
Klepac et al 2018	Chang et al 2012)	9.8	Did not contain population overlaps.	N/A
	Ha et al 2014	12.1		
	Hao et al 2016	12.2		
	Johnson et al 2016	12.0		
	Lavigne et al 2016	12.1		
	Padula et al 2014	12.1		
	Pereira et al 2014a	8.3		
	Wu et al 2011[17]	10.6		

[16] Population overlaps were assessed 'blind', that is, without considering the effect estimates reported in each meta-analysis

[17] Klepac et al 2018 included two separate estimates for Wu et al 2011

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [16]	Is the population overlap 'tolerable' or 'intolerable'?
	Wu et al 2011[17]	10.8		
Lamichhane et al 2015	Pereira et al 2014a	3.52	1. **Pereira et al 2014a and Hyder et al 2014** Location: Pereira et al 2014a: Connecticut, USA, Hyder et al 2014: Connecticut and Massachusetts, USA Study Period: both 2000-2006	1. Intolerable overlap. Full time period overlap, and Hyder et al 2014 only adds subjects from Massachusetts
	Ha et al 2014	6.83		
	Hyder et al 2014	6.95		
		Out of 17.29 (not 100)		
Li et al 2017 (per IQR)	Hyder et al 2014	26.3	1. **Pereira et al 2014a and Hyder et al 2014** Location: both Connecticut, USA Study Period: both 2000-2006	1. Intolerable overlap. Full geographical and time period overlap.
	Pereira et al 2014a	23.0	2. **Pereira et al 2014a and Pereira et al 2014c** Location: both Connecticut, USA Study Period: both 2000-2006	2. Intolerable overlap. Full geographical and time period overlap.
	Pereira et al 2014c	24.3	3. **Hyder et al 2014 and Pereira et al 2014c** Location: both Connecticut, USA Study Period: both 2000-2006	3. Intolerable overlap. Full time period overlap, and Hyder et al 2014 only adds subjects from Massachusetts.
	Ha et al 2014	26.5		
Sapkota et al 2012	Information was not provided in article and it was not possible to infer from other tables – we wrote to the authors, but they did not provide any additional information.			
Stieb et al 2012	Wilhelm and Ritz 2005	100	Did not contain population overlaps.	N/A
Sun et al 2015*	Chang et al 2015			

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [16]	Is the population overlap 'tolerable' or 'intolerable'?
	Ha et al 2014	Data not provided in the study.	1. **Hyder et al 2014 and Pereira et al 2014a** Location: Hyder et al 2014: Connecticut and Massachusetts, USA, Pereira et al 2014a: Connecticut, USA Study Period: both 2000-2006	1. Intolerable overlap. Full geographical and time period overlap.
	Hyder et al 2014			
	Pereira et al 2014a			
	Pereira et al 2014c			
Zhu et al 2015	Hannam et al 2014	36.24	Did not contain population overlaps.	N/A
	Jalaludin et al 2007	18.42		
	Pereira et al 2014a	45.34		

d. $PM_{2.5}$, PTB and trimester 3

Table A 7. Population overlap details meta-analyses concerning $PM_{2.5}$, preterm birth and trimester 3

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [18]	Is the population overlap 'tolerable' or 'intolerable'?
Klepac et al 2018	Chang et al 2012	6	Did not contain population overlaps.	N/A
	Ha et al 2014	21		
	Hao et al 2016	27.8		
	Lavigne et al 2016	27.2		
	Padula et al 2014	15.9		
	Pereira et al 2014a	2.1		
Lamichhane et al 2015	Wilhelm & Ritz 2005	4.14	1. **Pereira et al 2014a and Hyder et al 2014** Location: Pereira et al 2014a: Connecticut, USA, Hyder et al 2014: Connecticut and Massachusetts, USA Study Period: both 2000-2006 2. **Hyder et al 2014 and Kloog et al 2012** Location: Hyder et al 2014: Connecticut and Massachusetts, USA, Kloog et al 2012: Massachusetts, USA Study Period: Hyder et al 2014: 2000-2006, Kloog et al 2014: 2000-2008	1. Intolerable overlap. Full time period overlap, and Hyder et al 2014 only adds subjects from Massachusetts. 2. Intolerable overlap. Kloog et al 2012 is almost a complete subset of Hyder et al 2014, adding only 2 additional years of data for Massachusetts.
	Pereira et al 2014a	2.72		
	Ha et al 2014	6.45		
	Hyder et al 2014	6.96		
	Kloog et al 2012	6.85		
		Out of 27.13 (not 100)		

[18] Population overlaps were assessed 'blind', that is, without considering the effect estimates reported in each meta-analysis

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [18]	Is the population overlap 'tolerable' or 'intolerable'?
Li et al 2017 (per IQR)	Hyder et al 2014	34.7	1. **Pereira et al 2014a and Hyder et al 2014** Location: both Connecticut, USA Study Period: both 2000-2006	1. Intolerable overlap. Full geographical and time period overlap.
	Pereira et al 2014a	11.0	2. **Pereira et al 2014a and Pereira et al 2014c** Location: both Connecticut, USA Study Period: 2000-2006	2. Intolerable overlap. Full geographical and time period overlap.
	Pereira et al 2014c	13.1	3. **Hyder et al 2014 and Pereira et al 2014c** Location: both Connecticut, USA Study Period: both 2000-2006	3. Intolerable overlap. Full time period overlap, and Hyder et al 2014 only adds subjects from Massachusetts.
	Ha et al 2014	41.2		
Sapkota et al 2012	Information was not provided in article and it was not possible to infer from other tables – we wrote to the authors, but they did not provide any additional information.			
Stieb et al 2012	Gehring et al 2011	3.01	1. **Huynh et al 2006 and Wilhelm & Ritz 2005** Location: Huynh et al 2006: California, USA: Wilhelm and Ritz 2005: Los Angeles County, California, USA Study Period: Huynh et al 2006: 1999-2000, Wilhelm & Ritz 2005: 1994-2000	1. Tolerable overlap. Some crossover, but Huynh et al 2006 includes participants from the entire state of California, and Wilhelm & Ritz 2005 adds an additional 5 years of data for Los Angeles county.
	Huynh et al 2006	66.31		
	Jalaludin et al 2007	5.23		
	Wilhelm & Ritz 2005	25.46		
Sun et al 2015	Chang et al 2015	Data not provided in the study.		1. Intolerable overlap. Full geographical and time period overlap
	Ha et al 2014			
	Hyder et al 2014			

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [18]	Is the population overlap 'tolerable' or 'intolerable'?
	Jalaludin et al 2007		1. **Hyder et al 2014 and Pereira et al 2014a** Location: Hyder et al 2014: Connecticut and Massachusetts, USA, Pereira et al 2014a: Connecticut, USA Study Period: both 2000-2006 2. **Hyder et al 2014 and Kloog et al 2012** Location: Hyder et al 2014: Connecticut and Massachusetts, USA, Kloog et al 2012: Massachusetts, USA Study Period: Hyder et al 2014: 2000-2006, Kloog et al 2012: 2000-2008	2. Intolerable overlap. Kloog et al 2012 is almost a complete subset of Hyder et al 2014, adding only 2 additional years of data for Massachusetts.
	Kloog et al 2012			
	Pereira et al 2014a			
	Pereira et al 2014c			
	Rudra et al 2011			
	Wilhelm & Ritz 2005			
Zhu et al 2015	Rudra et al 2011	0.41	Did not contain population overlaps.	N/A
	Gehring et al 2011	4.18		
	Hannam et al 2014	10.79		
	Jalaludin et al 2007	11.07		
	Kloog et al 2012	49.19		
	Pereira et al 2014a	24.36		

e. PM_{10}, PTB and WP

Table A 8. Population overlap details meta-analyses concerning PM_{10}, preterm birth and whole pregnancy

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [19]	Is the population overlap 'tolerable' or 'intolerable'?
Klepac et al 2018	Dibben and Clemens 2015	3.2	Did not contain population overlaps.	N/A
	Hao et al 2016)	21.8		
	Padula et al 2014	23.3		
	Wu et al 2011[20]	14.4		
	Wu et al 2011[20]	16.9		
	Zhao et al 2015	20.4		
Lamichhane et al 2015	Hansen et al 2006	0.35	Did not contain population overlaps.	N/A
	Rogers and Dunlop 2006	1.01		
	Zhao et al 2015	0.09		
		(out of 1.44)		
Sapkota et al 2012	Brauer et al 2008	Not provided in the study.	Did not contain population overlaps.	N/A
	Jalaludin et al 2007			
	Leem et al 2006			
	Ritz et al 2000			

[19] Population overlaps were assessed 'blind', that is, without considering the effect estimates reported in each meta-analysis

[20] Wu et al 2011 present estimates for LA and OC separately

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [19]	Is the population overlap 'tolerable' or 'intolerable'?
	Rogers and Dunlop 2006			
	Sagiv et al 2005			
	Suh et al 2009			
	Wilhelm & Ritz 2005			
Stieb et al 2012	Rogers and Dunlop 2006	60.45	Did not contain population overlaps.	N/A
	Hansen et al 2006	24.60		
	Brauer et al 2008	14.95		

f. PM_{10}, PTB and trimester 1

Table A 9. Population overlap details meta-analyses concerning PM_{10}, preterm birth and trimester 1

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [21]	Is the population overlap 'tolerable' or 'intolerable'?
Klepac et al 2018	Hansen et al 2006	2.9		
	Hao et al 2016	12.6		
	Huang et al 2015	11.4		
	Jalaludin et al 2007	4.2		

[21] Population overlaps were assessed 'blind', that is, without considering the effect estimates reported in each meta-analysis

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [21]	Is the population overlap 'tolerable' or 'intolerable'?
	Lee et al 2013	4.8	1. **Wilhelm & Ritz 2005 and Padula et al 2014** Location: Wilhelm & Ritz 2005: California, USA, Padula et al 2014: California, (except Los Angeles and Orange Counties), USA Study Period: Wilhelm & Ritz 2005: 1994-2000, Padula et al 2014: 2000-2006 2. **Wilhelm & Ritz 2005 and Wu et al 2011** Location: Wilhelm & Ritz 2005: Los Angeles County, California, USA, Wu et al 2011: Los Angeles County & Orange County, California, USA (but each county was included separately in the meta-analysis, and it is impossible to discern which analysis relates to which county) Study Period: Wilhelm & Ritz 2005: 1994-2000, Wu et al 2011: 1997-2006	1. Tolerable overlap. Wilhelm & Ritz 2005 add an additional 6 years of data for California, and Padula et al 2014 adds an additional 6 years of data for California (except Los Angeles and Orange counties). 2. Tolerable overlap. Wilhelm & Ritz 2005 add an additional 3 years of data for Los Angeles County, and Wu et al 2011 adds an additional 6 years of data for Los Angeles and Orange counties.
	Lin et al 2015	13.4		
	Padula et al 2014	12.0		
	Wilhelm & Ritz 2005	12.6		
	Wu et al 2011a[22]	8.2		
	Wu et al 2011b[22]	6.0		
	Zhao et al 2015	11.9		
Lamichhane et al 2015	Hansen et al 2006	0.79	Did not contain population overlaps.	N/A
	Kim et al 2007	4.50		
	Ritz et al 2000	10.16		
	Suh et al 2009	7.69		

[22] Wu et al 2011 present estimates for LA and OC separately

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [21]	Is the population overlap 'tolerable' or 'intolerable'?
	Jalaludin et al 2007	2.62		
	Zhao et al 2015	0.53		
	Wilhelm & Ritz 2005	8.57		
	Lee et al 2013	2.22		
		(out of 37.09)		
Sapkota et al 2012 (62)	Information was not provided in article and it was not possible to infer from other tables – we wrote to the authors, but they did not provide any additional information.			
Stieb et al 2012	Hansen et al 2006	5.88	Did not contain population overlaps.	N/A
	Ritz et al 2000	24.71		
	Wilhelm & Ritz 2005	22.71		
	Suh et al 2009	23.52		
	Kim et al 2007	14.68		
	Jalaludin et al 2007	8.50		

g. PM_{10}, PTB and trimester 2

Table A 10. Population overlap details meta-analyses concerning PM_{10}, preterm birth and trimester 2

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [23]	Is the population overlap 'tolerable' or 'intolerable'?
Klepac et al 2018	Hao et al 2016	15.9	Did not contain population overlaps.	N/A
	Huang et al 2015	14.5		
	Lin et al 2015	16.5		
	Padula et al 2014	15.5		
	Wu et al 2011a[24]	9.8		
	Wu et al 2011b[24]	12.8		
	Zhao et al 2015	15.0		
Lamichhane et al 2015	Kim et al 2007	5.28	Did not contain population overlaps.	N/A
	Suh et al 2009	7.82		
	Zhao et al 2015	0.25		
	Wilhelm & Ritz 2005	9.58		
Sapkota et al 2012	Information was not provided in article and it was not possible to infer from other tables – we wrote to the authors, but they did not provide any additional information.			
Stieb et al 2012	Kim et al 2007	9.89	Did not contain population overlaps.	N/A
	Suh et al 2009	40.29		
	Wilhelm & Ritz 2005	49.81		

[23] Population overlaps were assessed 'blind', that is, without considering the effect estimates reported in each meta-analysis

[24] Wu et al 2011 present estimates for LA and OC separately

h. PM_{10}, PTB and trimester 3

Table A 11. Population overlap details meta-analyses concerning PM_{10}, preterm birth and trimester 3

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [25]	Is the population overlap 'tolerable' or 'intolerable'?
Klepac et al 2018	Hao et al 2016	3.7	Did not contain population overlaps.	N/A
	Huang et al 2015	0.8		
	Lin et al 2015	93.0		
	Padula et al 2014	1.3		
Lamichhane et al 2015	Jalaludin et al 2007	2.49	Did not contain population overlaps.	N/A
	Kim et al 2007	6.16		
	Ritz et al 2000	9.75		
	Suh et al 2009	9.29		
	Hansen et al 2006	0.98		
	Wilhelm & Ritz 2005	9.25		
	Zhao et al 2015	0.63		
		(out of 38.54)		
Sapkota et al 2012	Information was not provided in article and it was not possible to infer from other tables – we wrote to the authors, but they did not provide any additional information.			
Stieb et al 2012	Hansen et al 2006	0.84	Did not contain population overlaps.	N/A
	Suh et al 2009	22.90		
	Kim et al 2007	9.49		

[25] Population overlaps were assessed 'blind', that is, without considering the effect estimates reported in each meta-analysis

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [25]	Is the population overlap 'tolerable' or 'intolerable'?
	Ritz et al 2000	37.89		
	Wilhelm & Ritz 2005	27.29		
	Jalaludin et al 2007	1.60		

i. NO_2, PTB and WP

Table A 12. Population overlap details meta-analyses concerning NO_2, preterm birth and whole pregnancy

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [26]	Is the population overlap 'tolerable' or 'intolerable'?
Klepac et al 2018	Dibben et al 2015	11.5	1. **Lavigne et al 2016 and Stieb et al 2016b** Location: Lavigne et al 2016: Ontario, Canada, Stieb et al 2016b: Canada Study Period: Lavigne et al 2016: 2005-2012, Stieb et al 2016b: 1999-2008	1. Tolerable overlap. Some crossover, but Lavigne et al 2016 adds an additional 4 years of data for Ontario, and Stieb et al 2016b adds an additional 6 years of data for the whole of Canada.
	Hao et al 2016	15.5		
	Lavigne et al 2016	15.6		
	Maroziene and Grazuleviciene 2002)	3.4		
	Padula et al 2014	12.0		
	Stieb et al 2016b	16.0		
	Wu et al 2011a[27]	12.9		
	Wu et al 2011b[27]	13.1		

[26] Population overlaps were assessed 'blind', that is, without considering the effect estimates reported in each meta-analysis

[27] Wu et al 2011 present estimates for LA and OC separately

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [26]	Is the population overlap 'tolerable' or 'intolerable'?
Stieb et al 2012	Maroziene and Grazuleviciene 2002	18.79	Did not contain population overlaps.	N/A
	Hansen et al 2006	3.45		
	Brauer et al 2008	37.87		
	Jackson et al 2009	25.47		
	Gehring et al 2011	14.43		

j. NO_2, PTB and trimester 1

Table A 13. Population overlap details meta-analyses concerning NO_2, preterm birth and trimester 1

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [28]	Is the population overlap 'tolerable' or 'intolerable'?
Klepac et al 2018	Hansen et al 2006	1.3	1. **Ritz et al 2007 and Wu et al 2011** Location: Ritz et al 2007: 111 Los Angeles County zip codes, California, USA, Wu et al 2011: Los Angeles County & Orange County, California, USA (but each county was included separately in the meta-analysis, and it is impossible to discern which analysis relates to which county) Study Period: Ritz et al 2007: 2003, Wu et al 2011: 1997-2006	1. Intolerable overlap. Ritz et al 2007 is a complete subset of Wu et al 2011.
	Hao et al 2016	12.7		
	Huang et al 2015	3.7		
	Jalaludin et al 2007	7.0		
	Johnson et al 2016	12.2		
	Lin et al 2015	12.9		
	Maroziene and Grazuleviciene 2002	0.7		
	Padula et al 2014	9.4		
	Ritz et al 2007	8.5		
	Stieb et al 2016b	12.7		

[28] Population overlaps were assessed 'blind', that is, without considering the effect estimates reported in each meta-analysis

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [28]	Is the population overlap 'tolerable' or 'intolerable'?
	Wu et al 2011[29]	10.5		
	Wu et al 2011[29]	8.5		
Stieb et al 2012	Maroziene and Grazuleviciene 2002	6.66	Did not contain population overlaps.	N/A
	Liu et al 2003	25.38		
	Wilhelm & Ritz 2005	26.46		
	Hansen et al 2006	10.88		
	Gehring et al 2011	12.93		
	Jalaludin et al 2007	17.68		

k. NO_2, PTB and trimester 2

Table A 14. Population overlap details meta-analyses concerning NO_2, preterm birth and trimester 2

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [30]	Is the population overlap 'tolerable' or 'intolerable'?
Klepac et al 2018	Hao et al 2016	12.4	1. **Lavigne et al 2016 and Stieb et al 2016b** Location: Lavigne et al 2016: Ontario, Canada, Stieb et al 2016b: Canada Study Period: Lavigne et al 2016: 2005-2012, Stieb et al 2016b: 1999-2008	
	Huang et al 2015	6.3		
	Johnson et al 2016	12.1		
	Lavigne et al 2016	12.4		
	Lin et al 2015	12.6		

[29] Wu et al 2011 present estimates for LA and OC separately

[30] Population overlaps were assessed 'blind', that is, without considering the effect estimates reported in each meta-analysis

	Maroziene and Grazuleviciene 2002	1.5		1. Tolerable overlap. Some crossover, but Lavigne et al 2016 adds an additional 4 years of data for Ontario, and Stieb et al 2016b adds an additional 6 years of data for the whole of Canada.
	Padula et al 2014	9.9		
	Stieb et al 2016b	12.5		
	Wu et al 2011[31]	11.1		
	Wu et al 2011[31]	9.2		
Stieb et al 2012	Maroziene and Grazuleviciene 2002	11.07	Did not contain population overlaps.	N/A
	Wilhelm & Ritz 2005	88.93		

I. NO_2, PTB and trimester 3

Table A 15. Population overlap details meta-analyses concerning NO_2, preterm birth and trimester 3

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [32]	Is the population overlap 'tolerable' or 'intolerable'?
Klepac et al 2018	Hao et al 2016	19.6	1. **Lavigne et al 2016 and Stieb et al 2016b** Location: Lavigne et al 2016: Ontario, Canada, Stieb et al 2016b: Canada Study Period: Lavigne et al 2016: 2005-2012, Stieb et al 2016b: 1999-2008	
	Huang et al 2015	4.7		
	Lavigne et al 2016	19.3		
	Lin et al 2015	19.8		

[31] Wu et al 2011 present estimates for LA and OC separately

[32] Population overlaps were assessed 'blind', that is, without considering the effect estimates reported in each meta-analysis

	Maroziene and Grazuleviciene 2002	2.0		1. Tolerable overlap. Some crossover, but Lavigne et al 2016 adds an additional 4 years of data for Ontario, and Stieb et al 2016b adds an additional 6 years of data for the whole of Canada.
	Padula et al 2014	15.1		
	Stieb et al 2016b	19.6		
Stieb et al 2012	Mariozene and Grazuleviciene 2002	1.64	Did not contain population overlaps.	N/A
	Jalaludin et al 2007	4.21		
	Liu et al 2003	25.46		
	Hansen et al 2006	1.74		
	Wilhelm & Ritz 2005	63.51		
	Gehring et al 2011	3.45		

m. O_3, PTB and WP

Table A 16. Population overlap details meta-analyses concerning O_3, preterm birth and whole pregnancy

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [33]	Is the population overlap 'tolerable' or 'intolerable'?
Klepac et al 2018	Ha et al 2014	30.0	Did not contain population overlaps.	N/A
	Hao et al 2016	42.6		

[33] Population overlaps were assessed 'blind', that is, without considering the effect estimates reported in each meta-analysis

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [33]	Is the population overlap 'tolerable' or 'intolerable'?
	Lavigne et al 2016	20.7		
	Wu et al 2011[34]	4.0		
	Wu et al 2011[34]	2.7		
Stieb et al 2012	Hansen et al 2006	45.04	Did not contain population overlaps.	N/A
	Brauer et al 2008	54.96		

[34] Wu et al 2011 present estimates for LA and OC separately (unclear which is which)

n. O_3, PTB and trimester 1

Table A 17. Population overlap details meta-analyses concerning O_3, preterm birth and trimester 1

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [35]	Is the population overlap 'tolerable' or 'intolerable'?
Klepac et al 2018	Ha et al 2014	24.1	1. **Ritz et al 2007 and Wu et al 2011** Location: Ritz et al 2007: 111 Los Angeles County zip codes, California, USA, Wu et al 2011: Los Angeles County & Orange County, California, USA (but each county was included separately in the meta-analysis, and it is impossible to discern which analysis relates to which county) Study Period: Ritz et al 2007: 2003, Wu et al 2011: 1997-2006	1. Intolerable overlap. Ritz et al 2007 is a complete subset of Wu et al 2011.
	Hansen et al 2006	1.0		
	Hao et al 2016	28.3		
	Jalaludin et al 2007	4.9		
	Lee et al 2013	2.8		
	Lin et al 2015	29.2		
	Ritz et al 2007	3.9		
	Wu et al 2011[36]	2.2		
	Wu et al 2011[36]	3.6		
Stieb et al 2012	Hansen et al 2006	14.16	Did not contain population overlaps.	N/A
	Jalaludin et al 2007	24.63		
	Liu et al 2003	29.54		
	Wilhelm & Ritz 2005	31.67		

[35] Population overlaps were assessed 'blind', that is, without considering the effect estimates reported in each meta-analysis

[36] Wu et al 2011 present estimates for LA and OC separately (unclear which is which)

o. O_3, PTB and trimester 2

Table A 18. Population overlap details meta-analyses concerning O_3, preterm birth and trimester 2

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [37]	Is the population overlap 'tolerable' or 'intolerable'?
Klepac et al 2018	Ha et al 2014	16.8	Did not contain population overlaps.	N/A
	Hao et al 2016	20.0		
	Lavigne et al 2016	19.2		
	Lin et al 2015	20.1		
	Wu et al 2011[38]	10.9		
	Wu et al 2011[38]	13.0		
Stieb et al 2012	Wilhelm & Ritz 2005	100	Did not contain population overlaps.	N/A

[37] Population overlaps were assessed 'blind', that is, without considering the effect estimates reported in each meta-analysis

[38] Wu et al 2011 present estimates for LA and OC separately (unclear which is which)

p. O_3, PTB and trimester 3

Table A 19. Population overlap details meta-analyses concerning O_3, preterm birth and trimester 3

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [39]	Is the population overlap 'tolerable' or 'intolerable'?
Klepac et al 2018	Ha et al 2014	24.8	Did not contain population overlaps.	N/A
	Hao et al 2016	28.9		
	Lavigne et al 2016	16.6		
	Lin et al 2015	29.7		
Stieb et al 2012 (64)	Hansen et al 2006	2.70	Did not contain population overlaps.	N/A
	Wilhelm et al 2005	51.74		
	Jalaludin et al 2007	13.11		
	Liu et al 2003	32.45		

[39] Population overlaps were assessed 'blind', that is, without considering the effect estimates reported in each meta-analysis

2. Term low birth weight (TLBW)

a. $PM_{2.5}$, TLBW and WP

Table A 20. Population overlap details meta-analyses concerning $PM_{2.5}$, term low birth weight and whole pregnancy

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [40]	Is the population overlap 'tolerable' or 'intolerable'?
Li et al 2020	Hyder et al 2014	3.69	1. **Stieb et al 2016b and Lavigne et al 2016** Location: Stieb et al 2016b: Canada, Lavigne et al 2016: Ontario, Canada Study Period: Stieb et al 2016b: 1999-2008, Lavigne et al 2016: 2005-2012 2. **Stieb et al 2016b and Brauer et al 2008** Location: Stieb et al 2016b: Canada, Brauer et al 2008: British Columbia, Canada Study Period: Stieb et al 2016b: 1999-2008, Brauer et al 2008: 1999-2002 3. **Xiao et al 2018 and Yuan et al 2019** Location: Both Shanghai, China Study Period: Xiao et al 2018: 2011-2014, Yuan et al 2019: 2013-2016 4. **Harris et al 2014 and Bell et al 2007**	1. Tolerable overlap. Stieb et al 2016b adds an additional 6 years of data for Canada, and Lavigne et al 2016 adds an additional 4 years of data for Ontario.
	Stieb et al 2016b	5.31		
	Coker et al 2015	2.93		
	Lavigne et al 2016	2.75		
	Harris et al 2014	5.66		

[40] Population overlaps were assessed 'blind', that is, without considering the effect estimates reported in each meta-analysis

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [40]	Is the population overlap 'tolerable' or 'intolerable'?
	Chen et al 2018	1.24	Location: Harris et al 2014: Connecticut, Maine, Minnesota, New Jersey, New York, Utah and Wisconsin, USA, Bell et al 2007: Connecticut and Massachusetts, USA Study period: Harris et al 2014: 2001-2004, Bell et al 2007: 1999-2002 5. **Harris et al 2014 and Hyder et al 2014** Location: Harris et al 2014: Connecticut, Maine, Minnesota, New Jersey, New York, Utah and Wisconsin, USA, Hyder et al 2014: Connecticut and Massachusetts, USA Study period: Harris et al 2014: 2001-2004, Hyder et al 2014: 2000-2006 **6. Bell et al 2007 and Hyder et al 2014** Location: Both Connecticut and Massachusetts, USA Study period: Bell et al 2007: 1999-2002, Hyder et al 2014: 2000-2006 **7. Morello-Frosch et al 2010 and Laurent et al 2016b** Location: Both California, USA. Study period: Morello-Frosch et al 2010: 1996-2006, Laurent et al 2016b: 2001-2008 8. **Morello-Frosch et al 2010 and Coker et al 2015** Location: Morello-Frosch et al 2010: California, USA, Coker et al 2015: Los Angeles County, California, USA Study Period: Morello-Frosch et al 2010: 1996-2006, Coker et al 2015: 1995-2006 9. **Morello-Frosch et al 2010 and Laurent et al 2014** Location: Morello-Frosch et al 2010: California, USA, Laurent et al 2014: Los Angeles County, California, USA	2. Intolerable overlap. Brauer et al 2008 is a complete subset of Stieb et al 2016b. 3. Tolerable overlap. There is minimal overlap as Yuan et al 2019 recruited participants as part of a prospective cohort study whereas Xiao et al 2018 used birth registry data. In additional, Xiao et al 2018 adds an additional 2 years of data, as does Yuan et al 2019 4. Tolerable overlap. Harris et al 2014 adds an additional 2 years of data for Connecticut and 4 years of data for the other 6 states, and Bell et al 2007 add an additional 2 years od data for Connecticut and Massachusetts. 5. Tolerable overlap. Harris et al 2014 adds an additional 4 years of data for 6 states, and Hyder et al 2014 adds 6 years of data for Massachusetts and 3 years of data for Connecticut.
	Wu et al 2018	1.57		
	Balakrishnam et al 2018	6.57		
	Kirwa et al 2019	2.61		
	Yuan et al 2020	0.36		
	Ye et al 2018	5.20		
	Lin et al 2020	2.19		
	Pedersen et al 2013	1.78		
	Brauer et al 2008	0.74		
	Bell et al 2007	3.26		

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [40]	Is the population overlap 'tolerable' or 'intolerable'?
	Kumar et al 2016	2.91	Study Period: Morello-Frosch et al 2010: 1996-2006, Laurent et al 2014: 2001-2008 10. **Morello-Frosch et al 2010 and Laurent et al 2013** Location: Morello-Frosch et al 2010: California, USA, Laurent et al 2013: Los Angeles and Orange Counties, California, USA Study Period: Morello-Frosch et al 2010: 1996-2006, Laurent et al 2013: 1997-2006 11. **Laurent et al 2016b and Coker et al 2015** Location: Laurent et al 2016b: California, USA, Coker et al 2015: Los Angeles County, California, USA Study Period: Laurent et al 2016: 2001-2008, Coker et al 2015: 1995-2006 12. **Laurent et al 2016b and Laurent et al 2014** Location: Laurent et al 2016b: California, USA, Laurent et al 2014: Los Angeles County, California, USA Study Period: Both 2001-2008 13. **Laurent et al 2016b and Laurent et al 2013** Location: Laurent et al 2016b: California, USA, Laurent et al 2013: Los Angeles and Orange Counties, California, USA Study Period: Laurent et al 2016: 2001-2008, Laurent et al 2013: 1997-2006 14. **Coker et al 2015 and Laurent et al 2014**	6. Intolerable overlap. Bell et al 2007 is almost a complete subset of Hyder et al 2014 (only adds one additional year of data) 7. Tolerable overlap. Morello-Frosch et al 2010 adds an additional 5 years of data, and Laurent et al 2016b adds an additional 2 years of data 8. Intolerable overlap. Coker et al 2015 is almost a complete subset of Morello-Frosch et al 2010 (and only adds 1 year of data for Los Angeles County) 9. Tolerable overlap. Morello-Frosch et al 2010 adds an additional 5 years of data for California, and Laurent et al 2014 adds an additional 2 years of data for Los Angeles Country 10. Intolerable overlap. Laurent et al 2013 is a complete subset of Morello-Frosch et al 2010
	Laurent et al 2013	2.27		
	Laurent et al 2014	6.57		
	Laurent et al 2016b	6.11		
	Dadvand et al 2014	0.37		
	Xiao et al 2018	3.20		
	Smith et al 2017	1.61		
	Morello-Frosch et al 2010	6.40		

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [40]	Is the population overlap 'tolerable' or 'intolerable'?
	Kingsley et al 2017)	0.11	Location: Both Los Angeles County, California, USA Study Period: Coker et al 2015: 1995-2006, Laurent et al 2014: 2001-2008 15. **Coker et al 2015 and Laurent et al 2013** Location: Coker et al 2015: Los Angeles County, California, USA, Laurent et al 2013: Los Angeles and Orange Counties, California, USA Study Period: Coker et al 2015: 1995-2006, Laurent et al 2013: 1997-2006 16. **Laurent et al 2014 and Laurent et al 2013** Location: Laurent et al 2014: Los Angeles County, California, USA, Laurent et al 2013: Los Angeles and Orange Counties, California, USA Study Period: Laurent et al 2014: 2001-2008, Laurent et al 2013: 1997-2006	11. Tolerable overlap. Laurent et al 2016b adds an additional 2 years of data for California, and Coker et al 2015 adds an additional 6 years of data for Los Angeles county. 12. Intolerable overlap. Laurent et al 2014 is a complete subset of Laurent et al 2016 13. Tolerable overlap. Laurent et al 2016b adds an additional 2 years of data for California, and Laurent et al 2013 adds additional years of data for Los Angeles and Orange counties 14. Tolerable overlap. Coker et al 2015 adds an additional 6 years of data and Laurent et al 2014 adds an additional 2 years of data. 15. Tolerable overlap. Coker et al 2015 adds an additional 2 years of data for Los Angeles county and Laurent et al 2013 adds data for Orange county for 1997-2006
	Ha et al 2014	3.55		
	Choe et al 2019	6.44		
	Gray et al 2014	4.13		
	Tapia et al 2020	4.89		
	Hao et al 2016			

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [40]	Is the population overlap 'tolerable' or 'intolerable'?
				16. Tolerable overlap. Laurent et al 2014 adds an additional 2 years of data for Los Angeles county and Laurent et al 2013 adds an additional 4 years of data for Los Angeles (and covers Orange county for 1997-2006)
Li et al 2017 (per 10 µg/m^3)	Stieb et al 2016a	22.2	1. **Morello-Frosch et al 2010 and Coker et al 2015** Location: Morello-Frosch et al 2010: California, USA, Coker et al 2015: Los Angeles County, California, USA Study Period: Morello-Frosch et al 2010: 1996-2006, Coker et al 2015: 1995-2006	1. Intolerable overlap. Coker et al 2015 is almost a complete subset of Morello-Frosch et al 2010 (and only adds 1 year of data for Los Angeles County)
	Coker et al 2015	23.8		
	Fleischer et al 2014	28.3		
	Morello-Frosch et al 2010	25.8		
Li et al 2017 (60) (per IQR)	Hyder et al 2014	1.1	1. **Basu et al 2014 and Laurent et al 2014** Location: Basu et al 2014: California, USA, Laurent et al 2014: Los Angeles and Orange Counties, California, USA Study Period: Basu et al 2014: 2000-2006, Laurent et al 2014: 1997-2006	
	Ebisu et al 2012	8.5		

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [40]	Is the population overlap 'tolerable' or 'intolerable'?
	Bell et al 2007	1.9	2. **Hyder et al 2014 and Bell et al 2007** Location: Hyder et al 2014: Connecticut, USA, Bell et al 2007: Connecticut, and Massachusetts, USA Study Period: Hyder et al 2014: 2000-2006, Bell et al 2007: 1999-2002 3. **Hyder et al 2014 and Ebisu and Bell 2012** Location: Hyder et al 2014: Connecticut, USA, Ebisu and Bell 2012: Several northeast states (Inc. Connecticut and Massachusetts), USA Study Period: Hyder et al 2014: 2000-2006, Ebisu and Bell 2012: 2000-2007 4. **Bell et al 2007 and Ebisu and Bell 2012** Location: Bell et al 2007: Connecticut and Massachusetts, USA, Ebisu and Bell 2012: Several northeast states (Inc. Connecticut and Massachusetts), USA Study Period: Bell et al 2007: 1999-2002, Ebisu and Bell 2012: 2000-2007	1. Tolerable overlap. Basu et al 2014 adds data for the entire state of California, and Laurent et al 2014 adds an additional 3 years of data for Los Angeles and Orange counties 2. Tolerable overlap. Hyder et al 2014 adds an additional 4 years of data for Connecticut and Bell et al 2007 adds one year of data for Connecticut (and covers Massachusetts) 3. Intolerable overlap. Hyder et al .2014 is a subset of Ebisu and Bell 2012. 4. Intolerable overlap. Bell et al 2007 is almost a complete subset of Ebisu and Bell 2012 (and would only add 1 year of data for Connecticut and Massachusetts)
	Laurent et al 2014	77.8		
	Dadvand et al 2014	0.2		
	Basu et al 2014	5.2		
	Gray et al 2014	5.4		

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [40]	Is the population overlap 'tolerable' or 'intolerable'?
Sun et al 2016	Basu et al 2014 Bell et al 2007 Bell et al 2010 Brauer et al 2008 Brown et al 2015	Data not provided in the study.	Many studies with overlapping geographies and study periods, including (this may not be an exhaustive list): 1. **Basu et al 2014 and Laurent et al 2014** Location: Basu et al 2014: California, USA, Laurent et al 2014: Los Angeles and Orange Counties, California, USA Study Period: Basu et al 2014: 2000-2006, Laurent et al 2014: 1997-2006 2. **Harris et al 2014 and Bell et al 2007** Location: Harris et al 2014: Connecticut, Maine, Minnesota, New Jersey, New York, Utah and Wisconsin, USA, Bell et al 2007: Connecticut and Massachusetts, USA Study period: Harris et al 2014: 2001-2004, Bell et al 2007: 1999-2002 3. **Harris et al 2014 and Hyder et al 2014** Location: Harris et al 2014: Connecticut, Maine, Minnesota, New Jersey, New York, Utah and Wisconsin, USA, Hyder et al 2014: Connecticut and Massachusetts, USA Study period: Harris et al 2014: 2001-2004, Hyder et al 2014: 2000-2006 4. **Bell et al 2007 and Hyder et al 2014**	1. Tolerable overlap. Basu et al 2014 adds data for the entire state of California, and Laurent et al 2014 adds an additional 3 years of data for Los Angeles and Orange counties 2. Tolerable overlap. Harris et al 2014 adds an additional 2 years of data for Connecticut and 4 years of data for the other 6 states, and Bell et al 2007 add an additional 2 years of data for Connecticut and Massachusetts. 3. Tolerable overlap. Harris et al 2014 adds an additional 4 years of data for 6 states, and Hyder et al 2014 adds 6 years of data for Massachusetts and 3 years of data for Connecticut.

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [40]	Is the population overlap 'tolerable' or 'intolerable'?
	Dadvand et al 2014 Ebisu and Bell 2012) Fleischer et al 2014 Gray et al 2014 Ha et al 2014 Harris et al 2014		Location: Both Connecticut and Massachusetts, USA Study period: Bell et al 2007: 1999-2002, Hyder et al 2014: 2000-2006 5. **Morello-Frosch et al 2010 and Laurent et al 2014** Location: Morello-Frosch et al 2010: California, USA, Laurent et al 2014: Los Angeles County, California, USA Study Period: Morello-Frosch et al 2010: 1996-2006, Laurent et al 2014: 2001-2008 6. **Hyder et al 2014 and Ebisu and Bell 2012** Location: Hyder et al 2014: Connecticut, USA, Ebisu and Bell 2012: Several northeast states (Inc. Connecticut and Massachusetts), USA Study Period: Hyder et al 2014: 2000-2006, Ebisu and Bell 2012: 2000-2007 7. **Bell et al 2007 and Ebisu and Bell 2012** Location: Bell et al 2007: Connecticut and Massachusetts, USA, Ebisu and Bell 2012: Several northeast states (Inc. Connecticut and Massachusetts), USA Study Period: Bell et al 2007: 1999-2002, Ebisu and Bell 2012: 2000-2007 8. **Bell et al 2007 and Parker & Woodruff 2008** Location: Bell et al 2007: Connecticut and Massachusetts, USA, Parker & Woodruff 2008: All mainland USA Study Period: Bell et al 2007: 1999-2002, Parker & Woodruff 2008: 2001-2003	4. Intolerable overlap. Bell et al 2007 is almost a complete subset of Hyder et al 2014 (only adds one additional year of data) 5. Tolerable overlap. Morello-Frosch et al 2010 adds an additional 5 years of data for California, and Laurent et al 2014 adds an additional 2 years of data for Los Angeles Country 6. Intolerable overlap. Hyder et al 2014 is a subset of Ebisu and Bell 2012. 7. Intolerable overlap. Bell et al 2007 is almost a complete subset of Ebisu and Bell 2012 (and would only add 1 year of data for Connecticut and Massachusetts)

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [40]	Is the population overlap 'tolerable' or 'intolerable'?
	Hyder et al 2014 Kumar et al 2012 Laurent et al 2014 Madsen et al 2010 Morello-Frosch et al 2010 Parker and Woodruff 2008		9. **Morello-Frosch et al 2010 and Parker & Woodruff et al 2008** Location: Morello-Frosch et al 2010: California, USA, Parker & Woodruff et al 2008: All mainland USA Study Period: Morello-Frosch et al 2010: 1996-2006, Parker & Woodruff et al 2008: 2001-2003	8. Tolerable overlap. Bell et al 2007 adds an additional 2 years of data for Connecticut and Massachusetts, and Parker & Woodruff add data for the rest of mainland USA 9. Tolerable overlap. Morello-Frosch et al 2010 adds an additional 5 years of data for California and Parker & Woodruff add data for the rest of mainland USA

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [40]	Is the population overlap 'tolerable' or 'intolerable'?
	Pedersen et al 2013			
	Wilhelm et al 2012			
Zhu et al 2015	Brauer et al 2008	0.37	Did not contain population overlaps.	N/A
	Morello-Frosch et al 2010	89.17		
	Darrow et al 2011	2.72		
	Gehring et al 2011	0.04		
	Ebisu et al 2012	4.43		
	Gray et al 2014	3.26		
Sapkota et al 2012	Bell et al 2007			

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [40]	Is the population overlap 'tolerable' or 'intolerable'?
	Brauer et al 2008	Data not provided in the study.	1. **Bell et al 2007 and Parker & Woodruff 2008** Location: Bell et al 2007: Connecticut and Massachusetts, USA, Parker & Woodruff 2008: All mainland USA Study Period: Bell et al 2007: 1999-2002, Parker & Woodruff 2008: 2001-2003	1. Tolerable overlap. Bell et al 2007 adds an additional 2 years of data for Connecticut and Massachusetts, and Parker & Woodruff add data for the rest of mainland USA
	Parker and Woodruff 2008			
	Slama et al 2007			
Stieb et al 2012	Bell et al 2007	17.29	1. **Bell et al 2007 and Parker & Woodruff 2008** Location: Bell et al 2007: Connecticut and Massachusetts, USA, Parker & Woodruff 2008: All mainland USA Study Period: Bell et al 2007: 1999-2002, Parker & Woodruff 2008: 2001-2003	1. Tolerable overlap. Bell et al 2007 adds an additional 2 years of data for Connecticut and Massachusetts, and Parker & Woodruff add data for the rest of mainland USA
	Brauer et al 2008	0.79		
	Madsen et al 2010	32.38	2. **Morello-Frosch et al 2010 and Parker & Woodruff et al 2008** Location: Morello-Frosch et al 2010: California, USA, Parker & Woodruff et al 2008: All mainland USA Study Period: Morello-Frosch et al 2010: 1996-2006, Parker & Woodruff et al 2008: 2001-2003	2. Tolerable overlap. Morello-Frosch et al 2010 adds an additional 5 years of data for California and Parker & Woodruff add data for the rest of mainland USA
	Morello-Frosch et al 2010	31.37		
	Parker and Woodruff et al 2008	18.01		
	Slama et al 2007	0.15		

b. $PM_{2.5}$, TLBW and trimester 1

Table A 21. Population overlap details meta-analyses concerning $PM_{2.5}$, term low birth weight and trimester 1

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [41]	Is the population overlap 'tolerable' or 'intolerable'?
Li et al 2020	Stieb et al 2016b Liang et al 2019a Choe et al 2019 Xiao et al 2018 Ye et al 2018	Not provided in the study	1. **Xiao et al 2018 and Yuan et al 2019** Location: Both Shanghai, China Study Period: Xiao et al 2018: 2011-2014, Yuan et al 2019: 2013-2016 2. **Harris et al 2014 and Hyder et al 2014** Location: Harris et al 2014: Connecticut, Maine, Minnesota, New Jersey, New York, Utah and Wisconsin, USA, Hyder et al 2014: Connecticut and Massachusetts, USA Study period: Harris et al 2014: 2001-2004, Hyder et al 2014: 2000-2006 3. **Hao et al 2016 and Kumar et al 2016** Location: Hao et al 2016: Mainland USA, Kumar et al 2016: Chicago, USA Study Period: Hao et al 2016: 2002, Kumar et al 2016: 2000-2004 4. **Hao et al 2016 and Kingsley et al 2017** Location: Hao et al 2016: Mainland USA, Kingsley et al 2017: Rhode Island, USA. Study Period: Hao et al 2016: 2002, Kingsley et al 2017: 2002-2012 5. **Hao et al 2016 and Hyder et al 2014**	1. Tolerable overlap. There is minimal overlap as Yuan et al 2019 recruited participants as part of a prospective cohort study whereas Xiao et al 2018 used birth registry data. In addition, Xiao et al 2018 adds an additional 2 years of data, as does Yuan et al 2019

[41] Population overlaps were assessed 'blind', that is, without considering the effect estimates reported in each meta-analysis

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [41]	Is the population overlap 'tolerable' or 'intolerable'?
	Vinikoor-Imler et al 2014		Location: Hao et al 2016: Mainland USA, Hyder et al 2014: Connecticut and Massachusetts, USA Study Period: Hao et al 2016: 2002, Hyder et al 2016: 2000-2006 6. **Hao et al 2016 and Harris et al 2014** Location: Hao et al 2016: Mainland USA, Harris et al 2014: Connecticut, Maine, Minnesota, New Jersey, New York, Utah and Wisconsin, USA, Study Period: Hao et al 2016: 2002. Harris et al 2014: 2001-2004	2. Tolerable overlap. Harris et al 2014 adds an additional 4 years of data for 6 states, and Hyder et al 2014 adds 6 years of data for Massachusetts and 3 years of data for Connecticut. 3. Tolerable overlap. Hao et al 2016 covers all of mainland USA and Kumar et al 2016 adds an additional 4 years of data for Chicago 4. Tolerable overlap. Hao et al 2016 covers all of mainland USA and Kingsley et al 2017 adds an additional 10 years of data for Rhode Island.
	Yuan et al 2020			
	Wu et al 2018			
	Hao et al 2016			
	Chen et al 2018			
	Smith et al 2017			
	Kingsley et al 2017			

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [41]	Is the population overlap 'tolerable' or 'intolerable'?
	Hyder et al 2014			5. Tolerable overlap. Hao et al 2016 covers all of mainland USA and Hyder et al 2014 add an additional 6 years of data for Connecticut and Massachusetts 6. Tolerable overlap. Hao et al 2016 covers all of mainland USA and Harris et al 2014 adds an additional 3 years of data for several states.
	Dadvand et al 2014			
	Harris et al 2014			
	Ha et al 2014			
	Kumar et al 2016			
	Tapia et al 2020			
	Kirwa et al 2019			

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [41]	Is the population overlap 'tolerable' or 'intolerable'?
Li et al 2017l (per IQR)	Hyder et al 2014	41.7	Did not contain population overlaps.	N/A
	Vinikoor-Imler et al 2014	42.5		
	Dadvand et al 2014	15.8		
Sun et al 2016	Information was not provided in article – we wrote to the authors, but they did not provide any additional information.			
Sapkota et al 2012	Information was not provided in article and it was not possible to infer from other tables – we wrote to the authors, but they did not provide any additional information.			

c. $PM_{2.5}$, TLBW and trimester 2

Table A 22. Population overlap details meta-analyses concerning $PM_{2.5}$, term low birth weight and trimester 2

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [42]	Is the population overlap 'tolerable' or 'intolerable'?
Li et al 2020	Stieb et al 2016b	Not provided in the study.	1. **Xiao et al 2018 and Yuan et al 2019** Location: Both Shanghai, China Study Period: Xiao et al 2018: 2011-2014, Yuan et al 2019: 2013-2016	

[42] Population overlaps were assessed 'blind', that is, without considering the effect estimates reported in each meta-analysis

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [42]	Is the population overlap 'tolerable' or 'intolerable'?
	Lavigne et al 2016 Liang et al 2019a Choe et al 2019 Xiao et al 2018 Ye et al 2018 Vinikoor-Imler et al 2014		2. **Harris et al 2014 and Hyder et al 2014** Location: Harris et al 2014: Connecticut, Maine, Minnesota, New Jersey, New York, Utah and Wisconsin, USA, Hyder et al 2014: Connecticut and Massachusetts, USA Study period: Harris et al 2014: 2001-2004, Hyder et al 2014: 2000-2006 3. **Hao et al 2016 and Kumar et al 2016** Location: Hao et al 2016: Mainland USA, Kumar et al 2016: Chicago, USA Study Period: Hao et al 2016: 2002, Kumar et al 2016: 2000-2004 4. **Hao et al 2016 and Kingsley et al 2017** Location: Hao et al 2016: Mainland USA, Kingsley et al 2017: Rhode Island, USA. Study Period: Hao et al 2016: 2002, Kingsley et al 2017: 2002-2012 5. **Hao et al 2016 and Hyder et al 2014** Location: Hao et al 2016: Mainland USA, Hyder et al 2014: Connecticut and Massachusetts, USA Study Period: Hao et al 2016: 2002, Hyder et al 2016: 2000-2006 6. **Hao et al 2016 and Harris et al 2014** Location: Hao et al 2016: Mainland USA, Harris et al 2014: Connecticut, Maine, Minnesota, New Jersey, New York, Utah and Wisconsin, USA, Study Period: Hao et al 2016: 2002, Harris et al 2014: 2001-2004 **7. Hao et al 2016 and Laurent et al 2014** Location: Hao et al 2016: Mainland USA, Laurent et al 2014: California, USA Study Period: Hao et al 2016: 2002, Laurent et al 2014: 2001-2008	1. Tolerable overlap. There is minimal overlap as Yuan et al 2019 recruited participants as part of a prospective cohort study whereas Xiao et al 2018 used birth registry data. In addition, Xiao et al 2018 adds an additional 2 years of data, as does Yuan et al 2019. 2. Tolerable overlap. Harris et al 2014 adds an additional 4 years of data for 6 states, and Hyder et al 2014 adds 6 years of data for Massachusetts and 3 years of data for Connecticut. 3. Tolerable overlap. Hao et al 2016 covers all of mainland USA and Kumar et al 2016 adds an additional 4 years of data for Chicago

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [42]	Is the population overlap 'tolerable' or 'intolerable'?
	Yuan et al 2020		8. **Stieb 2016b and Lavigne 2016** Location: Stieb et al 2016b: Canada, Lavigne et al 2016: Ontario, Canada Study Period: Stieb et al 2016b: 1999-2008, Lavigne et al 2016: 2005-2012	4. Tolerable overlap. Hao et al 2016 covers all of mainland USA and Kingsley et al 2017 adds an additional 10 years of data for Rhode Island 5. Tolerable overlap. Hao et al 2016 covers all of mainland USA and Hyder et al 2014 add an additional 6 years of data for Connecticut and Massachusetts 6. Tolerable overlap. Hao et al 2016 covers all of mainland USA and Harris et al 2014 adds an additional 3 years of data for several states. 7. Tolerable overlap. Hao et al 2016 covers all of mainland USA and Laurent et al 2014 adds an additional 7 years of data for California
	Wu et al 2018			
	Hao et al 2016			
	Chen et al 2018			
	Smith et al 2017			
	Kingsley et al 2017			

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [42]	Is the population overlap 'tolerable' or 'intolerable'?
	Hyder et al 2014			8. Tolerable overlap. Stieb et al 2016b adds an additional 6 years of data for Canada, and Lavigne et al 2016 adds an additional 4 years of data for Ontario.
	Dadvand et al 2014			
	Laurent et al 2014			
	Harris et al 2014			
	Ha et al 2014			
	Kumar et al 2016			

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [42]	Is the population overlap 'tolerable' or 'intolerable'?
	Tapia et al 2020			
	Kirwa et al 2019			
Li et al 2017 (per IQR)	Hyder et al 2014	24.8	Did not contain population overlaps.	N/A
	Vinikoor-Imler et al 2014	24.3		
	Dadvand et al 2014	15.8		
	Ha et al 2014	26.1		
Sun et al 2016	Information was not provided in article – we wrote to the authors, but they did not provide any additional information.			
Sapkota et al 2012	Information was not provided in article and it was not possible to infer from other tables – we wrote to the authors, but they did not provide any additional information.			

d. $PM_{2.5}$, TLBW and trimester 3

Table A 23. Population overlap details meta-analyses concerning $PM_{2.5}$, term low birth weight and trimester 3

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [43]	Is the population overlap 'tolerable' or 'intolerable'?
Li et al 2020	Stieb et al 2016a Lavigne et al 2016 Liang et al 2019b Choe et al 2019 Xiao et al 2018 Ye et al 2018	Not provided in the study.	1. **Xiao et al 2018 and Yuan et al 2019** Location: Both Shanghai, China Study Period: Xiao et al 2018: 2011-2014, Yuan et al 2019: 2013-2016 2. **Harris et al 2014 and Hyder et al 2014** Location: Harris et al 2014: Connecticut, Maine, Minnesota, New Jersey, New York, Utah and Wisconsin, USA, Hyder et al 2014: Connecticut and Massachusetts, USA Study period: Harris et al 2014: 2001-2004, Hyder et al 2014: 2000-2006 3. **Hao et al 2016 and Kumar et al 2016** Location: Hao et al 2016: Mainland USA, Kumar et al 2016: Chicago, USA Study Period: Hao et al 2016: 2002, Kumar et al 2016: 2000-2004 4. **Hao et al 2016 and Kingsley et al 2017** Location: Hao et al 2016: Mainland USA, Kingsley et al 2017: Rhode Island, USA. Study Period: Hao et al 2016: 2002, Kingsley et al 2017: 2002-2012 5. **Hao et al 2016 and Hyder et al 2014** Location: Hao et al 2016: Mainland USA, Hyder et al 2014: Connecticut and Massachusetts, USA	1. Tolerable overlap. There is minimal overlap as Yuan et al 2019 recruited participants as part of a prospective cohort study whereas Xiao et al 2018 used birth registry data. In addition, Xiao et al 2018 adds an additional 2 years of data, as does Yuan et al 2019 2. Tolerable overlap. Harris et al 2014 adds an additional 4 years of data for 6 states, and Hyder et al 2014 adds 6 years of data for Massachusetts and 3 years of data for Connecticut.

[43] Population overlaps were assessed 'blind', that is, without considering the effect estimates reported in each meta-analysis

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [43]	Is the population overlap 'tolerable' or 'intolerable'?
	Vinikoor-Imler et al 2014 Yuan et al 2020 Wu et al 2018 Hao et al 2016 Chen et al 2018 Smith et al 2017 Kingsley et al 2017		Study Period: Hao et al 2016: 2002, Hyder et al 2016: 2000-2006 6. **Hao et al 2016 and Harris et al 2014** Location: Hao et al 2016: Mainland USA, Harris et al 2014: Connecticut, Maine, Minnesota, New Jersey, New York, Utah and Wisconsin, USA, Study Period: Hao et al 2016: 2002. Harris et al 2014: 2001-2004 **7. Hao et al 2016 and Laurent et al 2014** Location: Hao et al 2016: Mainland USA, Laurent et al 2014: California, USA Study Period: Hao et al 2016: 2002. Laurent et al 2014: 2001-2008 8. **Stieb 2016a and Lavigne 2016** Location: Stieb et al 2016a: Canada, Lavigne et al 2016: Ontario, Canada Study Period: Stieb et al 2016a: 1999-2008, Lavigne et al 2016: 2005-2012	3. Tolerable overlap. Hao et al 2016 covers all of mainland USA and Kumar et al 2016 adds an additional 4 years of data for Chicago 4. Tolerable overlap. Hao et al 2016 covers all of mainland USA and Kingsley et al 2017 adds an additional 10 years of data for Rhode Island 5. Tolerable overlap. Hao et al 2016 covers all of mainland USA and Hyder et al 2014 add an additional 6 years of data for Connecticut and Massachusetts

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [43]	Is the population overlap 'tolerable' or 'intolerable'?
	Hyder et al 2014			6. Tolerable overlap. Hao et al 2016 covers all of mainland USA and Harris et al 2014 adds an additional 3 years of data for several states. 7. Tolerable overlap. Hao et al 2016 covers all of mainland USA and Laurent et al 2014 adds an additional 7 years of data for California 8. Tolerable overlap. Stieb et al 2016a adds an additional 6 years of data for Canada, and Lavigne et al 2016 adds an additional 4 years of data for Ontario.
	Dadvand et al 2014			
	Laurent et al 2014			
	Harris et al 2014			
	Ha et al 2014			
	Kumar et al 2016			
	Tapia et al 2020			

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [43]	Is the population overlap 'tolerable' or 'intolerable'?
	Kirwa et al 2019			
Li et al 2017 (per IQR)	Hyder et al 2014	43.0	Did not contain population overlaps.	N/A
	Vinikoor-Imler et al 2014	49.0		
	Dadvand et al 2014	7.9		
Sun et al 2016	Information was not provided in article – we wrote to the authors, but they did not provide any additional information.			
Sapkota et al 2012	Information was not provided in article and it was not possible to infer from other tables – we wrote to the authors, but they did not provide any additional information.			

e. PM_{10}, TLBW and WP

Table A 24. Population overlap details meta-analyses concerning PM_{10}, term low birth weight and whole pregnancy

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [44]	Is the population overlap 'tolerable' or 'intolerable'?
Sapkota et al 2012	Bell et al 2007	Not provided in the study.	1. **Parker and Woodruff 2008 and Bell et al 2007** Location: Parker and Woodruff 2008: Mainland USA, Bell et al 2007: Connecticut and Massachusetts, USA Study Period: Parker and Woodruff 2008: 2001-2003, Bell et al 2007: 1999-2002	1. Tolerable overlap. Bell et al 2007 adds an additional 2 years of data for Connecticut and Massachusetts, and Parker & Woodruff add data for the rest of mainland USA.
	Brauer et al 2008			
	Chen et al 2002			
	Gouveia et al 2004			
	Lee et al 2003			
	Lin et al 2004			
	Maisonet et al 2001			
	Parker et al 2008			
	Rogers and Dunlop 2006			
	Salam et al 2005			
	Wilhelm & Ritz 2005			

[44] Population overlaps were assessed 'blind', that is, without considering the effect estimates reported in each meta-analysis

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [44]	Is the population overlap 'tolerable' or 'intolerable'?
Stieb et al 2012	Seo et al 2010 [45]	23.13	1. **Morello-Frosch et al 2010 and Salam et al 2005** Location: both California, USA Study Period: Morello-Frosch et al 2010: 1996-2006, Salam et al 2005: 1993-1996	1. Tolerable overlap. Morello-Frosch et al 2010 adds a further 10 years of data.
	Xu et al 2011	1.73		
	Salam et al 2005	0.62		
	Brauer et al 2008	0.12		
	Madsen et al 2010	17.77		
	Bell et al 2007	11.40		
	Lee et al 2003	22.58		
	Morello-Frosch et al 2010	22.64		
Li et al 2020	Lee et al 2003	12.09	**1. Choe et al 2019 and Seo et al 2010** Location: both Korea, Seoul Study period: Choe et al 2019: 2002-2012, Seo et al 2010: 2004-2004. **2. US studies** **a. Laurent et al 2013 and Bell et al 2007** Location: Laurent et al 2013: LA and Orange Counties, California, USA, Bell et al 2007: California, USA Study Period: Laurent et al 2013: 1997-2006, Bell et al 2007: 1999-2002	1. Tolerable overlap. There is only 1-year overlap for Seoul.
	Lu et al 2020	0.55		
	Seo et al 2010[46]	3.43		
	Seo et al 2010[46]	3.54		
	Seo et al 2010[46]	1.91		
	Seo et al 2010[46]	3.60		

[45] Seo et al 2010 contained separate estimates relating to 7 different cities, so the separate estimates were entered into the meta-analysis individually. Here, we present the combined weight of the separate estimates.

[46] Seo et al 2010 contained separate estimates relating to 7 different cities, so the separate estimates were entered into the meta-analysis individually.

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [44]	Is the population overlap 'tolerable' or 'intolerable'?
	Seo et al 2010[46]	2.36	**b. Laurent et al 2013 and Morello-Frosch et al 2010** Location: Laurent et al 2013: LA and Orange Counties, USA, Morello-Frosch et al 2010: California, USA. Study Period: Laurent et al 2013: 1997-2006, Morello-Frosch et al 2010: 1996-2006 **c. Morello-Frosch et al 2010 and Bell et al 2007** Location: Morello-Frosch et al 2010: California, USA, Bell et al 2007: California, USA. Study Period: Morello-Frosch et al 2010: 1996-2006, Bell et al 2007: 1999-2002	2a. Tolerable overlap. Overlap of 4 years for some of the orange county and LA portion of the population. Extra 6 years of data from Laurent et al 2013. Bell et al 2007 covers a wider population. b. Intolerable overlap. Laurent et al 2013 is a complete subset of Morello-Frosch et al 2010. c. Intolerable overlap. Bell et al 2007 is a subset of Morello-Frosch et al 2010.
	Seo et al 2010[46]	3.25		
	Seo et al 2010[46]	1.89		
	Ye et al 2018	8.59		
	Lin et al 2020	1.68		
	Araban et al 2012	0.01		
	Pedersen et al 2013	1.86		
	Bell et al 2007	7.89		
	Salam et al 2005	0.59		
	Kumar et al 2016	0.48		
	Laurent et al 2013	3.05		
	Wang et al 2019	11.09		
	Smith et al 2017	2.30		
	Morello-Frosch et al 2010	11.75		
	Choe et al 2019	6.36		
	Arroyo et al 2019b	10.14		
	Xu et al 2011	1.60		

f. PM_{10}, TLBW and trimester 1

Table A 25. Population overlap details meta-analyses concerning PM_{10}, term low birth weight and trimester 1

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [47]	Is the population overlap 'tolerable' or 'intolerable'?
Sapkota et al 2012 (62)	Information was not provided in article and it was not possible to infer from other tables – we wrote to the authors, but they did not provide any additional information.			
Stieb et al 2012	Xu et al 2011	6.03	1.**Salam et al 2005 and Wilhelm & Ritz 2005** Location: Salam et al 2005: California, USA, Wilhelm & Ritz 2005: Los Angeles county, California, USA Study Period: Salam et al 2005: 1993-1996, Wilhelm & Ritz 2005: 1994-2000 2. **Maisonet et al 2001 and Xu et al 2011** Location: Maisonet et al 2001: northern cities, USA (includes Pittsburgh, Pennsylvania), Xu et al 2001: Alleghany county, Pennsylvania Study Period: Maisonet et al 2001: 1994-1996, Xu et al 2011: 1994-2000	1. Tolerable overlap. Salam et al 2005 covers the whole of California, and Wilhelm & Ritz 2005 adds a further 4 years of data. Overlap for 2 years for LA County only. 2.Tolerable overlap. Partial overlap as Pittsburgh is in Alleghany county. However, Xu et al 2011 adds a further 4 years of data.
	Kim et al 2007	9.85		
	Dugandzic et al 2006	3.63		
	Wilhelm & Ritz 2005	20.99		
	Lee et al 2003	41.25		

[47] Population overlaps were assessed 'blind', that is, without considering the effect estimates reported in each meta-analysis

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [47]	Is the population overlap 'tolerable' or 'intolerable'?
	Salam et al 2005	3.71		
	Maisonet et al 2001	14.54		
Li et al 2020	Araban et al 2012	Not provided in the study	**1. Choe et al 2019 and Kim et al 2007** Location: both Korea, Seoul Study period: Choe et al 2019: 2002-2012, Kim et al 2007: 2001-2004 **2. Maisonet et al 2001 and Xu et al 2011** Location: Maisonet et al 2001: northern cities, USA (includes Pittsburgh, Pennsylvania), Xu et al 2001: Alleghany county, Pennsylvania, USA Study period: Maisonet et al 2001: 1994-1996, Xu et al 2011: 1994-2000,	1. Intolerable overlap. There is geographical overlap and Kim et al 2007 only includes 1 extra year. 2.Tolerable overlap. Partial overlap as Pittsburgh is in Alleghany county. However, Xu et al 2011 adds a further 4 years of data.
	Choe et al 2019			
	Ye et al 2018			
	Lee et al 2003			
	Smith et al 2017			
	Wang et al 2019			
	Xu et al 2011			

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [47]	Is the population overlap 'tolerable' or 'intolerable'?
	Kim et al 2007			
	Maisonet et al 2001			
	Dadvand et al 2014			
	Dugandzic et al 2006			
	Kumar et al 2016			
	Lu et al 2020			

g. PM_{10}, TLBW and trimester 2

Table A 26. Population overlap details meta-analyses concerning PM_{10}, term low birth weight and trimester 2

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [48]	Is the population overlap 'tolerable' or 'intolerable'?
Sapkota et al 2012	Information was not provided in article and it was not possible to infer from other tables – we wrote to the authors, but they did not provide any additional information.			
Stieb et al 2012	Xu et al 2011	4.47	**1. Salam et al 2005 and Wilhelm & Ritz 2005** Location: Salam et al 2005: California, USA and Wilhelm & Ritz 2005: Los Angeles county, California, USA Study Period: Salam et al 2005: 1993-1996, Wilhelm & Ritz 2005: 1994-2000 2. **Maisonet et al 2001 and Xu et al 2011** Location: Maisonet et al 2001: northern cities, USA (includes Pittsburgh, Pennsylvania), Xu et al 2001: Alleghany county, Pennsylvania, USA Study period: Maisonet et al 2001: 1994-1996, Xu et al 2011: 1994-2000	1. Tolerable overlap. Salam et al 2005 covers the whole of California, and Wilhelm & Ritz 2005 adds a further 4 years of data. Overlap for 2 years for LA County only. 2.Tolerable overlap. Partial overlap as Pittsburgh is in Alleghany county. However, Xu et al 2011 adds a further 4 years of data.
	Salam et al 2005	2.37		
	Kim et al 2007	5.25		
	Lee et al 2003	57.08		
	Wilhelm & Ritz 2005	18.67		

[48] Population overlaps were assessed 'blind', that is, without considering the effect estimates reported in each meta-analysis

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [48]	Is the population overlap 'tolerable' or 'intolerable'?
	Dugandzic et al 2006	2.33		
	Maisonet et al 2001	9.83		
Li et al 2020	Araban et al 2012	Not provided in the study	**1. Choe et al 2019 and Kim et al 2007** Location: both Korea, Seoul Study period: Choe et al 2019: 2002-2012, Kim et al 2007: 2001-2004 **2. Maisonet et al 2001 and Xu et al 2011** Location: Maisonet et al 2001: northern cities, USA (includes Pittsburgh, Pennsylvania), Xu et al 2001: Alleghany county, Pennsylvania, USA Study period: Maisonet et al 2001: 1994-1996, Xu et al 2011: 1994-2000	1. Intolerable overlap. There is geographical overlap and Kim et al 2007 only includes 1 extra year. 2.Tolerable overlap. Partial overlap as Pittsburgh is in Alleghany county. However, Xu et al 2011 adds a further 4 years of data.
	Choe et al 2019			
	Ye et al 2018			
	Lee et al 2003			
	Smith et al 2017			
	Wang et al 2019			
	Xu et al 2011			
	Kim et al 2007			
	Maisonet et al 2001			
	Dadvand et al 2014			
	Dugandzic et al 2006			

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [48]	Is the population overlap 'tolerable' or 'intolerable'?
	Kumar et al 2016			
	Lu et al 2020			

h. PM_{10}, TLBW and trimester 3

Table A 27. Population overlap details meta-analyses concerning PM_{10}, term low birth weight and trimester 3

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [49]	Is the population overlap 'tolerable' or 'intolerable'?
Sapkota et al 2012	Chen et al 2002	Not provided in the study.	1. **Salam et al 2005 and Wilhelm & Ritz 2005** Location: Salam et al 2005: California, USA and Wilhelm & Ritz 2005: Los Angeles county, USA Study Period: Salam et al 2005: 1993-1996, Wilhelm & Ritz 2005: 1994-2000	1. Tolerable overlap. Salam et al 2005 covers the whole of California, and Wilhelm & Ritz 2005 adds a further 4 years of data. Overlap for 2 years for LA County only.
	Gouveia et al 2004			
	Lee et al 2003			
	Lin et al 2004			
	Maisonet et al 2001			
	Salam et al 2005			
	Wilhelm & Ritz 2005			

[49] Population overlaps were assessed 'blind', that is, without considering the effect estimates reported in each meta-analysis

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [49]	Is the population overlap 'tolerable' or 'intolerable'?
Stieb et al 2012	Salam et al 2005	1.63	**1. Salam et al 2005 and Wilhelm & Ritz 2005** Location: Salam et al 2005: California, USA and Wilhelm & Ritz 2005: Los Angeles County, USA Study Period: Salam et al 2005: 1993-1996, Wilhelm & Ritz 2005: 1994-2000 2. **Maisonet et al 2001 and Xu et al 2011** Location: Maisonet et al 2001: Northern cities, USA (includes Pittsburgh, Pennsylvania), Xu et al 2001: Alleghany county, Pennsylvania Study period: Maisonet et al 2001: 1994-1996, Xu et al 2011: 1994-2000	1. Tolerable overlap. Salam et al 2005 covers the whole of California, and Wilhelm & Ritz 2005 adds a further 4 years of data. Overlap for 2 years for LA County only. 2.Tolerable overlap. Partial overlap as Pittsburgh is in Alleghany county. However, Xu et al 2011 adds a further 4 years of data.
	Xu et al 2011	3.05		
	Kim et al 2007	5.98		
	Wilhelm & Ritz 2005	14.04		
	Lee et al 2003	67.86		
	Maisonet et al 2001	6.16		
	Dugandzic et al 2006	1.28		
Li et al 2020	Araban et al 2012	Not provided in the study	**1. Choe et al 2019 and Kim et al 2007** Location: both Korea, Seoul Study period: Choe et al 2019: 2002-2012, Kim et al 2007: 2001-2004. **2. Maisonet et al 2001 and Xu et al 2011**	1. Intolerable overlap. There is geographical overlap and Kim et al 2007 only includes 1 extra year.
	Choe et al 2019			
	Ye et al 2018			
	Lee et al 2003			

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [49]	Is the population overlap 'tolerable' or 'intolerable'?
	Smith et al 2017		Location: Maisonet et al 2001: northern cities, USA (includes Pittsburgh, Pennsylvania), Xu et al 2001: Alleghany county, Pennsylvania, USA Study period: Maisonet et al 2001: 1994-1996, Xu et al 2011: 1994-2000	2.Tolerable overlap. Partial overlap as Pittsburgh is in Alleghany county. However, Xu et al 2011 adds a further 4 years of data
	Wang et al 2019			
	Xu et al 2011			
	Kim et al 2007			
	Maisonet et al 2001			
	Dadvand et al 2014			
	Dugandzic et al 2006			
	Kumar et al 2016			
	Lu et al 2020			

i. NO_2, TLBW and WP

Table A 28. Population overlap details meta-analyses concerning NO_2, term low birth weight and whole pregnancy

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [50]	Is the population overlap 'tolerable' or 'intolerable'?
Stieb et al 2012	Mariozene and Grazuleviciene 2002	0.19	1. **Morello-Frosch et al 2010 and Salam et al 2005** Location: both California, USA Study Period: Morello-Frosch et al 2010: 1996-2006, Salam et al 2005: 1993-1996	1. Tolerable overlap. Morello-Frosch et al 2010 adds a further 10 years of data.
	Slama et al 2007	0.09		

[50] Population overlaps were assessed 'blind', that is, without considering the effect estimates reported in each meta-analysis

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [50]	Is the population overlap 'tolerable' or 'intolerable'?
	Brauer et al 2008	1.38		
	Bell et al 2007	12.07		
	Morello-Frosch et al 2010	28.42		
	Lee et al 2003	21.27		
	Madsen et al 2010	29.53		
	Jackson et al 2009[51]	5.40		
	Jackson et al 2009[51]	0.90		
	Salam et al 2005	0.77		
Li et al 2020	Dedele et al 2017	0.06	**1. Stieb et al 2016a and Lavigne et al 2016** Location: Stieb et al 2016a: Canada, Lavigne et al 2016: Canada, Ontario. Study Period: Stieb et al 2016a: 1999-2008, Lavigne et al 2016: 2005-2012 **2. Stieb et al 2016a and Brauer et al 2008** Location: Stieb et al 2016a: Canada, Brauer et al 2008: British Columbia, Canada Study Period: Stieb et al 2016a: 1999-2008, Brauer et al 2008: 1999-2002 **3. USA studies** **a. Laurent et al 2013 and Ghosh et al 2012** Location: Laurent et al 2013: Los Angeles and Orange Counties, California, USA, Ghosh et al 2012: California, Los Angeles, USA Study Period: Laurent et al 2013: 1997-2006, Ghosh et al 2012: 1995-2006	1. Tolerable overlap. Stieb et al 2016a adds an additional 6 years of data for Canada, and Lavigne et al 2016 adds an additional 4 years of data for Ontario.
	Lee et al 2003	8.51		

[51] Jackson et al 2009 provides estimates using administrative data and data from the millennium cohort for England and Wales

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [50]	Is the population overlap 'tolerable' or 'intolerable'?
	Lu et al 2020	0.06	**b. Laurent et al 2013 and Bell et al 2007** Location: Laurent et al 2013: LA and Orange Counties, California, USA, Bell et al 2007: California, USA Study Period: Laurent et al 2013: 1997-2006, Bell et al 2007: 1999-2002 **c. Laurent et al 2013 and Morello-Frosch et al 2010** Location: Laurent et al 2013: LA and Orange Counties, California, USA, Morello-Frosch et al 2010: California, USA. Study Period: Laurent et al 2013: 1997-2006, Morello-Frosch et al 2010: 1996-2006. **d. Laurent et al 2013 and Laurent et al 2016b** Location: Laurent et al 2013: LA and Orange Counties, California, USA, Laurent et al 2016b: California, USA Study Period: Laurent et al 2013: 1997-2006, Laurent et al 2016b: 2001-2008 **e. Laurent et al 2014 and Laurent et al 2013** Location: Laurent et al 2014: Los Angeles County, California, USA, Laurent et al 2013: Los Angeles and Orange Counties, California, USA Study Period: Laurent et al 2014: 2001-2008, Laurent et al 2013: 1997-2006 **f. Bell et al 2007 and Morello-Frosch et al 2010** Location: Bell et al 2007: California, USA, Morello-Forsch et al 2010: California, USA Study Period: Bell et al 2007: 1999-2002, Morello-Frosch et al 2010: 1996-2006 **g. Bell et al 2007 and Laurent et al 2016b**	2. Intolerable overlap. Brauer et al 2008 is a complete subset of Stieb et al 2016a. 3a. Intolerable overlap. Both studies cover similar geographical regions and Ghosh et al 2012 only adds 2 extra years of data. b. Tolerable overlap. Overlap of 4 years for some of the orange county and LA portion of the population. Extra 6 years of data from Laurent et al 2013. Bell et al 2007 covers a wider population.
	Stieb et al 2016a	8.86		
	Lavigne et al 2016	9.70		
	Ghosh et al 2012	6.51		

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [50]	Is the population overlap 'tolerable' or 'intolerable'?
	Maroziene and Grazuleviciene 2002	0.17	Location: Both California Study Period: Bell et al 2007: 1999-2002, Laurent et al 2016b: 2001-2008 **h. Bell et al 2007 and Ghosh et al 2012** Location: Bell et al 2007: California, Ghosh et al 2012: Los Angeles, California Study Period: Bell 2007: 1999-2002, Ghosh 2012: 1995-2006 **i. Bell et al 2007 and Laurent et al 2014** Location: both California Study Period: Bell et al 2007: 1999-2002, Laurent et al 2014: 2001-2008 **j. Ghosh et al 2012 and Laurent et al 2014** Location: Ghosh et al 2012: Los Angeles, California, USA, Laurent et al 2014: California, USA Study Period: Ghosh et al 2012: 1995-2006, Laurent et al 2014: 2001-2008	c. Intolerable overlap. Laurent et al 2013 is a complete subset of Morello-Frosch et al 2010. d. Tolerable overlap. Laurent et al 2016b adds an additional 2 years of data for California, and Laurent et al 2013 adds additional years of data for Los Angeles and Orange counties
	Ye et al 2018	0.17		
	Araban et al 2012	0.17		
	Pedersen et al 2013	1.50		

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [50]	Is the population overlap 'tolerable' or 'intolerable'?
	Brauer et al 2008	1.64		e. Tolerable overlap. Laurent et al 2014 adds an additional 2 years of data for Los Angeles county and Laurent et al 2013 adds an additional 4 years of data for Los Angeles (and covers Orange county for 1997-2006) f. Intolerable overlap. Bell et al 2007 is a complete subset of Morello-Frosch et al 2010. g. Tolerable overlap. Only 1 year of overlap. Laurent et al 2016b adds an additional 6 years of data.
	Bell et al 2007	6.45		
	Salam et al 2005	0.70		
	Kumar et al 2012	3.69		

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [50]	Is the population overlap 'tolerable' or 'intolerable'?
	Laurent et al 2016b	8.22		h. Tolerable overlap. Bell et al 2007 covers a wider geographical region. Ghosh et al 2012 adds an additional 8 years of data. i. Tolerable overlap. Laurent et al 2014 adds an additional 6 years of data. j. Tolerable overlap. Laurent et al 2014 covers a wider geographical location. Ghosh et al 2012 adds an additional 6 years of data and Laurent et al 2012 adds an additional 2 years of data.
	Laurent et al 2013	3.70		
	Laurent et al 2014	8.76		
	Wang et al 2019	6.62		

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [50]	Is the population overlap 'tolerable' or 'intolerable'?
	Smith et al 2017	5.31		
	Morello-Frosch et al 2010	9.53		
	Yorifuji et al 2015	2.78		
	Michikawa et al 2017	0.97		

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [50]	Is the population overlap 'tolerable' or 'intolerable'?
	Arroyo et al 2019b	5.92		

j. NO_2, TLBW and trimester 1

Table A 29. Population overlap details meta-analyses concerning NO_2, term low birth weight and trimester 1

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [52]	Is the population overlap 'tolerable' or 'intolerable'?
Stieb et al 2012	Wilhelm & Ritz 2005	12.29	1. **Wilhelm & Ritz 2005 and Salam et al 2005** Location: both USA, California Study Period: Wilhelm & Ritz 2005: 1994-2000, Salam et al 2005: 1993-1996	1. Tolerable overlap.
	Lee et al 2003	83.89		
	Liu et al 2003	3.29		
	Salam et al 2005	0.51		
	Maroziene and Grazuleviciene 2002	0.02		
Li et al 2020	Lavigne et al 2016		No overlaps.	N/A
	Lu et al 2020			
	Wang et al 2019			

[52] Population overlaps were assessed 'blind', that is, without considering the effect estimates reported in each meta-analysis

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [52]	Is the population overlap 'tolerable' or 'intolerable'?
	Ye et al 2018	Not provided in the paper.		
	Araban et al 2012			
	Michikawa et al 2017			
	Lee et al 2003			
	Maroziene and Grazuleviciene 2002			
	Dadvand et al 2014			
	Smith et al 2017			
	Ghosh et al 2012			
	Salam et al 2005			
	Laurent et al 2014			

k. NO_2, TLBW and trimester 2

Table A 30. Population overlap details meta-analyses concerning NO_2, term low birth weight and trimester 2

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [53]	Is the population overlap 'tolerable' or 'intolerable'?
Stieb et al 2012	Wilhelm & Ritz 2005	10.70	1. **Wilhelm & Ritz 2005 and Salam et al 2005** Location: both USA, California Study Period: Wilhelm & Ritz 2005: 1994-2000, Salam et al 2005: 1993-1996	1. Tolerable overlap.
	Lee et al 2003	88.62		
	Salam et al 2005	0.64		
	Maroziene and Grazuleviciene 2002	0.04		
Li et al 2020 (71)	Lavigne et al 2016	Not provided in the paper.	No overlaps.	N/A
	Lu et al 2020			
	Wang et al 2019			
	Ye et al 2018			
	Araban et al 2012			
	Michikawa et al 2017			
	Lee et al 2003			
	Maroziene and Grazuleviciene 2002			
	Dadvand et al 2014			
	Smith et al 2017			

[53] Population overlaps were assessed 'blind', that is, without considering the effect estimates reported in each meta-analysis

	Ghosh et al 2012 Salam et al 2005 Laurent et al 2014			

I. NO_2, TLBW and trimester 3

Table A 31. Population overlap details meta-analyses concerning NO_2, term low birth weight and trimester 3

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [54]	Is the population overlap 'tolerable' or 'intolerable'?
Stieb et al 2012	Maroziene and Grazuleviciene 2002	0.79	1. **Wilhelm & Ritz 2005 and Salam et al 2005** Location: both USA, California Study Period: Wilhelm & Ritz 2005: 1994-2000, Salam et al 2005: 1993-1996	1. Tolerable overlap.
	Wilhelm & Ritz 2005	32.73		
	Lee et al 2003	40.22		
	Liu et al 2003	19.16		
	Salam et al 2005	7.10		
Li et al 2020 (71)	Chen et al 2018	Not provided in the paper.	No overlaps.	N/A
	Lavigne et al 2016			
	Lu et al 2020			
	Wang et al 2019			
	Ye et al 2018			
	Araban et al 2012			
	Michikawa et al 2017			
	Lee et al 2003			
	Maroziene and Grazuleviciene 2002			

[54] Population overlaps were assessed 'blind', that is, without considering the effect estimates reported in each meta-analysis

	Dadvand et al 2014			
	Smith et al 2017			
	Ghosh et al 2012			
	Salam et al 2005			
	Laurent et al 2014			

m. O_3, TLBW and WP

Table A 32. Population overlap details meta-analyses concerning O_3, term low birth weight and whole pregnancy

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [55]	Is the population overlap 'tolerable' or 'intolerable'?
Stieb et al 2012	Salam et al 2005	11.38	1. **Morello-Frosch et al 2010 and Salam et al 2005** Location: both California, USA Study Period: Morello-Frosch et al 2010: 1996-2006, Salam et al 2005: 1993-1996	1. Tolerable overlap. Morello-Frosch et al 2010 adds a further 10 years of data.
	Morello-Frosch et al 2010	77.76		
	Brauer et al 2008	10.85		
Li et al 2020	Lavigne et al 2016	8.96	**1. USA Overlaps** **a. Laurent et al 2013 and Morello-Frosch et al 2010** Location: Laurent et al 2013: LA and Orange Counties, California, USA, Morello-Frosch et al 2010: California, USA. Study Period: Laurent et al 2013: 1997-2006, Morello-Frosch et al 2010: 1996-2006. **b. Laurent et al 2013 and Laurent et al 2016b**	1a. Intolerable overlap. Laurent et al 2013 is a complete subset of Morello-Frosch et al 2010.
	Chen et al 2018	2.34		

[55] Population overlaps were assessed 'blind', that is, without considering the effect estimates reported in each meta-analysis

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [55]	Is the population overlap 'tolerable' or 'intolerable'?
	Araban et al 2012	0.01	Location: Laurent et al 2013: LA and Orange Counties, California, USA, Laurent et al 2016b: California, USA Study Period: Laurent et al 2013: 1997-2006, Laurent et al 2016b: 2001-2008 **c. Laurent et al 2014 and Laurent et al 2013** Location: Laurent et al 2014: Los Angeles County, California, USA, Laurent et al 2013: Los Angeles and Orange Counties, California, USA Study Period: Laurent et al 2014: 2001-2008, Laurent et al 2013: 1997-2006 **d. Laurent et al 2016b and Morello-Frosch et al 2010** Location: Both California, USA. Study period: Laurent et al 2016b: 2001-2008, Morello-Frosch et al 2010: 1996-2006 **e. Laurent et al 2016b and Laurent et al 2014** Location: Laurent et al 2016b: California, USA, Laurent et al 2014: Los Angeles County, California, USA Study Period: Both 2001-2008 **f. Morello-Frosch et al 2010 and Laurent et al 2014** Location: Morello-Frosch et al 2010: California, USA, Laurent et al 2014: Los Angeles County, California, USA Study Period: Morello-Frosch et al 2010: 1996-2006, Laurent et al 2014: 2001-2008	b. Tolerable overlap. Laurent et al 2016b adds an additional 2 years of data for California, and Laurent et al 2013 adds additional years of data for Los Angeles and Orange counties c. Tolerable overlap. Laurent et al 2014 adds an additional 2 years of data for Los Angeles county and Laurent et al 2013 adds an additional 4 years of data for Los Angeles (and covers Orange county for 1997-2006)
	Salam et al 2005	1.58		
	Kumar et al 2012	8.87		
	Laurent et al 2016b	11.54		
	Laurent et al 2013	7.34		
	Laurent et al 2014	11.45		

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [55]	Is the population overlap 'tolerable' or 'intolerable'?
	Smith et al 2017	8.43		d. Tolerable overlap. Morello-Frosch et al 2010 adds an additional 5 years of data, and Laurent et al 2016b adds an additional 2 years of data e. Intolerable overlap. Laurent et al 2014 is a complete subset of Laurent et al 2016b f. Intolerable overlap. Laurent et al 2014 is a complete subset of Morello-Frosch et al 2010.
	Morello-Frosch et al 2010	11.57		
	Ha et al 2014	9.35		
	Gray et al 2014	10.58		
	Yorifuji et al 2015	2.29		
	Michikawa et al 2017	5.68		

n. O_3, TLBW and trimester 1

Table A 33. Population overlap details meta-analyses concerning O_3, term low birth weight and trimester 1

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [56]	Is the population overlap 'tolerable' or 'intolerable'?
Stieb et al 2012	Liu et al 2003	24.64	1. **Salam et al 2005 and Wilhelm and Ritz 2005** Location: Salam et al 2005: California, USA, Wilhelm & Ritz 2005: Los Angeles county, California, USA Study Period: Salam et al 2005: 1993-1996, Wilhelm & Ritz 2005: 1994-2000	1. Tolerable overlap. Salam et al 2005 covers the whole of California, and Wilhelm & Ritz 2005 adds a further 4 years of data. Overlap for 2 years for LA County only.
	Salam et al 2005	5.48		
	Wilhelm & Ritz 2005	50.42		
	Dugandzic et al 2006	5.22		
	Ha et al 2001	14.24		
Li et al 2020	Chen et al 2018	Not provided in the study.	No population overlaps.	N/A
	Dugandzic et al 2006			
	Araban et al 2012			
	Michikawa et al 2017			
	Ha et al 2001			
	Smith et al 2017			
	Salam et al 2005			
	Ha et al 2014			

[56] Population overlaps were assessed 'blind', that is, without considering the effect estimates reported in each meta-analysis

	Vinikoor-Imler et al 2014			

o. O_3, TLBW and trimester 2

Table A 34. Population overlap details meta-analyses concerning O_3, term low birth weight and trimester 2

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [57]	Is the population overlap 'tolerable' or 'intolerable'?
Stieb et al 2012	Salam et al 2005	17.80	1. **Salam et al 2005 and Wilhelm and Ritz 2005** Location: Salam et al 2005: California, USA, Wilhelm & Ritz 2005: Los Angeles county, California, USA Study Period: Salam et al 2005: 1993-1996, Wilhelm & Ritz 2005: 1994-2000	1. Tolerable overlap. Salam et al 2005 covers the whole of California, and Wilhelm & Ritz 2005 adds a further 4 years of data. Overlap for 2 years for LA County only.
	Dugandzic et al 2006	19.43		
	Wilhelm & Ritz 2005	62.77		
Li et al 2020	Chen et al 2018	Not provided in the study.	No population overlaps.	N/A
	Dugandzic et al 2006			
	Lavigne et al 2016			
	Michikawa et al 2017			
	Smith et al 2017			
	Salam et al 2005			
	Ha et al 2014			
	Vinikoor-Imler et al 2014			

[57] Population overlaps were assessed 'blind', that is, without considering the effect estimates reported in each meta-analysis

p. O_3, TLBW and Trimester 3

Table A 35. Population overlap details meta-analyses concerning O_3, term low birth weight and trimester 3

Meta-analysis	Individual studies included in the forest plot	% Weight	Population overlap details [58]	Is the population overlap 'tolerable' or 'intolerable'?
Stieb et al 2012	Salam et al 2005	14.55	1. **Salam et al 2005 and Wilhelm and Ritz 2005** Location: Salam et al 2005: California, USA, Wilhelm & Ritz 2005: Los Angeles county, California, USA Study Period: Salam et al 2005: 1993-1996, Wilhelm & Ritz 2005: 1994-2000	1. Tolerable overlap. Salam et al 2005 covers the whole of California, and Wilhelm & Ritz 2005 adds a further 4 years of data. Overlap for 2 years for LA County only.
	Ha et al 2001	22.07		
	Dugandzic et al 2006	13.93		
	Liu et al 2003	23.27		
	Wilhelm & Ritz 2005	26.19		
Li et al 2020	Dugandzic et al 2006	Not provided in the study.	No population overlaps.	N/A
	Lavigne et al 2016			
	Araban et al 2012			
	Michikawa et al 2017			
	Ha et al 2001			
	Smith et al 2017			
	Salam et al 2005			
	Ha et al 2014			

[58] Population overlaps were assessed 'blind', that is, without considering the effect estimates reported in each meta-analysis

	Vinikoor-Imler et al 2014			

Appendix 10. The approach utilised to determine the level of commentary written for Preterm birth (PTB), term low birth weight (TLBW) and stillbirth and the pollutants ($PM_{2.5}$, PM_{10}, NO_2 and O_3)

Outcome	Pollutant	Have new studies been published since selected meta-analyses?	More than 4 acceptable and informative* studies?	More than 10 acceptable and informative* studies?	Endpoint Reached (as illustrated in Figure 6)
PTB	$PM_{2.5}$	Yes	Yes	Yes	D
	PM_{10}	Yes	Yes	Yes	D
	NO_2	Yes	Yes	Yes	D
	O_3	Yes	Yes	Yes	D
TLBW	$PM_{2.5}$	Yes	Yes	Yes	D
	PM_{10}	Yes	Yes	Yes	D
	NO_2	Yes	Yes	Yes	D
	O_3	Yes	Yes	Yes	D
Stillbirth	$PM_{2.5}$	Yes	Yes	No	C
	PM_{10}	Yes	Yes	No	C
	NO_2	Yes	Yes	No	C
	O_3	Yes	Yes	No	C

Appendix 11. Detailed results and summary characteristics from selected meta-analyses per pollutant and outcome pair

1. Preterm birth (PTB)

a. $PM_{2.5}$ and PTB

Table A 36. Summary results and study characteristics from meta-analyses examining preterm birth and $PM_{2.5}$ exposure.

Meta-analysis		Whole Pregnancy	trimester 1	trimester 2	trimester 3
Klepac et al 2018	**OR (95% CI)**	EXCLUDED	EXCLUDED	1.10 (0.96, 1.27)	1.05 (1.02, 1.09)
	Increment			Per 10 µg/m^3	Per 10 µg/m^3
	N studies			8 (9 estimates)	6
	I^2			97%	57%
	Study publication dates			2011-2016	2012-2016
	Individual studies (*adjusted for smoking)			Chang et al 2012* Ha et al 2014* Hao et al 2016* Johnson et al 2016* Lavigne et al 2016*) Padula et al 2014 Pereira et al 2014a* ΔWu et al 2011) ΔWu et al 2011	Chang et al 2012* Ha et al 2014* Hao et al 2016* Johnson et al 2016* Lavigne et al 2016* Padula et al 2014 Pereira et al 2014a*
	N studies adjusted for smoking			6/8 (6/9 estimates)	5/6
Li et al 2017	**OR (95% CI)**	1.02 (0.93, 1.12)	EXCLUDED	EXCLUDED	EXCLUDED
	Increment	Per 10 µg/m^3			
	N studies	6			
	I^2	97%			

Meta-analysis		Whole Pregnancy	trimester 1	trimester 2	trimester 3
	Study publication dates	2006-2016			
	Individual studies (*adjusted for smoking)	Fleischer et al 2014 Hannam et al 2014* Huynh et al 2006 Kloog et al 2012* Ritz et al 2007* Stieb et al 2016a			
	N studies adjusted for smoking	3/6			
Liu et al 2017	**OR (95% CI)**	1.15 (0.99, 1.33)	1.15 (1.05, 1.24)	NOT PERFORMED	NOT PERFORMED
	Increment	Per 10 µg/m^3	Per 10 µg/m^3		
	N studies	7	9		
	I^2	85%	33%		
	Study publication dates	2008-2016	2007-2016		
	Individual studies (*adjusted for smoking)	Brauer et al 2008* Gehring et al 2011* Ha et al 2014* Hannam et al 2014* Pereira et al 2014a* Pereira et al 2014b* Qian et al 2016*	DeFranco 2016* Gehring et al 2011* Ha et al 2014* Hannam et al 2014* Lee et al 2013* Pereira et al 2014a*) Pereira et al 2014b* Pereira et al 2016*) Ritz et al 2007*		
	N studies adjusted for smoking	7/7	9/9		
Zhu et al 2015	**OR (95% CI)**	1.10 (1.03, 1.18)	0.96 (0.77,1.21)	0.90 (0.79, 1.03)	0.97 (0.89, 1.05)
	Increment	Per 10 µg/m^3	Per 10 µg/m^3	Per 10 µg/m^3	Per 10 µg/m^3
	N studies	8	6	3	6
	I^2	52%	87.2%	0%	31.4%

Meta-analysis		Whole Pregnancy	trimester 1	trimester 2	trimester 3
	Study publication dates	2006-2014	2005-2014	2007-2014	2007-2014
	Individual studies (*adjusted for smoking)	Brauer et al 2008* Gehring et al 2011* Gray et al 2014* Hannam et al 2014* Huynh et al 2006 Kloog et al 2012* Pereira et al 2014a* Wu et al 2009*	Gehring et al 2011* Hannam et al 2014* Jalaludin et al 2007* Lee et al 2013* Pereira et al 2014a* Wilhelm & Ritz 2005	Hannam et al 2014* Jalaludin et al 2007* Pereira et al 2014a*)	Gehring et al 2011* Hannam et al 2014* Jalaludin et al 2007* Kloog et al 2012* Pereira et al 2014a* Rudra et al 2011*)
	N studies adjusted for smoking	7/8	5/6	3/3	6/6
Stieb et al 2012	**OR (95% CI)**	EXCLUDED	0.85 (0.60, 1.20)	0.66 (0.57, 0.77)	1.05 (0.98, 1.13)
	Increment		Per 20 µg/m^3	Per 20 µg/m^3	Per 20 µg/m^3
	N studies		4	1	4
	I^2		94.5%		33.2%
	Study publication dates		2005-2011	2005	2005-2011
	Individual studies (*adjusted for smoking)		Gehring et al 2011* Huynh et al 2006 Jalaludin et al 2007* Wilhelm & Ritz 2005	Wilhelm & Ritz 2005	Gehring et al 2011* Huynh et al 2006 Jalaludin et al 2007* Wilhelm & Ritz 2005
	N studies adjusted for smoking		2/4	0/1	2/4
Individual studies in common between meta-analyses?		Li & Liu=1 Li & Zhu=3 Liu & Zhu=4	Liu & Stieb=1 Liu & Zhu=4 Stieb & Zhu=3	Klepac & Zhu=1	Klepac & Zhu=1 Stieb & Zhu=2

b. PM_{10} and PTB

Table A 37. Summary results and study characteristics from meta-analyses examining preterm birth and PM_{10} exposure

Meta-analysis		Whole Pregnancy	trimester 1	trimester 2	trimester 3
Klepac et al 2018	**OR (95% CI)**	1.09 (1.03, 1.16)	1.04 (1.01, 1.08)	1.04 (0.98, 1.09)	1.00 (0.99, 1.00)
	Increment	Per 10 µg/m^3	Per 10 µg/m^3	Per 10 µg/m^3	Per 10 µg/m^3
	N studies	5 studies (6 estimates) Δ	10 (11 estimates) Δ	6 (7 estimates) Δ	5
	I^2	79%	91%	95%	0%
	Study publication dates	2011-2016	2011-2016	2011-2016	2014-2016
	Individual studies (*adjusted for smoking)	Dibben & Clemens 2015* Hao et al 2016* Padula et al 2014 Wu et al 2011 Wu et al 2011 Zhao et al 2015*	Hansen et al 2006 Hao et al 2016* Huang et al 2015 Jalaludin et al 2007* Lee et al 2013* Lin et al 2015* Padula et al 2014 Wilhelm & Ritz 2005 Wu et al 2011 Wu et al 2011 Zhao et al 2015*	Hao et al 2016* Huang et al 2015 Lin et al 2015* Padula et al 2014 Wu et al 2011 Wu et al 2011 Zhao et al 2015*	Hao et al 2016* Huang et al 2015 Lin et al 2015* Padula et al 2014 Zhao et al 2015*
	N studies adjusted for smoking	3/5 (3/6 estimates)	5/10 (5/11 estimates)	3/6 (3/7 estimates)	3/5
Sapkota et al 2012	**OR (95% CI)**	1.02 (0.99, 1.04)	EXCLUDED	NOT PERFORMED	EXCLUDED
	Increment	Per 10 µg/m^3			
	N studies	8			
	I^2	73%			
	Study publication dates	2000-2009			
	Individual studies (*adjusted for smoking)	Brauer et al 2008* Jalaludin et al 2007* Leem et al 2006 Ritz et al 2000* Rogers and Dunlop 2006* Sagiv et al 2005 Suh et al 2009			

Meta-analysis		Whole Pregnancy	trimester 1	trimester 2	trimester 3
		Wilhelm & Ritz 2005			
	N studies adjusted for smoking	4/8			
Stieb et al 2012	**OR (95% CI)**	1.35 (0.97, 1.90)	0.97 (0.87, 1.07)	0.95 (0.91, 0.99)	1.06 (1.03, 1.11)
	Increment	Per 20 µg/m^3	Per 20 µg/m^3	Per 20 µg/m^3	Per 20 µg/m^3
	N studies	3	6	3	6
	I^2	16.9%	85.3%	0%	20.1%
	Study publication dates	2006-2008	2000-2009	2005-2009	2000-2009
	Individual studies (*adjusted for smoking)	Brauer et al 2008* Hansen et al 2006) Rogers & Dunlop 2006*	Hansen et al 2006 Jalaludin et al 2007* Kim et al 2007 Ritz et al 2000* Suh et al 2009 Wilhelm & Ritz 2005	Kim et al 2007 Suh et al 2009 Wilhelm & Ritz 2005	Hansen et al 2006 Jalaludin et al 2007* Kim et al 2007 Ritz et al 2000* Suh et al 2009 Wilhelm & Ritz 2005
	N studies adjusted for smoking	2/3	2/6	0/3	2/6
Individual studies in common between meta-analyses?		Sapkota & Stieb= 2	Klepac & Stieb=2	No	No

c. NO_2 and PTB

Table A 38. Summary results and study characteristics from meta-analyses examining preterm birth and NO_2 exposure

Meta-analysis		Whole Pregnancy	trimester 1	trimester 2	trimester 3
Klepac et al 2018	**OR (95% CI)**	1.05 (0.99, 1.11)	EXCLUDED	1.02 (0.97, 1.08)	1.02 (0.96, 1.08)
	Increment	Per 10 ppb		Per 10 ppb	Per 10 ppb
	N studies	8		10	7
	I^2	95%		96%	97%
	Study publication dates	2002-2016		2002-2016	2002-2016
	Individual studies (*adjusted for smoking)	Dibben et al 2015 Hao et al 2016* Lavigne et al 2016* Maroziene and Grazulevicie ne 2002 Padula et al 2014 Stieb et al 2016a Wu et al 2011 Wu et al 2011		Hao et al 2016* Huang et al 2015 Johnson et al 2016 Lavigne et al 2016* Lin et al 2015* Marozienne and Grazulevicien e 2002 Padula et al 2014 Stieb et al 2016a Wu et al 2011 Wu et al 2011	Hao et al 2016* Haung et al 2015 Lavigne et al 2016* Lin et al 2015* Maroziene and Grazulevicien e 2002 Padula et al 2014 Stieb et al 2016a
	N studies adjusted for smoking	2/8		3/10	3/7
Stieb et al 2012	**OR (95% CI)**	1.16 (0.83, 1.63)	0.87 (0.64, 1.17)	1.03 (0.77, 1.39)	1.06 (0.96, 1.18)
	Increment	Per 20 ppb	Per 20 ppb	Per 20 ppb	Per 20 ppb
	N studies	5	6	2	6
	I^2	53.3%	89.1%	21.6%	19.5%
	Study publication dates	2002-2011	2002-2011	2002-2005	2002-2011
	Individual studies (*adjusted for smoking)	Brauer et al 2008* Gehring et al 2011 Hansen et al 2006 Jackson et al 2009	Gehring et al 2011 Hansen et al 2006 Jalaludin et al 2007* Liu et al 2003 Maroziene and Grazulevicien e 2002	Maroziene and Grazulevicien e 2002 Wilhelm & Ritz 2005	Gehring et al 2011 Hansen et al 2006 Wilhelm and Rtiz 2005 Jalaludin et al 2007* Liu et al 2003

Meta-analysis		Whole Pregnancy	trimester 1	trimester 2	trimester 3
		Maroziene and Grazulevicie ne 2002	Wilhelm & Ritz 2005		Maroziene and Grazulevicien e 2002
	N studies adjusted for smoking	1/5	1/6	0/2	1/6
Individual studies in common between meta-analyses?					

d. O_3 and PTB

Table A 39. Summary results and study characteristics from meta-analyses examining PTB and O_3 exposure

Meta-analysis		Whole Pregnancy	trimester 1	trimester 2	trimester 3
Klepac et al 2018	OR (95% CI)	1.03 (1.01, 1.04)	EXCLUDED	1.12 (1.05, 1.19)	1.01 (0.99, 1.03)
	Increment	Per 10 ppb		Per 10 ppb	Per 10 ppb
	N studies	5		6	4
	I^2	0%		98%	89%
	Study publication dates	2011-2016		2011-2016	2014-2016
	Individual studies (*adjusted for smoking)	Ha et al 2014* Hao et al 2016* Lavigne et al 2016* Wu et al 2011 Wu et al 2011		Ha et al 2014* Hao et al 2016* Lavigne et al 2016* Lin et al 2015* Wu et al 2011 Wu et al 2011	Ha et al 2014* Hao et al 2016* Lavigne et al 2016* Lin et al 2015*
	N studies adjusted for smoking	3/5		4/6	4/4
Stieb et al 2012	OR (95% CI)	1.92 (0.38, 9.76)	1.22 (0.91, 1.64)	0.94 (0.88, 1.00)	0.97 (0.86, 1.10)
	Increment	Per 20 ppb	Per 20 ppb	Per 20 ppb	Per 20 ppb
	N studies	2	4	1	4
	I^2	88.5%	89.8%	N/A	44.5%
	Study publication dates	2006-2008	2003-2007	2005	2003-2007
	Individual studies (*adjusted for smoking)	Brauer et al 2008* Hansen et al 2006	Hansen et al 2006 Jalaludin et al 2007* Liu et al 2003 Wilhelm & Ritz 2005	Wilhelm & Ritz 2005	Hansen et al 2006 Wilhelm & Ritz 2005) Jalaludin et al 2007* Liu et al 2003
	N studies adjusted for smoking	1/2	1/4	0/1	1/4
Individual studies in common between meta-analyses?		No	No	No	No

2. Term low birth weight (TLBW)

a. $PM_{2.5}$ and TLBW

Table A 40. Summary results and study characteristics from meta-analyses examining term low birth weight and $PM_{2.5}$ exposure

Meta-analysis		Whole Pregnancy	trimester 1	trimester 2	trimester 3
Li et al 2020	OR (95% CI)	EXCLUDED	1.031 (0.972, 1.093)	1.031 (0.982, 1.081)	1.053 (1.010, 1.097)
	Increment		Per 10 µg/m^3	Per 10 µg/m^3	Per 10 µg/m^3
	N studies		19	20	20
	I^2		95.1%	91.5%	92%
	Study publication dates		1999-2017	1999-2017	1999-2017
	Individual studies (*adjusted for smoking)		Chen et al 2018* Choe et al 2019 Dadvand et al 2014* Ha et al 2014[59] Hao et al 2016* Harris et al 2014* Hyder et al 2014* Kingsley et al 2017* Kirwa et al 2019 Kumar et al 2016 Liang et al 2019a Smith et al 2017[60] Stieb et al 2016a Tapia et al 2020 Vinikoor-Imler et al 2014* Wu et al 2018 Xiao et al 2018 Ye et al 2018 Yuan et al 2020*	Chen et al 2018* Choe et al 2019 Dadvand et al 2014* Ha et al 2014[59] Hao et al 2016* Harris et al 2014* Hyder et al 2014* Kingsley et al 2017* Kirwa et al 2019) Kumar et al 2016 Lavigne et al 2016* Liang et al 2019a Smith et al 2017[60] Stieb et al 2016a Tapia et al 2020 Vinikoor-Imler et al 2014* Wu et al 2018 Xiao et al 2018 Ye et al 2018	Chen et al 2018* Choe et al 2019 Dadvand et al 2014*) Ha et al 2014[59] Hao et al 2016* Harris et al 2014* Hyder et al 2014* Kingsley et al 2017* Kirwa et al 2019 Kumar et al 2016 Lavigne et al 2016* Liang et al 2019a Smith et al 2017[60] Stieb et al 2016a Tapia et al 2020 Vinikoor-Imler et al 2014* Wu et al 2018 Xiao et al 2018 Ye et al 2018

59 Unclear if controlled for smoking or not. Controlled for any confounders that were associated with both exposure and outcome but which these were unspecified. Smoking is mentioned as being associated with the outcome but unclear whether also associated with the exposure.

60 Smoking status assessed by tobacco expenditure (continuous).

Meta-analysis		Whole Pregnancy	trimester 1	trimester 2	trimester 3
				Yuan et al 2020*	Yuan et al 2020*
	N studies adjusted for smoking		9/19	10/20	10/20
Zhu et al 2015	OR (95% CI)	1.05 (1.02–1.07)	n/a	n/a	n/a
	Increment	Per 10 µg/m^3			
	N studies	6			
	I^2	39.7%			
	Study publication dates	2008-2014 (literature search to March 2014)			
	Individual studies (*adjusted for smoking)	Brauer et al 2008* Darrow et al 2011* Ebisu et al 2012* Gehring et al 2011* Gray et al 2014* Morello-Frosch et al 2010			
	N studies adjusted for smoking	5/6			
Stieb et al 2012	OR (95% CI)	1.05 (0.99, 1.12)	n/a	n/a	n/a
	Increment	Per 10 µg/m^3			
	N studies	6			
	I^2	85.5%			
	Study publication dates	2007-2010 (literature search Jan 1980 – Jan 2011)			
	Individual studies (*adjusted for smoking)	Bell et al 2007* Brauer et al 2008* Madsen et al 2010* Morello-Frosch et al 2010 Parker et al 2008* Slama et al 2007*			
	N studies adjusted for smoking	5/6			

Meta-analysis	Whole Pregnancy	trimester 1	trimester 2	trimester 3
Individual studies in common between meta-analyses?	Two (Brauer et al 2008; Morello-Frosch et al 2010) for Whole Pregnancy. n/a trimester specific estimates			

1. Unclear if controlled for smoking or not. Controlled for any confounders that were associated with both exposure and outcome but which these were unspecified. Smoking is mentioned as being associated with the outcome but unclear whether also associated with the exposure.

b. PM_{10} and TLBW

Table A 41. Summary results and study characteristics from meta-analyses examining term low birth weight and PM_{10} exposure

Meta-analysis		Whole Pregnancy	trimester 1	trimester 2	trimester 3
Li et al 2020	**OR (95% CI)**	EXCLUDED	1.022 (0.998, 1.047)	1.011 (1.005, 1.017)	1.003 (0.995, 1.011)
	Increment		Per 10 µg/m^3	Per 10 µg/m^3	Per 10 µg/m^3
	N studies		13	13	13
	I^2		71.5%	28.2%	20.6%
	Study publication dates		1988-2016	1988-2016	1988-2016
	Individual studies (*adjusted for smoking, †examined LBW not term LBW)		Araban et al 2012 Choe et al 2019 Dadvand et al 2014* Dugandzic et al 2006* Kim et al 2007 Kumar et al 2016 Lee et al 2003 Lu et al 2020 Maisonet et al 2001 Smith et al 2017[61] Wang et al 2019 Xu et al 2011* Ye et al 2018)	Araban et al 2012 Choe et al 2019 Dadvand et al 2014* Dugandzic et al 2006* Kim et al 2007 Kumar et al 2016 Lee et al 2003 Lu et al 2020 Maisonet et al 2001 Smith et al 2017[61] Wang et al 2019 Xu et al 2011* Ye et al 2018	Araban et al 2012) Choe et al 2019 Dadvand et al 2014*) Dugandzic et al 2006*) Kim et al 2007 Kumar et al 2016 Lee et al 2003 Lu et al 2020 Maisonet et al 2001 Smith et al 2017[61] Wang et al 2019 Xu et al 2011* Ye et al 2018
	N studies adjusted for smoking		4/13	4/13	4/13
Sapkota et al 2012	**OR (95% CI)**	1.02 (0.99, 1.05)	EXCLUDED	NOT PERFORMED	1.00 (0.99, 1.01)
	Increment	Per 10 µg/m^3			Per 10 µg/m^3
	N studies	11			7
	I^2	54.5%			Not provided
	Study publication dates	2001-2008			2001-2005

61 Smoking status assessed by tobacco expenditure (continuous).

Meta-analysis		Whole Pregnancy	trimester 1	trimester 2	trimester 3
	Individual studies (*adjusted for smoking, †examined LBW not term LBW)	Bell et al 2007*† Brauer et al 2008* Chen et al 2002* Gouveia et al 2004 Lee et al 2003 Lin et al 2004 Maisonet et al 2001* Parker et al 2008* Rogers and Dunlop 2006*† Salam et al 2005* Wilhelm & Ritz 2005			Chen et al 2002* Gouveia et al 2004 Lee et al 2003 Lin et al 2004 Maisonet et al 2001* Salam et al 2005* Wilhelm & Ritz 2005
	N studies adjusted for smoking	7/11	-	-	3/7
Stieb et al 2012 (64)	**OR (95% CI)**	1.10 (1.05, 1.15)	1.03 (0.95, 1.11)	1.02 (0.96, 1.09)	EXCLUDED
	Increment	Per 20 µg/m^3	Per 20 µg/m^3	Per 20 µg/m^3	
	N studies	8 (14 estimates) $^{\Delta}$	7	7	
	I^2	68.4%	41.6%	22.8%	
	Study publication dates	2003-2011	2001-2011	2001-2011	
	Individual studies (*adjusted for smoking, †examined LBW not term LBW)	Bell et al 2007*† Brauer et al 2008* Lee et al 2003 Madsen et al 2010* Morello-Frosch et al 2010) Salam et al 2005* Seo et al 2010 Xu et al 2011*	Dugandzic et al 2006* Kim et al 2007 Lee et al 2003 Maisonet et al 2001* Salam et al 2005* Wilhelm & Ritz 2005 Xu et al 2011*	Dugandzic et al 2006* Kim et al 2007) Lee et al 2003 Maisonet et al 2001* Salam et al 2005* Wilhelm & Ritz 2005 Xu et al 2011*	
	N studies adjusted for smoking	5/8 (5/14 estimates)	4/7	4/7	
Individual studies in common between meta-analyses?		Yes (4)	-	-	Yes (2)

c. NO_2 and TLBW

Table A 42. Summary results and study characteristics from meta-analyses examining TLBW and NO_2 exposure

Meta-analysis		Whole Pregnancy	trimester 1	trimester 2	trimester 3
Li et al 2020	**OR (95% CI)**	EXCLUDED	1.022 (1.009, 1.035)	1.013 (0.988, 1.038)	1.012 (0.969, 1.058)
	Increment		Per 10 ppb	Per 10 ppb	Per 10 ppb
	N studies		12	13	13
	I^2		10.6%	74.9%	78.1%
	Study publication dates		1975-2016	1975-2016	1975-2016
	Individual studies (*adjusted for smoking)		Araban et al 2012 Dadvand et al 2014* Ghosh et al 2012 Laurent et al 2014 Lee et al 2003 Lu et al 2020 Maroziene and Grazulevicien e 2002* Michikawa et al 2017 Salam et al 2005) Smith et al 2017[62] Wang et al 2019 Ye et al 2018	Araban et al 2012 Dadvand et al 2014* Ghosh et al 2012) Laurent et al 2014 Lavigne et al 2016*) Lee et al 2003 Lu et al 2020 Maroziene and Grazulevicien e 2002* Michikawa et al 2017 Salam et al 2005 Smith et al 2017[62] Wang et al 2019 Ye et al 2018	Araban et al 2012 Dadvand et al 2014* Ghosh et al 2012 Laurent et al 2014 Lavigne et al 2016* Lee et al 2003 Lu et al 2020 Maroziene and Grazulevicien e 2002* Michikawa et al 2017 Salam et al 2005 Smith et al 2017[62] Wang et al 2019 Ye et al 2018
	N studies adjusted for smoking		3/12	4/12	4/12
Stieb et al 2012	**OR (95% CI)**	1.05 (1.00, 1.09)	1.03 (0.99, 1.06)	1.04 (1.01, 1.08)	0.98 (0.87, 1.10)
	Increment	Per 20 ppb	Per 20 ppb	Per 20 ppb	Per 20 ppb
	N studies	10	5	4	5
	I^2	78.4%	0%	0%	69.7%
	Study publication dates	2002-2010	2002-2005	2002-2005	2002-2005

62 Smoking status assessed by tobacco expenditure (continuous).

Meta-analysis		Whole Pregnancy	trimester 1	trimester 2	trimester 3
	Individual studies (*adjusted for smoking)	Bell et al 2007) Brauer et al 2008* Jackson et al 2009 Lee et al 2003 Madsen et al 2010 Maroziene and Grazulevicie ne 2002 Morello-Frosch et al 2010 Salam et al 2005 Slama et al 2007	Lee et al 2003) Liu et al 2003 Maroziene and Grazulevicien e 2002 Salam et al 2005 Wilhelm & Ritz 2005	Lee et al 2003 Maroziene and Grazulevicien e 2002) Salam et al 2005 Wilhelm & Ritz 2005	Lee et al 2003 Liu et al 2003 Maroziene and Grazulevicien e 2002 Salam et al 2005 Wilhelm & Rtiz 2005
	N studies adjusted for smoking	1/9	0/5	0/4	0/5

d. O_3 and TLBW

Table A 43. Summary results and study characteristics from meta-analyses examining term low birth weight and O_3 exposure

Meta-analysis		Whole Pregnancy	trimester 1	trimester 2	trimester 3
Li et al 2020	**OR (95% CI)**	EXCLUDED	0.996 (0.947, 1.046)	1.015 (0.948, 1.087)	1.093 (0.992, 1.204)
	Increment		Per 10 ppb	Per 10 ppb	Per 10 ppb
	N studies		9	8	9
	I^2		78.5%	87.4%	95.8%
	Study publication dates		1988-2013	1988-2013	1988-2013
	Individual studies (*adjusted for smoking, †examined LBW not term LBW)		Araban et al 2012 Chen et al 2018* Dugandzic et al 2006* Ha et al 2001 Ha et al 2014* Michikawa et al 2017[63] Salam et al 2005* Smith et al 2017[64] Vinikoor-Imler et al 2014*	Chen et al 2018* Dugandzic et al 2006* Ha et al 2014* Lavigne et al 2016* Michikawa et al 2017[63] Salam et al 2005* Smith et al 2017[64] Vinikoor-Imler et al 2014*	Araban et al 2012 Dugandzic et al 2006* Ha et al 2001 Ha et al 2014* Lavigne et al 2016* Michikawa et al 2017[63] Salam et al 2005* Smith et al 2017[64] Vinikoor-Imler et al 2014*
	N studies adjusted for smoking		6/9	7/8	6/9
Stieb et al 2012	**OR (95% CI)**	1.01 (0.82, 1.25)	0.99 (0.91, 1.08)	0.95 (0.79, 1.15)	1.03 (0.84, 1.26)
	Increment	Per 20 ppb	Per 20 ppb	Per 20 ppb	Per 20 ppb
	N studies	3	5	3	5
	I^2	24.9%	0%	33.5%	75.6%
	Study publication dates	2005-2010	2001-2006	2005-2006	2001-2006
	Individual studies (*adjusted for smoking, †examined LBW not term LBW)	Brauer et al 2008* Morello-Frosch et al 2010 Salam et al 2005*	Dugandzic et al 2006 Ha et al 2001 Liu et al 2003† Salam et al 2005* Wilhelm & Ritz 2005	Dugandzic et al 2006 Salam et al 2005* Wilhelm & Ritz 2005	Dugandzic et al 2006 Ha et al 2001 Liu et al 2003† Salam et al 2005* Wilhelm & Ritz 2005

[63] All studies included in these meta-analyses analysed LBW restricted to term births, except for Michikawa et al 2017 for which the estimate included in the meta-analysis is for 'adverse birth weight at term' defined as 'LBW in addition to SGA at term'

64 Smoking status assessed by tobacco expenditure (continuous).

Meta-analysis	Whole Pregnancy	trimester 1	trimester 2	trimester 3
N studies adjusted for smoking	2/3	1/5	1/3	1/5
Individual studies in common between meta-analyses?	N/A	Yes (3)	Yes (2)	Yes (3)

3. Stillbirth

a. $PM_{2.5}$ and Stillbirth

Table A 44. Summary results and study characteristics from meta-analyses examining Stillbirth and $PM_{2.5}$ exposure

Meta-analysis		Whole Pregnancy	trimester 1	trimester 2	trimester 3
Siddika al. 2016	**OR (95% CI)+**	1.021 (0.996-1.046)	1.042 (0.920-1.180)	1.040 (0.940-1.152)	1.00 (0.981-1.020)
	Increment	Per 4 µg/m^3	Per 4 µg/m^3	Per 4 µg/m^3	Per 4 µg/m^3
	N studies	2	2	2	2
	I²	0%	57.4%	47.9%	0%
	Study publication dates	Start -April 2015	Start -April 2015	Start -April 2015	Start -April 2015
	Individual studies (*adjusted for smoking)	Faiz et al 2012* Green et al 2015	Faiz et al 2012* Green et al 2015)	Faiz et al 2012* Green et al 2015	Faiz et al 2012* Green et al 2015
	N studies adjusted for smoking	1/2	1/2	1/2	1/2

b. PM_{10} and Stillbirth

Table A 45. Summary results and study characteristics from meta-analyses examining Stillbirth and PM_{10} exposure

Meta-analysis		Whole Pregnancy	trimester 1	trimester 2	trimester 3
Siddika al. 2016	**OR (95% CI)⁺**	1.014 (0.948-1.085)	0.998 (0.936-1.064)	1.005 (0.905-1.116)	1.021 (0.919-1.134)
	Increment	Per 10 µg/m^3	Per 10 µg/m^3	Per 10 µg/m^3	Per 10 µg/m^3
	N studies	2	2	2	2
	I²	85.0%	54.1%	81.2%	90.9%
	Study publication dates	Start -April 2015	Start -April 2015	Start -April 2015	Start -April 2015
	Individual studies (*adjusted for smoking)	Hwang et al 2011 Kim et al 2007	Hwang et al 2011 Kim et al 2007	Hwang et al 2011 Kim et al 2007	Hwang et al 2011 Kim et al 2007
	N studies adjusted for smoking	0/2	0/2	0/2	0/2

c. NO_2 and Stillbirth

Table A 46. Summary results and study characteristics from meta-analyses examining Stillbirth and NO_2 exposure

Meta-analysis		Whole Pregnancy	trimester 1	trimester 2	trimester 3
Siddika al. 2016	**OR (95% CI)⁺**	1.066 (0.965-1.178)	1.035 (0.983-1.089)	1.007 (0.948-1.071)	1.015 (0.980-1.051)
	Increment	Per 10 ppb	Per 10 ppb	Per 10 ppb	Per 10 ppb
	N studies	3	3	3	3
	I^2	79.6%	54.8%	65.7%	0%
	Study publication dates	Start -April 2015	Start -April 2015	Start -April 2015	Start -April 2015
	Individual studies (*adjusted for smoking)	Faiz et al 2012* Green et al 2015 Hwang et al 2011	Faiz et al 2012* Green et al 2015 Hwang et al 201	Faiz et al 2012* Green et al 2015 Hwang et al 2011	Faiz et al 2012* Green et al 2015 Hwang et al 2011)
	N studies adjusted for smoking	1/3	1/3	1/3	1/3

d. O_3 and Stillbirth

Table A 47. Summary results and study characteristics from meta-analyses examining Stillbirth and O_3 exposure

Meta-analysis		Whole Pregnancy	trimester 1	trimester 2	trimester 3
Siddika al. 2016	**OR (95% CI)[1]**	1.002 (0.971-1.034)	1.001 (0.983-1.020)	0.991 (0.944-1.040)	1.012 (0.966-1.060)
	Increment	Per 10 ppb	Per 10 ppb	Per 10 ppb	Per 10 ppb
	N studies	2	2	2	2
	I^2	19.6%	0%	68.6%	63.2%
	Study publication dates	Start -April 2015	Start -April 2015	Start -April 2015	Start -April 2015
	Individual studies (*adjusted for smoking)	Green et al 2015 Hwang et al 2011	Green et al 2015 Hwang et al 2011)	Green et al 2015) Hwang et al 2011	Green et al 2015 Hwang et al 2011
	N studies adjusted for smoking	0/2	0/2	0/2	0/2

Appendix 12. Selected key studies of interest not included in core meta-analyses

1. Typical UK levels

One of the considerations for selecting key studies of interest was how the air pollutant exposures compared with typical UK levels. To avoid repeating text in many tables, we summarise this information here.

a. $PM_{2.5}$

UK annual average urban $PM_{2.5}$ concentrations in 2019 were around 10 µg/m^3 (95% confidence interval around 9 to 11 µg/m^3) (averaged across all monitoring sites with more than 75% data capture). Urban background and roadside sites were analysed separately but fairly similar – concentrations at urban background sites were just below 10 µg/m^3, with tighter confidence intervals than those above[65]. Concentrations will be higher in some cities for example, around 13 µg/m^3 annual average in 2016 according to modelled data, with inner London areas around 14 to18 µg/m^3.[66]

b. PM_{10}

UK annual average urban PM_{10} concentrations in 2019 were around 18 µg/m^3 (95% confidence interval around 17 to 19 µg/m^3) (averaged across all monitoring sites with more than 75% data capture). Urban background and roadside sites were analysed separately – concentrations at urban background sites were around 15 µg/m^3, with tighter confidence intervals.[65]

c. NO_2

UK annual average NO_2 concentrations are higher at roadside monitoring sites than at urban background sites. The former averaged around 31 µg/m^3 (95% confidence interval around 29 to 33 µg/m^3) in 2019 (annual average averaged across all urban traffic monitoring sites with more than 75% data capture). Urban background and roadside sites were analysed separately – urban background sites were around 20 µg/m^3, with 95% confidence intervals around 19 to 21 µg/m^3.[65]

d. O_3

UK ozone concentrations, measured as the annual average of the daily maximum 8-hour mean, are around 72 µg/m^3 at rural background sites (95% confidence intervals around 71 to 73 µg/m^3) and around 61 µg/m^3 (95% confidence intervals around 60 to 62 µg/m^3) at urban background sites. It is

[65] Taken from the annual statistics as published here: ENV02 - Air quality statistics - GOV.UK

[66] https://www.london.gov.uk/sites/default/files/pm2.5_in_london_october19.pdf

higher at rural background sites because it is formed from chemical reactions in the atmosphere that take time to occur as polluted air from urban areas moves into rural areas. Ozone is also suppressed in urban areas by reaction with nitric oxide from vehicle emissions (and would thus be even lower at traffic sites, were ozone measured there).[65]

2. Preterm birth

a. $PM_{2.5}$ and PTB

Table A 48. Summary results and study characteristics for selected key studies of interest not included in core meta-analyses for $PM_{2.5}$ and preterm birth

Author	Region/ Location/ Time Period	Adjustment for smoking	Population size	Study design	Exposure assessment	Mean $PM_{2.5}$ ($\mu g/m^3$)	Results (all ORs are from adjusted models)
Laurent et al 2016a	USA California 2000-2008	Y	1,326,942	State-wide nested case-control study	Monitoring stations (4.4km)	Not reported	OR (95% CI) per IQR 6.45 $\mu g/m^3$ increase in $PM_{2.5}$: WP: 1.133 (1.118,1.148)
Sheridan et al 2019	USA California 2005-2010	N	2,293,218	Retrospective cohort study using routine birth records	Environmental monitoring with inverse distance weighting	13.5	Per 10 $\mu g/m^3$ increase in $PM_{2.5}$: WP: HR = 1.12 (1.09, 1.14) T1: HR= 0.99 (0.98, 1.00) T2: HR= 1.04 (1.03, 1.06) T3: HR= 1.05 (1.04, 1.07)
Chen et al 2018	Australia Brisbane 1st July 2003 - 31 December 2013	Y	173,720	Retrospective cohort	Environmental monitoring using Differential Optical Absorbance Spectroscopy	6.21	HR (95% CI) per IQR 1.32 $\mu g/m^3$ increase in $PM_{2.5}$: WP: 1.05 (1.02, 1.08) T1, T2, T3: Not reported, except to say, 'for all air pollutants, highest HRs were present at trimester 3 and the lowest at trimester 1.'

Author	Region/ Location/ Time Period	Adjustment for smoking	Population size	Study design	Exposure assessment	Mean $PM_{2.5}$ ($\mu g/m^3$)	Results (all ORs are from adjusted models)
							Multi-pollutant models (HR (95% CI): +SO_2: 1.01 (0.98, 1.05) +NO_2: 1.05 (1.01, 1.08) +O_3: 1.05 (1.02, 1.09) +SO_2 +NO_2 +O_3: 1.04 (1.01, 1.08)
Giorgis-Allemand et al 2017	Europe (11 countries) The Netherlands; Hungary; Sweden; England; Denmark; France; Italy; Spain; Lithuania; Norway; Greece 1994-2011	Y	71, 493	Pooled study of birth cohorts European Study of Cohorts for Air Pollution Effects (ESCAPE)	Modelling: LUR models and traffic density on street nearest maternal home and total traffic load on major roads within 100m distance estimated	15 (Roughly estimated from plots as the raw figures are not provided)	OR (95% CI) per 5 $\mu g/m^3$ increase in $PM_{2.5}$: WP: 0.96 (0.87, 1.04) T1: 0.98 (0.91, 1.05) T2: 0.96 (0.90, 1.03) Previous week:1.00 (0.96, 1.03) Previous month: 0.97 (0.91, 1.02)
Smith et al 2020	UK London/Greater London 2006-2010	Y (but using an area-level proxy for smoking)	581,774	Retrospective cohort using routine birth records	Modelled (Dispersion modelling to 20x20 m grid resolution)	14.3 (mean of trimester averages)	OR (95% CI) per IQR 3.2 $\mu g/m^3$ increase in $PM_{2.5}$: T1: OR= 0.93 (0.91, 0.95) OR (95% CI) per IQR 3.6 $\mu g/m^3$ increase: T2: OR= 0.85 (0.83, 0.87)

Author	Region/ Location/ Time Period	Adjustment for smoking	Population size	Study design	Exposure assessment	Mean $PM_{2.5}$ ($\mu g/m^3$)	Results (all ORs are from adjusted models)
							Multi-pollutant models: T1: +O_3: 0.96 (0.93, 0.98) T2: +O_3: 0.87 (0.85, 0.89)
Liang et al 2019a	China Pearl River Delta (9 cities - Guangzhou, Shenzhen, Zhuhai, Dongguan, Foshan, Jiangmen, Zhongshan, Huizhou and Zhaoqing), China 2014-2017	N	1,455,026	Retrospective cohort using routine birth records	Environmental monitoring Daily concentration from 50 monitoring stations	T1: 36.57 T2: 34.41 T3: 33.7	HR (95% CI) per 10 $\mu g/m^3$ increase in $PM_{2.5}$: T1: 1.01 (1.00, 1.02) T2: 1.06 (1.05, 1.07) T3: 1.07 (1.06, 1.08) Multi-pollutant models: T1: $PM_{2.5}$+O_3: 1.01 (0.99, 1.02) $PM_{2.5}$+NO_2: 1.01 (1.00, 1.02) $PM_{2.5}$+SO_2: 0.99 (0.98, 1.01) T2: $PM_{2.5}$+O_3: 1.06 (1.05, 1.07) $PM_{2.5}$+NO_2: 1.07 (1.05, 1.08) $PM_{2.5}$+SO_2: 1.05 (1.03, 1.07) T3: $PM_{2.5}$+O_3: 1.07 (1.05, 1.08) $PM_{2.5}$+NO_2: 1.07 (1.06, 1.09) $PM_{2.5}$+SO_2: 1.08 (1.07, 1.10)

Author	Region/ Location/ Time Period	Adjustment for smoking	Population size	Study design	Exposure assessment	Mean $PM_{2.5}$ ($\mu g/m^3$)	Results (all ORs are from adjusted models)
Guo et al 2018	China National January 2014 - December 2014	Y	426,246	Retrospective cohort study	Daily average per city from the China National Environmental Monitoring Centre from 1619 monitors	Not reported	HR (95% CI) per 10 $\mu g/m^3$ increase in $PM_{2.5}$: WP: very PTB: 1.12 (1.11,1.14) PTB: 1.06 (1.05,1.06) PTB: T1: 1.04 (1.03–1.04) T2: 1.02 (1.02–1.02) T3: 1.06 (1.06–1.07)

b. PM_{10} and PTB

Table A 49. Summary results and study characteristics for selected key studies of interest not included in core meta-analyses for PM_{10} and preterm birth

Authors	Region/ Location/ Time Period	Adjustment for smoking	Population size	Study design	Exposure assessment	Mean PM_{10} ($\mu g/m^3$)	Results
Smith et al 2020)	UK London/Greater London 2006-2010	Y (but using an area-level proxy for smoking)	581,774	Retrospective cohort using routine birth records	Modelled (Dispersion modelling to 20x20 m grid resolution)	T1: 23.6 T2: 23.6 Last 3 months: 22.5	OR (95% CI) per IQR 4.0 $\mu g/m^3$ increase in PM_{10} T1: 0.88 (0.87-0.90) OR (95% CI) per IQR 4.5 $\mu g/m^3$ increase in PM_{10} T2: 0.83 (0.82, 0.85)
Giorgis-Allemand et al 2017	Europe (11 countries)	Y	71, 493	Pooled study of birth cohorts	Modelling:	Approx. 18-35	OR (95% CI) per 20 $\mu g/m^3$ increase in PM_{10} WP: 0.97 (0.87-1.07) T1: 0.98 (0.90-1.07)

Authors	Region/ Location/ Time Period	Adjustment for smoking	Population size	Study design	Exposure assessment	Mean PM_{10} ($\mu g/m^3$)	Results
	The Netherlands; Hungary; Sweden; England; Denmark; France; Italy; Spain; Lithuania; Norway; Greece 1994-2011			European Study of Cohorts for Air Pollution Effects (ESCAPE)	LUR models and traffic density on street nearest maternal home and total traffic load on major roads within 100m distance estimated	(Roughly estimated from plots as the raw figures are not provided)	T2: 0.98 (0.90-1.06) Previous week: 0.99 (0.95, 1.04) Previous month: 0.97 (0.91, 1.03) Additional Adj temperature & atmospheric pressure: WP: 0.97 (0.87-1.07) T1: 0.98 (0.90-1.07) T2: 0.98 (0.90-1.06)
Hannam et al 2014	UK Northwest England 2004-2008	Y	265,163	Retrospective cohort using routine birth records	Spatio-temporal (S-T) model alongside a traditional nearest stationary monitor technique (NSTAT)	S-T Model: 40.99 NSTAT: 53.07	OR (95% CI) S-T model WP: 0.90 (0.74-1.11) T1: 1.06 (0.92-1.21) T2: 0.95 (0.83-1.09) T3: 0.97 (0.84-1.11) OR (95% CI) NSTAT WP: 0.90 (0.74-1.11)
Mendola et al 2016	USA 12 centers (19 hospitals) States of: CA, UT, IL, TX, IN, OH, MA, NY, DE, MD & Washington DC, FL 2002-2008	Y	223,502	Retrospective cohort using electronic medical records	Modelling: modified version of the Community Multiscale Air Quality model, a 3-dimensional, regional air quality model developed by the EPA	WP: 22.10	OR (95% CI) in PM_{10} results for total PTB in women without asthma WP: 1.02 (1.00-1.04) weeks 1-7: 1.04 (1.02, 1.07) weeks 8-14: 0.98 (0.96, 1.00) weeks 15-21: 1.00 (0.98, 1.02) weeks 22-28: 1.00 (0.98, 1.02) Last 6 weeks of pregnancy: 1.01 (0.99, 1.03)

Authors	Region/ Location/ Time Period	Adjustment for smoking	Population size	Study design	Exposure assessment	Mean PM_{10} ($\mu g/m^3$)	Results
	2002-2008						OR (95% CI) in PM_{10} results for total PTB in women with asthma WP: 1.05 (0.99, 1.10) weeks 1-7: 1.08 (1.02, 1.15) weeks 8-14: 1.02 (0.96, 1.08) weeks 15-21: 1.03 (0.98, 1.09) weeks 22-28: 1.02 (0.97, 1.08) Last 6 weeks of pregnancy: 1.09 (1.03, 1.15)
Poirier et al 2015	Canada Halifax, Nova Scotia, Canada 2008-2012	Y	14,415	Retrospective cohort using routine birth records	Modelling: LUR	3.3	OR (95% CI) per IQR for PM_{10} adjusted for maternal age, parity, smoking WP: 0.94 (0.89-1.00) OR (95% CI) per IQR for PM_{10} adjusted for maternal age, parity, smoking, neighborhood income WP: 0.93 (0.88-0.98)

c. NO_2 and PTB

Table A 50. Summary results and study characteristics for selected key studies of interest not included in core meta-analyses for NO_2 and preterm birth

Authors	Region/ Location/ Date	Adjustment for smoking	Population size	Study design	Exposure assessment	Mean NO_2 (ppb) (unless otherwise stated)	Results (all ORs are from adjusted models)
Poirier et al 2015	Canada Halifax, Nova Scotia, Canada 2008-2012	Y	13,400	Retrospective cohort using routine birth records	Modelling: LUR	5.0	OR (95% CI) per IQR adjusted for mother's age, parity, and smoking. WP: 1.00 (0.90, 1.09) OR (95% CI) per IQR adjusted for mother's age, parity, smoking, and neighborhood income. WP: 0.94 (0.85, 1.04)
Giorgis-Allemand et al 2017	Europe (11 countries) The Netherlands; Hungary; Sweden; England; Denmark; France; Italy; Spain; Lithuania; Norway; Greece 1994-2011	Y	71, 493	Pooled study of birth cohorts European Study of Cohorts for Air Pollution Effects (ESCAPE)	Modelling: LUR models and traffic density on street nearest maternal home and total traffic load on major roads within 100m distance estimated	22 µg/m^3 (Roughly estimated from plots as the raw figures are not provided)	OR (95% CI) per a 10 µg/m^3 increase in NO_2: WP: 0.96 (0.91,1.01) T1: 0.97 (0.92,1.02) T2: 0.96 (0.92,1.01) Previous week: 0.98 (0.94,1.01) Previous month: 0.96 (0.92, 1.00)

Authors	**Region/ Location/ Date**	**Adjustment for smoking**	**Population size**	**Study design**	**Exposure assessment**	**Mean NO_2** (ppb) (unless otherwise stated)	**Results (all ORs are from adjusted models)**
Smith et al 2020	UK London/Greater London 2006-2010	Y (but using an area-level proxy for smoking)	581,774	Retrospective cohort using routine birth records	Modelled (Dispersion modelling to 20x20 m grid resolution)	40.5 µg/m^3	OR (95% CI) per IQR 14.0 µg/m^3 in NO_2 T1: 0.92 (0.90,0.94) OR (95% CI) per IQR 14.4 µg/m^3 in NO_2 T2: 0.88 (0.86,0.90) Multi-pollutant models: T1: +O_3: 1.00 (0.97, 1.03) T2: +O_3: 0.92 (0.89, 0.95)
Hannam et al 2014	UK Northwest England 2004-2008	Y	265,163	Retrospective cohort using routine birth records	Spatio-temporal (S-T) model alongside a traditional nearest stationary monitor technique (NSTAT)	S-T Model: 54.63 µg/m^3 NSTAT: 74.01 µg/m^3	OR (95% CI) S-T model: WP: 1.10 (0.82, 1.48) T1: 1.08 (0.92, 1.27) T2: 1.07 (0.91, 1.25) T3: 1.01 (0.86, 1.19) OR (95% CI) NSTAT: WP= 0.90 (0.79,1.04)
Chen et al 2018	Australia Brisbane 1 July 2003 - 31 December 2013	Y	173,720	Retrospective cohort	Environmental monitoring Environmental monitoring using Differential Optical Absorbance Spectroscopy	6.74	HRs (95% CIs) of PTB associated with an IQR (1.19 ppb) increase in NO_2 WP: 1.07 (1.01,1.13) (and remained borderline significant only at moderate temperature when stratified by temperature, null for low/high temp) T1, T2, T3: Not reported, except to say, 'for all air pollutants, **highest HRs**

Authors	**Region/ Location/ Date**	**Adjustment for smoking**	**Population size**	**Study design**	**Exposure assessment**	**Mean NO_2** (ppb) (unless otherwise stated)	**Results (all ORs are from adjusted models)**
							were present at trimester 3 and the lowest at trimester 1.' Multi-pollutant models: +$PM_{2.5}$: 1.01 (0.98, 1.03) +SO_2: 1.00 (0.98, 1.02) +O_3: 1.02 (0.96, 1.09) +$PM_{2.5}$ +SO_2 +O_3: 0.93 (0.87, 1.00)
Ji X et al 2019	China Shanghai 2014-2015	N	25,493	Retrospective cohort study + matched case control study	Modelling: LUR model 100x100m of 55 monitoring sites	48.2 µg/m^3	OR (95% CI) for 10 µg/m^3 increase in NO_2 WP: 1.03 (0.96, 1.10) T1: 1.00 (0.95, 1.07) T2: 1.01 (0.96, 1.07) T3: 1.07 (1.03, 1.13) Last month: 1.10 (1.04, 1.15) Last week: 1.05 (1.00, 1.09)
Liang et al 2019b	China Pearl River Delta (7 cities Guangzhou, Shenzhen, Zhuhai, Dongguan, Foshan, Jiangmen, and Zhongshan), China	N	628,439 Local residents: 308,201 Migrants: 320,238	Retrospective birth cohort	Environmental monitoring. 48 stations	WP: 42.8 µg/m^3 T1: 43.3 µg/m^3 T2: 42.6 µg/m^3 T3: 42.6 µg/m^3	HR (95% CI) per 10 µg/m^3 increase Local residents: WP: 0.96 (0.94,0.98) T1: 0.97 (0.96, 0.99) T2: 0.98 (0.96, 0.99) T3: 1.01 (0.99, 1.02) Migrants WP: 0.95 (0.94,0.97) T1: 1.01 (0.99, 1.03) T2: 1.01 (0.99, 1.02) T3: 1.01 (0.99, 1.02)

Authors	Region/ Location/ Date	Adjustment for smoking	Population size	Study design	Exposure assessment	Mean NO_2 (ppb) (unless otherwise stated)	Results (all ORs are from adjusted models)
	2015-2017						Multi pollutant models: Local residents: NO_2+$PM_{2.5}$: 0.95 (0.94, 0.97) NO_2+O_3: 0.97 (0.96, 0.99) NO_2+SO_2: 0.96 (0.94, 0.98) Migrants: NO_2+$PM_{2.5}$: 0.91 (0.90, 0.93) NO_2+O_3: 0.98 (0.96, 1.00) NO_2+SO_2: 0.96 (0.94, 0.97)

d. O_3 and PTB

Table A 51. Summary results and study characteristics for selected key studies of interest not included in core meta-analyses for O_3 and preterm birth

Author	Region/ Location/ Date	Adjustment for smoking	Population size	Study design	Exposure	Mean O_3 (ppb unless otherwise stated)	Results (all ORs are from adjusted models)
Huang et al 2018	North America California 2009-2012	N	1,448,600	Retrospective cohort using routine birth records and hospital records	Environmental monitoring (with IDW) for census tracts with centre <50km from nearest monitor. Nearest monitor approach for census tracts with centres > 50 km from the nearest monitor.	Means NR Range O_3: 0-0.07 ppm	OR (95% CI) per 0.02ppm IQR increase in O_3: WP: 0.9945 (0.9803, 1.0089)

Author	Region/ Location/ Date	Adjustment for smoking	Population size	Study design	Exposure	Mean O_3 (ppb unless otherwise stated)	Results (all ORs are from adjusted models)
Mendola et al 2016	USA 12 centers (19 hospitals) States of: CA, UT, IL, TX, IN, OH, MA, NY, DE, MD & Washington DC, FL 2002-2008 2002-2008	Y	223,502	Retrospective cohort using electronic medical records	Modelling: modified version of the Community Multiscale Air Quality model, a 3-dimensional, regional air quality model developed by the EPA	29.33	OR (95% CI) per IQR increase in O_3: results for total PTB in women without asthma: WP: 1.08 (1.03, 1.13) weeks 1-7: 0.98 (0.96, 1.00) weeks 8-14: 1.04 (1.01, 1.06) weeks 15-21: 1.06 (1.04, 1.09) weeks 22-28: 1.03 (1.00, 1.06) Last 6 weeks of pregnancy: 1.01 (0.99, 1.04) OR (95% CI) per IQR increase in O_3: Results for total PTB in women with asthma: WP: OR 1.02 (0.91, 1.15) Weeks 1-7: 0.97 (0.91, 1.04) Weeks 8-14: 1.03 (0.95, 1.10) Weeks 15-21: 1.05 (0.98, 1.13) Weeks 22-28: 0.99 (0.92, 1.07) Last 6 weeks of pregnancy: 0.97 (0.91, 1.03)

Author	Region/ Location/ Date	Adjustment for smoking	Population size	Study design	Exposure	Mean O_3 (ppb unless otherwise stated)	Results (all ORs are from adjusted models)
Gray et al 2014	North America USA, North Carolina 2002-2006	Y	457,642	Retrospective cohort study using routine birth records	Modelled (Bayesian space-time downscaling fusion model at census tract level)	43.2	OR (95% CI) for O_3 per IQR increase: WP: 1.02 (0.99, 1.04)
Smith et al 2020	UK London/Greater London 2006-2010	Y (but using an area-level proxy for smoking)	581,774	Retrospective cohort using routine birth records	Modelled (Dispersion modelling to 20x20 m grid resolution)	31.8 µg/m^3	OR (95% CI) per IQR 17.6 µg/m^3 increase in O_3 T1: 1.15 (1.11, 1.18) OR (95% CI) per IQR 19.1 µg/m^3 increase in O_3 T2: 1.15 (1.11, 1.18) Multi-pollutant models: T1: OR (95% CI) per IQR 17.6 µg/m^3 increase in O_3 +NO_2: 1.15 (1.10, 1.19) +NO_X: 1.16 (1.11, 1.20) +$PM_{2.5}$ traffic exhaust: 1.14 (1.10, 1.18) +$PM_{2.5}$ traffic non-exhaust: 1.18 (1.14, 1.22) +$PM_{2.5}$: 1.12 (1.09, 1.16) +PM_{10}: 1.09 (1.06, 1.12 T2: OR (95% CI) per IQR 19.1 µg/m^3 increase in O_3 +NO_2: 1.07 (1.02, 1.11) +NO_X: 1.09 (1.04, 1.14)

Author	Region/ Location/ Date	Adjustment for smoking	Population size	Study design	Exposure	Mean O_3 (ppb unless otherwise stated)	Results (all ORs are from adjusted models)
							+PM2.5 traffic exhaust: 1.08 (1.05, 1.12) +$PM_{2.5}$ traffic non-exhaust: 1.13 (1.09, 1.18) +$PM_{2.5}$: 1.07 (1.04, 1.11) +PM_{10}: 1.03 (1.00, 1.06)
Chen et al 2018	Australia Brisbane 1st July 2003 - 31 December 2013	Y	173,720	Retrospective cohort using routine birth records	Environmental monitoring using Differential Optical Absorbance Spectroscopy	16.82	HRs (95% CIs) of PTB associated with an IQR (1.55 ppb) increase in O_3 WP: 1.13 (1.10, 1.16) (and remained significant at all levels when stratified by temperature) T1, T2, T3: Not reported, except to say, '**for all air pollutants, highest HRs were present at trimester 3 and the lowest at trimester 1**.' Multi-pollutant models (WP O_3 exposure): +$PM_{2.5}$: 1.03 (1.02, 1.04) +SO_2: 1.03 (1.02, 1.04) +NO_2: 1.03 (1.03, 1.04) +$PM_{2.5}$ +SO_2 +NO_2: 1.09 (1.05, 1.12)
Liang et al 2019b	China	N	628,439	Retrospective birth cohort	Environmental monitoring. 48 stations	WP: 25.2 T1: 25.3 T2: 25.0	HRs (95% CIs) of PTB associated with a 10 ppb increase in O_3

Author	Region/ Location/ Date	Adjustment for smoking	Population size	Study design	Exposure	Mean O_3 (ppb unless otherwise stated)	Results (all ORs are from adjusted models)
	Pearl River Delta (7 cities Guangzhou, Shenzhen, Zhuhai, Dongguan, Foshan, Jiangmen, and Zhongshan), China 2015-2017		Local residents: 308,201 Migrants: 320,238			T3: 25.3	Local residents: WP: 1.18 (1.12, 1.23 T1: 1.01 (0.97, 1.05) T2: 0.92 (0.89, 0.96) T3: 1.08 (1.04, 1.13) Migrants WP: 1.32 (1.27, 1.39) T1: 1.01 (0.98, 1.05) T2: 1.03 (0.99, 1.06) T3: 1.08 (1.04, 1.12) Multi pollutant models (WP): Local Residents: O_3+$PM_{2.5}$: 1.18 (1.12, 1.24) O_3+NO_2: 1.15 (1.09, 1.21) O_3+SO_2: 1.18 (1.12, 1.23) Migrants: O_3+$PM_{2.5}$: 1.15 (1.09, 1.20) O_3+NO_2: 1.31 (1.25, 1.37) O_3+SO_2: 1.23 (1.18, 1.29)
Qian et al 2016)	Asia China, Wuhan 2011-2013	Y	95,911	Retrospective cohort using routine electronic hospital and birth records	Environmental monitoring (nearest monitor approach)	75 µg/m^3	OR (95% CI) per 10 µg/m^3 WP: 1.05 (1.02, 1.07) T3: 1.04 (1.02, 1.06) Does not report T1 or T2 results. Reports only that T3 demonstrates strongest association. 2-pollutant models for WP per 10 µg/m^3: O_3 + $PM_{2.5}$: 1.04 (1.02, 1.07)

Author	Region/ Location/ Date	Adjustment for smoking	Population size	Study design	Exposure	Mean O_3 (ppb unless otherwise stated)	Results (all ORs are from adjusted models)
				(although authors incorrectly describe this as a 'Population-based prospective cohort study',			O_3 + PM_{10}: 1.05 (1.03, 1.08) O_3 + SO_2: 1.04 (1.01, 1.07) O_3 + NO_2: 1.05 (1.01, 1.08) O_3 + CO: 1.07 (1.05, 1.10)

3. Term low birth weight (TLBW)

a. $PM_{2.5}$ and TLBW

Table A 52. Summary results and study characteristics for selected key studies of interest not included in core meta-analyses for $PM_{2.5}$ and term low birth weight

Authors	Region/ Location/ Date	Adjustment for smoking	Population size	Study design	Exposure Assessment	Mean $PM_{2.5}$ ($\mu g/m^3$)	Results
Pedersen et al 2013	Europe ESCAPE: Norway, Sweden,	Y	50,151	Prospective cohort	Modelled data: LUR	16.5	OR (95%) per 5 $\mu g/m^3$ increase in $PM_{2.5}$ WP: 1.18 (1.06–1.33) T1: 1.12 (1.02-1.23)

Authors	Region/ Location/ Date	Adjustment for smoking	Population size	Study design	Exposure Assessment	Mean $PM_{2.5}$ ($\mu g/m^3$)	Results
	Denmark, Lituania, England, Netherlands, Germany, France, Hungary, Italy, Greece 11 February 1994- 2 June 2011						T2: 1.13 (1.02–1.26) T3: 1.15 (1.04-1.27) Multi-pollutant model: $PM_{2.5}$ adjusted for NO_2: 1.20 (1.074, 1.35)
Brown et al 2015	North America New York State excluding New York City Sep 2001 – Dec 2006	Y	480,430	Retrospective Cohort	Daily estimates on 12x12km grid scale	Not reported	OR (95% CI) for $PM_{2.5}$ WP: <25th percentile (4.49-9.74 $\mu g/m^3$): REF >25th-50th percentile (9.75-10.91 $\mu g/m^3$): 0.95 (95% CI: 0.89,1.01) >50th-75th (10.92-12.11 $\mu g/m^3$): 0.92 (0.86,0.98) >75th (12.12-18.07 $\mu g/m^3$): 0.97 (0.91,1.04)

Authors	Region/ Location/ Date	Adjustment for smoking	Population size	Study design	Exposure Assessment	Mean $PM_{2.5}$ ($\mu g/m^3$)	Results
							T1: <25th percentile (3.66-9.49 $\mu g/m^3$): REF >25th-50th percentile (9.50-10.9 $\mu g/m^3$): 0.97 (0.91, 1.03) >50th-75th (10.91-12.49 $\mu g/m^3$): 0.94 (0.88, 1.00) >75th (12.50-23.97 $\mu g/m^3$): 0.97 (0.91, 1.04) T2: <25th percentile (3.69-9.50 $\mu g/m^3$): REF >25th-50th percentile (9.51-10.83 $\mu g/m^3$): 0.97 (0.91, 1.04) >50th-75th (10.84-12.32 $\mu g/m^3$): 0.96 (0.90, 1.02) >75th (12.33-19.62 $\mu g/m^3$): 0.99 (0.93, 1.06)

Authors	Region/ Location/ Date	Adjustment for smoking	Population size	Study design	Exposure Assessment	Mean $PM_{2.5}$ ($\mu g/m^3$)	Results
							T3: <25th percentile (0.51-9.45 $\mu g/m^3$): REF >25th-50th percentile (9.46-10.77 $\mu g/m^3$): 0.97 (0.91, 1.03) >50th-75th (10.78-12.22 $\mu g/m^3$): 0.94 (0.88, 1.00) >75th (12.23-47.47 $\mu g/m^3$): 0.97 (0.91, 1.04)
Ha et al 2017	North America Hospital sites in Baltimore; Chicago; Cleveland); Houston; Indianapolis; Los Angeles; Miami; Newark; New York; Salt Lake City; Springfield, Massachusetts; Washington, DC. 2002-2008	Y	220,572	Consortium on Safe Labour (CSL) observational cohort study, medical records	Modified Community Multiscale Air Quality model. Concentrations assigned to hospital region	WP: 12.0 T1: 12.0 T2: 11.9 T3: 12.0	RR (95% CI) per IQR of 5.5 $\mu g/m^3$ in $PM_{2.5}$(WP: 4.7, T1: 5.3; T2: 5.0; T3: 5.3) WP:1.10 (0.97,1.26) T1: 1.08 (0.99,1.17) T2: 1.01 (0.93, 1.10) T3: 0.93 (0.86, 1.01)

Authors	Region/ Location/ Date	Adjustment for smoking	Population size	Study design	Exposure Assessment	Mean $PM_{2.5}$ ($\mu g/m^3$)	Results
Poirier et al 2015	Canada Halifax, Nova Scotia, Canada 2008-2012	Y	14,415	Retrospective cohort using routine birth records	Modelling: LUR	1.1	OR (95%) per IQR (quartile concentrations not given) for $PM_{2.5}$ adjusted for mother's age, parity, and smoking WP: 1.01 (0.88, 1.16) OR (95%) per IQR adjusted for mother's age, parity, smoking, and neighbourhood income. WP: 0.96 (0.83, 1.11)

b. PM_{10} and TLBW

Table A 53. Summary results and study characteristics for selected key studies of interest not included in core meta-analyses for PM_{10} and term low birth weight

Authors	Region/ Location/ Date	Adjustment for smoking	Population size	Study design	Exposure assessment	Mean PM_{10} ($\mu g/m^3$)	Results
Enders et al 2019	USA California 2002-2013	Y	2,719,596	Retrospective cohort using routine birth records	Environmental monitoring.	15.92 ($PM_{2.5-10}$ rather than PM_{10})	OR (95% CI) of TLBW by quartile of $PM_{10-2.5}$ (after controlling for $PM_{2.5}$) exposure compared to first quartile exposure:

Authors	Region/ Location/ Date	Adjustment for smoking	Population size	Study design	Exposure assessment	Mean PM_{10} ($\mu g/m^3$)	Results
					Daily $PM_{10-2.5}$ measurements were calculated by subtracting daily $PM_{2.5}$ measurements from concurrent PM_{10} measurements for each co-located monitor.		WP: 2nd quartile: 1.00 (0.98, 1.03) 3rd quartile: 1.03 (1.00, 1.06) 4th quartile: 1.04 (1.01, 1.07) T1: 2nd quartile 1.02 (0.99, 1.04) 3rd quartile: 1.04 (1.01, 1.06) 4th quartile: 1.03 (1.00, 1.06) T2: 2nd quartile: 0.99 (0.97, 1.02) 3rd quartile: 1.03 (1.00, 1.06) 4th quartile: 1.02 (0.99, 1.05) T3: 2nd quartile: 1.00 (0.97, 1.02): 3rd quartile: 1.03 (1.00, 1.06) 4th quartile: 1.04 (1.01, 1.07) The inner quartile cut-offs for $PM_{10-2.5}$ exposure was 11.4 $\mu g/m^3$, 14.3 $\mu g/m^3$ and 18.5 $\mu g/m^3$ for the second, third, and fourth quartile respectively.

Authors	Region/ Location/ Date	Adjustment for smoking	Population size	Study design	Exposure assessment	Mean PM_{10} ($\mu g/m^3$)	Results
							Adding NO_2 and CO increased Bayesian information criterion & did not change the point estimates for $PM_{10-2.5}$
Ha et al 2017	North America Hospital sites in Baltimore; Chicago; Cleveland); Houston; Indianapolis; Los Angeles; Miami; Newark; New York; Salt Lake City; Springfield, Massachusetts; Washington, DC. 2002-2008	Y	220,572	Consortium on Safe Labour (CSL) observational cohort study, medical records	Modified Community Multiscale Air Quality model. Concentrations assigned to hospital region	WP: 22.2 T1: 22.1 T2: 21.9 T3: 22.2	RR (95% CI) per IQR increase in PM_{10} (WP: 4.4, T1: 6.3; T2: 6.4; T3: 7.4) WP: 1.08 (1.01, 1.16) T1: 1.03 (0.96, 1.09) T2: 1.06 (1.00, 1.13) T3: 0.99 (0.92, 1.05)
Poirier et al 2015	Canada Halifax, Nova Scotia, Canada 2008-2012	Y	13,400	Retrospective cohort using routine birth records	Modelling: LUR	3.3	OR (95% CI) per IQR adjusted in PM_{10} for mother's age, parity, and smoking. WP: 1.09 (0.94, 1.26)

Authors	Region/ Location/ Date	Adjustment for smoking	Population size	Study design	Exposure assessment	Mean PM_{10} ($\mu g/m^3$)	Results
							OR (95% CI) per IQR for PM_{10} adjusted for mother's age, parity, smoking, and neighborhood income. WP: 1.04 (0.90, 1.19)
Pedersen et al 2013	Europe ESCAPE: Norway, Sweden, Denmark, Lituania, England, Netherlands, Germany, France, Hungary, Italy, Greece 11 February 1994- 2 June 2011	Y	50,151	Prospective cohort	Modelled data: LUR	25.4	OR (95% CI) for 10 $\mu g/m^3$ increase in PM_{10} WP: 1.16 (1.00–1.35) T1: 1.13 (1.00-1.28) T2: 1.12 (0.97-1.28) T3: 1.10 (0.96-1.27) Two pollutant models: PM_{10} adjusted for NO_2: 1.19 (1.00-1.42)

c. NO_2 and TLBW

Table A 54. Summary results and study characteristics for selected key studies of interest not included in core meta-analyses for NO_2 and term low birth weight

Authors	Region/ Location/ Dates	Adjustment for smoking	Population size	Study design	Exposure assessment	Mean NO_2 (ppb unless otherwise stated)	Results (all ORs are from adjusted models)
Kumar et al 2012	USA Chicago 1 January 2000 - 30 December 2004	Y	398,120	Retrospective cohort using routine birth records	Environmental Monitoring: Pollutants monitored at spatially dispersed sites and daily estimates acquired from U.S. EPA.	3 miles: 24.9 6 miles: 25.2 9 miles: 25.0 12 miles: 25.0 County: 25.1	OR (95% CI) per 1ppb NO_2: WP: 3 miles: 1.0045 (0.981, 1.028) 6 miles: 0.9892 (0.972, 1.007) 9 miles: 0.9958 (0.984, 1.008) 12 miles: 1.0010 (0.990, 1.012) County Level: 1.0050 (0.996, 1.014) T1: 3 miles: 0.9984 (0.978, 1.019) 6 miles: 0.9898 (0.974, 1.006) 9 miles: 0.9967 (0.985, 1.008) 12 miles: 1.0041 (0.993, 1.015) County Level: 1.0063 (0.998, 1.015) T2: 3 miles: 1.0010 (0.980, 1.023) 6 miles: 0.9898 (0.974, 1.006) 9 miles: 0.9918 (0.981, 1.003) 12 miles: 0.9958 (0.986, 1.006) County Level: 1.0015 (0.994, 1.009)

Authors	**Region/ Location/ Dates**	**Adjustment for smoking**	**Population size**	**Study design**	**Exposure assessment**	**Mean NO_2** (ppb unless otherwise stated)	**Results (all ORs are from adjusted models)**
							T3: 3 miles: 1.0166 (0.991, 1.043) 6 miles: 0.9933 (0.974, 1.013) 9 miles: 0.9942 (0.981, 1.008) 12 miles: 0.9961 (0.985, 1.008) County Level: 1.0010 (0.992, 1.010)
Poirier et al 2015	Canada Halifax, Nova Scotia, Canada 2008-2012	Y	13,400	Retrospective cohort using routine birth records	Modelling: LUR	5.0	OR (95% CI) per IQR in NO_2 adjusted for mother's age, parity, and smoking. WP: 1.15 (0.95, 1.40) OR (95% CI) per IQR in NO_2 adjusted for mother's age, parity, smoking, and neighbourhood income. WP: 1.02 (0.83, 1.26)
Pedersen et al 2013	Europe ESCAPE: Norway, Sweden, Denmark, Lituania, England, Netherlands, Germany, France, Hungary, Italy, Greece	Y	50,151	Prospective cohort	Modelled data: LUR	26.2 µg/m^3	OR (95% CI) per 10 µg/m^3 increment in NO_2 WP: 1.09 (1.00,1.19) T1: 1.08 (1.00, 1.17) T2: 1.08 (1.00, 1.17) T3: 1.07 (0.98, 1.16) Two pollutant models (WP): NO_2 adjusted for $PM_{2.5}$: 1.01 (0.91, 1.11)

Authors	**Region/ Location/ Dates**	**Adjustment for smoking**	**Population size**	**Study design**	**Exposure assessment**	**Mean NO_2** (ppb unless otherwise stated)	**Results (all ORs are from adjusted models)**
	11 February 1994- 2 June 2011						
Guo P et al 2020	China 21 Cities, 123 districts of the Guangdong province, China 1 January 2014 - 31 December 2017	Y	2,567,457	Retrospective population-based cohort	Environmental monitoring with inverse distance weighting interpolation	WP: 26.28 $\mu g/m^3$ T1: 26.56 $\mu g/m^3$ T2: 26.12 $\mu g/m^3$ T3: 26.17 $\mu g/m^3$	OR (95% CI) per IQR increase in NO_2. (IQR increments: WP: 14.56 $\mu g/m^3$, T1: 14.57 $\mu g/m^3$, T2: 15.94 $\mu g/m^3$, T3: 15.73 $\mu g/m^3$) WP: 0.97 (0.92, 1.03) T1: 0.98 (0.94 – 1.02) T2: 0.98 (0.93 – 1.03) T3: 0.97 (0.92 – 1.03) Adjustment for other pollutants: NO_2 adjusted for CO: 0.96 (0.90, 1.01) NO_2 adjusted for O_3: 0.97 (0.91, 1.03) NO_2 adjusted for PM_{10}: 0.96 (0.90, 1.02) NO_2 adjusted for $PM_{2.5}$: 0.97 (0.91, 1.03) NO_2 adjusted for SO_2: 0.98 (0.93, 1.04)

d. O_3 and TLBW

Table A 55. Summary results and study characteristics for selected key studies of interest not included in core meta-analyses for O_3 and term low birth weight

Author	Region/ Location/ Date	Adjustment for smoking	Population size	Study design	Exposure assessment	Mean O_3 (ppb unless otherwise stated)	Results (all ORs are from adjusted models)
Brown et al 2015	North America New York State excluding New York City Sep 2001 – Dec 2006	Y	480,430	Retrospective Cohort	Daily estimates on 12x12km grid scale	Range WP: 15.52-60.35 T1: 10.60-66.33 T2: 11.31-65.73 T3: 5.08-99.69 *Extreme values in T3 due to PTB being included in the average exposure calculation.	OR (95% CI) for O_3 WP: ≤25th percentile 15.52-35.61 ppb: Ref >25th-50th percentile 35.62-38.77 ppb: 0.88 (0.83-0.944) >50th-75th percentile 38.78-42.03 ppb: 0.86 (0.81-0.923) >75th percentile 42.04-60.35 ppb: 0.98 (0.92-1.04) T1: ≤25th percentile 10.60-29.51 ppb: Ref >25th-50th percentile 29.51-39.28 ppb: 1.02 (0.96-1.09) >50th-75th percentile 39.29-47.35 ppb: 1.04 (0.98-1.11) >75th percentile 47.36-66.33 ppb: 1.07 (0.99-1.14) T2: ≤25th percentile 11.31-30.11 ppb: Ref

Author	Region/ Location/ Date	Adjustment for smoking	Population size	Study design	Exposure assessment	Mean O_3 (ppb unless otherwise stated)	Results (all ORs are from adjusted models)
							>25th-50th percentile 30.12-40.12 ppb: 0.98 (0.92-1.05) >50th-75th percentile 40.13-47.43 ppb: 0.97 (0.91-1.03) >75th percentile 47.44-65.73 ppb: 0.98 (0.92-1.05) T3: ≤25th percentile 5.08-30.35 ppb: Ref >25th-50th percentile 30.36-39.57 ppb: 0.95 (0.90-1.02) >50th-75th percentile 39.58-46.74 ppb: 0.90 (0.84-0.96) >75th percentile 46.75-99.69 ppb: 0.95 (0.90-1.02)
Ha et al 2017	North America Hospital sites in Baltimore; Chicago; Cleveland); Houston; Indianapolis; Los Angeles; Miami; Newark; New York; Salt Lake City; Springfield, Massachusetts;	Y	220,572	Consortium on Safe Labour (CSL) observational cohort study, medical records	Modified Community Multiscale Air Quality model. Concentrations assigned to hospital region	WP: 28.8 T1: 28.5 T2: 28.9 T3: 29.3	RR (95% CI) per IQR increase for O_3 in ppb (WP: 7.8, T1: 12.3, T2: 11.9, T3: 12.7) WP: 1.01 (0.91, 1.11) T1: 1.00 (0.92, 1.07) T2: 1.03 (0.95, 1.11) T3: 1.00 (0.93, 1.08)

Author	Region/ Location/ Date	Adjustment for smoking	Population size	Study design	Exposure assessment	Mean O_3 (ppb unless otherwise stated)	Results (all ORs are from adjusted models)
	Washington, DC. 2002-2008						
Kumar et al 2012	USA Chicago 1 January 2000 - 30 December 2004	Y	398,120	Retrospective cohort using routine birth records	Environmental Monitoring: Pollutants monitored at spatially dispersed sites and daily estimates acquired from U.S. EPA.	3 miles: 22.8 6 miles: 22.9 9 miles: 22.8 12 miles: 22.8 County level: 22.6	Null for association between O_3 and TLBW at all distance intervals, for all exposure windows: T1, T2, T3, WP. OR (95% CI) per 1ppb O_3 WP: 3 miles: 1.0061 (0.993, 1.019) 6 miles: 1.0010 (0.992, 1.010) 9 miles: 1.0011 (0.993, 1.009) 12 miles: 1.0006 (0.994, 1.008) County Level: 1.0037 (0.997, 1.010) T1: 3 miles: 1.0032 0.993, 1.014) 6 miles: 0.9980 (0.990, 1.006) 9 miles: 1.0006 (0.994, 1.007) 12 miles: 1.0013 (0.996, 1.007) County Level: 1.0011 (0.997, 1.005) T2: 3 miles: 0.9985 (0.988, 1.009) 6 miles: 1.0007 (0.993, 1.009) 9 miles: 0.9991 (0.993, 1.005) 12 miles: 1.0000 (0.994, 1.006)

Author	Region/ Location/ Date	Adjustment for smoking	Population size	Study design	Exposure assessment	Mean O_3 (ppb unless otherwise stated)	Results (all ORs are from adjusted models)
							County Level: 1.0013 (0.997, 1.005) T3: 3 miles: 1.0019 (0.991, 1.012) 6 miles: 1.0012 (0.994, 1.008) 9 miles: 1.0011 (0.996, 1.007) 12 miles: 1.0000 (0.995, 1.005) County Level: 1.0021 (0.998, 1.006)
Guo P et al 2020	China 21 Cities, 123 districts of the Guangdong province, China 1 January 2014 - 31 December 2017	Y	2,567,457	Retrospective population-based cohort	Environmental monitoring with inverse distance weighting interpolation	WP: 54.6 μg/m^3 T1: 54.25 μg/m^3 T2: 54.34 μg/m^3 T3: 55.06 μg/m^3	OR (95% CI) per IQR increase in O_3. (IQR increments: WP: 12.15 μg/m^3, T1: 15.82 μg/m^3, T2: 14.79 μg/m^3, T3: 14.36 μg/m^3) WP: 1.05 (0.99, 1.10) T1: 1.02 (0.98, 1.06) T2: 1.04 (0.98, 1.11) T3: 0.99 (0.94, 1.05) Adjustment for other pollutants (WP): O_3 adjusted for CO: 1.05 (1.00, 1.11) O_3 adjusted for NO_2: 1.04 (0.98, 1.10) O_3 adjusted for PM_{10}: 1.04 (0.98, 1.09) O_3 adjusted for $PM_{2.5}$: 1.04 (0.98, 1.10)

Author	Region/ Location/ Date	Adjustment for smoking	Population size	Study design	Exposure assessment	Mean O_3 (ppb unless otherwise stated)	Results (all ORs are from adjusted models)
							O_3 adjusted for SO_2: 1.03 (0.98, 1.09)
Le et al 2012	North America Detroit 1990-2001	Y	139,951 (for term births analyses)	Retrospective cohort using routine birth records	Environmental monitoring 3 fixed site monitors	Mean 8hours average from monitor: 41.99 Month average: 69.78 Mean trimester average: 76.56	OR (95% CI) for TLBW per quartile of O_3 exposure – for all seasons. T1: Q2: 1.00 (0.88-1.15) Q3: 0.91 (0.79-1.04) Q4: 0.92 (0.80-1.05) T2: Q2: 0.97 (0.84-1.11) Q3: 0.96 (0.84-1.11) Q4: 0.90 (0.78-1.03) T3: Q2: 0.94 (0.83-1.05) Q3: 1.02 (0.91-1.14) Q4: 0.93 (0.83-1.05) Also, presented for high season only, but no statistically significant results.

4. Stillbirth

a. All Pollutants ($PM_{2.5}$, PM_{10}, NO_2 and O_3) and Stillbirth

Table A 56. Summary results and study characteristics for selected key studies of interest not included in core meta-analyses for All Pollutants ($PM_{2.5}$, PM_{10}, NO_2 and O_3) and Stillbirth

Author	Region/ Location/ Date	Adjustment for smoking	Population size	Study design	Exposure	Mean O_3 (ppb unless otherwise stated)	Results (all ORs are from adjusted models)
Xue et al 2019	Africa 33 Countries 1 January 1998 – 31 December 2016	Y (only in a sensitivity analyses)	67,566	Retrospective cohort	Satellite remote sensing technology	Case: 25.59 Control: 26.02	$PM_{2.5}$ per 10 µg/m^3 WP: 1.09 (1.05, 1.14)
Yang et al 2018	China Wuhan 10 June 2011 – 9 June 2013	N	95,354	Prospective birth cohort	Monitoring stations	$PM_{2.5}$: 70.6 PM_{10}: 101.9 NO_2: 58.8 O_3: 74.9	$PM_{2.5}$ per 10 µg/m^3 WP: 1.13 (1.05 1.21) T1: 1.06(1.01, 1.12) T2: 1.07 (1.03, 1.11) T3: 1.02 (0.98, 1.07) PM_{10} per 20 µg/m^3 WP: 1.18 (0.91, 1.51) T1: 0.96 (0.84, 1.10) T2: 1.09 (0.96, 1.23) T3: 1.20 (1.01, 1.44)

Author	Region/ Location/ Date	Adjustment for smoking	Population size	Study design	Exposure	Mean O_3 (ppb unless otherwise stated)	Results (all ORs are from adjusted models)
							NO_2 per 10 µg/m^3 WP: 1.00 (0.88, 1.13) T1: 0.95 (0.87, 1.02) T2: 1.00 (0.93, 1.08) T3: 1.02 (0.92, 1.13) O_3 per 10 µg/m^3 WP: 0.99(0.95, 1.03) T1: 0.99 (0.95, 1.03) T2: 1.01 (0.98, 1.04) T3: 0.96 (0.92, 1.000)
Zang et al 2019	China Yancheng City 1 January 2015 – 31 December 2017	N	59,868	Prospective birth cohort	Monitoring stations	$PM_{2.5}$: 44.3 PM_{10}: 77.6 NO_2: 24.2 O_3: 78.2	$PM_{2.5}$ per 10 µg/m^3 WP:1.14 (1.08, 1.20) T1:1.14 (1.06, 1.22) T2: 1.11 (1.06, 1.16) T3: 1.15 (1.09, 1.21) PM_{10} per 10µg/m^3 WP: 1.04 (0.96, 1.12) T1: 1.09 (1.04, 1.14) T2: 1.05 (0.94, 1.15) T3: 1.02 (0.97, 1.17)

Author	Region/ Location/ Date	Adjustment for smoking	Population size	Study design	Exposure	Mean O_3 (ppb unless otherwise stated)	Results (all ORs are from adjusted models)
							NO_2 per 10 µg/m^3 WP: 1.01 (0.98, 1.04) T1: 1.03 (0.96, 1.10) T2: 1.01 (0.95, 1.07) T3: 1.03 (0.97, 1.09) O_3 per 10 µg/m^3 WP: 1.03 (0.97, 1.09) T1: 1.05 (1.01, 1.09) T2: 1.01 (0.96, 1.06) T3: 1.04 (1.00, 1.08)
Smith et al 2020	UK London/Greater London 1 January 2006 – 31 December 2010	Y (but using an area-level proxy for smoking)	581,774	Retrospective cohort using routine birth records	Modelled (Dispersion modelling to 20x20 m grid resolution)	T1: $PM_{2.5}$: 14.6 PM_{10}: 23.6 NO_2: 41.1 O_3: 31.1 T2: $PM_{2.5}$: 14.5 PM_{10}: 23.2 NO_2: 40.6 O_3: 32.0	$PM_{2.5}$ per IQR (3.2 µg/m^3) T1: 0.88 (0.82, 0.93) $PM_{2.5}$ per IQR (3.6 µg/m^3) T2: 0.91 (0.85, 0.99) PM_{10} per IQR (4.0 µg/m^3) T1: 0.84 (0.79, 0.88) PM_{10} per IQR (4.5 µg/m^3) T2: 0.91 (0.86. 0.97) NO_2 per IQR (14.0 µg/m^3) T1: 0.93 (0.87, 1.00) NO_2 per IQR (14.4 µg/m^3) T2: 0.92 (0.86, 0.99) O_3 per IQR (17.6 µg/m^3) T1: 1.17 (1.07, 1.27)

Author	Region/ Location/ Date	Adjustment for smoking	Population size	Study design	Exposure	Mean O_3 (ppb unless otherwise stated)	Results (all ORs are from adjusted models)
							O_3 per IQR (14.4 µg/m^3) T2: 1.20 (1.09, 1.32)
Mendola et al 2017	USA 12 Clinical Sites 2002-2008	Y	228,438 deliveries	Retrospective cohort	Modelling: Community Multiscale Air Quality model	WP: $PM_{2.5}$: 11.9 PM_{10}: 22.1 O_3: 29.0 ppb T1: $PM_{2.5}$: 11.9 PM_{10}: 22.0 O_3: 28.7 ppb	$PM_{2.5}$ per IQR (4.7 µg/m^3) WP: 1.08 (0.75, 1.55) $PM_{2.5}$ per IQR (5.3 µg/m^3) T1: 0.91 (0.77, 1.07) PM_{10} per IQR (4.4 µg/m^3) WP: 1.04 (0.86, 1.25) PM_{10} per IQR (6.3 µg/m^3) T1: 0.97 (0.85, 1.12) NO_2 per IQR (24.2 ppb) WP: 0.96 (0.62, 1.49) NO_2 per IQR (30.2 ppb) T1: 0.85 (0.66, 1.09) O_3 per IQR (7.8 ppb) WP: 1.39 (1.05, 1.84) O_3 per IQR (12.3 ppb) T1: 1.18 (1.00, 1.39)
Ebisu et al 2018	USA California 1 January 2002 – 31 December 2009	N	32,262	Nested case control	Monitoring stations	$PM_{2.5}$ WP: 17.8	$PM_{2.5}$ per IQR (7.23 µg/m^3) WP: 1.06 (1.01, 1.11)

Author	Region/ Location/ Date	Adjustment for smoking	Population size	Study design	Exposure	Mean O_3 (ppb unless otherwise stated)	Results (all ORs are from adjusted models)
Rammah et al 2019b	USA Texas, Harris County 2008-2013	Y	358,366	Retrospective birth cohort	Monitoring with inverse distance weighting	O_3 WP: range 26.33 – 52.50 ppb	O_3 per IQR (3.6 ppb) WP: 1.09 (1.01, 1.18)

Appendix 13. Summary tables of studies investigating underlying mechanisms

1. Oxidative stress – Epidemiological studies with a mechanistic component

Reference – Cohort	Population location & size	Pollutant – assessment & concentration	Biomarker (sample)	Birth outcome assessed	Main findings
Saenen et al 2016 ENVIRONAGE	Genk, Belgium (n=330)	$PM_{2.5}$ (Mean μg/m^3) WP: 16.1 NO_2 (Mean μg/m^3) WP: 20.5 Black carbon (BC) (Mean μg/m^3) WP: 0.97 Modelled	3-nitrotyrosine (3-NTp) (Placenta)	Birth weight Birth length	Positive & significant correlation between placental 3-NTp & WP exposure to $PM_{2.5}$ & BC but not NO_2 For an IQR increment in $PM_{2.5}$ & BC exposure over WP, placental 3-NTp increased by 35.0% (95% CI: 13.9, 60.0; P < 0.0006). & 13.9% (95% CI: − 0.21, 29.9; P = 0.05) respectively Birth outcome results No significant associations between placental 3-NT & birth weight or birth length
Grevendonk et al 2016 ENVIRONAGE	Genk, Belgium (n=224-293)	$PM_{2.5}$ (Medium μg/m^3) WP 21.4 PM_{10} (Medium μg/m^3) WP: 16.6 Modelled	Mitochondrial 8-hydroxy-2′ -deoxyguanosine (8-OHdG) (Maternal blood at delivery (n=224) & cord blood (293))	Birth weight SGA	PM_{10} exposure during WP was positively associated with mitochondrial 8-OHdG in maternal blood (18.3, 95 % CI: 5.6 to 33.4 %, p = 0.004 for an IQR increment in PM10 exposure) $PM_{2.5}$ exposure during WP was positively associated with mitochondrial 8-OHdG in maternal blood (13.9, 95 % CI: 0.4 to 29.4 %, p = 0.04 for an IQR increment in $PM_{2.5}$ exposure) No significant association between cord blood 8-OHdG levels & PM_{10} or $PM_{2.5}$ exposure during WP

Reference – Cohort	Population location & size	Pollutant – assessment & concentration	Biomarker (sample)	Birth outcome assessed	Main findings
					PM_{10} (but not $PM_{2.5}$) exposure during T1 & T2 was positively associated with mitochondrial 8-OHdG in cord blood (23.0, 95 % CI: 5.9 to 42.8 %, p = 0.007 & 16.6%, 95% CI: 1.8 to 33.5%, p=0.03 respectively for an IQR increment in PM10 exposure) Birth outcome results Neither PM exposure nor 8-OHdG levels in cord & maternal blood were associated with birth weight or SGA
Ambroz et al 2016	Czech Republic (n=170) Ceske Budejovice (CB): a locality with a clean air Karvina: a locality with high air pollution	$PM_{2.5}$ (Mean; µg/m^3) CB summer: 9.45 CB winter: 27.96 Karvina summer: 10.41 Karvina winter: 53.67 BaP (Mean; ng/m^3) CB summer: 0.16 CB winter: 1.45 Karvina summer: 1.16 Karvina winter: 4.20 Environmental monitoring	8-OHdG (Maternal & newborn urine) 5-F2t-isoprostane (15-F2t-IsoP) (Maternal & newborn blood &)	None	8-oxodG in newborns were elevated in the Karvina samples compared to the CB samples in winter season only Overall: $PM_{2.5}$ associated with maternal 8-oxodG levels No association of $PM_{2.5}$ with newborn 8-oxodG No association of $PM_{2.5}$ with maternal or fetal 15-F2t-IsoP No association of BaP with either 8-oxodG or 15-F2t-IsoP in newborns or mothers In separate Karvina analysis: $PM_{2.5}$ concentrations associated with increased 8-oxodG excretion & blood 15-F2t-IsoP in newborns $PM_{2.5}$ concentrations associated with increased 8-oxodG excretion in mothers BaP concentrations associated with increased blood 15-F2t-IsoP in newborns (but not 8-OhdG)
Rossner et al 2011	Czech Republic (n=891)	$PM_{2.5}$ (pregnancy monthly means 28.5-32.1 µg/m^3) B[a]P	8-OHdG (Placenta)	IUGR TLBW	No difference in 8-oxodG between the districts, despite higher levels of air pollutants in the Teplice region Using univariate analysis: 8-OxodG levels in placental DNA was not affected by Ba $PM_{2.5}$ exposure during the second month of pregnancy was significantly associated with 8-oxodG levels in placental DNA.

Reference – Cohort	Population location & size	Pollutant – assessment & concentration	Biomarker (sample)	Birth outcome assessed	Main findings
	Prachatice: low levels of air pollution Teplice: at the time one of the most polluted districts of Czech Republic	Cannot access suppl material file for BaP concentrations Environmental monitoring			Birth outcome results Borderline statistically significant elevated 8-OHdG levels found in IUGR group versus non-IUGR group (p=0.06) Slightly elevated levels of 8-oxodG found in TLBW group versus non-TLBW group ($p<0.05$) $PM_{2.5}$ concentrations not associated with IUGR but in 1st month of pregnancy was associated with risk of TLBW BaP exposure affected neither IUGR nor LBW.
Maciel-Ruiz et al 2019	Mexico City, Mexico (n=106)	$PM_{2.5}$ PM_{10} O_3 Values not given but statement made: 'majority of pregnant women were exposed to higher PM levels than the Mexican Guideline values for the annual mean of $PM_{2.5}$ (12 μg/m^3) & PM_{10} & (40 μg/m^3) Environmental monitoring	8-iso-prostaglandin F2α (8-iso-PGF2α) (Maternal [at labour] & cord blood)		Maternal & cord blood 8-iso-PGF2α correlated significantly ($\rho = 0.265$, 95% CI (−0.039, 0.524), $P = 0.039$) Cord blood levels were significantly higher ($P < 0.0001$) Significant association between maternal 8-iso-PGF2α & O_3 only during last month of pregnancy No association between the levels of cord serum 8-iso-PGF2α & O_3 concentrations No association with $PM_{2.5}$ or PM_{10} levels Birth outcome results No clear association between PM &/or adduct levels & bodyweight/length.

Reference – Cohort	Population location & size	Pollutant – assessment & concentration	Biomarker (sample)	Birth outcome assessed	Main findings
Janssen et al 2012 ENVIRONAGE	Genk, Belgium (n=174-176)	PM_{10} (Mean μg/m^3) WP: 22.7 T1: 21.5 T2: 22.3 T3: 24.4 Last month: 25.6 Modelled Residential distance to major road; mean: 207 m	Mitochondrial DNA (mtDNA) content (Placenta & cord blood)	None	In a multivariate-adjusted analysis 10-μg m^{-3} increase in PM_{10} exposure during the last month & T3 of pregnancy associated with decreases of 16.1% &17.4% in placental mtDNA respectively Each doubling in residential distance to major roads associated with 4% increase in placental mtDNA content (95% CI: 0.4, 7.8%, p = 0.03) No association between cord blood mtDNA content & distance to major roads or PM_{10} concentrations
Rosa et al 2017 PROGRESS Study	Mexico (n=456)	$PM_{2.5}$ (Mean μg/m^3) WP: 23.1 Modelled	mtDNA (Cord blood)	None	Significant associations between increased concentration of $PM_{2.5}$ (during late pregnancy: 35-40 weeks) & decrease in mtDNA in cord blood
Brunst et al 2018 PRISM Study	Boston, USA (n=162)	$PM_{2.5}$ (Mean μg/m^3) WP: 7.95/7.96 Modelled	mtDNA copy number (cn) (Placenta & cord blood)	None	Increased $PM_{2.5}$ concentrations across WP was associated with decreased mtDNAcn in cord blood (cumulative effect estimate = −0.78; 95% CI −1.41, −0.16) While in the same direction, an effect of $PM_{2.5}$ on placental mtDNAcn did not reach significance ((cumulative effect estimate = −0.07; 95% CI −0.23, 0.04)
Clemente et al 2016 ENVIRONAGE & INMA	Belgium & Spain ENVIRONAGE (n=550) INMA (n=376)	NO_2 (ENVIRONAGE: 21.1 μg/m^3; INMA: 25.5 μg/m^3; ENVIRONAGE + INMA: 22.7 μg/m^3) Modelled	mtDNA content (Placenta)	Birth weight	10 μg/m^3 increment in average NO_2 exposure during pregnancy was associated with a 4.9% decrease in placental mtDNA content Birth outcome results 10 μg/m^3 increment in average NO_2 exposure during pregnancy was associated with a 48g decrease in BW Placental mtDNA content was positively & significantly associated with birth weight in both cohorts & in the pooled analysis

Reference – Cohort	Population location & size	Pollutant – assessment & concentration	Biomarker (sample)	Birth outcome assessed	Main findings
					IQR increase in mtDNA content was significantly associated with a 66g increase in birth weight in boys but there was no significant difference in girls
Lavigne et al 2018	Ontario, Canada (n=196,171)	$PM_{2.5}$ (Mean, $\mu g/m^3$) T1: 9.0 T2: 9.1 T3: 9.0 $PM_{2.5}$: environmental monitoring OP: city level	Oxidative potential (OP) of particulate matter	PTB TLBW Birth weight	Most associations of $PM_{2.5}$ and GSH- and AA-related oxidative burden with TLBW and PTB, were not statistically significant However, an IQR increase in $PM_{2.5}$ GSH-related oxidative burden was more strongly associated with TLBW than an IQR increase in $PM_{2.5}$ mass concentration.
Suh et al 2008	Seoul, Korea (n=235)	PM_{10} (Medium $\mu g/m^3$) T1: 61.12 T2: 59.05 T3: 54.64 Environmental monitoring	+/- GSTM1 null genotype (Maternal blood)	PTB	Birth outcome results Risk of PTB conferred by exposure to high levels of PM_{10} (that is,≥75th percentile) was significant in T3 (OR = 2.33, 95% CI = 1.33–4.80, p = 0.022), but not in T1 or T2 (p > 0.05). Risk of PTB was associated with the GSTM1 null genotype but not GSTM1 positive genotype Exposure to high levels (≥75th percentile) of PM10 during T3 in the presence of the GSTM1 null genotype is significantly associated with risk of PTB
Zhao et al 2021	Taiyuan, China (n=436)	Chinese Air Quality Index (AQI)	Single nucleotide polymorphisms in: Glutathione peroxidase 4 (GPx4) Glutaredoxin (GLRX) endothelial growth factor A (EGF)	PTB	Birth outcome results Higher levels of maternal AQI exposure associated with a higher risk of PTB among women who carried polymorphisms in genes for GPX4, GLRX, EGF & IL-1α alpha After adjustment for multiple comparisons, only GPX4-rs376102 and AQI interaction remained statistically significant

Reference – Cohort	Population location & size	Pollutant – assessment & concentration	Biomarker (sample)	Birth outcome assessed	Main findings
			IL-1α (Maternal blood)		

2. Oxidative stress – Animal studies

Reference	Animal model	Exposure conditions (concentration, time)	Biomarker (sample)	Main findings
Veras et al 2012	BalbC mice	Unfiltered São Paulo ambient air (mean $PM_{2.5}$ 32.8 μg/m^3 or filtered air (6.4 μg/m^3) Throughout mating & pregnancy	15-F2t-isoprostane (15-F2t-IsoP) Endothelin A receptor (ET_AR) Endothelin B receptor (ET_BR) (Umbilical vessels)	Greater volumes of regions immunostained for 15-F2t-IsoP, ET_AR & ET_BR.
Weldy et al 2014	C57Bl/6J mice	Diesel exhaust (single cylinder Yanmar engine: ~300 μg/m^3 $PM_{2.5}$, 6 hrs/day, 5 days/week) or filtered air Embryonic day 0.5 to 17.5	3-nitrotyrosine (3-NT) (Placenta)	Elevated 3-NT protein modification, predominantly within perivascular regions in the fetal labyrinth layer

3. Inflammation – Epidemiological studies with a mechanistic component

Reference – Cohort	Population Location & number	Pollutant – assessment & concentration	Biomarker (sample)	Birth outcome assessed	Main findings
Nachman et al 2016	Boston, USA (n=5,059)	$PM_{2.5}$ 5.54 to 29.00 $\mu g/m^3$	Intrauterine inflammation (IUI)	None	Comparing highest with the lowest $PM_{2.5}$ exposure quartiles, multi-adjusted association with IUI was significant for all exposure periods: 3 mo before conception (OR = 1.52; 95% CI: 1.22, 1.89) T1 (OR = 1.93; 95% CI: 1.55, 2.40) T2 (OR = 1.67; 95% CI: 1.35, 2.08) T3 (OR = 1.53; 95% CI: 1.24, 1.9) WP (OR = 1.92; 95% CI: 1.55, 2.37)
Van den Hooven et al 2012b Generation R	Rotterdam, the Netherlands (n=5,067/4,450)	Mean for the prior week to maternal blood draw PM_{10}: 30.6 $\mu g/m^3$ NO_2: 40.3 $\mu g/m^3$ Mean week prior to delivery PM_{10}: 29.6 $\mu g/m^3$ NO_2: 39.5 $\mu g/m^3$ Monitoring data & dispersion l techniques	Elevated high C-reactive protein (hs CRP) concentrations (> 8ng/ml) Maternal blood: median, 13.2 weeks of pregnancy n=5067) Cord blood (n=4450)	None	*Maternal hs-CRP* Compared with the lowest quartile, higher PM_{10} exposure levels for the prior 1 & 2 weeks were associated with elevated maternal hs CRP levels in the first trimester NO_2 exposure not associated with elevated maternal hs CRP levels *Newborn hs-CRP* Higher WP PM_{10} & NO_2 concentrations associated with elevated fetal hs CRP levels at delivery PM_{10} and NO_2 exposure levels for 1, 2 and 4 weeks preceding delivery were not consistently associated with fetal hs CRP levels at delivery
Lee et al 2011	Allegheny County, PA, USA (n=1,696)	0-7 day before blood collection $PM_{2.5}$ 16.4 $\mu g/m^3$ PM_{10} 26.1 $\mu g/m^3$	Elevated hs CRP concentrations (≥ 8ng/ml) Maternal blood (before 11 weeks of pregnancy)	None	Maternal hs-CRP Positive associations between $PM_{2.5}$, PM_{10} & O_3 & elevated concentrations of hsCRP during early pregnancy No associations for NO_2

Reference – Cohort	Population Location & number	Pollutant – assessment & concentration	Biomarker (sample)	Birth outcome assessed	Main findings
Prenatal Exposures and Preeclampsia Prevention Study		NO_2 18.8 ppb O_3 29.9 ppb			
Friedman et al 2021 Healthy Start Study	Denver, Colorado (n=515)	Median full pregnancy average: $PM_{2.5}$ 7.51 μg/m^3 O_3 44.0 ppb	CRP Interleukin-6 (IL-6) Tumour necrosis factor α (TNFα) (Maternal blood during mid-pregnancy and cord blood)	Birth weight Gestational age	*Maternal CRP* $PM_{2.5}$ & O_3: No significant associations *Maternal IL-6* $PM_{2.5}$: Positive associations with T1& prior 30-day concentrations O_3: Negative associations with prior 30- and 7-day concentrations & positive associations with T1 concentrations *Maternal TNFα* $PM_{2.5}$: Positive associations with T1, T2 & prior 30-day concentrations O_3: Positive associations with T1 & T2 concentrations *Cord blood CRP, IL-6, TNFα* Patterns inconsistent for associations between $PM_{2.5}$ and O_3 Maternal CRP positively associated with infant gestational age at birth (β; 1.30, 95% CI; 0.40–2.21); no associations between maternal CRP, IL-6 & TNFα & birth weight.

4. Inflammation – Animal studies

Reference	Animal model	Exposure conditions (concentration, time)	Main findings
Fujimoto et al 2005	BalbC mice	300- 3000 μg/m^3 diesel exhaust particles (Isuzu Motor Co) From 2 until 13 days pc; 12 h/day, 7 days/week	Elevation in mRNA levels of IL-1β, IL-2α, IL-2β, IL-5, IL-6 & TNF-α & other inflammatory markers in placenta of mice at day 14 post coitum
Weldy et al 2014	C57Bl/6J mice	Diesel exhaust (single cylinder Yanmar engine; ~300 μg/m^3 $PM_{2.5}$ or filtered air from 6 hrs/day, 5 days/week from embryonic day 0.5 to 17.5	Increased CD45+ cells within the decidua layer in placentas
de Melo et al 2015	Wistar rats	Filtered air or concentrated ambient particles (CAPs) (600 μg/m^3 $PM_{2.5}$;) 5 x/week during the 3 weeks before pregnancy &/or once per day (30-90 min), 7 days a week during pregnancy, starting on the 6th day of gestation) CAPs collected from urban areas of São Paolo where vehicular emissions are the most relevant source $PM_{2.5}$	IL-4 content elevated in the fetal portion of the placenta (but not in serum) No differences observed for placental or serum IL-1β, IL-4, IL-6, IL-10, INF-γ, TNF-α, TLR4

5. DNA damage: Epidemiological studies with a mechanistic component

Reference – Cohort	Population location & number	Pollutant – assessment & concentration	Adduct (Sample)	Birth outcome assessed	Main findings
Topinka et al 1997	Prachatice & Teplice, Czech Republic (n=98)	Comparison of low (rural Prachatice) & high (industrial Teplice) pollution districts	PAH-adducts Pooled: chrysene, benzo[a]anthracene, Benzo[b]fluoranthene (B[b]F), benzo[k]fluoranthene, benzo[a]pyrene (BaP), dibenz[a,h]anthracene (DBA), indeno[1,2,3-c,d]pyrene & benzo[g,h,i]perylene) (Placenta) +/- GSTM1 null genotype	None	DNA adduct levels in the placentas of mothers from both districts were found to be very low Significant difference in placental DNA adduct levels between 2 districts was found for the winter sampling period only (1.49 vs 0.96 adducts per 10^8 nucleotides; p=0.023) Higher DNA adduct levels detected in the group of mothers with GSTM1-null genotype (p=0.009) DNA adduct levels in placentas of mothers with GSTM1-null genotype living in the polluted district of Teplice were higher than those in Prachatice (p =0.05)
Perera et al 1999	Krakow & Limanowa, Poland (n=160)	PM_{10} Environmental monitoring Plus, comparison of 2 districts with high (industrial Krakow) versus low (agricultural Limanowa) concentrations of air pollution	PAH adducts (Placenta, maternal & cord blood) +/- GSTM1 null genotype	Birth weight Birth length Head circumference	No difference in PAH-DNA adduct levels in placenta or maternal/cord blood from the 2 districts PAH-DNA adduct non-significantly higher in cord blood than maternal blood Adduct level in placenta higher than maternal blood (p<0.01) In Krakow: PM_{10} significantly associated with the amount of PAH-DNA adducts in maternal & cord blood (p≤0.05) No association between maternal adduct concentration & GSTM1 genotype *Birth outcome results*

Reference – Cohort	Population location & number	Pollutant – assessment & concentration	Adduct (Sample)	Birth outcome assessed	Main findings
					Newborns with elevated PAH-DNA adducts had significantly decreased birth weight, birth length & head circumference compared to those with lower adducts
Tang et al 2006	Tongliang coal-fired plant, China (n=150)	Distance from power plant	BaP adducts (Maternal (at delivery) & cord blood)	Birth weight Birth length Head circumference	Mean cord blood adduct level non-significantly higher than maternal adduct level Cord adduct level was negatively but not significantly correlated with distance from power plant (p=0.13) *Birth outcome results* High PAH–DNA adduct levels (above the median of detectable adduct level) associated with decreased birth head circumference (p = 0.057) Longer distance from the power plant was significantly associated with greater birth length (p = 0.03)
Tang et al 2014	Tongliang coal-fired plant, China (n=58)	Cohort 1: recruited prior to power plant shutdown (2002) Cohort 2: recruited following shutdown (2005)	B[a]P adducts (Cord blood)	Head circumference	Decreased ambient BaP from 13.14 ng/m^3 in 2002 to 9.14 ng/m^3 in 2005 Reduced PAH-DNA adducts in the 2005 cohort (0.33 ± 0.140 adducts/10^{-8} in 2002 vs. 0.20 ± 0.08 in 2005 p ≤0.001) *Birth outcome results* Cohort 2 versus cohort 1 Greater head circumference (p = 0.001) Reduced levels of cord blood PAH-DNA adducts in (p < 0.001) Reduced levels of ambient PAHs (p = 0.01)
Suter et al 2019	Harris County, Texas, USA	Living near	PAH adducts Plus, concentrations of: BaP	PTB Birth weight	Bulky PAH-DNA adducts were more prevalent in placental tissue from individuals residing near Superfund sites

Reference – Cohort	Population location & number	Pollutant – assessment & concentration	Adduct (Sample)	Birth outcome assessed	Main findings
	(n=104)	Superfund sites versus those living in non-Superfund site areas	B[b]F DBA (Placenta)		Concentrations of DBA in the placenta significantly higher than those of BaP & BbF *Birth outcome results* Concentrations of B[a]P, B[b]F & DBA higher in placentae from preterm deliveries compared with term deliveries in women living near Superfund sites (not in those living in non-Superfund site areas) DBA inversely correlated with gestational age at delivery & birth weight
Maciel-Ruiz et al 2019	Mexico City, Mexico (n=106)	$PM_{2.5}$ PM_{10} Values not given but statement made: 'majority of pregnant women were exposed to higher PM levels than the Mexican Guideline values for the annual mean of $PM_{2.5}$ (12 μg/m3) & PM_{10} & (40 $\mu g/m^3$) Environmental monitoring	PAH-DNA adducts *BaP diol-epoxide modified DNA standard used.* (Maternal (at delivery) & cord blood) +/- GSTM1 null genotype	Birth weight Birth length	Adducts significantly (≈ 20%) higher in cord blood compared to maternal blood Levels of maternal DNA adducts positively associated with cord blood levels Adducts significantly associated with $PM_{2.5}$ and PM_{10} levels, mainly during T3 Associations between $PM_{2.5}$ or PM_{10} & adduct levels always higher in cord blood Association between DNA adducts in cord blood & $PM_{2.5}$ showed a linear dose-response trend No association between bulky DNA adducts in maternal blood & GSTM1 genotype *Birth outcome results* No clear association between PM &/or adduct levels & bodyweight/length.

6. DNA methylation: Epidemiological studies with a mechanistic component

Reference – Cohort	Population location & number	Pollutant – assessment & concentration	Methylated assessed (Sample)	Birth outcome assessed	Main findings
Janssen et al 2013 ENVIRONAGE	Genk, Belgium (n=240)	$PM_{2.5}$ (Mean $\mu g/m^3$) WP: 17.4 T1: 16.7 T2: 17.4 T3: 18.2 Modelled	Global DNA methylation (Placenta)	None	Placental global DNA methylation: Inversely associated with WP $PM_{2.5}$ concentrations Relatively decreased by 2.19% (95% CI: -3.65, -0.73%, p = 0.004) for each 5 $\mu g/m^3$ $PM_{2.5}$ increase In a multi-lag model only $PM_{2.5}$ concentrations during T1 was significantly associated with lower global DNA methylation (– 2.13% per 5 μ g/m^3 increase, 95% CI: -3.71, -0.54%, p = 0.009) Analysing shorter time windows within T1: Lower placental DNA methylation at birth during all implantation stages but exposure during the implantation range (6-21d) was strongest associated (– 1.08% per 5 μ g/m^3 increase, 95 CI: -1.80, -0.36%, p = 0.004)
Janssen et al 2015 ENVIRONAGE	Genk, Belgium (n=381)	$PM_{2.5}$ (Mean $\mu g/m^3$) WP: 16.7 T1: 16.0 T2: 16.9 T3: 17.3 Modelled	Mitochondrial (mt)DNA methylation in displacement D-loop control region & 12S rRNA (Placenta)	LBW SGA	IQR increment in $PM_{2.5}$ concentrations over WP was positively associated with mtDNA methylation (MT-RNR1: +0.91%, P = 0.01 and D-loop: +0.21%, P = 0.05) Inverse association between placental mtDNA methylation & content Neither methylation levels of the MT-RNR1 nor the D-loop region were associated with LBW or small for gestational age

Reference – Cohort	Population location & number	Pollutant – assessment & concentration	Methylated assessed (Sample)	Birth outcome assessed	Main findings
Vos et al 2020 ENVIRONAGE	Genk, Belgium (n=40)	High ($PM_{2.5}$: 16.0 ± 1.4 μg/m^3, black carbon: 1.8 ± 0.3 μg/m^3) versus low ($PM_{2.5}$: 10.6 ±1.7 μg/m^3, black carbon: 0.9 ±0.1 μg/m^3) Modelled	mtDNA methylation in: D-loop & LDLR2 regions (Placenta)	Birth weight	Absolute D-loop methylation levels higher for mothers highly exposed to air pollutants (+0.47%, 95% CI: 0.20% to 0.73%) D-loop methylation levels correlated with placental mtDNA content (r = −0.40, p = 0.002) Birth outcome results D-loop methylation levels associated with birth weight (−106.98 g, 95% CI: −209.60 g to −4.36 g for an IQR increase in D-loop methylation)
Saenen et al 2017 ENVIRONAGE	Genk, Belgium (n=361)	$PM_{2.5}$ (Mean μg/m^3) T1: 15.7 T2: 15.5 T3: 17.2 Modelled	DNA methylation of leptin (LEP) promoter (Placenta)	None	Placental LEP methylation was 1.4% lower (95% CI: –2.7, –0.19%, p = 0.02) with an IQR increment in T2 $PM_{2.5}$ concentrations (7.5 μg/m^3) but not T1 or T3 Doubling in placental 3-NTp content at birth was associated with a significantly lower overall methylation level of the LEP region evaluated (–0.43%; 95% CI: –0.85, –0.02%, p = 0.04)
Liu X et al 2019	Guangzhou, China	$PM_{2.5}$ (Mean μg/m^3) WP: 32.07 PM_{10} (Mean μg/m^3) WP: 48.67 PM_1 (Mean μg/m^3) WP: 27.93 Environmental monitoring	DNA methylation of long interspersed nucleotide element (LINE)-1 (Maternal & cord blood)	PTB	Significant negative associations of PM_{10}, $PM_{2.5}$ & PM_1 concentrations with cord blood LINE-1 methylation levels Negative association between PM_1 concentration & maternal LINE-1 methylation level Birth outcome results Risk of PTB positively associated with $PM_{2.5}$ and PM_1 during gestational weeks 12-20 Risk of PTB negatively associated with maternal blood LINE-1 methylation levels & positively correlated with cord blood LINE-1 methylation levels but findings not significant

Reference – Cohort	Population location & number	Pollutant – assessment & concentration	Methylated assessed (Sample)	Birth outcome assessed	Main findings
Cai et al 2017	Shanghai, China (n=181)	PM_{10} (Median $\mu g/m^3$) WP: 63.51 T1: 62.74 T2: 68.07 T3: 65.29	DNA methylation of LINE-1, HSD11B2 & NR3C1 (Placenta)	Fetal growth restriction	Placental LINE-1 DNA methylation reversely associated with T1 PM_{10} concentrations Placental HSD11B2 DNA methylation was associated with T1 & T2 PM_{10} Birth outcome results Above associations were much more evident in FGR newborns
Kingsley et al 2016 Rhode Island Child Health Study	Rhode Island, USA (n=471)	Residential proximity to roadway (defined as living ≤ 150 m from a primary highway or primary road or ≤ 50 m from a secondary road)	DNA methylation of repetitive elements (LINE-1 and AluYb8) (Placenta)	Birth weight SGA	Living close to major roadway associated with: 175.9 g (95% CI: -319.4, -32.5; p=0.016) lower birth weight Lower mean placental LINE-1 methylation levels Differential methylation of 7 CpG sites (in n=215 sub-cohort) Additional adjustment for placental methylation did not attenuate the association between roadway proximity & birth weight
Gruzieva et al 2017 MeDALL (pooled from INMA Spain, EDEN France, BAMSE Sweden; PIAMA the Netherlands) Generation R Study (the Netherlands) CHS (Children's Health Study; USA)	Meta-analysis of 4 N American & European studies (n=1,508)	NO_2 (Percentiles 25th, 50th, 75th (minimum - maximum) $\mu g/m^3$: Medall pooled, EDEN & INMA: 19.0, 37.4, 47.1 (9.0-89.9) Generation R: 36.0, 38.7, 41.8 (28.6-55.9) CHS: 23.0, 32.1, 41.8 (7.5-51.0)	Genome-wide DNA methylation plus DNA methylation at CpG sites of antioxidant defence genes (catalase (CAT) & thyroid peroxidase (TPO)) (Cord blood)	None	Associations (positive & negative) between NO_2 concentration during pregnancy & epigenome-wide DNA methylation for: 3 CpG sites in genes involved in mitochondrial function (cg12283362 (LONP1), cg24172570 (3.8 kbp upstream of HIBADH), cg08973675 (SLC25A28)) Link between prenatal NO_2 exposure and methylation of CpG loci in CAT & TPO

Reference – Cohort	Population location & number	Pollutant – assessment & concentration	Methylated assessed (Sample)	Birth outcome assessed	Main findings
MoBa (the Mother and Child Cohort Study; Norway)		MoBa: 7.5, 10.3, 12.9 (0.01-27.6)) Modelled & environmental monitoring			
Ladd-Acosta et al 2019 EARLI	Philadelphia, Baltimore, San Francisco Bay Area, and Sacramento, USA (n=175)	NO_2 (24-h average during pregnancy; ppb) 12.27-12.25 O_3 (Average daily 8-h max during pregnancy; ppb) 37.42-37.72 Environmental monitoring	Global DNA methylation (Placenta (n=133) & cord blood (n=175))	None	Increased O_3 concentrations associated with significant decrease in DNA methylation in cord blood at open sea regions (p=0.0016) & in placenta at shelf regions of the genome (p=0.00028) Suggestive (p=0.003) decrease in placental (but not cord blood) global DNA methylation in association with NO_2 concentrations Differentially methylated regions identified included those located at genes involved in immune & inflammatory responses
He et al 2018	Zhengzhou, China (n=527)	PM_{10} (Mean $\mu g/m^3$) WP: 106.66 T1: 104.88 T2: 107.57 T3: 107.27 NO_2 (Mean $\mu g/m^3$) WP: 44.16 T1: 44.71 T2: 44.65 T3: 43.10	DNA methylation of H19 in promoter region & DMR (Maternal & cord blood)	Birth weight Birth length	PM_{10} & NO_2 concentrations during WP associated with lower H19 DMR methylation & higher H19 methylation in cord blood respectively Birth outcome results Statistically significant associations between NO_2 & PM_{10} exposures on birth weight & length but these varied by different trimesters & were sometimes positively associated & sometimes negatively associated No mediating effects of H19/DMR methylation in maternal or cord blood were found in the relationship between air pollutants exposure during pregnancy and birth sizes

Reference – Cohort	Population location & number	Pollutant – assessment & concentration	Methylated assessed (Sample)	Birth outcome assessed	Main findings
		Environmental monitoring			
Maghbooli et al 2018 A nested case-control of a birth cohort study	Tehran, Iran n=92	Comparison of polluted versus non-polluted regions of Tehran WP mean µg/m^3 Polluted region $PM_{2.5}$: 37.12 PM_{10}: 91.45 Non-polluted region $PM_{2.5}$: 25.18 PM_{10}: 70.43 Environmental monitoring	Global DNA methylation (Placenta)	Gestational age Birth weight	No significant difference was observed in placental global DNA methylation levels between mothers who lived in polluted & non-polluted regions Positive correlations between PM_{10} and $PM_{2.5}$ & increased global methylation for T1 No significant association between birth outcomes &: Polluted versus non-polluted regions $PM_{2.5}$ & PM_{10} concentrations Placental global DNA methylation levels
Herbstman et al 2012 CCCEH	NYC (n=164)	Total PAH Mean: 2.66 ng/m^3 Personal monitoring	Global DNA methylation (Cord blood) Plus, B[a]P–DNA adducts	None	PAH exposure associated with lower global methylation in cord blood (p=0.05) Associations similar for BaP and pyrene Global methylation levels were positively associated with the presence of detectable adducts in cord blood (p = 0.01) NB No association between PAH exposure and BaP adducts

Abbreviations. BaP: benzo[*a*]pyrene; BO: birth outcome; BW: birth weight; CAT: catalase; CB: carbon black; CI: confidence interval; DMR: differentially methylated regions; NO_2: nitrogen dioxide; PAH: polycyclic aromatic hydrocarbons; $PM_{2.5}$: particular matter < 2.5 µm in diameter; PM_{10}: particulate matter < 10 µm in diameter; T: trimester; TPO: thyroid peroxidase

7. miRNA expression: Epidemiological studies with a mechanistic component

Reference – Cohort	Population location & number	Pollutant – assessment & concentration	Candidate miRNA (Sample)	Birth outcome assessed	Main findings
Tsamou et al 2018 ENVIRONAGE	Genk, Belgium (n=210)	Mean (μg/m^3) $PM_{2.5}$ T1: 15.99 T2: 16.38 T3: 16.74 NO_2 T: 19.97 T2: 20.69 T3: 20.91 Modelled	miR-16 miR-20a miR-21 miR-34a miR-146a miR-222 (Placenta)	None	Placental expression of miR-21, miR-146a & miR-222 inversely associated with $PM_{2.5}$ exposure during T2 Placental expression of miR-20a & miR-21 was positively associated with T1 concentration Tumour suppressor phosphatase & tensin homolog (PTEN) identified as a common target of the miRNAs significantly associated with $PM_{2.5}$ concentrations Placental PTEN expression was strongly & positively associated with T3 $PM_{2.5}$ concentrations

8. DNA methylation – Animal studies

Reference	Animal model	Exposure conditions (concentration, time)	Main findings
Dang et al 2018	Sprague-Dawley rats	Xi'an $PM_{2.5}$ High: >150 μg/m^3 Low: 75 -150 μg/m^3 Clean: ≤ 75 μg/m^3 Throughout pregnancy	Compared with Clean group: Birth weight & endometria vascular diameter maximum in Low & High groups significantly lower Methylation levels of the CpG sites in the estrogen receptor (ER) promoter region of uterus were significantly elevated

Reference	Animal model	Exposure conditions (concentration, time)	Main findings
			Findings accompanied by decreased mRNA & protein expression of the ERα (but not ERβ) & eNOs in the uterus.

9. Feto-placental haemodynamics/placental growth: Epidemiological studies with a mechanistic component

Reference – Cohort	Population location and number	Pollutant – assessment & concentration	Placentation and hemodynamic measurements	Birth outcome assessed	Main findings
van den Hooven et al 2012c Generation R Study	Rotterdam, The Netherlands City characterised by high emissions of air pollution from road traffic, shipping, households & industry (n=7,801)	PM_{10} (Mean μg/m^3) WP: 30.3 NO_2 (Mean μg/m^3) WP: 39.9 Monitoring data & dispersion modelling techniques	PlGF (cord blood) sFlt-1 (cord blood) Umbilical artery PI (T2&T3) Uterine artery PI (T2&T3) Uterine artery notching (T2&T3)	Placental weight Birth weight	PM_{10} & NO_2 averaged over WP associated with higher sFlt-1 (35.8%; 95% CI: 25.6, 45.9%; and 8.9%; 95% CI: 0.6, 17.3 per 10-μg/m3 increase in PM_{10} and NO_2, respectively) & lower PlGF– -16.3%; 95% CI: -21.9, -10.7 and -14.6%; 95% CI: –19.3, –10.0 per 10μg/m^3 increase in PM_{10} & NO_2 respectively) levels in cord blood PM_{10} & NO_2 not consistently associated with T2 or T3 placental resistance *Birth outcome results* PM_{10} & NO_2 associated with lower placenta weight (-11.8 g; 95% CI: -20.9, -2.7 & -10.7 g; 95% CI: -19.0, -2.4, respectively per 10μg/m^3 increase in the prior 2 months but not during the prior 2 weeks or over WP but not with placenta to birth weight ratio PM_{10} & NO_2 exposures during all 3 time periods associated with significant reductions in BW (–34.6; 95% CI: –66.3, –2.9; and –39.3; 95% CI: –69.1, –9.6 per 10-μg/m$_3$ increase in PM_{10} and NO_2 during total pregnancy, respectively)
Carvalho et al 2016	São Paulo, Brazil	NO_2 (mean μg/m^3)	T3:	Fetal weight	O_3 in T2 was associated with higher umbilical artery PI values

Reference – Cohort	Population location and number	Pollutant – assessment & concentration	Placentation and hemodynamic measurements	Birth outcome assessed	Main findings
Procriar	(n=366)	T1: 42.9 T2: 39.2 T3: 37.2 O_3 (mean µg/m^3) T1: 8.2 T2: 8.3 T3: 8.1 (Personally monitored; 3 times in each trimester of pregnancy)	Umbilical artery PI Middle cerebral artery PI Uterine artery PI Fetal weight	Birth weight	O_3 in T3 was related to lower umbilical artery PI values NO_2 did not influence umbilical artery PI values Fetal weight had no association with NO_2 or O_3 *Birth outcome results* Birth weight had no association with NO_2 or O_3
Contreras et al 2018 Behaviour in Pregnancy Study	Los Angeles, USA (n=360)	NO_2 (mean µg/m^3) WP: 43.4 NO_x (Mean µg/m^3) T1: 119.8 T2: 111.0 T3: 108.9 Modelled	Uterine vascular resistance: RI, PI, S/D ratio using Doppler velocimetry Visits: 1: 18–20 w gestation 2: 28–30 w gestation 3: 35–37 w gestation		Associations between NO_2 & NO_x & increased uterine vascular resistance indices (OR = 4.16, 95% CI: 1.73, 10.02 and β = 0.30, 95% CI: 0.01, 0.59 respectively) at 3rd visit

Abbreviations. PI: pulsatility index; PIGF: Placental growth factor; RI: resistance index, S/D: systolic/diastolic; sFlt-1: soluble fms-like tyrosine kinase 1

10. Placenta/umbilical cord morphology & uterine vascular resistance: Animal studies

Reference	Animal model	Exposure conditions (concentration, time)	Main findings
Veras et al 2008	Balb C mice	Unfiltered São Paulo ambient air (mean $PM_{2.5}$ 27.5 μg/m^3) or filtered air (6.5 μg/m^3) Throughout pregnancy until 18th day of gestation	Reduced volumes, calibers & surface areas of maternal blood spaces Greater fetal capillary surfaces & diffusive conductances Significant reduction in fetal weight
Weldy et al 2014	C57Bl/6J mice	Diesel exhaust (single cylinder Yanmar engine; ~300 μg/m^3 $PM_{2.5}$) or filtered air 6 hours/day, 5 days/week from embryonic day 0.5 to 17.5	Embryo resorption No effect on fetal weight Reduced placental weight Placental injury
Valentino et al 2016	New Zealand Rabbits	1000 μg/m^3 filtered diesel exhaust (25KVA Loxam engine) or clean air 2 hours/day, 5 days/week for 20 days of a 31-day gestation	Reduced placental efficiency (fetal/placental weight), placental blood flow & fetal vessel volume Increased volume of maternal blood space Mid gestation: decreased head length & umbilical pulse Near term, reduced fetal head length & plasma insulin & insulin-like growth factor 1 concentrations
Veras et al 2012	BalbC mice	Unfiltered São Paulo ambient air (mean $PM_{2.5}$ 32.8 μg/m^3 or filtered air (6.4 μg/m^3)	Umbilical cords of mice from non-filtered chambers 22% smaller in volume due to 28% loss of mucoid connective tissue & 60% decrease in volume of collagen

Reference	Animal model	Exposure conditions (concentration, time)	Main findings
		Throughout mating & pregnancy	
Yue et al 2019	C57BL/6 mice	Taiyuan $PM_{2.5}$ collected during typical heating period in Northern China (3 mg/kg) or saline by oropharyngeal aspiration Throughout gestation until embryonic day 13.5 (E13.5), 15.5 or 18.5	Reduced the area of the labyrinth layer at E18.5 but not at E13.5 or E15.5. Decreased levels of *Ascl2* and *Hand1* genes at E13.5 and E18.5 respectively. Decreased areas of gestational blood sinuses and fetal blood vessels at E18.5. Partial activation of MAPK signaling pathway.
Miller et al 2017	Long-Evans rats	O_3 (0.4 or 0.8 ppm) * or filtered air 4 hours/day) during implantation (gestational days 5 & 6)	Maternal uterine arterial flow measures (resistance index & mean velocity) indicated increased resistance between GD 15 & GD 21 following 0.8 ppm O_3 Absolute values similar in all groups on GD 21 Lower fetal weight in litters exposed to O_3 but no significant effect on fetal length
Miller et al 2019a	Long-Evans rats	O_3 (0.8 ppm) or filtered air 4 hours/day) during implantation (gestational days 5 & 6)	Air exposed rats showed significant reduction of resistance index, pulsatility index & systole/diastole ratio in uterine artery from GD 15 to GD21. Reduction in resistance index not evident in O_3 exposed rats No difference in placenta weight Reduction in placental free fatty acids and triglycerides in male placentas from O_3 exposed dams ($p<0.05$ vs. controls) Lower fetal weight in litters exposed to O_3 but no significant effect on fetal length
Miller et al 2019b	Long-Evans rats	O_3 (0.4 ppm) or filtered air 4 hours/day) during implantation (gestational days 5 & 6)	Exposure to ozone reduced uterine arterial resistance at GD 15 with no further reduction by GD 21. Exposure to O_3 reduced male (but not female) fetal weight at GD 21

*These concentrations were selected based on evidence that there is a 4- to 5-fold difference in relative lung ozone dosimetry between resting rats and humans exercising heavily on an intermittent basis, as determined by heavy (18O-labeled) reaction product concentrations in the lung (Hatch et al 1994). Thus, the 0.4 and 0.8 ppm concentrations in rats correspond approximately to 0.08 and 0.16 ppm exposure concentrations, respectively, in humans.

11. Metabolomics – Epidemiological studies with a mechanistic component

Reference – Cohort	**Population Location & number**	**Pollutant – assessment & concentration**	**Birth outcome assessed**	**Approach/sample** **Metabolites identified**	**Pathway enrichment**
Robinson et al 2018 ENVIRONAGE INMA Piccolipiu Rhea	Belgium, Spain, Italy, Greece (n=481)	PM_{10} $PM_{2.5}$ $PM_{0.1}$ Modelled	Birth weight	Untargeted approach; cord blood 68 metabolites significantly associated with birth weight, including vitamin A, progesterone, docosahexaenoic acid, indolelactic acid, & multiple acylcarnitines & phosphatidylcholine	Tryptophan metabolism, prostaglandin formation, C21-steroid hormone signalling, carnitine shuttle & glycerophospholipid metabolism *Birth outcome results* Association of vitamin A with both maternal smoking and birth weight
Yan et al 2019	Central Valley of California, USA (n=160)	TRAP (CO, nitric oxides, $PM_{2.5}$) Modelled High exposed (T1 mean) CO: 49.9 ppb NO_x: 6.39 ppb $PM_{2.5}$: 0.24 mg/m^3	None	Untargeted approach; maternal serum (T2) Approx 200 metabolic features were identified that discriminated between high (n=98) & low exposed (n=62) groups	24 metabolic pathways including linoleate, leukotriene and prostaglandin metabolism

Reference – Cohort	Population Location & number	Pollutant – assessment & concentration	Birth outcome assessed	Approach/sample Metabolites identified	Pathway enrichment
		Low exposed (T1 mean) CO: 3.66 ppb NO_x: 0.34 $PM_{2.5}$: 0.02 mg/m^3			

12. Thyroid hormones – Epidemiological studies with a mechanistic component

Reference – Cohort	Population Location & number	Pollutant – assessment & concentration	Biomarker (sample)	Birth outcome assessed	Main findings
Janssen et al 2017b ENVIRONAGE	Genk, Belgium (n=499/431)	Third trimester $PM_{2.5}$ Mean: 16 μg/m^3 Mean levels similar throughout (data not shown)	Free thyroid hormone (FT_3 & FT_4) Thyroid stimulating hormone (TSH) Maternal & cord blood collected at delivery	Birth weight	IQR increment (8.2 μg/m^3) in T3 $PM_{2.5}$ inversely associated with: Cord blood TSH levels (–11.6%; 95% CI: –21.8, –0.1) FT4/FT3 ratio (–62.7%; 95% CI: –91.6, –33.8) *Birth outcome results* 10th–90th percentile decrease in cord blood FT4 levels associated with a 56 g decrease in mean birth weight (95% CI: –90, –23)

Appendix 14. BaP concentrations in the UK

The European Environment Agency's (EEA's) mapping of BaP concentrations, presented in their 2020 Air Quality in Europe report shows that concentrations in the UK in 2018 were low (EEA, 2020). Only one site (Scunthorpe) reported concentrations in excess of the EU's target value of 1.0 ng/m^3) owing to local steel production. However, the REVIHAAP report (WHO, 2013) noted that this target value, based on past literature, was associated with a lifetime cancer risk of approximately 1 × 10^{-4} (one chance in 10,000), which was regarded as too high. The UK air quality objective for BaP is 0.25 ng/m^3 thus 2 other sites reporting exceedances are; (i) Port Talbot, South Wales located at an entrance to the steel works and (ii) in Northern Ireland (Derry/ Rosemount), likely to be linked to domestic burning of solid fuel. Several other sites reported values above the RL (reference level corresponding to a lifetime risk of 1 in 100,000 for developing lung cancer) of 0.12 ng/m^3.

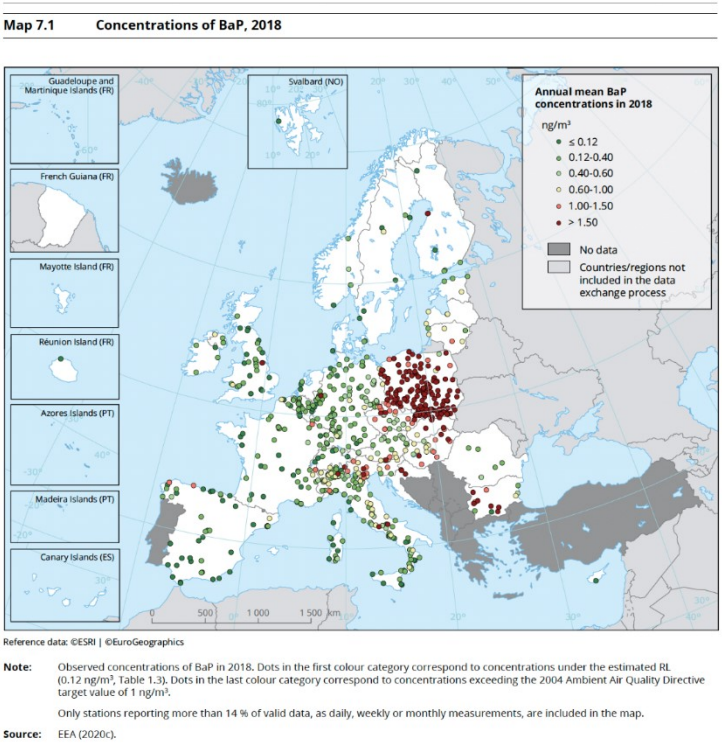


Figure 13. Concentrations of BaP, 2018 across Europe. Reproduced from the European Environment Agency Air quality in Europe – 2020 report (EEA Report No 09/2020)
This shows varying concentrations across Europe, as described in the text.

Values above 1.0 ng/m^3 are predominant in central and Eastern Europe with the highest concentrations recorded at many stations in Poland, which, together with the Czech Republic, were locations for the early molecular epidemiology studies that measured associations between air pollution and PAH-DNA adduct

formation in the placenta, maternal and cord blood (section 3.3.4). It is unknown whether PAHs are as potent in eliciting adverse birth outcomes as they are for cancer risk and thus, whether these findings are of potential concern at UK levels of PAHs.

Future epidemiological studies, incorporating a mechanistic component, into effects of PAHs on birth outcomes may not be a priority, or indeed enlightening owing to slight concentration gradients, in the UK alone. However, whilst a consistent decrease in emissions occurred between 1990 and 2002, they have subsequently stabilised and even increased in recent years, predominantly owing to increased emissions from domestic combustion (Defra, 2020). More 'attractive' locations for study, possibly in partnership with the UK, include Poland and Northern Italy. Such studies could inform interpretation of PAH-DNA adducts as a potential mechanistic pathway for PM, at least in historical epidemiological studies with higher PAH and PM concentrations.