IMPERIAL

Assessing and improving the evidence synthesis of gestational ambient air pollution exposure and the risk of preterm birth: A systematic review, meta-analysis and health impact assessment

Shawn Yu-Cheng Lee

Environmental Research Group

Imperial College London

November 2024

A thesis submitted in fulfilment of requirements for the degree of

Doctor of Philosophy of Imperial College London

Abstract

Systematic reviews and meta-analyses of gestational air pollution exposure and the risk of PTB have methodological shortcomings and inconsistencies in their findings, which hindered the use of their findings as concentration-response functions (CRF) in health impact assessments (HIA). A comprehensive update for the effect of gaseous pollutants was also needed. The aim of this thesis was to conduct a high quality, rigorous, up-to-date systematic review and meta-analysis of gestational exposure to $PM_{2.5}$, NO_2, O_3, and the risk of PTB, with methodological improvements made in areas such as quality assessment and analysis of study dependencies. Whole pregnancy $PM_{2.5}$ exposure was associated with an increase in the risk of PTB (OR: 1.10 [1.06-1.14], per 10 μg/m^3 increment, n=34, I^2=92%), moderate-to-late PTB, and very PTB. Trimester 3 NO_2 exposure was associated with an increase in the risk of PTB (OR: 1.20 [1.07-1.34], per 10 μg/m^3 increment, n=16, I^2=99%). No associations were found for O_3. Factors such as exposure assessment method were partial drivers of the high heterogeneity observed. Coupled with the findings from the narrative synthesis, the relationship between $PM_{2.5}$ exposure and the risk of PTB was deemed “likely to be causal”, whereas that of NO_2 exposure and the risk of PTB was “suggestive of, but not sufficient to infer” a causal relationship. Using the increase in risk of PTB associated with whole pregnancy $PM_{2.5}$ exposure as a CRF, HIA was conducted to quantify the impact of $PM_{2.5}$ reduction caused by Net-Zero policies in England. Approximately 27,000 PTBs would be averted if the Balanced Net Zero Pathway instead of a business-as-usual scenario was adopted, and it will result in a monetised benefit of £6,000,000,000 from 2019 to 2154. The findings in this thesis reaffirms the necessity of high-quality evidence synthesis and encourage the inclusion of PTB as an outcome in future HIAs.

Declarations

I hereby declare that the work in this thesis is my own original research and that I have appropriately cited any work that is not my own.

The research presented in this thesis required collaboration and expertise from other people. The contributions made by the collaborators are stated below:

1. Systematic review: The screening protocols and the quality assessment tool were developed with guidance from my supervisors, as well as inputs from second reviewers, Sharon Sridhara and Katherine Mcenery, as they were involved in the dual screening processes. Sharon Sridhara was the second reviewer for the titles and abstracts, whereas Katherine Mcenery was the second reviewer in the full-text screening process. Dual data extraction and quality assessment were conducted with Katherine Mcenery. Conflicts were resolved by a third reviewer which was Heather Walton, Rachel B Smith or Karen Exley.
2. Health impact assessment: The modelling and projection of $PM_{2.5}$ levels in the emissions scenarios used were provided by the Air Pollution Modelling team at the Environmental Research Group.

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-Non Commercial 4.0 International Licence (CC BY-NC). Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose. When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes. Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.'

Acknowledgements

I am very grateful to my supervisors Dr Heather Walton, Dr Rachel B Smith, Dr Karen Exley for their guidance and support throughout my PhD, without whom this work would not be possible. Heather, for being a kind and supportive supervisor with amazing ideas. Rachel, for pushing me when I need it and always working hard to ensure my work is as good as it can be. Karen, for providing thoughtful advice, guidance and encouragement.

I would like to thank members of the Air Pollution Epidemiology Team, in particular Dr Dimitris Evangelopolous, Dr Dylan Wood, and Professor Klea Katsouyanni for their support and help to develop skills that have helped along the way.

I would like to thank Dr David Dajnak and Dr Andrew Beddows at the ERG Air Pollution Modelling team in providing the modelled projections of $PM_{2.5}$ levels for the emission scenarios used, and their guidance in critiquing models of air pollution, Professor Philip Bennet for his guidance on the clinical risk factors of preterm birth, Dr Mike Holland and Professor Laure de Preux for their advice on the methods of economic evaluation.

I would like to thank the COMEAP that provided the forthcoming adverse birth outcomes report and opportunity for me to observe their meetings.

I would like to thank the Imperial Library Services for retrieving the Chinese articles used in the review, and the following authors of the primary studies used in this thesis that have replied to my enquiry: Dr Yue Ba, Professor Howard Chang, Dr Gongbo Chen, Professor Emily DeFranco, Professor Linda Dodds, Dr Keita Ebisu, Dr Simone Gary, Professor Bin-Fang Hwang, Dr Sindana D Ilango, Dr Markey Johnson, Dr Sarah Johnson, Dr Ling Liu, Dr Yang Liu, Dr Gavin Pereira, Dr Kirsten Rappazzo, Professor Carole Rudra, Professor David Savitz, Dr David Stieb, Dr Jun Tu, Dr Kari Weber, Dr Qingyang Xiao, and Dr Guoyu Zhou.

I would like to thank the NIHR HPRU in Environmental Exposures and Health, and the UKHSA for funding my PhD studentship.

I would like to thank the wider ERG for the warm and supportive atmosphere, particularly Ms Valentina Lotti, that have looked after me since day one.

I would like to thank my friends for the excellent conversations and shared laughs that distracted me from my research when I needed it. I particularly thank Kai Yee Eng for proofreading my thesis.

I am always grateful to my family, particularly my mum, who have provided unwavering support throughout my education, without whom this journey would not be possible.

Table of Contents

Abstract *2*

Declarations *3*

Acknowledgements *4*

List of Figures *12*

List of Tables *17*

List of abbreviations *20*

Chapter 1. Background *23*

1.1. Air Pollution **24**

1.1.1. Particulate Matter 24

1.1.2. Nitrogen dioxide 25

1.1.3. Ozone 26

1.2. Assessing air pollution exposure **27**

1.2.1. The use of biomarkers 27

1.2.2. Extrapolations from air quality monitor(s) 28

1.2.3. Modelled approaches 28

1.2.4. Personal monitoring 29

1.2.5. Combined approaches 29

1.3. Preterm birth **30**

1.3.1. Epidemiology and aetiology of preterm birth 30

1.3.2. Air pollution exposure and the risk of PTB 31

1.3.2.1 Mechanistic evidence 31

1.3.2.2. Epidemiological evidence 33

1.3.2.3. General limitations of the epidemiological evidence 33

1.4. Evidence synthesis **36**

1.4.1. Methods of evidence synthesis and causality determination in observational studies 36

1.4.2. Systematic reviews and meta-analyses in environmental health 38

1.4.3. Practices and shortcomings: using systematic reviews and meta-analysis of air pollution exposure and PTB as an example 39

1.5. Applying findings from evidence synthesis **40**

1.5.1. Steps of a health impact assessment 40

1.5.2. Concentration-response functions (CRF) for the effects of air pollution 42

1.6. Policies linking health impacts of air pollution reduction and climate change **43**

1.6.1. UK's commitment to achieve Net-Zero by 2050 43

1.6.2. Existing evidence on the effects of air pollution reduction through climate change policies in the UK 44

1.6.3. Opportunities from a UK project 45

1.7. Gaps in the literature and research needs **45**

Chapter 2. Aims and objectives *47*

Chapter 3. Analysis of preceding systematic reviews and meta-analyses 49

3.1. Aim and scope 49

3.2. Methods 49

3.2.1. Coverage of available literature and reporting of study selection process 49

3.2.2. Quality assessment 50

3.2.3. Study dependencies 50

3.3. Results and discussion 51

3.3.1. Overall characteristics 51

3.3.2. Coverage of available literature and reporting of study selection process 58

3.3.3. Quality assessment 61

3.3.3.1. Limited reporting of methods and findings 61

3.3.3.2 Score-based approaches 63

3.3.3.3 Domain-based approaches 65

3.3.3.4 Recommendations 65

3.3.4. Study dependencies / Population overlap 66

3.4. Interim conclusion 68

Chapter 4. Systematic review 70

4.1. Introduction 70

4.2. Methodology 70

4.2.1. PECOS statement 70

4.2.2. Search strategy 71

4.2.2.1. Development 71

4.2.2.2. Validation 73

4.2.3. PROSPERO protocol 73

4.2.4. Screening protocols 74

4.2.5. Data extraction 75

4.2.6. Quality assessment 76

4.2.6.1. WHO Air Quality Guideline Risk of Bias Assessment Instrument 76

4.2.6.2. Determining the critical confounders 76

4.2.6.3. Details of the modifications made 77

4.2.7. Addressing gaps in reporting 79

4.2.8. Narrative synthesis 79

4.3. Results 80

4.3.1. Study characteristics 80

4.3.2. Quality assessment 127

4.3.3. Narrative synthesis and interpretation of findings 127

4.3.3.1. Exposure assessment 128

4.3.3.2. Environmental factors 131

4.3.3.3. Other factors 140

4.3.3.4. Studies that used different designs or methods 144

4.4. Discussion 149

4.4.1. Summary of findings 149

4.4.2. Interpretation and implication of findings 150

4.4.2.1. Exposure assessment, residential mobility, and time-activity pattern 152

4.4.2.2. Maternal and clinical factors 152
4.4.2.3. Quality assessment 153
4.4.3. Strength and limitations 154
4.4.4. Interim conclusion 154

Chapter 5. Meta-analysis 156

5.1. Introduction 156

5.2. Methods 156
5.2.1. Conversion of effect estimates 156
5.2.2. Measurement metric of O_3 158
5.2.3. Selection of effect estimates 158
5.2.4. Analysing study dependencies and selection of studies for inclusion in meta-analysis 159
5.2.4.1. Sorting study dependencies: an example using studies in California 163
5.2.5. Meta-analytical method of choice 165
5.2.6. Leave-one-out analysis 165
5.2.7. Investigating drivers of heterogeneity 166
5.2.7.1. Risk of bias 166
5.2.7.2. Location 168
5.2.7.3. Exposure assessment methodology 169
5.2.7.4. Concentration levels 169
5.2.7.5. Source of data 170
5.2.8. Publication bias 171
5.2.9. Interpreting the results of meta-analysis 171
5.2.10. Interpreting the results of the stratified analysis 173
5.2.11. Causality discussion based on the Bradford Hill criteria 174

5.3. Results 175
5.3.1. Preconception exposure 175
5.3.2. Effect of $PM_{2.5}$ exposure 175
5.3.2.1. Trimester 1 $PM_{2.5}$ exposure (OR) 175
5.3.2.2. Trimester 2 $PM_{2.5}$ exposure (OR) 176
5.3.2.3. Trimester 3 $PM_{2.5}$ exposure (OR) 177
5.3.2.4. Whole pregnancy $PM_{2.5}$ exposure (OR) 178
5.3.2.5. Trimester 1 $PM_{2.5}$ exposure (HR) 180
5.3.2.6. Trimester 2 $PM_{2.5}$ exposure (HR) 180
5.3.2.7. Trimester 3 $PM_{2.5}$ exposure (HR) 181
5.3.2.8. Whole pregnancy $PM_{2.5}$ exposure (HR) 182
5.3.3. Effect of NO_2 exposure 183
5.3.3.1. Trimester 1 NO_2 exposure (OR) 183
5.3.3.2. Trimester 2 NO_2 exposure (OR) 184
5.3.3.3. Trimester 3 NO_2 exposure (OR) 185
5.3.3.4. Whole pregnancy NO_2 exposure (OR) 186
5.3.3.5. Trimester 1 NO_2 exposure (HR) 187
5.3.3.6. Trimester 2 NO_2 exposure (HR) 188
5.3.3.7. Trimester 3 NO_2 exposure (HR) 189
5.3.3.8. Whole pregnancy NO_2 exposure (HR) 190
5.3.4. Effect of O_3 exposure 191
5.3.4.1. Trimester 1 O_3 exposure (OR) 191
5.3.4.2. Trimester 2 O_3 exposure (OR) 192

5.3.4.3. Trimester 3 O_3 exposure (OR) 193
5.3.4.4. Whole pregnancy O_3 exposure (OR) 194
5.3.4.5. Trimester 1 O_3 exposure (HR) 195
5.3.4.6. Trimester 2 O_3 exposure (HR) 196
5.3.4.7. Trimester 3 O_3 exposure (HR) 197
5.3.4.8. Whole pregnancy O_3 exposure (HR) 198
5.3.5. Summary of the effects of $PM_{2.5}$, NO_2, O_3 exposures on the risk of PTB 199
5.3.6. The effects of $PM_{2.5}$, NO_2, O_3 exposures on PTB subgroups 199
5.3.7. Investigating the drivers of heterogeneity 201
5.3.7.1. Risk of bias 201
5.3.7.2. Location 209
5.3.7.3. Exposure assessment methodology 216
5.3.7.4. Concentration levels 222
5.3.7.5. Data source 227
5.3.8. Publication bias and funnel plots 232

5.4. Discussion 233
5.4.1. Summary of findings of the main meta-analyses 233
5.4.2. Summary of findings of the stratified analyses 234
5.4.3. Implication of findings and comparison with preceding work 235
5.4.3.1. $PM_{2.5}$ 235
5.4.3.2. NO_2 and O_3 237
5.4.3.3. Stratification by risk of bias 240
5.4.3.4. Stratified analysis by location 241
5.4.3.5. Stratified analysis by exposure assessment method 242
5.4.3.6. Stratified analysis by concentration levels 243
5.4.3.7. Stratified analysis by source of data 244
5.4.4. Findings from ORs and HRs: advantages and shortfalls 245
5.4.5. Causality discussion based on the Bradford Hill criteria 246
5.4.5.1. Temporality 246
5.4.5.2. Specificity 247
5.4.5.3. Experiment 247
5.4.5.4. Plausibility 248
5.4.5.5. Strength 249
5.4.5.6. Coherence 249
5.4.5.7. Analogy 250
5.4.5.8. Consistency 251
5.4.5.9. Biological gradient 252
5.4.5.10. Result of the causality discussion 252
5.4.6. Strength and limitations 253
5.4.7. Second interim conclusion 255

Chapter 6. Health impact assessment 256

6.1. Introduction 256

6.2. Methods 257
6.2.1. Health impact assessment (HIA) 257
6.2.1.1. Base year $PM_{2.5}$ levels and its projections 257
6.2.1.2. Birth data 259
6.2.1.3. Birth projection 262

6.2.2. Economic evaluation 262
6.2.2.1. Overall methodology 262
6.2.2.2. Healthcare costs 263
6.2.2.3. Lost productivity 263
6.2.2.4. Lost utility 264
6.2.3. Method of quantification 265
6.2.3.1. Health impact assessment (HIA) 265
6.2.3.2. Economic evaluation 267

6.3. Results 269
6.3.1. Characteristics of HIA inputs 269
6.3.1.1. Characteristics of births at base year (2019) 269
6.3.1.2. Projected trends in $PM_{2.5}$ concentration 270
6.3.2. Number of PTBs averted 271
6.3.3. Economic evaluation 273
6.3.3.1. Medical costs 273
6.3.3.2. Lost productivity 274
6.3.3.3. Lost utility 275
6.3.3.4. Total monetised benefits and sensitivity analyses 276

6.4. Discussion 278
6.4.1. Summary of findings 278
6.4.2. Implications of findings 278
6.4.2.1. Comparison with other health outcomes 278
6.4.2.2. Comparison with literature on PTB 279
6.4.3. Limitations 282
6.4.3.1. Health impact assessment (HIA) 282
6.4.3.2. Economic evaluation 284
6.4.4. Future work and future perspectives 284

Chapter 7. Wider discussion and concluding remarks 286

7.1. Outline 286

7.2. Summary of key findings and consistency with recent literature 286

7.3. Wider discussion 287
7.3.1. Comparison of findings and the approach of causality determination 287
7.3.2. Insights from the meta-analyses of $PM_{2.5}$ exposure and different types of PTB 288
7.3.3. General limitations of this thesis 289
7.3.4. Implications of the findings for policymakers and pregnant women 291

7.4. Implications of this thesis for future research 293
7.4.1. Methodological decisions are critical in systematic reviews and should be well-informed 293
7.4.2. Comprehensive coverage of literature in the narrative synthesis improves causality determination 294
7.4.3. High quality evidence synthesis is necessary 296
7.4.4. PTB should be one of the health outcomes considered in HIA for air pollution 298

References 299

Appendices 337

Appendix 1. Snapshot of the matrix showing the individual studies included in previous systematic reviews and/or meta-analyses (SR/MAs). 337
Appendix 2. Search terms used in the respective databases and the number of records retrieved. 338
Appendix 3. The title and abstract screening protocol 341
Appendix 4. The full-text screening protocol 343
Appendix 5. The data extraction form. 344
Appendix 6. The modified WHO Air Quality Guideline Risk of Bias assessment instrument for studies of gestational air pollution exposure and the risk of preterm birth. 346
Appendix 7. Guidance for using the modified framework 348
Appendix 8. Studies that were excluded from the meta-analysis and the justifications 355
Appendix 9. Studies that have conducted multipollutant analyses that were only graphically presented, on other pollutants or sources/constituents of $PM_{2.5}$. 363
Appendix 10. Studies that were included in the narrative synthesis that have provided results from single-pollutant models (SPM) and two/multipollutant models (MPMs) for whole pregnancy $PM_{2.5}$, NO_2, O_3 exposure. 364
Appendix 11. Sorting of studies eligible to be included in the meta-analyses of O_3 exposure by measurement metric. 367
Appendix 12. Studies with estimates in population subgroups that can be meta-analysed using fixed-effect model. 368
Appendix 13. Information given by authors of primary studies used in the narrative synthesis or meta-analysis. 370
Appendix 14. The process of sorting study dependencies in the main meta-analysis conducted. . 375
Appendix 15. Forest plots of the association between $PM_{2.5}$ exposure and the risk of PTB subgroups for studies reporting ORs. 389
Appendix 16. Forest plots of the association between $PM_{2.5}$ exposure and the risk of PTB subgroups for studies reporting HRs. 392
Appendix 17. Forest plots of the association between NO_2 exposure and the risk of PTB subgroups for studies reporting HRs. 395
Appendix 18. Results of the leave-one-out analysis conducted for the eligible exposure-outcome pairs. 397
Appendix 19. Funnel plots for the meta-analyses of the association between air pollution exposure and the risk of PTB with more than 10 studies. 400
Appendix 20. Results of the Egger's test conducted to assess funnel plot asymmetry of the meta-analyses conducted for the association between gestational air pollution exposure and the risk of PTB (for exposure-outcome pairs with >10 studies). 404
Appendix 21. Results of the stratified analyses conducted by quartiles of mean/median $PM_{2.5}$ level of the studies. 405
Appendix 22. Link to online platform storing the individual forest plots of the stratified analyses. . 406
Appendix 23. Forest plot showing the results from the meta-analyses conducted in this thesis, Ju et al. (2023), and Yu et al. (2022) for the association between $PM_{2.5}$ exposure and the risk of PTB subgroups. 407
Appendix 24. Description for the weight of evidence available for the different conclusions made in the causality determination process (reproduced from Owens et al., 2017). 408
Appendix 25. Local authorities (LA) with months with no information reported (excluding those that were abolished). 409
Appendix 26. The reason for missing data to be present for each of the local authorities (LA) with more than 50% of missing data. 410
Appendix 27. Correspondence with NHS Digital 411

Appendix 28. $PM_{2.5}$ concentrations at (a) base year 2019, (b) year 2050 under the business-as-usual (BAU) scenario, and (c) year 2050 under the Balanced Net-Zero Pathways (BNZP) scenario. 412
Appendix 29. Reduction in $PM_{2.5}$ concentrations at year 2050 if the business-as-usual (BAU) scenario was implemented (compared to the base year 2019). 413
Appendix 30. Cumulative medical cost (in £2022, millions) derived from Hua et al. (2023) for the comparison between different scenarios of interest from 2019 up to 2154..... 414
Appendix 31. Monetised value of the lost utility (in £2022, millions) based on the value of statistical life (VSL) from the UK Department of Transport and value of statistical case (VSC) from Ščasný et al. (2023). 415
Appendix 32. Monetised benefit from the PTBs averted (in £2022, millions) based on the different values of medical cost, lost productivity, and lost utility for births that would have been born preterm from 2019 to 2154, in different comparison of scenarios...... 416

List of Figures

Figure 1. Direct acyclic graph (DAG) produced for the relationship between gestational air pollution exposure and preterm birth. 77

Figure 2. PRISMA diagram showing the study selection process for this systematic review and meta-analysis.. 82

Figure 3. Primary studies that have provided estimates from both non-modelled and modelled approaches. 129

Figure 4. Primary studies that have provided estimates for continuous exposure for participants that have changed address or spent more than a certain amount of time at home daily during pregnancy. 130

Figure 5. Estimates from studies that have compared results from SPM and MPM with $PM_{2.5}$ and NO_2 exposure..... 132

Figure 6. Results from studies that have provided estimates for the risk of PTB associated with continuous $PM_{2.5}$ exposure from traffic. 134

Figure 7. The studies that have provided estimates for whole pregnancy exposure to $PM_{2.5}$ components and the risk of PTB based on continuous exposure. 136

Figure 8. Studies that have conducted a stratified analysis by season of birth or conception for air pollution exposure and the risk of PTB, based on continuous exposure. 139

Figure 9. Studies that have conducted a stratified analysis by presence or absence of certain conditions for the risk of PTB associated with air pollution exposure. 141

Figure 10. Studies that have conducted a stratified analysis by trigger of delivery, type of delivery, and method of conception. 144

Figure 11. A flowchart showing how study dependencies between studies were identified, and how a study(ies) was selected to be included in the meta-analysis. 162

Figure 12. The location, start and end date of studies reporting ORs in the US for whole pregnancy $PM_{2.5}$ exposure and risk of PTB. 164

Figure 13. A flowchart showing the process of sensitivity analysis conducted for the meta-analysis for each main result, if more than two studies were available. 167

Figure 14. The interpretation of the result of a meta-analysis regarding the presence, magnitude, and robustness of the association. 172

Figure 15. Forest plot for the meta-analysis of the association between trimester 1 $PM_{2.5}$ exposure and the risk of PTB from studies using ORs. 176

Figure 16. Forest plot for the meta-analysis of the association between trimester 2 $PM_{2.5}$ exposure and the risk of PTB from studies using ORs. 177

Figure 17. Forest plot for the meta-analysis of the association between trimester 3 $PM_{2.5}$ exposure and the risk of PTB from studies using ORs. 178

Figure 18. Forest plot for the meta-analysis of the association between whole pregnancy $PM_{2.5}$ exposure and the risk of PTB from studies using ORs. 179

Figure 19. Forest plot for the meta-analysis of the association between trimester 1 $PM_{2.5}$ exposure and the risk of PTB from studies using HRs. 180

Figure 20. Forest plot for the meta-analysis of the association between trimester 2 $PM_{2.5}$ exposure and the risk of PTB from studies using HRs. 181

Figure 21. Forest plot for the meta-analysis of the association between trimester 3 $PM_{2.5}$ exposure and the risk of PTB from studies using HRs. 182

Figure 22. Forest plot for the meta-analysis of the association between whole pregnancy $PM_{2.5}$ exposure and the risk of PTB from studies using HRs. 183

Figure 23. Forest plot for the meta-analysis of the association between trimester 1 NO_2 exposure and the risk of PTB from studies using ORs. 184

Figure 24. Forest plot for the meta-analysis of the association between trimester 2 NO_2 exposure and the risk of PTB from studies using ORs. 185

Figure 25. Forest plot for the meta-analysis of the association between trimester 3 NO_2 exposure and the risk of PTB from studies using ORs.. 186

Figure 26. Forest plot for the meta-analysis of the association between whole pregnancy NO_2 exposure and the risk of PTB from studies using ORs. 187

Figure 27. Forest plot for the meta-analysis of the association between trimester 1 NO_2 exposure and the risk of PTB from studies using HRs. 188

Figure 28. Forest plot for the meta-analysis of the association between trimester 2 NO_2 exposure and the risk of PTB from studies using HRs.. 189

Figure 29. Forest plot for the meta-analysis of the association between trimester 3 NO_2 exposure and the risk of PTB from studies using HRs. 190

Figure 30. Forest plot for the meta-analysis of the association between whole pregnancy NO_2 exposure and the risk of PTB from studies using HRs. 191

Figure 31. Forest plot for the meta-analysis of the association between trimester 1 O_3 exposure and the risk of PTB from studies using ORs.. 192

Figure 32. Forest plot for the meta-analysis of the association between trimester 2 O_3 exposure and the risk of PTB from studies using ORs. 193

Figure 33. Forest plot for the meta-analysis of the association between trimester 3 O_3 exposure and the risk of PTB from studies using ORs. 194

Figure 34. Forest plot for the meta-analysis of the association between whole pregnancy O_3 exposure and the risk of PTB from studies using ORs. 195

Figure 35. Forest plot for the meta-analysis of the association between trimester 1 O_3 exposure and the risk of PTB from studies using HRs. 196

Figure 36. Forest plot for the meta-analysis of the association between trimester 2 O_3 exposure and the risk of PTB from studies using HRs. 197

Figure 37. Forest plot for the meta-analysis of the association between trimester 3 O_3 exposure and the risk of PTB from studies using HRs.. 197

Figure 38. Forest plot for the meta-analysis of the association between whole pregnancy O_3 exposure and the risk of PTB from studies using HRs. 198

Figure 39. Forest plot for the summary estimates of the main meta-analyses. 199

Figure 40. Forest plot for the summary estimates of the meta-analyses conducted for the association between $PM_{2.5}$ exposure at trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP) and PTB, mlPTB, and vPTB. 200

Figure 41. Forest plot for the summary estimates of the meta-analyses conducted for the association between NO_2 exposure at trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP) and PTB and vPTB. 201

Figure 42. Forest plot for the summary estimates of the meta-analyses conducted for the association between $PM_{2.5}$ exposure at trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP), and the risk of PTB from studies using ORs, stratified by the risk of bias (RoB) levels of the studies used (Low, Moderate, High), and a further stratification by the extent of confounder adjustment (All, Not All).....202

Figure 43. Forest plot for the summary estimates of the meta-analyses conducted for the association between $PM_{2.5}$ exposure at trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP), and the risk of PTB from studies using HRs, stratified by the risk of bias (RoB) levels (Low, Moderate, High) of the studies used, and a further stratification by the extent of confounder adjustment (All, Not all). 204

Figure 44. Forest plot for the summary estimates of the meta-analyses conducted for the association between NO_2 exposure at trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP), and the risk of PTB from studies using ORs, stratified by the risk of bias (RoB) levels (Low, Moderate, High) of the studies used, and a further stratification by the extent of confounder adjustment (All, Not all).....206

Figure 46. Forest plot for the summary estimates of the meta-analyses conducted for the association between O_3 exposure at trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP), and the risk of PTB from studies using ORs, stratified by the risk of bias (RoB) levels (Low, Moderate, High) of the studies used, and a further stratification by the extent of confounder adjustment (All, Not all).....208

Figure 47. Forest plot for the summary estimates of the meta-analyses conducted for the association between $PM_{2.5}$ exposure at trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP), and the risk of PTB from studies using ORs, stratified in the levels of WHO region (Americas, Europe, Western Pacific), country (China, US, Canada), and climatic regions within the China and the US. 210

Figure 48. Forest plot for the summary estimates of the meta-analyses of the association between trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP) $PM_{2.5}$ exposure and the risk of PTB from studies using HRs, stratified by WHO region (Americas, Europe, Western Pacific), country (China, US, Canada), and climatic regions within the China and the US. 211

Figure 49. Forest plot for the summary estimates of the meta-analyses conducted for the association between NO_2 exposure at trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP), and the risk of PTB from studies using ORs, stratified in the levels of WHO region (Americas, Europe, Western Pacific), country (China, US, Canada), and climatic regions within the China and the US. 213

Figure 50. Forest plot for the summary estimates of the meta-analyses conducted for the association between NO_2 exposure at trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP), and the risk of PTB from studies using HRs, stratified according to WHO region of Western Pacific, China, and the hot summer region within China. 214

Figure 51. Forest plot for the summary estimates of the meta-analyses conducted for the association between O_3 exposure at trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP), and the risk of PTB from studies using ORs, stratified in the levels of WHO region (Americas, Europe, Western Pacific), country (China, US, Canada), and climatic regions within the China and the US. 215

Figure 52. Forest plot for the summary estimates of the meta-analyses conducted for the association between $PM_{2.5}$ exposure at trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP), and the risk of PTB from studies using ORs, stratified by different exposure assessment methods.....216

Figure 53. Forest plot for the summary estimates of the meta-analyses conducted for the association between $PM_{2.5}$ exposure at trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP), and the risk of PTB from studies using HRs, stratified by different exposure assessment methods.....217

Figure 54. Forest plot for the summary estimates of the meta-analyses conducted for the association between NO_2 exposure at trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP), and the risk of PTB from studies using ORs, stratified by different exposure assessment methods. The summary effect sizes (per 10 μg/m^3 increase in exposure) and 95%CIs, I^2 statistics, and number of estimates (n) are presented. The forest plots which show the individual studies included in the sensitivity analyses are available in Appendix X. IDW: Inverse distance weighting; LUR: Land-use regression; CTM: Chemical transport models.....219

Figure 55. Forest plot for the summary estimates of the meta-analyses conducted for the association between NO_2 exposure at trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP), and the risk of PTB from studies using HRs, with restriction to studies that have extrapolated exposure from nearest monitors only.....220

Figure 56. Forest plot for the summary estimates of the meta-analyses conducted for the association between O_3 exposure at trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP), and the risk of PTB from studies using ORs, stratified by different exposure assessment methods.....221

Figure 57. Forest plot for the summary estimates of the meta-analyses conducted for the association between O_3 exposure at whole pregnancy (WP) and the risk of PTB from studies using HRs, with restriction to studies that have extrapolated exposure from nearest monitors only.....221

Figure 58. Forest plot for the summary estimates of the meta-analyses conducted for the association between $PM_{2.5}$ exposure at trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP), and the risk of PTB from studies using ORs, stratified by brackets of $PM_{2.5}$ set using the WHO Air Quality Guideline levels. 223

Figure 59. Forest plot for the summary estimates of the meta-analyses conducted for the association between $PM_{2.5}$ exposure at trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP), and the risk of PTB from studies using HRs, stratified by brackets of $PM_{2.5}$ set using the WHO Air Quality Guideline levels. 224

Figure 60. Forest plot for the summary estimates of the meta-analyses conducted for the association between NO_2 exposure at trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP), and the risk of PTB from studies using ORs, stratified by brackets of NO_2 set using the WHO Air Quality Guideline levels.....225

Figure 61. Forest plot for the summary estimates of the meta-analyses conducted for the association between O_3 exposure at trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP), and the risk of PTB from studies using ORs, stratified by brackets of O_3 set using the WHO Air Quality Guideline levels.....226

Figure 62. Forest plot for the summary estimates of the meta-analyses conducted for the association between O_3 exposure at trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP), and the risk of PTB from studies using ORs, stratified by the quartiles of the distribution of O_3 levels in the primary studies used in the meta-analyses. 227

Figure 63. Forest plot for the summary estimates of the meta-analyses conducted for the association between $PM_{2.5}$ exposure at trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP), and the risk of PTB from studies using ORs, stratified by the source of data of the study 228

Figure 64. Forest plot for the summary estimates of the meta-analyses conducted for the association between $PM_{2.5}$ exposure at trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP), and the risk of PTB from studies using HRs, stratified by the source of data of the study.....229

Figure 65. Forest plot for the summary estimates of the meta-analyses conducted for the association between NO_2 exposure at trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP), and the risk of PTB from studies using ORs, stratified by the source of data of the study. 230

Figure 66. Forest plot for the summary estimates of the meta-analyses conducted for the association between NO_2 exposure at trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP), and the risk of PTB from studies using HRs, stratified by the source of data of the study.....231

Figure 67. Forest plot for the summary estimates of the meta-analyses conducted for the association between O_3 exposure at trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP), and the risk of PTB from studies using ORs, stratified by the source of data of the study. 231

Figure 68. Forest plot for the summary estimates of the meta-analyses conducted for the association between O_3 exposure at trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP), and the risk of PTB from studies using HRs, stratified by the source of data. 232

Figure 69. Forest plot of the results from preceding meta-analyses and this work for exposure periods that found an increase in risk of PTB associated with per 10 $\mu g/m^3$ increase in$PM_{2.5}$ exposure at trimester 1 (T1), 2 (T2), 3 (T3), or whole pregnancy (WP) 236

Figure 70. Forest plot of the results from preceding meta-analyses and this work for exposure periods that found an increase in risk of PTB associated with per 10 $\mu g/m^3$ increase in NO_2 exposure at trimester 1 (T1), 2 (T2), 3 (T3), or whole pregnancy (WP) 238

Figure 71. Forest plot of the results from preceding meta-analyses and this work for exposure periods that found an increase in risk of PTB associated with per 10 $\mu g/m^3$ increase in O_3 exposure at trimester 1 (T1), 2 (T2), 3 (T3), or whole pregnancy (WP). 239

Figure 72. The level of missingness for each local authority in England by month in the base year (2019). 261

Figure 73. Averaged number of births (left) and prevalence of preterm birth (right) in the 317 LAs in England at 2019 (averaged across all months). 269

Figure 74. Projected $PM_{2.5}$ concentrations for the counterfactual, BAU, and BNZP scenarios from 2019 to 2154. 270

Figure 75. The reduction in $PM_{2.5}$ concentrations at 2050 if the BNZP scenario was implemented (compared to the base year 2019). 271

Figure 76. Spatial distribution of the PTBs averted across the LAs of England over the whole study period (135 years). 273

List of Tables

Table 1. Definition of the adverse birth outcomes (ABOs) mentioned (continued on next page). 30

Table 2. Biases in perinatal epidemiology. 35

Table 3. Examples of methods used in evidence synthesis (adapted from Cook et al., 2017). 36

Table 4. The Bradford Hill Criteria (adapted from Fedak et al., 2015; Hill, 1965) 38

Table 5. CRF recommended for long-term $PM_{2.5}$ exposure and asthma incidences in children (adapted from Forastiere et al., 2024) 42

Table 6. Examples of different levels of recommendation by COMEAP regarding CRFs for exposure-outcome pairs (COMEAP, 2023). 43

Table 7. Characteristics of the systematic reviews (and/or meta-analyses) retrieved. 53

Table 8. Classification of preceding SR/MAs retrieved according to their aims. 59

Table 9. The primary studies of $PM_{2.5}$ exposure retrieved in the three systematic reviews and meta-analyses (SR+MAs) without restriction to specific groups that were published within the literature cut-off date of Li et al. (2017). 60

Table 10. Primary studies included in 3 meta-analyses of whole pregnancy $PM_{2.5}$ exposure and the risk of PTB published with literature cut-off date within 3 years of each other. 61

Table 11. Sources of bias assessed in the self-set criteria for Shah et al. (2011), Stieb et al. (2012), and Khader et al. (2016). 62

Table 12. Examples of primary studies that were critiqued by 4 or more SR/MAs using the NOS, and the methodological issues that are not captured (continued on next page). 64

Table 13. Strategies adopted by systematic reviews (and/or meta-analyses) to reduce study dependencies 67

Table 14. The Population, Exposure, Control, Outcome, Study design (PECOS) statement of the review conducted in this chapter. 70

Table 15. Databases used to retrieve potential eligible primary studies and their reasons for selection. 72

Table 16. The questions in the full-text screening protocol for determining a primary study's eligibility to be included in the meta-analyses 75

Table 17. Modifications made to the WHO Air Quality Guidelines Risk of Bias Assessment Instrument (continued on next page). 78

Table 18. Characteristics of the primary studies included in this systematic review. 83

Table 19. The distribution of studies and those that have reported an increase in risk of PTB or subgroups of PTB using ORs/RRs/HRs in each pollutant-exposure period-outcome relationship, among those that were eligible to be included in the narrative synthesis. 128

Table 20. Studies that have studied how temperature or heatwaves modify the association between air pollution exposure and the risk of PTB. 138

Table 21. Summary of findings from narrative synthesis 149

Table 22. Conversion factors used for converting between parts per billion (ppb) to micrograms per cubic metres ($\mu g/m^3$), at 20°C and 1013 mb. 157

Table 23. The statistical method-pollutant-exposure period-outcome combinations of interest. 161

Table 24. The definitions for a factor not being a driver, being a partial driver, or a driver of the heterogeneity observed. 173

Table 25. P-values obtained from Q-test of subgroup differences for stratification by risk of bias and confounder adjustment for meta-analyses of studies of $PM_{2.5}$ exposure reporting ORs. 203

Table 26. P-values obtained from Q-test of subgroup differences for stratification by risk of bias and confounder adjustment for meta-analyses of studies of $PM_{2.5}$ exposure reporting HRs. 203

Table 27. P-values obtained from Q-test of subgroup differences for stratification by risk of bias and confounder adjustment for meta-analyses of studies of NO_2 exposure reporting ORs.. 205

Table 28. P-values obtained from Q-test of subgroup differences for stratification by risk of bias and confounder adjustment for meta-analyses of studies of O_3 exposure reporting ORs. 207

Table 29. P-values obtained from Q-test of subgroup differences for stratification by location for studies reporting ORs for the association between $PM_{2.5}$ exposure and the risk of PTB. 209

Table 30. P-values from Q-test of subgroup differences for stratification by location for studies reporting HRs for the association between $PM_{2.5}$ exposure and the risk of PTB. 212

Table 31. P-values from Q-test of subgroup differences for stratification by location for studies reporting ORs for the association between NO_2 exposure and the risk of PTB.. 212

Table 32. P-values obtained from Q-test of subgroup differences for stratification by location for meta-analyses of studies of NO_2 exposure reporting HRs. 214

Table 33. P-values obtained from Q-test of subgroup differences for stratification by location for meta-analyses of studies of O_3 exposure reporting ORs. 215

Table 34. P-values obtained from Q-test of subgroup differences for stratification by exposure assessment for meta-analyses of studies of $PM_{2.5}$ exposure reporting ORs. 217

Table 35. P-values obtained from Q-test of subgroup differences for stratification by exposure assessment for meta-analyses of studies of $PM_{2.5}$ exposure reporting HRs. 218

Table 36. P-values obtained from Q-test of subgroup differences for stratification by exposure assessment for meta-analyses of studies of NO_2 exposure reporting ORs. 218

Table 37. P-values obtained from Q-test of subgroup differences for stratification by exposure assessment for meta-analyses of studies of O_3 exposure reporting ORs 220

Table 38. P-values obtained from Q-test of subgroup differences for stratification by pollution levels for meta-analyses of studies of $PM_{2.5}$ exposure reporting ORs. 222

Table 39. P-values obtained from Q-test of subgroup differences for stratification by pollution levels for meta-analyses of studies of $PM_{2.5}$ exposure reporting HRs. 224

Table 40. P-values obtained from Q-test of subgroup differences for stratification by pollution levels for meta-analyses of studies of NO_2 exposure reporting ORs. 225

Table 41. P-values obtained from Q-test of subgroup differences for stratification by pollution levels for meta-analyses of studies of O_3 exposure reporting ORs. 226

Table 42. P-values obtained from Q-test of subgroup differences for stratification by data source for meta-analyses of studies of $PM_{2.5}$ exposure reporting ORs. 228

Table 43. P-values obtained from Q-test of subgroup differences for stratification by data source for meta-analyses of studies of $PM_{2.5}$ exposure reporting HRs. 228

Table 44. P-values obtained from Q-test of subgroup differences for stratification by data source for meta-analyses of studies of NO_2 exposure reporting ORs. 229

Table 45. P-values obtained from Q-test of subgroup differences for stratification by data source for meta-analyses of studies of NO_2 exposure reporting HRs. 230

Table 46. Summary of the results of the stratified analyses, according to the potential strength of a factor being a driver of the heterogeneity observed. 234

Table 47. Summary of the results of the causality discussion based on Bradford Hill criteria. 246

Table 48. Comparison of the analysis presented in this Chapter and the linked project. 257

Table 49. Percentage of LA with missing data present in each month of the base year (2019) and the type of information that was missing. 260

Table 50. Hospital admission costs for respective PTB types and term births over different time horizons. 263

Table 51. The Willingness-to-pay (WTP) values used in this analysis and their respective sources. 264

Table 52. Steps of quantification used to derive the number of PTBs averted 265

Table 53. The sources of medical cost, lost productivity, and lost utility in main results and sensitivity analyses. 268

Table 54. Number of PTBs averted based on the concentration differences between the scenarios of choice from 2019 up to 2050. 272

Table 55. Number of PTBs averted based on the concentration differences between the scenarios of choice from 2019 up to 2154. 272

Table 56. Cumulative hospital costs (in £2022, millions) at different time horizons of the life of a PTB based on comparison between the scenarios of interest (with 5 $\mu g/m^3$ cut-off) 274

Table 57. Monetised value of the lost productivity (in £2022, millions) due to delivery of PTBs based on comparison between the scenarios of interest from 2019 to 2154. 275

Table 58. Monetised value of lost utility (in £2022, millions) derived from PTBs averted from 2019 to 2154. 276

Table 59. The sum of monetised benefits (in £2022, millions) from its respective components in the main and sensitivity analysis (comparison between the BAU and BNZP scenarios) from 2019 to 2154. 277

List of abbreviations

ADMS	Atmospheric Dispersion Modelling System
AOD	Aerosol optical depth
BAU	Business-as-usual
BMI	Body mass index
BNZP	Balanced net-zero pathway
CALINE	California Line Source Dispersion Model
CARB	California Air Resources Board
CBA	Cost-benefit analysis
CI	Confidence intervals
CMAQ	Community Multiscale Air Quality (modelling)
CINAHL	Cumulative Index to Nursing and Allied Health Literature
CNKI	Chinese National Knowledge Infrastructure
CO	Carbon monoxide
COMEAP	Committee on the Medical Effects of Air Pollutants
CRF	Concentration-response function
CSTJ	Chinese Science and Technology Journal Database
CTM	Chemical Transport Model
DEFRA	Department for Environmental, Food and Rural Affairs
DNA	Deoxyribonucleic acid
DoT	Department of Transport
DUACAG	Dalhousie University Atmospheric Composition Analysis Group
ESCAPE	The European Study of Cohorts for Air Pollution Effects
EU	European Union
ExPTB	Extremely preterm birth
FSD	Statistically Fused Air and Deposition Surface database
GBD	Global burden of disease
GBP	Great British Pound
GeoDHOC	Geospatial Determinants of Health Outcomes Consortium
GEOS-Chem	Goddard Earth Observing System-Chemistry
HIA	Health impact assessment
HR	Hazard ratio

HRAPIE	Health Risk of Air Pollution in Europe project
IDW	Inverse distance weighting
IL	Interleukins
IQR	Interquartile range
LA	Local authority
LAEI	London atmospheric emissions inventory
LUR	Land use regression
MA	Meta-analysis
MAIAC	Multi-Angle Implementation of Atmospheric Correction
MATES	Multiple Air Toxics Exposure Study
MDA	Malondialdehyde
MeSH	Medical Subject Heading
MIDIRS	Maternity and Infant Care Database
MIPTB	Moderate-to-late preterm birth
NAEI	National atmospheric emission inventory
NHS	National Health Services
NOAA	National Oceanic and Atmospheric Administration
NO_2	Nitrogen dioxide
NO_x	Nitrogen oxides
NOS	Newcastle-Ottawa Scale
O_3	Ozone
OECD	Organisation for Economic Co-operation and Development
ONS	Office of National Statistics
OR	Odds ratio
PM	Particulate matter
$PM_{0.1}$	Particulate matter with an aerodynamic diameter less than 0.1μm
$PM_{2.5}$	Particulate matter with an aerodynamic diameter less than 2.5μm
PM_{10}	Particulate matter with an aerodynamic diameter less than 10μm
PRISMA	Preferred Reporting Items for Systematic reviews and Meta-Analyses
PROSPERO	International prospective register of systematic reviews
PTB	Preterm birth
PTD	Preterm delivery

QA	Quality assessment
QALY	Quality-adjusted life year
RIS	Research Information Systems
RR	Risk ratio
SA	Sensitivity analysis
SCAQMD	South Coast Air Quality Management District
SES	Socioeconomic status
SILAM	System for Integrated Modelling for Atmospheric Compositions
SMOKE	Sparse Matrix Operator Kernel Emissions
SO_2	Sulphur dioxide
SR	Systematic review
SR/MAs	Systematic reviews and/or meta-analyses
TLBW	Term low birth weight
TNF-α	Tumour necrosis factor alpha
TSP	Total suspended particulates
T1	Trimester 1
T2	Trimester 2
T3	Trimester 3
UFP	Ultrafine particulate
UK	United Kingdom
US	United States
USEPA	United States Environmental Protection Agency
USEPA AIRS	USEPA Aerometric Information Retrieval System
USEPA AQPS	USEPA Fused Air Quality Prediction Surface
USEPA AQS	USEPA Air Quality System
VLBW	Very low birth weight
VPTB	Very preterm birth
VSC	Value of a statistical case
VSL	Value of a statistical life
VOCs	Volatile organic compounds

Chapter 1. Background

Air pollution is one of the greatest environmental and public health challenges faced in the 21st century. Air pollutants are generated from natural and anthropogenic sources, such as wildfires, transport, or secondary chemical reactions (Manisalidis et al., 2020). Air pollutants that have received most attention in the past decades include particulate matter (PM), particularly PM with an aerodynamic diameter less than 2.5 μm ($PM_{2.5}$), ozone (O_3), nitrogen dioxide (NO_2), sulphur dioxide (SO_2), and carbon monoxide (CO) (United States Environmental Protection Agency (USEPA), 2020).

There are strong motivations to reduce air pollution. First, adverse health effects of ambient air pollution exposure are well-established, with the strongest evidence linking it to all-cause mortality and cardiorespiratory health outcomes (Brook et al., 2010; Samoli et al., 2016; World Health Organization (WHO) for Europe, 2013a). These health outcomes predominantly affect the older population, but the evidence regarding its effect on adverse birth outcomes (ABOs), such as preterm birth (PTB), are also accumulating due to known vulnerability of pregnant women and the foetus against environmental stressors (Chief Medical Officer (CMO), 2022; Royal College of Physician, 2016). One-third of PTBs in the world were estimated to be attributable to the effect of ambient air pollution exposure in 2019 (Ghosh et al., 2021), and PTB is associated with increased risk of mortality and conditions with lifelong implications (Kelly & Griffith, 2020; Swamy et al., 2008).

Second, air pollution is interconnected with climate change, with common pathways between how it exerts negative consequences on the ecosystem and public health (Richardson et al., 2023). There is keen interest in building the connection between the effect of climate change mitigation and adaptation actions, for instance, through emissions and hence air pollution reduction, and their health benefits (Romanello et al., 2022). The associated economic losses averted through the health benefits realised are likely immense (Reis et al., 2022). It is thus important to conduct health impact assessments (HIA) and economic evaluation of climate change mitigation and adaptation policies via reductions in air pollution.

A robust concentration-response function (CRF) is needed to accurately derive the number of cases averted for a given condition that is attributable to air pollution reduction. It is often derived from syntheses of epidemiological evidence using systematic reviews and/or meta-analyses (SR/MAs) that show a consensus in the direction and magnitude of the association, strength of the evidence, and areas that require further research (Lasserson et al., 2021). For PTB, there are numerous SR/MAs available. However, concerns remain due to inconsistencies in findings and shortcomings in methodological choices of the preceding SR/MAs (Committee of the Medical Effects of Air Pollutants (COMEAP), forthcoming).

Therefore, COMEAP decided not to recommend the use of their findings as concentration-response functions (CRF) in health impact assessments (HIAs) as they were not sufficiently robust.

The following sections will explore these aforementioned issues in greater depth, explaining how they informed the formulation of the research problem and hypotheses presented at the end of the chapter.

1.1. Air Pollution

Air pollutants can be broadly divided into two groups: particulates and gases. Section 1.1.1-1.1.3 will focus on three pollutants that are used in the analysis presented in the following chapters: $PM_{2.5}$, NO_2, and O_3.

1.1.1. Particulate Matter

Particulate matter (PM) is an umbrella term for solid particles and liquid droplets suspended in the air (USEPA, 2009). It is commonly subcategorised into three different groups according to their aerodynamic diameter: less than 10 μm (PM_{10}), less than 2.5 μm ($PM_{2.5}$), and less than 0.1 μm ($PM_{0.1}$, also known as ultrafine particles (UFP)). Their precise definitions can be found in the European Union (EU) Directive[1] (EU, 2008). There is an extensive number of constituents that could make up PM, for instance, organic carbonaceous compounds such as elemental carbon (EC), black carbon (BC), or polycyclic aromatic hydrocarbons (PAH), ions such as nitrates, or trace elements such as iron. The exact constituent profile of PM varies greatly across different locations, seasons, or climates. PM can also be grouped into primary or secondary PM. Primary PM is generated directly from natural or anthropogenic sources, such as sea salt, volcanic eruptions, road transport, or fossil fuel combustion. Secondary PM is generated from reactions of atmospheric precursor gases. For example, it can be generated through the oxidation of sulphur dioxide to sulphate, or ammonia from soil or animal excretes (USEPA, 2009).

During inhalation, PM of a smaller aerodynamic diameter can penetrate deeper into the respiratory system. $PM_{0.1}$ and smaller components of $PM_{2.5}$ can reach as deep as the air-blood barrier, whereas larger components of $PM_{2.5}$ and PM_{10} can reach the terminal bronchioles and primary bronchi respectively (Kelly & Fussell, 2012). After entering the

[1] For $PM_{2.5}$, it was defined as “PM which passes through a size-selective inlet as defined in the reference method for the sampling and measurement of $PM_{2.5}$, EN 14907, with a 50 % efficiency cut- off at 2.5 μm aerodynamic diameter”.

bloodstream, PM can reach all organs, with BC being shown to reach even the foetus of pregnant women, passing the placental barrier (Bové et al., 2019; Pryor et al., 2022). However, there are debates concerning the proportion of ultrafine particles that could translocate across to the bloodstream, as it was suggested to be very small and thus the level of exposure for organs, especially those with additional barriers e.g. the brain, would be even smaller, casting doubts on the toxicological significance of such a finding (COMEAP, 2022). Recent evidence suggests that $PM_{0.1}$ can also enter the central nervous system by directly travelling up the olfactory nerve, bypassing the blood-brain and air-blood barriers (Bongaerts et al., 2022; You et al., 2022).

PM promotes an array of adverse physiological changes, for instance, systemic inflammation, endothelial dysfunction, and increased heart rate variability (Schraufnagel, 2020; Thangavel et al., 2022; USEPA, 2019). The adverse health effects of short- and long-term exposure to ambient PM_{10} and $PM_{2.5,}$ through increased risk of all-cause mortality, adverse respiratory outcomes such as lung cancer incidences, chronic obstructive pulmonary disease (COPD), asthma exacerbations, as well as cardiovascular outcomes such as ischaemic heart diseases, heart failure, and cardiovascular-related hospital admissions, are well-supported by ample evidence in the literature (CMO, 2022; Atkinsons et al., 2014; Chen & Hoek, 2020; Yu et al., 2021).

Emerging evidence also suggests that the adverse health effects could include an increased risk of neurodegenerative diseases including dementia, and metabolic disorders such as diabetes. (Powers et al., 2016; Yang et al., 2020). Exposure to particulate matter pollution is estimated to be the top contributor to the global burden of disease (GBD) among all risk factors analysed in the GBD, Injuries, and Risk factor Study 2021 (Brauer et al., 2024). It is responsible for more than 232 million disability-adjusted life years (DALYs), with a risk-attributable burden greater than high systolic blood pressure, or high BMI (Brauer et al., 2024).

1.1.2. Nitrogen dioxide

Oxides of nitrogen (NO_x) found in the ambient air are mainly produced from fuel combustion that occurs in vehicles, power stations, or industrial processes, although natural processes such as lightning strikes, and agricultural activities such as fertiliser use can also lead to their formation. The majority of oxides of nitrogen are emitted as nitric oxide (NO), but they oxidise readily in the air when oxidants such as oxygen, ozone (O_3), or volatile organic compounds (VOCs) are present to form nitrogen dioxide (NO_2) (WHO, 2010). Vehicle

emissions are the primary source of ambient NO_2. Hence, levels of NO_2 in the air are often treated as a marker for traffic-related air pollution (COMEAP, 2018a).

After being inhaled, NO_2 encounters the epithelial lining fluid (ELF) of the respiratory tract. NO_2 is a strong oxidant that undergoes reactive absorption, where it reacts with various substances such as antioxidants glutathione (GSH) in the ELF while being absorbed. Direct absorption of NO_2 can also occur, particularly in the deeper alveolar region, where diffusion to the underlying tissues occurs very rapidly. Reactive oxygen species are generated from the process, and they trigger a cascade of physiological changes leading to events such as damage to the epithelial barrier in the respiratory tract, initiation of inflammation, and modification to the immune system (USEPA, 2016).

Short- and long-term exposure to ambient NO_2 share similarities to that of $PM_{2.5}$ in terms of the negative physiological consequences described in the previous section. However, the evidence available is generally less consistent. Long-term NO_2 exposure is associated with a variety of adverse respiratory outcomes, such as increased risk of respiratory-related hospital admissions, reduced lung function, pneumonia, and asthma exacerbation (Anenberg et al., 2022; Doiron et al., 2019; MacIntyre et al., 2014; Mills et al., 2014). Concentrations of NO_2 and $PM_{2.5}$ often highly correlate with each other, which makes disentangling their individual effects complicated (Health Protection Agency (currently known as UKHSA), 2011; Eeftens et al., 2012). Recent evidence has suggested that the effect of ambient NO_2 exposure may extend beyond the respiratory system. For example, an increase in risk was found for cardiovascular-related hospital admissions, with the evidence from triggering of myocardial infarction being the strongest (CMO, 2022). It could also be associated with increased risk of cognitive decline, and the development of obesity in children (CMO, 2022; Zheng et al., 2024)

1.1.3. Ozone

Ozone (O_3) is naturally present in the stratosphere. In the troposphere, O_3 is formed from reactions between its precursor gases in the presence of with solar ultraviolet radiation, which could be NO_x, CO, and VOCs. Levels of O_3 are determined by the abundance of its precursor gases, as well as meteorological conditions such as temperature, wind speed, or sunlight intensity. For instance, in urban areas with high levels of traffic, the concentration of O_3 tends to be lower as it is rapidly depleted by NO, but it can be regenerated in downwind areas of such urban areas, resulting in a higher concentration. Hence, levels of O_3 could be seen as negatively correlating with concentrations of NO or NO_2 (USEPA, 2020).

O_3 encounters the ELF after inhalation in a manner to that of NO_2. It is a stronger oxidant than NO_2, thus share similar cascade of events that lead to the negative physiological changes mentioned previously in **section 1.1.2** (USEPA, 2013). The adverse effects of short-term O_3 exposure are most evident for respiratory outcomes, such as an increase in risk of airway hyperresponsiveness, asthma-related hospital admissions, and infections (CMO, 2022; Mumby et al., 2019). However, there is also evidence suggestive of an association between long-term ambient O_3 exposure and a variety of health endpoints, such as increased risk of cardiovascular mortality, type 2 diabetes mellitus (T2DM), or cognitive decline (Cleary et al., 2018; Liu et al., 2022a; Yu et al., 2023). Furthermore, the negative effects of O_3 pollution on ecological systems are extensive, as it can reduce vegetation, decrease the quality and yield of agriculture, and damage the photosynthetic processes of trees (Ashmore, 2005; USEPA, 2020; Wittig et al., 2007).

Tropospheric O_3 exhibits a strong diurnal pattern, and its concentration also depends on location and altitude (Williams et al., 2019). Different averaging metrics can be used for O_3. Some commonly used ones include 1-hour maximum, 8-hour maximum, and 24-hour averages. Therefore, depending on the metric used to average O_3 concentration, drastic differences in average concentrations and, thus, the magnitude of effect estimates can be found (Lange, 2018; Rombout et al., 1986).

1.2. Assessing air pollution exposure

1.2.1. The use of biomarkers

Air pollution exposure occurs when a human encounters an air pollutant. The products generated during these processes mentioned in **section 1.1** can be used as biomarkers of exposure. Measuring these biomarkers gives insight to the actual uptake of the pollutant and past exposures. However, biomarkers are not available for all pollutants. For pollutants with biomarkers available, the biomarkers may have other physiological relevance, be non-specific to the effect of one pollutant, or be specific to one pollutant, but have other routes of exposure (Zou et al., 2009; Suhaimi & Jalaludin, 2015). In addition, biomarker monitoring for a larger study population is labour-intensive, expensive for the research team, and intrusive and burdensome for study participants. Therefore, in epidemiological studies, ambient concentrations of air pollutants are measured and/or modelled and used in most cases.

1.2.2. Extrapolations from air quality monitor(s)

Ambient concentrations of air pollutants are primarily obtained from reference-grade monitors that can provide continuous measurements in high temporal resolution (WHO, 2023a). The concentrations are typically averaged over a particular time period, then assigned to a study participant according to the participant's distance from the monitor(s), either directly, or inverse distance weighting (IDW), which assumes that participants at locations closer to the monitor will have more similar pollutant concentrations than those at locations further away (Shepard, 1968).

While extensively used in earlier research into the health effects of air pollution exposure as they are easy to apply with readily available data, these methods have notable methodological shortfalls. They do not consider factors that could affect the distribution and, thus, concentration of pollutants, such as land use, topography, or atmospheric chemistry. By overlooking these factors, complex spatial patterns are not captured, and potential variations in concentrations are missed, especially when large interpolation distances are used (Sexton & Ryan 1988; WHO 2023a). This pitfall is less severe for regional pollutants with less spatial variations, such as particulate matter or O_3, but would be more severe for local-varying or traffic pollutants, such as NO_2, underlining the need for more advanced and comprehensive approaches in health research.

1.2.3. Modelled approaches

Modelled approaches were developed in light of the shortcomings of direct extrapolation from monitor(s), as it allows factors in addition to distance to be considered when the concentration of pollutants at a particular location is derived. Models can be roughly grouped into categories according to the primary data source used to develop the model, such as land-use regression models (LUR), chemical transport models (CTMs), or dispersion models. LUR involves the incorporation of land-use parameters such as distance to the nearest roads statistically into the model, dispersion models simulate the transport and dispersion of pollutants across the area, and CTMs incorporate complex atmospheric chemistry and physical processes affecting pollutant levels (Appel et al., 2017; Beelen et al., 2013; Johnson, 2022). In recent years, satellite remote sensing technologies have been adopted to estimate pollutant concentrations using aerosol optical depth (AOD) (van Donkelaar et al., 2010).

Despite the increasing complexity and comprehensiveness of the models, there are inherent limitations. They require validation from monitor concentrations and have their respective shortcomings. For instance, LUR models often have a limited ability to capture temporal

variations (Mercer et al., 2011). Such shortcomings can be tackled by the use of multiple models and data sources. Hence, this approach has increased in popularity (Baklanov & Zhang, 2020). Nonetheless, some limitations remain, such as the inability to capture the time-activity pattern of the participants, not being readily available for pollutants of emerging concerns, and the lack of information on exposures in indoor or workplace environments (Baxter et al., 2013).

1.2.4. Personal monitoring

To overcome the limitations shown above, there is growing interest in using personal monitors when assessing an individual's exposure to air pollution (Larkin & Hystad, 2017). Personal monitoring involves using lightweight, portable monitors to assess exposure to air pollutants over the study period in single or multiple measurements. Through the use of travel diaries or global positioning systems, the time-activity pattern and the corresponding exposure to air pollutants can be constructed. However, deploying personal monitors at a large scale has been limited for several reasons. First, it is more disruptive to study participants than monitored/modelled approaches, as the participants need to keep the monitors around them. The presence and use of the monitor by itself could alter the participants' behaviour and introduce biases. Second, a high cost is involved in procuring, maintaining, and retrieving the monitors. Third, there is a limit on how long the data can be collected before the monitors are retrieved, thereby limiting how long the study can be conducted. Last, using personal exposure introduces new confounders that could be difficult to deal with. For instance, women with trickier pregnancies with a higher risk of PTB would tend to stay home more often, potentially lowering the exposure estimate for those more susceptible to PTB (Jedrychowski et al., 2010; Saxton & Ryan, 1988; WHO, 2023a).

1.2.5. Combined approaches

As all the described approaches have their respective limitations, recent studies have attempted to overcome them by utilising several of these approaches. For instance, the **C**ognitive Deve**L**opment in the **U**rban **E**nvironment (CLUE) study have used modelled exposure for the whole study population. However, a subset of the population undertook biomonitoring and used personal monitors, which was subsequently used to validate the modelled exposure of the wider population (Thompson et al., 2024).

1.3. Preterm birth

Preterm birth (PTB) refers to the birth of an infant with a gestational age of less than 37 weeks. It is one of the adverse birth outcomes (ABOs) alongside low birth weight (LBW), small for gestational age (SGA), miscarriage, stillbirth, and more. The definition of the ABOs named is shown in **Table 1**. According to the WHO, PTB can be subcategorised into extremely PTB (less than 28 gestational weeks), very PTB (28 to 32 gestational weeks), and moderate-to-late PTB (33 to 36 gestational weeks) (WHO, 2023b). Overlaps between the ABOs are possible. Variation in definitions of ABOs is also present. For instance, the cut-off of gestational age for a birth to be classified as stillbirth instead of miscarriage is 24 weeks in the UK but 20 weeks in the US.

Table 1. Definition of the adverse birth outcomes (ABOs) mentioned.

Adverse birth outcomes	Definitions
Preterm birth (PTB)	Birth of an infant with fewer than 37 weeks gestation.
Low birth weight (LBW)	Birth of an infant with weight less than 2500g.
Small for gestational age (SGA)	Birth of an infant with weight less than 10th percentile for gestational age and sex (Royal College of Obstetrics & Gynaecologists, 2013)
Miscarriage	The natural, non-medically induced loss of pregnancy before the 24th week of gestation (National Institute of Health and Care Excellence (NICE), 2020)
Stillbirth	The delivery of a dead infant after 24 completed weeks of gestation (National Health Services (NHS), 2021)

1.3.1. Epidemiology and aetiology of preterm birth

Globally, around 1 in 10 babies are born preterm. The actual proportion could be higher, as the main burden of PTB comes from developing countries where the number of PTBs captured by health registers may not be representative of the population (Chawanpaiboon et al., 2019). Six countries (China, India, Indonesia, Nigeria, Pakistan, United States) account for over half of the PTBs (Walani et al., 2020). In the past decade, there were no noticeable changes in global rates of PTB. A fall in the total number of PTBs was found. However, it was likely attributable to the reduced total number of births in most regions (Ohuma et al., 2023).

Most preterm infants survive well into adulthood without major complications. However, they are at increased risk of conditions in early life, such as cerebral palsy, learning disabilities, and behavioural problems (Moster et al., 2008; Saigal & Doyle, 2008). In addition, evidence suggests that preterm birth is associated with an increased risk of mortality as well as various conditions in later life, ranging from greater decline in lung function, hypertension, and diabetes to cognitive decline and chronic kidney disease (Crump, 2020; Li et al., 2014; Morniorli et al., 2023; Parkinson et al., 2013). It demonstrates that PTBs can result in a more significant healthcare burden, and being born preterm can negatively impact the quality of life of individuals due to the increased likelihood of chronic health conditions.

A variety of factors can influence the risk of PTB. Some of the strongest risk factors are at an individual level, including smoking, amphetamine use, and isolated single umbilical artery (Kim et al., 2017; Ladhani et al., 2011; Liu et al., 2020). The risk of PTB associated with higher age and parity, albeit smaller than those mentioned earlier, is well-established (Fuchs et al., 2018; Koullali et al., 2018). In addition, pregnancy complications, such as gestational diabetes, preeclampsia, placenta previa, and premature rupture of membrane (PROM), commonly predispose to PTB (Norman et al., 2009).

The effect of social and environmental determinants and the risk of PTB are more complex. Socioeconomic factors, such as maternal education of less than 10 years, are associated with an increased risk of PTB, but there are uncertainties regarding the strength of the association (Morgen et al., 2008; Ruiz et al., 2015). In the UK, Black and Asian ethnic groups are shown to have higher rates of PTB compared to White ethnic groups, which have rates lower than the national average (Office of National Statistics (ONS), 2023). The pattern persists in the US, with non-Hispanic blacks shown to have a greater risk of PTB than non-Hispanic whites in the US (Manuck, 2017). Many environmental risk factors are associated with preterm birth, such as extreme heat, phthalates, lead, and cadmium exposures (Kumar et al., 2017; McElroy et al., 2022). However, it might be difficult for pregnant women or those preparing for pregnancy to take action to reduce exposure to certain environmental toxins (Etzel et al., 2020).

1.3.2. Air pollution exposure and the risk of PTB

1.3.2.1 Mechanistic evidence

As discussed in **section 1.1**, various physiological responses are elicited upon entry of air pollutants (or products of their reactions with physiological substrates) into the body. Four major interlinking pathways may lead from air pollutant absorption to PTB: inflammation, oxidative stress, epigenetic alterations, and endocrine disruption. However, to date, human

mechanistic studies have not been able to clearly map out how air pollution leads to these physiological changes and thus leads to PTB. Therefore, the findings are associative (Fussell et al., 2024). Moreover, similar to epidemiological evidence, less research has been conducted on gaseous pollutants (Fussell et al., 2024). A summary of the pathways is given below.

The entry of air pollutants into the body triggers inflammatory responses, leading to the release of proinflammatory cytokines and interleukins (IL), such as tumour necrosis factor-alpha (TNF-α) and IL-6 (Tripathy et al., 2021). Similarly, various processes in pregnancy involve changes in the immune response and releases of inflammatory biomarkers (Robertson et al., 2018). Hence, it is possible that exposure to air pollution causes disruption and leads to unfavourable inflammatory responses, resulting in systemic and intrauterine inflammation, or even early initiation of labour leading to PTB (Fussell et al., 2024).

At birth, preterm infants express higher levels of various biomarkers of oxidative stress in the umbilical cord blood, such as plasma malondialdehyde (MDA) and 8-hydroxy-2-deoxyguanosine (8-OHdG), compared to term infants (Martin et al., 2018). It suggests the presence of oxidative stress-related damage. Such damage related to air pollution could be attributable to disturbances in the balances of ROS and oxidants during pregnancy, leading to placental malperfusion and further damage to the developing foetus (Fussell et al., 2024). Systemic oxidative stress can also be induced by inhalation of air pollutants and particulate matter constituents, particularly trace metals that have oxidative properties (Vadillo-Ortega et al., 2014).

Substantial epigenetic changes occur during pregnancy to sustain the development of the foetus and prepare for delivery. DNA methylation is one of the key changes that regulate gene expression by adding methyl groups at the CpG nucleotide. Ambient air pollution exposure is linked to decreased global DNA methylation in the placenta, indicated by lower expression of long interspersed nuclear element-1 (LINE-1) or genes such as *HSD11B2* and *H19* (Andrawus et al., 2022; Fussell et al., 2024). Reduced expression of *HSD11B2* is associated with increased maternal glucocorticoid exposure, which could lead to growth inhibition, increased risk of intrauterine infection and trigger preterm delivery (Demendi et al., 2012).

Air pollution exposure is associated with endocrine disruption during pregnancy. It inversely correlates with cord blood thyroid-stimulating hormone (TSH) levels and alters the foetal thyroid hormone 3 and 4 (FT3/FT4) ratio (Janssen et al., 2017). Changes in TSH and FT3/FT6 levels are linked to reduced foetal growth and neurodevelopmental effects, which could trigger PTB (Howe et al., 2018).

1.3.2.2. Epidemiological evidence

Epidemiological research into the effect of air pollution exposure during pregnancy and the risk of PTB has flourished since the early 1990s (Bobak & Leon, 1999; Xu *et al*., 1995). Since then, the number of studies available has increased substantially (Stieb et al., 2012; Ghosh et al., 2021). However, the majority of the recent syntheses of the evidence were on particulate matter rather than gaseous pollutants (details in Chapter 3). One of the first epidemiological studies demonstrating an increase in the risk of PTB following gestational air pollution exposure was by Ritz and colleagues. Using a large cohort in California, they found an increase in the risk of PTB associated with last 6-week CO and PM_{10} exposure (Ritz et al., 2000). Since then, multiple studies have found an increase in the risk of PTB associated with higher air pollution exposure, with the findings for $PM_{2.5}$ and PM_{10} being most consistent (Ghosh et al., 2021; Klepac et al., 2018; Lamicchane et al., 2015; Stieb et al., 2012), but associations with NO_2, O_3, SO_2 and $PM_{2.5}$ constituents were also found (Brauer et al., 2008; Cassidy-Bushrow et al., 2020; Chen et al., 2018; Ju et al., 2021; Siddika et al., 2019). The size of the association found was mostly smaller than some of the strongest risk factors mentioned, such as smoking and amphetamine use. However, the studies still varied greatly in the exposure assessment method, choice of exposure window, statistical methods, confounder factors adjusted, and association size. The limitations of these studies are discussed in the following section. These could contribute to the inconsistent findings in systematic reviews and meta-analyses published in the area (details in **section 1.4.3** and **Chapter 3**).

1.3.2.3. General limitations of the epidemiological evidence

First, the effect of household or indoor air pollution is often not considered. Pregnant women spend most of their time at home (Yi et al., 2022), and their time spent at home increases as pregnancy progresses (Wu et al., 2013). As levels of pollutants in ambient and indoor environments are driven by different factors, the correlation between them depends on the presence of certain contributors of indoor sources of pollutants, such as gas cookers (Kornartit et al., 2010). In addition, pollutant levels in the household environment could be higher and of different compositions compared to the ambient environment (Leung, 2015). To the best of my knowledge, no studies to date have considered the effect of both types of air pollution on the risk of PTB. For those that have considered household air pollution, their limitations lie in the fact that the pollutant mixture in different household environments could be substantially heterogeneous, and notable shortcomings are present in the methodologies

used, such as the use of questionnaires or repeated sampling that lasts only for a short period (Pan et al., 2022; Shezi et al., 2022).

Second, there is a risk of both differential and non-differential exposure misclassification. While the effect of differential exposure misclassification on the risk estimate is unclear, non-differential exposure misclassification generally results in an underestimation. Non-differential exposure misclassification can occur when exposure is measured using (1) broad spatial resolution such as census tracts, which masks local variations in air pollution, or (2) relatively long exposure periods, for instance, the whole pregnancy that could last up to 42 weeks, thus not capturing the most relevant period for PTB risk which could be for a narrow, specific window. The use of annual averages would also lead to exposure after birth being included in a mother's exposure. Thus, the use of 9-month/whole pregnancy averages should be the minimum requirement. Differential exposure misclassification can occur due to residential mobility, which refers to a change of residential address during pregnancy. Birth certificates or registers, a common source of information for birth outcomes studies, often have limited information on residential mobility during pregnancy, as only the maternal residential address at birth is available (ONS, 2022). Research on this issue has obtained mixed findings, with some studies suggesting a minimal effect and others showing that the impact may depend on the distance between the old and new addresses (Hodgson et al., 2015; Lupo et al., 2010; Warren et al., 2017). Therefore, differential exposure misclassification may occur as residential mobility can vary by socioeconomic status, and there could be tendencies for individuals to move to less deprived areas that may have lower levels of pollution (Hodgson et al., 2009).

Third, some of the usual biases in perinatal epidemiological studies need to be considered, namely fixed cohort bias and live birth bias. Their definitions can be found in **Table 2**. PTB is an outcome which is determined by the length of gestation. Fixed cohort bias results in fewer PTB being found for the beginning and end of the cohort, making early as well as late exposure appear protective, especially for exposures that are time-varying and exhibit strong seasonal patterns e.g. O_3 (Strand et al., 2011; Neophytou et al., 2020). Live birth bias leads to a "depletion of susceptibles", where pregnant women with foetuses more susceptible to the effect of air pollution are less likely to be included in the study population (Raz et al., 2018). This results in a negative bias, as supported by simulation studies (Goin et al., 2021; Leung et al., 2021; Liew et al., 2015).

Fourth, constituents of $PM_{2.5}$, such as PAHs, can originate from sources other than air pollution. For instance, PAHs can be produced from grilling or barbecuing food and can enter the body by ingestion of food and water, and could be linked to PTB by inducing

oxidative stress and inflammation (El-Sikaily et al., 2023; Wang et al., 2024). Exposure to PAH is measured via PAH adducts in maternal blood (a measure of both oral and ambient exposure) or ambient levels of PAH. Therefore, it is difficult to ascertain the effect of PAH from ambient sources on the risk of ABOs. Studies investigating the impact of ambient and/or serum PAH levels on the risk of ABOs have yielded mixed findings, with inconsistencies in the method of exposure assessment and measure of association (categorical/continuous levels of exposure) (Freije et al., 2022; Jiang et al., 2022; MoghaddamHosseini et al., 2023).

Table 2. Biases in perinatal epidemiology.

Biases	Definition
Fixed cohort bias	When the population is defined based on infant's date of birth rather than the date of conception, the population at risk (pregnant women) at the beginning of the study will miss pregnancies with shorter gestation (as pre-term births occurred before the time period based on births started). In contrast, the population at risk at the end of the study will miss pregnancies with longer gestations (as the birth will not yet have occurred when the time period based on births stopped) (from **Appendix 7**) (Strand et al., 2011; Neophytou et al., 2021).
Live-birth bias	Most studies of birth outcomes use birth cohorts that are formed retrospectively, and thus were restricted to live births due to data availability. However, the true population-at-risk of the association (pregnant mothers) can have pregnancies resulting in late miscarriages or stillbirths, which are competing risks of PTB. Restricting to live births results in a "left truncation", as some pregnancies which are at risk of delivering a PTB were excluded, and thus not allowed to reach the timepoint where the risk can occur (Howards et al., 2007; Lisonkova & Joseph, 2015)

1.4. Evidence synthesis

1.4.1. Methods of evidence synthesis and causality determination in observational studies

Primary research addresses a specific research question by analysing original or primary data. Evidence synthesis analyses data from primary studies to provide a comprehensive overview of available epidemiological evidence within the scope of the research question. Using systematic approaches in data collection and analysis, it aims to summarise and show the current state of the evidence, identify gaps in knowledge and methodological limitations, and inform policy decisions and development of best practice guidelines (Karlsson & Takhashi, 2017). Evidence synthesis can be conducted using different methods, depending on the type of studies and time available, and can have varying degrees of comprehensiveness (**Table 3**).

Table 3. Examples of methods used in evidence synthesis (adapted from Cook et al., 2017).

Methods	Definitions
Vote counting	Summarise and count votes in favour of or opposing the research question
Narrative review	Qualitative review of the primary studies available of the research question
Rapid review	Rapid evaluation of the results from primary studies which may not necessary follow all the steps of a systematic review
Conceptual model	Showing the state of the evidence and how different types of evidence relate to each other within a system
Systematic map	Showing the state of the evidence in a matrix so gaps in evidence can be visualised
Systematic review	Summarise the primary studies available in a systematic, transparent, and reproducible manner
Meta-analysis	Quantitatively summarise the results from comparable primary studies to test if an overall direction of evidence is present

The traditional “hierarchy of evidence” favours randomised controlled trials (RCTs) over cohort studies, as randomisation of exposure/intervention is involved to allow comparison with standard practice or placebo (Wallace et al., 2022). Yet, a fundamental difference

between evidence synthesis in clinical medicine and air pollution epidemiology is the lack of RCT available in systematic reviews and meta-analyses that can directly address the question: is exposure to air pollution causing an adverse health outcome? The gap exists due to several reasons.

Air pollution exposure is ubiquitous, making randomisation of exposure impossible. Exposing RCT participants to high levels of air pollution is also unethical. Thus, available RCTs focus on interventions to reduce air pollution exposure, such as using air filters or replacing biomass cookstoves (Jack et al., 2015; Ulziikhuu et al., 2022). However, these RCTs demonstrated the efficacy of such intervention in reducing exposure and the associated health effects. It does not prove that a causal relationship exists between the exposure and the interest, which can only be achieved through a RCT of increased exposure that would raise ethical concerns due to the potential intention to harm. Even if a RCT of increased exposure was approved, air pollution exposure is determined by numerous factors ranging from proximity to sources, seasonal variations to time-activity pattern and mode of transport used. It is an enormous challenge to adjust for these factors to observe the effect of the intervention (for instance, controlled exposure in a chamber for 30 minutes every month). In addition, it would be difficult and impractical to conduct RCTs for outcomes that have long latencies, which some of the adverse outcomes of air pollution require, such as lung cancer.

Therefore, causality has to be inferred in air pollution epidemiology. Evidence synthesis in the area often does so through the use of causality criteria such as the Bradford Hill Criteria (**Table 4**) (Hill, 1965). However, there are notable shortcomings in the Bradford Hill Criteria. For instance, there is no definition for how “strong” a relationship has to be to satisfy the “strength” component, and specificity is very rare in environmental epidemiology (Shimonovich et al., 2021).

To overcome the limitations mentioned, various regulating agencies developed more comprehensive frameworks for certainty of evidence assessment, which requires expert knowledge and transparency in the judgment process. Some examples of this process include the Office of Health Assessment and Translation (OHAT) framework and the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) framework (Boogaard et al., 2023).

Table 4. The Bradford Hill Criteria (adapted from Fedak et al., 2015; Hill, 1965)

Component	Definition
Strength	The stronger the observed relationship between the exposure and the outcome, the more likely the relationship is causal
Consistency	The more the relationship is observed across different studies, locations, populations by different researchers, the more likely the relationship is causal
Specificity	The exposure is associated with the outcome but not others.
Temporality	The exposure precedes the outcome
Biological gradient	A dose-response relationship can be observed
Plausibility	Biological plausibility must be demonstrated for the relationship
Coherence	The relationship observed should be consistent with knowledge available
Experiment	There are presence of experimental evidence supporting the presence of a causal relationship
Analogy	The likelihood of the presence of a causal relationship increases if similar causes are known for similar outcomes

1.4.2. Systematic reviews and meta-analyses in environmental health

Systematic reviews and meta-analyses of observational studies are common methods for synthesising evidence in environmental health epidemiology. In the past decade, there has been a substantial increase in the number of systematic reviews and meta-analyses published, but they are often redundant and conflicting, raising concerns regarding the rigour, quality, and reproducibility of future syntheses of evidence (Ioannidis, 2016; Ioannidis, 2018).

As accumulating evidence suggests an adverse effect of air pollution on adverse birth outcomes (details in **section 1.3.2**), the UK's COMEAP set out to undertake a review of meta-analyses concerning how air pollution exposure during pregnancy affects the risk of adverse birth outcomes (ABOs), intending to find suitable CRFs for quantifying the associated health effects. The review discovered similar issues in the literature on the air pollution's effect on ABOs. The following section discusses some of the findings from the report.

1.4.3. Practices and shortcomings: using systematic reviews and meta-analysis of air pollution exposure and PTB as an example

The COMEAP report on adverse birth outcomes identified nine relevant meta-analyses (Guo et al., 2019; Lamichhane et al., 2015; Li et al., 2017; Liu et al., 2017; Klepac et al., 2018; Sapkota et al., 2012; Stieb et al., 2012; Sun et al., 2015; Zhu et al., 2015), of which one was subsequently retracted due to errors made in transformation of estimates (Guo et al., 2019). Some methodological issues were observed in the eight remaining meta-analyses. In Chapter 3 of this thesis, an updated and more in-depth analysis of the available systematic reviews and meta-analyses of gestational air pollution exposure and risk of PTB is conducted. The paragraphs below briefly summarise the findings from the COMEAP report.

First, there was low overlap in the primary studies used in the review and/or meta-analyses retrieved, and there was a lack of transparency in how primary studies were selected for the review and/or meta-analysis. Thus, it is impossible to determine if primary studies used in preceding SR/MAs but were missing in superseding ones had been actively excluded or had not been captured by the literature search. Therefore, more recent SR/MAs could not build on the findings of preceding work, and their conclusions were not based on a synthesis of all the evidence available (COMEAP, forthcoming).

Second, most meta-analyses did not provide adequate assessment of study dependencies. Study dependencies (referred to as population overlap in the report) occur when more than one primary study, conducted using overlapping populations, locations and time periods, is included in the same meta-analysis. It is not problematic in the narrative synthesis stage of systematic review. However, in a meta-analysis, it violates the assumption of independent observations (Deeks et al., 2022), leading to meta-analytical results that are biased towards the effect in locations, time periods, or populations studied more often.

Other methodological shortcomings were also present. For instance, some meta-analyses lack *a priori* protocol, and some used general quality assessment frameworks, which were not designed for studies of environmental health or specific studies of air pollution exposure and ABOs. It hinders effective and comprehensive critique of the available studies to be conducted (COMEAP, forthcoming).

Based on the findings, COMEAP recommended an up-to-date systematic review and meta-analysis for gestational air pollution exposure and the risk of adverse birth outcomes, including PTB, particularly for $PM_{2.5}$, PM_{10}, NO_2, and O_3. Some relevant methodological recommendations were also made and listed below:

a. Ensure transparency and reproducibility of the methodology and study selection process by providing clear records, e.g. the reasons for studies being excluded from the meta-analysis.
b. Develop and use tailored quality assessment frameworks to allow adequate critique of study quality.
c. Comprehensively assess study dependencies before pooling effect estimates.

These recommendations lay the foundation for the work conducted in this thesis, particularly **Chapter 4** and **Chapter 5**.

1.5. Applying findings from evidence synthesis

1.5.1. Steps of a health impact assessment

Health impact assessments (HIA) are conducted to quantify the impact of an intervention, e.g. a policy, on the health of a population. It helps inform decision-making by guiding resource allocation and compliance with regulations and understanding if certain populations are disproportionately affected (Lock et al., 2000). However, HIAs are not without limitations. First, although there are general methods of conducting a HIA, there is no standardised methodology, which can lead to difficulty comparing the findings to other works. Second, it is often hard to validate the predictions of a HIA or accurately measure its impact, as follow-ups or comparisons with the actual health impact are rarely, or in some cases, impossible to be conducted (Parry & Stevens, 2001).

After formulating the research question, HIA is conducted using the following steps (Hassan Bhat et al., 2021).

(1) ***Estimating the population-at-risk***. The population-at-risk is the population among those exposed that could develop the outcome of interest. For instance, the population-at-risk for childhood asthma incidences would be children aged 0-18. For HIA involving projections into future years, population projections (lifetables) are used to derive how the population-at-risk changes over the time period of interest.
(2) ***Air quality information and forecasting***. Air quality information can be obtained from air quality monitors or models at a global or regional level. If the HIA is about a reduction in air pollution, the reduction would then be incorporated into the forecasts for air quality in future years.
(3) ***Choosing the appropriate concentration-response function (CRF).*** CRFs can be from high-quality meta-analyses or cohort studies if results from meta-analyses are

unavailable or cohort studies are large and can provide CRFs more specific to the location and population of interest. For the health effects of air pollution, CRFs are usually expressed as a proportional increase in the incidence/prevalence of a health outcome for a given increase in concentration.

(4) ***Estimating the baseline rate of the outcome of interest.*** The baseline rates of an outcome could be obtained directly from sources such as government statistical offices. Alternatively, the baseline numbers can be calculated from the baseline rate per unit population combined with the population at risk exposed to air pollution change.

(5) ***Presenting the results of the HIA.*** Results of the HIA can be presented using different metrics. The most direct way of presenting it would be population attributable fraction (PAF), i.e. the fraction of the outcome of interest attributable to air pollution, which could be expressed using the formula (1), where p is the fraction of the population exposed, and RR is the relative risk (indicated by the CRF). The number of cases attributable to air pollution can then be derived using formula (2).

$$(1): PAF = \frac{p(RR-1)}{p(RR-1)+1}$$

$$(2): Number\ of\ cases\ (C) = PAF \times baseline\ rate \times population\ at\ risk$$

For mortality or morbidity outcomes with case-fatality risk available, the results can also be expressed using disability-adjusted life years (DALYs), which is the sum of years of life lost due to premature death (YLL) and years of life lived with disability (YLD), derived using formula (3) and (4) respectively.

$$(3): YLL = D \times LE$$

$$(4): YLD = C \times Dw \times Y$$

D refers to the number of deaths resulting from the outcome of interest, LE refers to the life expectancy at the age of death in years, Dw refers to the disability weights assigned to the outcome, Y is the years of life the individual is expected to live with the disease.

(6) ***Evaluating the economic consequences of the health impact.*** Economic consequences of the health impacts can be quantified using the willingness-to-pay (WTP) method or the quality-adjusted life year (QALY) method. For WTP, by estimating the WTP of people using methods such as choice experiments, WTP for reducing fatality and non-fatal consequences can be derived. The WTP to avoid fatality is referred to as the value of a statistical life (VSL), whereas the WTP to avoid

non-fatal consequences is the value of a statistical case (VSC) (Andersson et al., 2020; Robinson et al., 2022). QALY is the quantification of life year in the range of 0 to 1, where 1 and 0 indicate perfect health and death respectively. For mortality, years of life gained is calculated using lifetable analysis. For morbidities, improvements in quality of life are derived using the product of years lived with the condition and QALY weights (Whitehead & Ali, 2010).

1.5.2. Concentration-response functions (CRF) for the effects of air pollution

One of the ways findings from evidence synthesis of epidemiological studies of environmental health, in particular meta-analyses, can inform policymaking is to provide high-quality, robust, and up-to-date concentration-response functions (CRF). CRFs are not presented in isolation but rather with accompanying information and guidance for them to be properly used. For instance, the health outcome (could be specified by ICD-10 code), population-at-risk (defined by age, sex, or others), concentration range where the CRF applies, source of the CRF, and pollutant metric should be included (Héroux et al., 2015). **Table 5** shows an example of the CRF of the relative risk (RR) and its 95% confidence interval (95% CI) for annual $PM_{2.5}$ exposure and asthma incidences in children.

Table 5. CRF recommended for long-term $PM_{2.5}$ exposure and asthma incidences in children (adapted from Forastiere et al., 2024)

Concentration-response function (CRF) **RR (95% CI): 1.34 (1.10-1.63), per 10 μg/m³ increase in $PM_{2.5}$**	
Pollutant metric	$PM_{2.5}$, annual mean
Health outcome	Asthma in children (ICD-10 J45)
Population-at-risk	Individuals aged 18 or below
Mean exposure range	5-38 μg/m³
Source	Khreis et al. (2017)

The evidence on mortality-related outcomes such as all-cause mortality is more established than morbidity outcomes (details in **section 1.1** and **1.3**), with a greater likelihood of a causal effect (using criteria discussed in **section 1.4.1**). Therefore, in the recommendations by COMEAP, USEPA, or WHO for CRFs quantifying the health effects of air pollution, higher confidence is usually assigned to mortality-related outcomes compared to morbidities (**Table 6**). The core analysis for HIAs thus often does not consider morbidity outcomes. However, including only the impact on mortality but not morbidity outcomes will likely underestimate

the health impacts, as air pollution is shown to be linked with a wide range of health effects which may not lead to premature mortality but lead to lower quality of life. Therefore, some morbidities with emerging evidence are recommended for sensitivity analysis to estimate its potential impact on the conclusion of the HIA (COMEAP, 2023).

Table 6. Examples of different levels of recommendation by COMEAP regarding CRFs for exposure-outcome pairs (COMEAP, 2023).

Exposure-outcome pair	Recommendation
Long-term $PM_{2.5}$ exposure – all cause mortality	Recommended for core analysis
Long-term PM_{10} exposure – chronic bronchitis	Recommended for sensitivity analysis
Long-term O_3 exposure - mortality	Not recommended

For reproductive outcomes such as PTB, the CRFs available from meta-analyses are not recommended for use so far. For instance, the USEPA classified the association between $PM_{2.5}$ exposure and ABOs as "suggestive of, but insufficient to infer, a causal relationship" (USEPA, 2019). Similarly, the COMEAP's summary of quantifications and WHO Health Risk of Air Pollution in Europe (HRAPIE) project did not include ABOs (COMEAP, 2023; WHO, 2013b). Therefore, with the accumulating evidence suggesting an association between air pollution exposure and PTB (details in **section 1.3.2**), depending on the results of an up-to-date meta-analysis, it may be warranted to include PTB in a health impact assessment as a core or sensitivity analysis. The economic consequences for outcomes with suggestive evidence could be more significant than those with established evidence due to a greater impact on quality of life (McGartland et al., 2017). The following section considers areas of policy to which a health impact assessment including the effects of air pollution on PTB might be applied.

1.6. Policies linking health impacts of air pollution reduction and climate change

1.6.1. UK's commitment to achieve Net-Zero by 2050

The Paris Agreement is a landmark United Nations treaty aiming to limit the global average temperature increase to well below two degrees Celsius, respond to adverse effects of climate change, and provide adequate financing for such efforts (United Nations Framework Conventions on Climate Change, 2015). It introduces the concept of nationally determined contributions (NDCs), where countries must develop plans for climate change mitigation and

adaptation. The review of NDCs submitted by countries found that despite significant progress in integrating health actions into climate targets, there is a pressing need to strengthen the evidence base for implementing climate mitigation and adaptation policies based on health effects (WHO, 2023c).

In 2019, the UK Government pledged to achieve net-zero emissions by 2050, implying that the UK will have to drastically reduce emissions to the extent that greenhouse gases (GHGs) emitted will be equal to or less than GHGs removed from the atmosphere by 2050 (HM Government, 2021). Subsequently, the UK Government submitted its NDCs which were subsequently updated in 2022, using recommendations from the Sixth Carbon Budget produced by the UK's Climate Change Committee (CCC) (CCC, 2020; Department for Business, Energy & Industrial Strategy (BEIS), 2022). The update shows the UK's multisectoral strategy in achieving its NDC targets, including but not limited to industrial decarbonisation, deployment of heat pumps and electrification of heating in the building sector, encouraging electric vehicle uptake, and increase in the use of renewable energy (BEIS, 2022).

1.6.2. Existing evidence on the effects of air pollution reduction through climate change policies in the UK

Using information regarding how climate change policies will affect emissions and, hence, future air quality from sources such as IPCC reports, governing bodies, or the Sixth Carbon Budget, researchers have begun to forecast how future air quality will change accordingly and model the corresponding health impacts.

The evidence was based on policies published in different years and at the national or global levels. Some of the earlier commitments at the national level being analysed include the UK's Climate Change Act (CCA), published in 2008 (Williams et al., 2018). More recent strategies include the Department of Transport (DoT) Road to Zero strategy published in 2018 (Mehlig et al., 2021). At a multinational or global level, analysis using the IPCC Representative Concentration Pathways (RCP), Near-Term Climate Forcers mitigation scenarios (NTCFs), and NCDs were conducted (Fenech et al., 2021; Hamilton et al., 2021; Markandya et al., 2018; Turnock et al., 2023).

More recent evidence was able to utilise the Sixth Carbon Budget to analyse the health impacts, which is the most updated recommendation by the CCC for the UK Government (ApSimon et al., 2023; Baldo et al., 2023; Dajnak et al., 2022; Milner et al., 2023a; Milner et al., 2023b). However, most of the studies quantified the health impacts using only mortality (number of deaths avoided), based on premature mortality or a limited range of morbidities

such as cancer diagnosis and heart diseases. As discussed in section 1.5, this would likely underestimate the health impacts. Hence, some recent studies have attempted to include a broader range of morbidities associated with air pollution (Dajnak et al., 2023; Walton et al., forthcoming).

The interest in including children's health outcomes, such as asthma incidences and ABOs, is also emerging (Milner et al., 2023a). This is because the current morbidity outcomes used, such as diabetes, heart disease, dementia and COPD, are more relevant for the old. Thus, health outcomes for the younger populations are overlooked. Yet, children's health outcomes could have lifelong consequences, and avoiding associated adverse endpoints could have substantial health and economic benefits (Perera et al., 2020)

1.6.3. Opportunities from a UK project

The UK's air quality health and economic costs and benefits of a zero-carbon pathway (Net-Zero project) is a project funded by the National Institute of Health and Care Research (NIHR) aiming to understand the health and economic impacts of several policy scenarios of climate mitigating measures in 2030, 2040, and 2050 across different groups up to 2154 (NIHR, 2020). How air pollution levels would change in different policy scenarios was modelled using established, high-resolution air quality models, and a variety of mortality and morbidity outcomes were included using ward-level incidence data (Beevers et al., forthcoming; Walton et al., forthcoming). Nonetheless, ABOs such as PTB were not included in the list of morbidities. The UK also has data on birth outcomes at relatively fine spatial and temporal resolution (NHS Digital, 2024). Thus, this offers a precious opportunity to test the influence of incorporating PTB as one of the morbidity outcomes in the health impact calculations, if an association between air pollution exposure and PTB is confirmed.

1.7. Gaps in the literature and research needs

The preceding sections showed a clear gap in the literature, which could be summarised as follows:

(1) Methodological shortcomings in the available synthesis of evidence of gestational exposure and the risk of PTB.
(2) Lack of appropriate and robust CRF to be used for quantifying the effect on the number of PTBs.

(3) PTB (as well as children and birth outcomes in general) are not considered in available health impact assessments of the effect of air pollution reduction associated with climate change.

(1) and (2) are interconnected. To address the existing knowledge gap, the methodological shortcomings in the available evidence must be understood clearly. The lack of an appropriate and robust CRF implies that a new systematic review and meta-analysis needs to be conducted. Thus, methodological improvements, which are based on the analysis of shortcomings of preceding works, are required to ensure that the new systematic review and meta-analysis are robustly conducted. When (2) is achieved, then it can be applied to health impact assessments to quantify the effect of air pollution reduction policies on PTBs.

Chapter 2. Aims and objectives

There is emerging evidence on the effect of gestational air pollution exposure and the risk of PTB. However, no high-quality, rigorously conducted systematic review and meta-analysis is available to provide a robust concentration-response function (CRF) to be used in health impact assessments. Climate change mitigation and adaptation policies, which have widespread health, economic, social and ecological impacts, offer a unique opportunity to use the CRF to quantify the impact of associated air pollution reduction on PTB and the corresponding economic evaluation. Therefore, the aims of this thesis are as follows:

1. Conduct a high-quality, rigorous, up-to-date systematic review and meta-analysis of gestational exposure to $PM_{2.5}$, NO_2, O_3 and the risk of PTB.

2. If a positive association is found, conduct a health impact assessment in England using the concentration-response function (CRF) derived from (1) and evaluate the economic consequences.

The specific objectives to achieve the two aims listed above are shown below.

Objectives for aim (1)

1. Review how previous systematic reviews and/or meta-analyses were conducted to understand the practices, shortcomings, and potential ways to overcome them.
2. Develop methods to overcome the shortcomings found in (1) to be incorporated in the subsequent objectives.
3. *A priori,* develop a systematic review protocol and submit it to PROSPERO.
4. Conduct a systematic literature search for studies of gestational exposure to $PM_{2.5}$, NO_2, O_3 and the risk of PTB.
5. Systematically review studies retrieved, including data extraction and quality assessment.
6. Conduct a narrative synthesis of studies retrieved from the literature search focusing primarily on those not eligible to be included in the meta-analysis.
7. Conduct meta-analyses to quantify the relationship between gestational exposure to $PM_{2.5}$, NO_2, O_3 and the risk of PTB.
8. Investigate the potential drivers for heterogeneity in the meta-analyses conducted.
9. Discuss the reliability and possibility of causality of the evidence to determine if the findings, if any, are suitable to be incorporated in a health impact assessment.

Objectives for aim (2)

1. Identify and extract the data sources (air quality forecast, baseline rate, population-at-risk, CRF) required for conducting the health impact assessment in England.

2. Carry out the health impact assessment to derive the number of PTBs averted associated with air pollution reduction from climate change policies.

3. Quantify the economic consequences associated with the number of PTBs averted.

4. Conduct sensitivity analyses to determine if different sources of economic cost data will drastically change the final evaluation.

Chapter 3. Analysis of preceding systematic reviews and meta-analyses

3.1. Aim and scope

The work presented in this Chapter serves to achieve the first objective of aim (1). It outlines the methods for retrieving the previous systematic reviews and/or meta-analyses (SR/MAs) and analyses them to understand their general practices and methodological shortcomings. It is not an umbrella review which analyses the characteristics of all the preceding SR/MAs, and the associations found between the different pollutants and PTBs, as it is beyond the scope of this thesis and has been conducted by others (COMEAP, 2018b; Nyadanu et al., 2022). The forthcoming COMEAP report has conducted an analysis on 9 previous meta-analyses (described in **section 1.4.2)**. This analysis was based on the broad findings from the report, namely transparency in reporting, quality assessment, and study dependencies. Afterwards, the results of the analysis that informed the methods used in the following chapters are described. Lastly, a conclusion was made regarding their quality, if any of the preceding work was suitable for providing a concentration-response function (CRF), and the recommendations for the following chapters.

3.2. Methods

A systematic literature search was conducted to retrieve relevant studies on gestational exposure to $PM_{2.5}$, NO_2, and O_3 and the risk of PTB published from any time up to 22nd April 2022. Details of the methods are available in Chapter 4. By filtering the search results from the literature search from Ovid Medline, the previous SR/MAs conducted were retrieved. An additional search was conducted on May 2024 online to find SR/MAs published after the cut-off date. Afterwards, the overall characteristics of the retrieved SR/MAs were described, and then their transparency in reporting, quality assessment methods, and assessment of study dependencies were investigated.

3.2.1. Coverage of available literature and reporting of study selection process

The retrieved SR/MAs were divided into two types: The first is the ones that aim to retrieve all relevant primary studies, and the second type is those restricted to certain groups of

studies. Subsequently, the list of studies included among SR/MAs of the first type, specifically those that conducted both systematic review and meta-analysis (SR+MA), were compared. While search strategies of systematic reviews usually strive to retrieve all available eligible studies, meta-analyses may be more restrictive as not all primary studies would satisfy the requirements of being in a meta-analysis. Whether these SR+MAs were able to use all available studies was explored by comparing the list of primary studies retrieved by earlier and more recent SR+MAs. Then, if the issue persists in meta-analyses, it was investigated by comparing the primary studies included in the meta-analyses published in a similar time period.

3.2.2. Quality assessment

The retrieved SR/MAs were divided according to their broad characteristics: self-set criteria, score-based approaches, and domain-based approaches. Then, the extent of reporting and the ability to critique studies of gestational air pollution and the risk of PTB were analysed and discussed.

3.2.3. Study dependencies

The definition of study dependency is available in Chapter 1 section 1.4.3. The strategies for dealing with study dependencies adopted by the SR/MAs retrieved were extracted and analysed. Then, examples found from preceding SR/MAs, as well as the forthcoming COMEAP adverse birth outcomes report were discussed to understand if the existing strategies could be improved.

3.3. Results and discussion

3.3.1. Overall characteristics

After filtering the result of the systematic literature search to only literature reviews, 115 records were retrieved, and subsequent screening retrieved 23 SR/MAs published in the area prior to the cut-off date (22nd April 2022). The additional search retrieved seven additional SR/MAs published after the cut-off date, leading to a total of 30 SR/MAs being analysed. Their characteristics are shown in **Table 7**. The earliest SR/MA was published in 2004, with December 2001 as the cut-off of their literature search (Glinianaia et al., 2004). There was a six-year gap before the second SR/MA was published (Bonzini et al., 2010). However, the number of SR/MAs published rose rapidly since then, with 5 SR/MAs having a cut-off date in 2015. The past SR/MAs searched four databases on average (range: 1 to 9). Nine SR/MAs have conducted searches in both Chinese and English, and the rest only searched for English literature. All SR/MAs were published in English except one, which was published in Chinese (Chen et al., 2016). Ten meta-analyses were published without a systematic review being conducted, and ten were systematic reviews without a meta-analysis. $PM_{2.5}$ is most studied (n = 24), followed by PM_{10} (n = 14), and NO_2 (n = 11).

A variety of quality assessment methods were used, most commonly the Newcastle Ottawa Scale (NOS) (Wells et al., 2013) (n = 13), either on its own (n = 9) or with other quality assessment tools (n=4), such as the Agency for Healthcare Research and Quality (AHRQ) risk of bias assessment tool (Viswanathan et al., 2013). Three studies used criteria set by the authors themselves, and five did not include any quality assessment of the literature in the published article.

Twelve studies did not incorporate any strategies to reduce dependencies of studies included in the same meta-analysis. Ten studies have provided a rationale for study selection when more than one study overlaps temporally, geographically, or used the same population. Only one SR/MA requested additional information from the authors (Stieb et al., 2012).

Some general methodological issues were found. Li et al. (2017) conducted a separate analysis for studies using per IQR and 10 μg/m^3 increase, but effect estimates per IQR increment can be converted to a per 10 μg/m^3 increment. Similarly, Wang et al. (2023) only included studies that provided estimates in per 10 μg/m^3 increment. Thus, these methodological choices reduced the total number of studies available for the meta-analysis. Hung et al. (2023) included only studies comparing exposures in the highest vs. lowest category. The effect estimates from these primary studies are likely incomparable due to

different pollutant distributions. Thus, the summary estimate pooled from these studies likely offers little insight.

Most SR/MAs have searched the literature databases for articles published from any time up to the cut-off date (shown in **Table 7**), but there were three that restricted the search by start date (Bonzini et al., 2011; Khader et al., 2016; Yuan et al., 2019), and two that did not provide a clear start and cut-off date of the literature search (Dzekem et al., 2024; Liu et al., 2017). No justification was provided by these studies.

Table 7. Characteristics of the systematic reviews (and/or meta-analyses) retrieved.

Study	Cut-off*	Search language	SR	MA	Pollutant(s)	Quality assessment method	Tackled study dependencies?	Notable findings
Glinianaia et al., 2004	31/12/2001	English	✓	✗	PM_{10}, $PM_{2.5}$, TSP, dust	NS	N/A	Retrieved 3 studies on PTB. Findings were inconsistent with small magnitude of increase in risk found.
Bonzini et al., 2010	31/12/2008	English	✓	✗	PM_{10}, $PM_{2.5}$, NO_2, O_3, CO	N/A	N/A	Retrieved 8 studies on PTB. ORs were extracted but not described.
Sapkota et al., 2012	31/07/2009	English	✗	✓	PM_{10}, $PM_{2.5}$	NS	✗	Increase in risk of PTB associated with per 10 $\mu g/m^3$ increase in whole pregnancy PM_{10} exposure (OR: 1.15 [1.14-1.16]) (n=8)
Shah et al., 2011	15/10/2010	English	✓	✗	PM_{10}, $PM_{2.5}$, NO_2, SO_2, O_3, CO, TSP	Self-defined criteria	N/A	Increase in risk of PTB found for SO_2
Stieb et al., 2012	31/01/2011	English	✓	✓	PM_{10}, $PM_{2.5}$, NO_2, SO_2, O_3, CO	Self-defined criteria	✓	Increase in risk of PTB associated with per 10 $\mu g/m^3$ increase in trimester 3 PM_{10} (OR: 1.06 [1.03-1.11]) (n=7) and whole pregnancy $PM_{2.5}$ (OR: 1.16 [1.07-1.26]) (n=4) exposure.
Khader et al., 2016	31/12/2014	English	✓	✗	NS	Self-defined criteria	N/A	Retrieved 1 study on air pollution and risk of PTB in the Eastern Mediterranean Region, hence insufficient evidence
Lamichhane et al., 2015	30/04/2015	English	✗	✓	PM_{10}, $PM_{2.5}$	Down & Black, 1998	✓	Increase in risk of PTB associated with per 10 $\mu g/m^3$ increase in whole pregnancy PM_{10} (OR: 1.23 [1.04-1.41]) (n=3) and $PM_{2.5}$ (OR: 1.14 [1.06-1.22]) (n=7) exposure.

Sun et al., 2015	31/12/2015	English, Chinese	✗	✓	$PM_{2.5}$	NOS	✓	Increase in risk of PTB associated with per 10 μg/m^3 increase in whole pregnancy $PM_{2.5}$ exposure (OR: 1.13 [1.03-1.24]) (n=13). Significant heterogeneity observed. Association remained in retrospective studies, prospective studies, and studies in the US.
Zhu et al., 2015	01/03/2015	English	✗	✓	$PM_{2.5}$	N/A	✓	Increase in risk of PTB associated with per 10 μg/m^3 increase in whole pregnancy $PM_{2.5}$ exposure (OR: 1.10 [1.03-1.18]) (n=8).
Chen et al., 2016	01/06/2015	English, Chinese	✗	✓	PM_{10}, $PM_{2.5}$, NO_2, SO_2, CO	NOS	✗	Increase in risk of PTB found for NO_2, PM_{10}, and SO_2 exposure
Jacobs et al., 2017	31/12/2015	English, Chinese	✓	✗	PM_{10}, $PM_{2.5}$, NO_2, SO_2, O_3, CO	Berman & Parker, 2002	✗	Retrieved 11 studies on PTB in China. Increase in risk mainly found for SO_2
Klepac et al., 2018	30/06/2016	English	✗	✓	NS	NOS	✗	Increase in risk of PTB associated with per 10 μg/m^3 increase in whole pregnancy PM_{10} (OR: 1.09 [1.04-1.16]) (n=6) and $PM_{2.5}$ (OR: 1.24 [1.08-1.41]) (n=13) exposure, and per 10 ppb increase in O_3 exposure (1.03 [1.01-1.06]) (n=5).
Li et al., 2017	31/07/2016	English, Chinese	✓	✓	$PM_{2.5}$	NOS & AHRQ	✓	Separate analysis for studies presented per 10 μg/m^3 and IQR increase. Increase in risk of PTB associated with per 10 μg/m^3 increase in whole pregnancy $PM_{2.5}$ exposure (OR: 1.03 [1.01-1.05]) (n=8) for studies using IQR.
Liu et al., 2017	NS	English, Chinese	✗	✓	$PM_{2.5}$	NOS	✗	Increase in risk of PTB associated with per 10 μg/m^3 increase in trimester 1 $PM_{2.5}$ exposure (OR: 1.15 [1.05-1.24]) (n=9).

Yuan et al., 2019	31/12/2017	English, Chinese	✓	✗	$PM_{2.5}$	N/A	N/A	Retrieved 18 studies on PTB. Increase in risk found in half of them.
Melody et al., 2019b	31/01/2018	English	✓	✗	NS	NOS	N/A	Some evidence for association between acute air pollution exacerbation events and increased risk of PTB
Thayamballi et al., 2021	30/06/2018	English	✓	✓	PM_{10}, $PM_{2.5}$, PM_c	N/A	✓	Explored the effect modification of race/ethnicity. No consistent association found
Bekkar et al., 2020	30/04/2019	English	✓	✗	$PM_{2.5}$, O_3	N/A	✗	$PM_{2.5}$ or O_3 exposure was associated with increased risk of PTB in 19 out of 24 studies
Health Effects Institute (HEI), 2022	31/07/2019	English	✓	✓	TRAP	Modified WHO risk of bias assessment instrument (WHO, 2020)	✓	Low confidence in presence of association. Increase in risk of PTB associated with per 10 $\mu g/m^3$ increase in trimester 3 NO_2 exposure (RR: 1.05 [1.02-1.08]) (n=6).
Luo et al., 2021	31/12/2019 (from 2007)	English, Chinese	✓	✗	NS	NOS	N/A	Used studies that have used land-use regression model to assess exposure. Only 1 out of 7 study reported increase in risk of PTB
Simoncic et al., 2020	04/01/2020	English	✓	✓	$PM_{2.5}$, NO_2	Doi & Thalib, 2008	✗	Increase in risk (that was not statistically significant) from NO_2 exposure found

Ju et al., 2021	31/10/2020	English	✗	✓	PM_{10}, $PM_{2.5}$, NO_2, SO_2, O_3, CO, NO_x	NOS	✗	$PM_{2.5}$: Increase in risk of PTB associated with per 10 μg/m^3 increase in trimester 3 (RR: 1.03 [1.00-1.07]) (n=23) and whole pregnancy (RR: 1.07 [1.05-1.10], n=31) exposure. NO_2: Increase in risk of PTB associated with per 10 μg/m^3 increase in trimester 3 (RR: 1.07 [1.03-1.10]) (n=15) exposure. O_3: Increase in risk of PTB associated with per 10 ppb increase in trimester 1 (RR: 1.04 [1.02-1.05]) (n=11), trimester 2 (RR: 1.02 [1.00-1.04]) (n=8), trimester 3 (RR: 1.04 [1.01-1.07]) (n=8), and whole pregnancy (RR: 1.03 [1.02-1.05]) (n=11) exposure.
Rappazzo et al., 2021	31/01/2021	English	✓	✓	O_3	Modified OHAT framework (Rooney et al., 2014)	✓	Increase in risk of PTB associated with per 10 ppb increase in O_3 exposure in trimester 1 (RR: 1.06 [1.03-1.10]) (n=17) and trimester 2 (RR: 1.05 [1.02-1.08]) (n=15).
Ghosh et al., 2021	04/04/2021	English	✗	✓	$PM_{2.5}$	AHRQ, OHAT	✓	12% greater risk found per 10 μg/m^3 increase in whole pregnancy $PM_{2.5}$ exposure (1.12, 95% UI: 1.07-1.16)
Ju et al., 2023	31/10/2021	English, Chinese	✗	✓	$PM_{2.5}$, PM_{10}	NOS	✗	Increase in risk of PTB associated with per 10 μg/m^3 increase in $PM_{2.5}$ exposure in trimester 2 (RR: 1.03 [1.01-1.06]) (n=25), trimester 3 (RR: 1.03 [1.01-1.06]) (n=24), and whole pregnancy (RR: 1.09 [1.06-1.12]) (n=28).

Yu et al., 2022	31/03/2022	English	✓	✓	$PM_{2.5}$, PM_{10}	NOS, OHAT	✓	Increase in risk of PTB associated with per 10 μg/m^3 increase in $PM_{2.5}$ exposure in trimester 2 (RR: 1.02 [1.00-1.04]) (n=38), trimester 3 (RR: 1.02 [1.01-1.03]) (n=31), and whole pregnancy (RR: 1.08 [1.06-1.11]) (n=43). Increase in risk for PTB subgroups in trimester 2
Wang et al., 2023	30/04/2022	English, Chinese	✓	✓	PM_{10}, $PM_{2.5}$, NO_2, SO_2, O_3	NOS	✗	Increase in risk of PTB associated with per 10 μg/m^3 increase in whole pregnancy $PM_{2.5}$ exposure (RR: 1.04 [1.01-1.07]) (n=38) and trimester 3 O_3 exposure (RR: 1.01 [1.01-1.02]) (n=4)
Hung et al., 2023	26/07/2022	English	✓	✓	$PM_{2.5}$, NO_2	NOS + AHRQ	✗	Used studies that dichotomised exposures NO_2: OR=1.16 (1.04-1.26), highest vs. lowest category
Blanc et al., 2023	31/10/2022	English	✓	✓	PM_{10}, $PM_{2.5}$, NO_2, SO_2, O_3	NOS	✗	Retrieved 3 studies on PTB for preconception exposure. No consistent association found.
Dzekem et al., 2024	NS	English	✓	✗	NS	NIH, 2024	N/A	Retrieved 8 studies on the impact of race/ethnicity on risk of PTB in the US. Evidence suggests an increase in risk among Black and Hispanics compared to non-Hispanic Whites.

NIH: National Institute of Health
NS: Not shown
PM_c: coarse fraction of PM
TRAP: Traffic-related air pollutant (traffic-derived NO_2, PM_{10}, $PM_{2.5}$ and its constituents, CO, UFP, and other pollutants and traffic measures)
TSP: Total suspended particulates
* Cut-off date of the literature search

3.3.2. Coverage of available literature and reporting of study selection process

It was often unclear how primary studies retrieved were selected for the review and/or the meta-analysis. Among the 30 SR/MAs retrieved, 16 of them did not restrict to certain groups of studies, whereas 14 of them were restricted to a certain location (n=3), study design (n=3), maternal factors (n=2), exposure period (n=2), exposure assessment (n=1), exposure length (n=1), source of pollutant (n=1), and method of analysis (n=1) (**Table 8**). Among the 16 SR/MAs without restriction to specific groups of studies, there are three that were both systematic reviews and meta-analyses (SR+MAs) (Li et al., 2017; Stieb et al., 2012; Yu et al., 2022).

Among the 3 SR+MAs, Stieb et al. (2012) covered pollutants including $PM_{2.5}$, NO_2, and O_3, yet Li et al. (2017) and Yu et al. (2022) only covered $PM_{2.5}$. Moreover, most of the preceding SR/MAs published from 2017 onwards were restricted to certain groups of studies. While these SR/MAs can provide a more in-depth understanding of that particular group of studies, their findings must be interpreted in conjunction with the results without restriction to understand if they are generalisable to other studies, which are unavailable for some pollutants. Hence, it shows the need for a comprehensive update of the literature on gaseous pollutants such as NO_2 and O_3, in addition to $PM_{2.5,}$ which most primary studies were based on.

The primary studies of the 3 SR+MAs published before the cut-off date of Li et al. (2017) were compared (**Table 9**, page 60). These 3 SR+MAs were expected to build on findings from preceding publications by including all primary studies used previously. However, 8 of the 19 primary studies were absent in the review of either Li et al. (2017) or Yu et al. (2022), and neither publication offered any explanation. This implies that these studies were not retrieved or excluded in the study selection process but was not reported. Moreover, it indicates that preceding SR/MAs have paid little attention to the findings of preceding work, which is an issue also found in other fields of work (Helfer et al., 2015).

The same was observed among meta-analyses without restriction to certain groups of studies that were published within similar time periods, where they were often found to miss some primary studies in preceding meta-analyses. An example is shown in **Table 10** (page 61). For instance, Brauer et al. (2008) was a study excluded by Lamichhane et al. (2015) and not mentioned in Li et al. (2017), but was included in the meta-analysis of Stieb et al. (2012). It was unclear why the exclusion was made, or if the study was retrieved. A similar issue is evident in several other studies in the table. Stieb et al. (2012) contacted the authors

for additional information, and SR/MAs superseding the work would have captured Brauer et al. (2008) if Stieb et al. (2012) had been examined.

Table 8. Classification of preceding SR/MAs retrieved according to their aims.

For all types of studies	For certain groups of studies	
Glinianaia et al. (2004) Bonzini et al. (2010)* Sapkota et al. (2012) Shah et al. (2011) Stieb et al. (2012) Khader et al. (2016) Lamicchane et al. (2015) Sun et al. (2015) Zhu et al. (2015) Chen et al. (2016) Klepac et al. (2018) Li et al. (2017) Liu et al. (2017) Bekkar et al. (2020) Ghosh et al. (2021) Yu et al. (2022)	By location	Jacobs et al. (2017) (China) Simoncic et al. (2020) (Europe) Wang et al. (2023) (China)
	By study design	Yuan et al. (2019) (Cohort studies) Ju et al. (2021) (Cohort studies) Ju et al. (2023) (Cohort studies)
	By exposure assessment method	Luo et al. (2021) (Land-use regression model)
	By exposure length	Melody et al. (2019) – Short/Medium exposure
	By exposure period	Rappazzo et al. (2021) – Trimester 1 and 2 Blanc et al. (2023) - Preconception
	By pollutant source	HEI (2022) – Traffic-related air pollution (TRAP)
	By maternal factors	Thayamballi et al. (2021) – SES/Ethnicity Dzekem et al. (2024) - Ethnicity
	By analysis method	Hung et al. (2022) – Dichotomous evaluation

Systematic reviews are highlighted in blue, where meta-analyses are highlighted in purple.
Studies in green with highlight: systematic review and meta-analysis. *: Studies that covered <5 years of literature

Several things can be considered to overcome the shortcomings shown. First, despite the high number of preceding SR/MAs available, most recent SR/MAs were restricted to certain groups of studies. Hence, a comprehensive update of the literature is needed, particularly for gaseous pollutants such as NO_2 and O_3. Second, the literature search conducted in the coming chapter must be able to retrieve all relevant primary studies. One way to ensure this is to cross-check the search results with the list of primary studies included in preceding publications. Third, there must be a clear record of the reasoning behind why a study was excluded from the review while it is included in preceding publications, which could be in the form of supplementary materials as done in Blanc et al. (2023).

Table 9. The primary studies of $PM_{2.5}$ exposure retrieved in the three systematic reviews and meta-analyses (SR+MAs) without restriction to specific groups that were published within the literature cut-off date of Li et al. (2017).

Primary studies for $PM_{2.5}$ exposure and risk of PTB	SR+MAs		
	Stieb et al. (2012)	Li et al. (2017)	Yu et al. (2022)
Wilhelm & Ritz (2005)*	✓	✗	✓
Huynh et al. (2006)	✓	✓	✓
Jalaludin et al. (2007)*	✓	✗	✓
Ritz et al. (2007)	✓	✓	✓
Brauer et al. (2008)*	✓	✗	✓
Wu et al. (2009)*	✓	✓	✗
Gehring et al. (2011a)*	✓	✓	✗
Kloog et al. (2012)		✓	✓
Pereira et al. (2014a)		✓	✓
Pereira et al. (2014b)		✓	✓
Fleischer et al. (2014)*		✓	✗
Hannam et al. (2014)		✓	✓
Hyder et al. (2014)		✓	✓
Gray et al. (2014)		✓	✓
Huang et al. (2015)		✓	✓
Stieb et al. (2016a)		✓	✓
Stieb et al. (2016b)*		✓	✗
Pereira et al. (2016)*		✗	✓
Lavigne et al. (2016)		✓	✓

*Primary studies not retrieved in 1 of the 3 SR+MAs
Grey highlight: studies published after the cut-off date of Stieb et al. (2012)

Table 10. Primary studies included in 3 meta-analyses of whole pregnancy $PM_{2.5}$ exposure and the risk of PTB published with literature cut-off date within 3 years of each other.

Primary studies	Time period	Location	Meta-analyses		
			Zhu et al., 2015	Sun et al., 2015	Li et al., 2017
Wilhelm & Ritz, 2005	1994 - 2000	LA, CA, US			
Huynh et al., 2006	1999 – 2000	CA, US			
Ritz et al., 2007	2003	LA, CA, US			
Jalaludin et al., 2007	1998 – 2000	SYD, AU			
Brauer et al., 2008	1999 – 2006	VBC, CAN			
Wu et al., 2009	1997 – 2006	LA+OC, CA, US			
Gehring et al., 2011	1996 – 1997	NL			
Rudra et al., 2011	1996 – 2006	WA, US			
Kloog et al., 2012	2000 – 2008	MA, US			
Lee et al., 2013	1997 – 2002	AC, PA, US			
Hyder et al., 2014	2000 – 2006	CT+MA, US			
Fleischer et al., 2014	2004 – 2008	Global	/		
Pereira et al., 2014a	2000 – 2006	CT, US			
Pereira et al., 2014b	1997 – 2007	PER, AU	/		
Ha et al., 2014	2004 – 2005	FL, US	/		
Gray et al., 2014	2002 – 2006	NC, US			
Hannam et al., 2014	2004 - 2008	NW, ENG			
Stieb et al., 2016a	1999 – 2008	CAN	/	/	
Lavigne et al., 2016	2005 – 2012	ON, CAN	/	/	

/: Study was not published before the meta-analysis' literature cut-off date. Blocks coloured grey indicate that the primary study was included in the meta-analysis, while those coloured white were not. Orange, yellow and blue highlights are the two groups of primary studies in overlapping locations and time periods. The study Chang et al. (2014) was a study that was included by Sun et al. (2015), but it was not shown in this table as it was unclear which publication Sun et al. (2015) was exactly referring to.

3.3.3. Quality assessment

3.3.3.1. Limited reporting of methods and findings

The three SR/MAs that used self-set criteria to assess the quality of primary studies tend to be published earlier (Khader et al., 2016; Shah et al., 2011; Stieb et al., 2012). Details of the criteria are shown in **Table 11**. Khader et al. (2016) did a qualitative discussion on the study quality instead of assigning a score or level (for instance low, moderate, or high). Shah et al.

(2011) was more comprehensive and have assigned a moderate risk of bias for most studies. It has included the rating for the respective components of its quality assessment framework, the result of the quality assessment, and the quality assessment tool in the appendix. Stieb et al. (2012) made no comments on study quality despite having the criteria shown. Among the three SR/MAs that used self-set criteria, exposure assessment and confounder adjustment were issues that were mentioned most frequently. However, they all lack details regarding what exposure assessment methods or confounders were deemed appropriate (similar to unmodified score-based approaches in **section 3.3.3.2**).

For SR/MAs that did not include or mention any quality or risk of bias assessment in the published articles, a potential cause might be due to journal constraints, such as shorter article length and supplementary materials being less of a common practice in the past. However, it poses a major methodological limitation as it is one of the essential steps of conducting a SR/MA (Higgins et al., 2024).

Table 11. Sources of bias assessed in the self-set criteria for Shah et al. (2011), Stieb et al. (2012), and Khader et al. (2016).

Study	Sources of biases
Shah et al., 2011	Study selection, exposure assessment, outcome assessment, confounder adjustment, analytical bias, attrition
Stieb et al., 2012	Study design, exposure assessment, confounder adjustment
Khader et al., 2016	Number of studies, exposure assessment, confounder adjustment, sample size

The application of a variety of methods, such as the use of Berman and Parker (2002) by Jacobs et al. (2017), the NIH risk tool by Dzekem et al. (2024), or the AHRQ checklist by Ghosh et al. (2021), Li et al. (2017), and Hung et al. (2022), was poorly reported. In particular, Ghosh et al. (2021), Li et al. (2017), and Hung et al. (2022) all used the AHRQ checklist in conjunction with other quality assessment tools such as the NOS. However, it was unclear how the tools were used in conjunction and if any modifications were made to incorporate some specific considerations needed for studies of air pollution exposure and PTB (more details in next section). This lack of transparency in reporting has made it difficult for authors of subsequent SR/MAs to use comparable methods if appropriate.

3.3.3.2 Score-based approaches

The Newcastle-Ottawa Scale (NOS) is the most used quality assessment tool. There is no established rule for interpreting the score from the NOS, but generally, studies having a score of 7 or above are classified as high quality (Gierisch et al., 2014). 12 out of the 13 SR/MAs assigned a mean score of 7 or above to the primary studies included, indicating that majority of the primary studies were deemed high quality by them (Blanc et al., 2023; Chen et al., 2016; Hung et al., 2022; Ju et al., 2021; Ju et al., 2023; Klepac et al., 2018; Li et al., 2017; Liu et al., 2017; Luo et al., 2021; Melody et al., 2019b; Sun et al., 2015; Yu et al., 2022).

The finding is consistent with the result of a methodological survey conducted for systematic reviews published in the field of public health (Xun et al., 2023). The rise in popularity could be due to the perception of it being "easy to use" (Deek et al., 2003). Nevertheless, when used without modifications, it suffers from problems of lack of detailed instruction for usage and low interrater reliability between reviewers (Hartling et al., 2013; Luchini et al., 2021). None of the SR/MAs that used NOS mentioned contacting the authors for additional information. However, it is recommended as a low agreement was observed between the authors of the studies and reviewers in how NOS was applied and thus the score resulted (Lo et al., 2014).

None of the SR/MAs retrieved that used NOS introduced modifications to make it more relevant for studies of air pollution exposures and birth outcomes. However, there are different methodological preferences and concerns depending on the specific area of research, which are not captured by the general questions in the NOS (Stang, 2010). For instance, exposure misclassification is a major issue in environmental epidemiology studies (Blair et al., 2007). Out of the eight questions in the NOS, there was only one question on exposure ascertainment (Wells et al., 2013):

> "**Ascertainment of exposure:** (a) secure records (eg surgical records) (b) structured interview (c) written self report (d) no description"

For studies of environmental exposures, the accuracy of exposure ascertainment depends on the spatiotemporal resolution of the exposure assessment method and how exposure was assigned, i.e. based on maternal residence or broader spatial regions. The question's ability to critique these concerns is thus minimal.

The question on confounder adjustment (referred as compatibility) in the NOS is as follows:

"**Comparability of cohorts on the basis of the design or analysis:**

(a) study controls for _____ (select the most important factor)

(b) study controls for any additional factor (This criteria could be modified to indicate specific control for a second important factor."

Although there is some consensus regarding what the important confounders for the association between gestational air pollution exposure and the risk of PTB are, it would be impractical to infer that confounder adjustment is sufficient based on the adjustment of only 2 confounders. For instance, if age and smoking are selected as the two most important factors, other important factors, such as parity and season of conception, would be missed. In addition, the authors should be transparent about which confounders were chosen and provide justifications for the choice. Furthermore, the question on outcome assessment is of less concern for studies of pregnancy outcomes, as the outcome is ascertained at birth and recorded in birth certificates and thus assumed to be accurate in most cases. Hence, using questions without adaptation will likely inflate the scores assigned by the NOS.

Score-based approaches, including the NOS and the checklist by Down & Black (1998) or the formula proposed by Doi & Thalib (2008), assign weights to different aspects of the research conducted to derive a total score. While it is convenient for gauging the overall study quality, it is likely problematic as the weights assigned are often equal or somewhat arbitrary when not. Thus, it overlooks the fact that a particular bias could have varied degrees of impact on different studies, as demonstrated in **Table 12**. Moreover, there is a possibility that the biases counteract each other (Arroyave et al., 2021). Thus, the checklist by Down & Black (1998) was found to be ineffective in identifying studies of higher risk of bias compared to domain-based approaches (O'Connor et al., 2015).

Table 12. Examples of primary studies that were critiqued by 4 or more SR/MAs using the NOS, and the methodological issues that are not captured (continued on next page).

Primary study	Range of NOS score assigned by ≥4 SR/MAs	Potential methodological issue overlooked
Huynh et al. (2006)	7-8	Use of nearest monitor approach, lack of adjustment of season of birth or conception
Jalaludin et al. (2007)	7-9	Exposure assigned at a city level

Guo et al. (2018)	7-9	Exposure assigned at a city level
Liang et al. (2019)	6-8	Use of nearest monitor approach, coupled with assignment of exposure at district level

3.3.3.3 Domain-based approaches

Two SR/MAs have adopted a domain-based approach and provided details of the modified framework and guidance on using it (HEI, 2022; Rappazzo et al., 2021). It may show that sharing such information is becoming encouraged in recent literature. Rappazzo et al. (2021) used a modified OHAT framework with seven domains (participant selection, exposure, outcome, confounding, analysis, selective reporting, sensitivity) to analyse the primary studies to incorporate specific considerations such as spatial resolution of the model. HEI (2022) has used the WHO Risk of Bias Assessment Instrument for Systematic Reviews informing the WHO Global Air Quality Guideline, which is specifically developed for assessing the quality of studies of air pollution and health outcomes (WHO, 2020). It was published in late 2020 and was unavailable for SR/MAs published before that. The framework itself recommends the incorporation of tailor-made changes for specific exposure-outcome relationships. HEI (2022) has used the framework and introduced modifications for each exposure-outcome relationship investigated and has produced a comprehensive assessment of the quality of the studies.

3.3.3.4 Recommendations

Based on the findings, it is essential to use or create a quality assessment framework that can:

i. reliably assess the studies of gestational air pollution exposure and the risk of preterm birth,
ii. allow quality assessment to be conducted in a similar manner by users of the tool in the future,
iii. avoid the use of score-based approaches.

The WHO Risk of Bias Assessment Instrument for Systematic Reviews informing the WHO Global Air Quality Guideline satisfies the recommendations listed as it uses a domain-based instead of score-based approach, has demonstrated its validity through being adopted in numerous systematic reviews informing the Air Quality Guidelines, and allow incorporation of

specific considerations. In Chapter 4, details of how the framework was modified to incorporate specific considerations regarding the specific type of study are explained.

3.3.4. Study dependencies / Population overlap

Among the 22 preceding works with a meta-analysis component, ten have acknowledged the issue of study dependencies and introduced various strategies to tackle it (**Table 13**). This implied that the summary estimates derived from the meta-analyses of the ten MAs without considering the issue could be suffering from varying degrees of bias towards locations studied more often. There is no consensus regarding how study dependencies should be analysed and tackled, but most SR/MAs selected the study with the largest sample size or longest study period. Three of the ten meta-analyses that have addressed the issue are shown in Table 10 (Li et al., 2017; Sun et al., 2015; Zhu et al., 2015). Sun et al. (2015) and Zhu et al. (2015) used the most common strategy mentioned, whereas Li et al. (2017) selected the study with the most recent population. However, it is evident that studies that are dependent on each other, such as Ritz et al. (2007) and Huynh et al. (2006), were still included. Different reasons could cause such findings. These relatively simple strategies may be ineffective at minimising the risk of study dependencies. On the other hand, the authors could have deemed the extent of study dependency acceptable and thus chose to keep dependent studies in the meta-analysis. Nevertheless, there were no explanations in the 3 SR/MAs analysed, raising concerns about limited reporting of this methodological aspect as well.

The COMEAP report on adverse birth outcomes classified study dependencies as “tolerable” or “intolerable”, depending on the similarity of the population, the extent of overlap in the time period, and the added value of the study(ies), in order to assess and describe the extent of the issue, and reach judgements about robustness of MAs (COMEAP, 2018b). An example can be made using two pairs of studies that were identified by the COMEAP report and in **Table 10**. The first is Hyder et al. (2014) and Kloog et al. (2012), conducted in Massachusetts, United States. The second is Huynh et al. (2006) and Wu et al. (2009), conducted in California, United States. The former was deemed “intolerable” as they were conducted in the exact same time period (2000-2006) and location. The latter is deemed “tolerable” because one study has eight additional years of study (COMEAP, forthcoming).

Table 13. Strategies adopted by systematic reviews (and/or meta-analyses) to reduce study dependencies

SR/MAs	Strategy for reducing study dependencies
Stieb et al. (2012)	Select study with largest sample size or longest study period
Zhu et al. (2015)	Select study with the largest sample size or longest study period
Lamicchane et al. (2015)	Select study with the largest sample size
Sun et al. (2015)	Select study with the most recent population or most information.
Li et al. (2017)	Select study with the more recent population data and/or information.
Thayamballi et al. (2021)	Select study with the “most inclusive population”. No further details provided.
HEI (2022)	Allowed population overlap unless “the same population and exposure assessment method was used in several publication on the same exposure-outcome pair”. Studies fitting the criteria was selected in the order of (1) population size, (2) confounder adjustment, (3) publication date.
Rappazzo et al. (2021)	Select study with the largest sample size
Ghosh et al. (2021)	Select study with the largest sample size or longest study period
Yu et al. (2022)	Select study with the largest sample size or longest study period

It becomes more complex when the analysis is extended to studies covering areas of various sizes. For instance, in **Table 10** Brauer et al. (2008) and Stieb et al. (2016) were both conducted in Canada in almost the same time period. However, one used the population in Vancouver (single city) and another used a national population. Whether this can be deemed tolerable or intolerable would heavily depend on whether Stieb et al. (2016)’s population covers that of Brauer et al. (2008) or are of different origins, and if they do overlap, if geographical coverage is valued over the potential added value of Brauer et al. (2008).

One of the criteria that can determine the added value of the study is its risk of bias, as it could be an important factor to consider when deciding if one study should be selected over another where they are not independent. The selection of studies based purely on sample size and time period may lead to the inclusion of studies with a higher risk of bias when studies of lower risk of bias with similar sample sizes and in a similar time period were available. In all the SR/MAs listed in **Table 13**, none of the approaches involved using risk of bias as a selection criterion. It could be because studies were screened for dependencies before being quality-assessed.

From what was found in the example of Brauer et al. (2008) and Stieb et al. (2016), an approach of selecting dependent studies covering overlapping geographical areas of different sizes and hence population must also be developed. The approach should be systematic and reproducible for authors of future SR/MAs. Therefore, in the meta-analyses conducted in Chapter 5, study dependencies were first sorted by risk of bias. Afterwards, as a larger sample size correlates with better coverage of the population in the study region, the studies were sorted by sample size. Rules for allowing “tolerable” overlaps in population for studies covering different but overlapping geographical areas were also introduced (more details in Chapter 5 section 5.2.2)

3.4. Interim conclusion

Based on what was discussed in the preceding sections, it was evident that various methodological shortcomings were present in the available work. There were several works that were conducted in a robust manner that could be utilised in health impact assessments, such as HEI (2022) and Stieb et al. (2012). However, they were either outdated or focused on only part of the literature. There were no meta-analyses conducted alongside a systematic review to synthesise all available published literature since 2017 for all of the pollutants of interest. Therefore, all of the results provided in the preceding work were not appropriate to be used in health impact assessments. Based on the results of the analysis presented in this chapter, several recommendations for improving the rigour and robustness of the methodologies used for the forthcoming Chapters were found below:

1. A comprehensive update of the literature is needed, using a systematic literature search that can retrieve all eligible studies. One of the ways that this can be achieved is through comparing the primary studies included in preceding SR/MAs and the results of the literature search.

2. Transparency in reporting is not just important, it's essential in the process of study retrieval, selection, quality assessment, and assessment of study dependencies. The results should be made available so the processes can be reproducible, instilling confidence in the research process.
3. A comprehensive quality assessment framework should be developed incorporating of considerations specific for PTB as an outcome, with domain-based approaches being a possible approach due to its comprehensiveness. This can be based on adjusting the WHO Risk of Bias Assessment Instrument for Informing the WHO Air Quality Guidelines, with.
4. Analysis of study dependencies should go beyond the relatively simplistic criteria used previously, particularly to deal with the point that dependent studies can be overlapping to different extents. Risk of bias should be a criterion for determining the choice of the study to include.

Chapter 4. Systematic review

4.1. Introduction

The limitations found in preceding SR/MAs and the recommendations made to improve them were shown and discussed in **Chapter 3**. Based on the recommendations, the work shown in this Chapter serves to achieve the aim (1) by presenting the work conducted for objectives 2, 3, and 4[2]. To achieve objectives 2 and 3, the search strings used for retrieving relevant literature were validated. Then, a bespoke quality assessment framework for epidemiological studies of gestational air pollution exposure and risk of PTB was developed. Afterwards, the strategy for analysing and selecting dependent studies was established. Authors were also contacted to address gaps in reporting that may be present. The results of the narrative synthesis, mainly focusing on studies not eligible to be meta-analysed, are presented in this Chapter, whereas the results of the meta-analyses conducted are presented in the next Chapter (**Chapter 5**).

4.2. Methodology

4.2.1. PECOS statement

Table 14. The Population, Exposure, Control, Outcome, Study design (PECOS) statement of the review conducted in this chapter (continued on next page).

Criteria	Description
Population	The population of interest is populations of pregnant women that have a singleton live birth. The populations can be formed prospectively through recruitment and followed up till birth, or retrospectively through birth registration or routinely collected data. Multiple births, miscarriages, and stillbirths are excluded.
Exposure	Long-term non-occupational ambient air pollution exposure during pregnancy, including exposure to $PM_{2.5}$, NO_2, and O_3 are included. Exposure due to natural events that could be accidental, for instance volcanic eruption or wildfire are considered if a quantitative analysis of exposure to air pollutants was generated and the risk of preterm birth was conducted. These pollutants are chosen because they are commonly modelled in outdoor air pollution policy analysis, and have previously been shown to have an adverse impact on birth outcomes. The exposure can be measured during the whole pregnancy, in different trimesters, or

[2] Objective 2: Develop methods to address the shortcomings found in Chapter 3.
Objective 3: Conduct a systematic literature search to retrieve related studies
Objective 4: Conduct a narrative synthesis for studies not eligible to be meta-analysed

	in specific periods during the pregnancy. No exposure assessment methodology is excluded *a priori,* and methods that use maternal or foetal serum level air pollutant concentration, personal air quality monitors, modelled concentrations derived from area-based air pollution monitors, or satellite-based estimation techniques are accepted.
Control	The results will be presented per 10 μg/m^3 increments in exposure, allowing a comparison to be made between those who had higher levels of ambient air pollution exposure with those with lower levels.
Outcome	Singleton live preterm births with fewer than 37 weeks of gestation.
Study design	Studies that are case-reports, case-series, reviews, meta-analyses or only available as abstracts are excluded. Studies that use non-human participants are excluded. Studies that present the risk of preterm birth using odds ratios (OR), risk ratios (RRs), or hazard ratios (HRs) are included.

4.2.2. Search strategy

4.2.2.1. Development

Preliminary searches were conducted to gauge the number of primary studies that would be retrieved from database searching. The English language literature databases selected were Medline, PubMed, Embase, Scopus, Cumulative Index to Nursing and Allied Health Literature (CINAHL), Maternity and Infant Care Database (MIDIRS), and GreenFILE. In addition, three Chinese literature databases were selected: Wan Fang Data Knowledge Platform (Wan Fang), China Science and Technology Journal Database (CSTJ), and China National Knowledge Infrastructure (CNKI). These databases were chosen to maximise the range of articles that could be retrieved. Specific reasons for why these databases were used are shown in **Table** 15. Databases used to **retrieve potential eligible primary studies and their reasons for selection.**

Ovid Medline was used as the primary database to test the search terms, which were subsequently translated for use in other databases. Chinese search terms were compiled separately using the best possible match of available Medical Subject Heading (MeSH) terms.

Table 15. Databases used to retrieve potential eligible primary studies and their reasons for selection.

Database(s)	Description
Ovid Medline PubMed	Ovid Medline and PubMed search the Medline database published by the US National Library of Medicine, but PubMed also searches a broader variety of citations, for instance, in-process citations and 'Ahead of print' citations. However, search strings written for PubMed are less structured in general, hence Ovid Medline is still used to conduct a structured, systematic literature search (National Library of Medicine, 2021).
Ovid Embase	The Embase database is published by Elsevier, which contains everything that can be found in Medline and more journals from the EU and other countries. Thus, it was selected to complement the use of Ovid Medline and PubMed, which mainly contain US journals and articles in English (Elsevier, 2022).
Scopus	Scopus is a multidisciplinary database published by Elsevier. Compared to the databases above, Scopus also provides broad coverage for non-English publications and interdisciplinary journals (Scopus, 2022)
MIDIRS CINAHL GreenFILE	MIDIRS and CINAHL are both databases that may contain articles related to outcome of interest (PTB). GreenFILE is a database that contain titles about human impacts such as pollution, to the environment and health. They are therefore different from databases like Medline that have a biomedical focus (EBSCO, 2024a; EBSCO, 2024b; MIDIRS, 2022).
WanFang CSTJ CNKI	These three databases are the largest scientific databases for Chinese articles. Due to the differences in how Chinese articles are indexed, they also contain other types of publications such as theses and patents. Chinese articles are also not well-indexed in other databases such as Medline, therefore searching directly on Chinese databases could mitigate this issue (Fung, 2008)

4.2.2.2. Validation

As discussed in Chapter 3, the reasons behind the discrepancies between included studies in preceding SR/MAs were unclear, reflecting a lack of transparency in the study selection process. The search strategies developed in this thesis aimed to retrieve all studies included in preceding SR/MAs. To achieve this aim, all primary studies (n=118) included in seventeen (sixteen written in English, one written in Chinese) SR/MAs used in the analysis presented in Chapter 3, were extracted and recorded in a matrix. Search results from preliminary searches were combined with duplicates removed. Then, the matrix was used to reveal how many primary studies of preceding SR/MAs were retrieved in the search. Part of the matrix is shown in **Appendix 1**.

In addition, the matrix was used to understand how changes in search terms impact the search results. Initial searches retrieved around 6,000 articles. Afterwards, alterations to the search terms were made, and the results were recorded in Microsoft Excel along with the date on which the search was conducted. The exclusion of search terms related to SO_2 and CO greatly reduced the number of articles retrieved from approximately 6,000 to 2,500, while retrieving close to all studies listed in the matrix. Sub-searching using the excluded search terms showed that most excluded studies were related to carbon monoxide poisoning or non-ambient exposure to SO_2, which is outside the scope of this review. Hence, the two pollutants were dropped from the search string for Ovid Medline. On the contrary, dropping the two pollutants in the search strings for the other databases did not result in substantial changes in the search results and thus were kept. The finalised search strings are shown in **Appendix 2**. The finalised search terms were able to capture all 118 studies shown in the matrix.

4.2.3. PROSPERO protocol

The purpose of registering the review protocol on PROSPERO is to ensure transparency and replicability of this review (Steward et al., 2012). The PROSPERO protocol for this review was published online on 13th April 2022 under the registration number CRD 42022301743 (Lee et al., 2022a). Further details that were not mentioned in the protocol are as follows:

1. Database GREENFile was searched but was not included in the list of databases used.
2. Analysis of studies using hazard ratios (HRs) was done but was not mentioned in the data analysis section.

3. The review was originally designed to include a longer list of pollutants ($PM_{2.5}$, PM_{10}, $PM_{0.1}$, SO_2, CO, NO, NO_2, O_3), and did not restrict the analysis to total PM2.5 but not its constituents or specific sources. However, due to time constraints, the work presented in this thesis is on $PM_{2.5}$, NO_2, and O_3.
4. The review of studies with time-series design, studies in other languages, and meta-analysis of results from multi-pollutant models were planned but not conducted due to time constraints.
5. The protocol did not provide any restriction on the exposure period analysed. However, due to time constraints the analysis was restricted to studies analysing preconception (PC), trimester 1 (T1), trimester 2 (T2), trimester (T3), and whole pregnancy (WP) exposures.

4.2.4. Screening protocols

A protocol for title and abstract screening and another protocol for full-text screening were developed independently by SL. Before being used, the review team members reviewed the protocols of that time (SS, HW, RS, KE). Two reviewers (SL and SS/KM) independently carried out the screening process. Disagreements were discussed in regular weekly meetings and resolved by consensus or consultation with a third reviewer (HW/RS/KE). The protocols were adopted on Covidence, an online systematic review workflow platform (Covidence, 2024a). Duplicate removal, title and abstract screening, and full-text screening were conducted on Covidence by two reviewers (SL and SS/KM). The finalised search results were exported from the literature databases in XML or RIS format and subsequently imported to Covidence.

The title and abstract screening protocol aimed to rapidly screen out irrelevant titles and abstracts. It consisted of ten questions, which allowed reviewers to categorise screened articles into three categories: “Yes” (include in full-text screening), “No” (exclude from full-text screening), or “Maybe” (require discussion between reviewers). When both reviewers categorised an article as “Maybe”, it was included in full-text screening (Covidence, 2024b). Based on the ten questions in the protocol, the reviewers could select one of five reasons for excluding the article from full-text screening. A full version of the protocol with an explanation regarding the grouping of the questions for the reason of exclusion is available in **Appendix 3**.

Full-texts of articles that passed title and abstract screening were further screened using the full-text screening protocol. The article must satisfy the criteria set by the first six questions

set in the protocol to be included in the systematic review. The full-text screening protocol is available in **Appendix 4**. Kappa statistics (κ) was calculated to estimate the interrater reliability at both stages of screening (Belur et al. 2018; McHugh, 2012). Four additional questions were asked to determine if a study is eligible to be included in the meta-analysis (**Table 16**). The study was eligible if all questions were answered yes. If not, the study is either excluded or additional information was requested from the authors. If additional information was received, the study would be screened again from the question where screening was stopped to determine its eligibility.

Table 16. The questions in the full-text screening protocol for determining a primary study's eligibility to be included in the meta-analyses

Question	Decision
Is information on gestational age at individual level, or in the form of summary statistics available for the study population?	Yes – Continue No –Write to authors
Were quantitative estimates and their 95%CIs (not only graphical representation) available for the pollutants of interest?	Yes – Continue No – Write to authors
Is information required to convert the quantitative estimates to a common unit available?	Yes – Continue No – Write to authors

4.2.5. Data extraction

The data extraction form aimed to include all the relevant information necessary to synthesise the evidence and undertake a meta-analysis if appropriate. SL independently developed the bespoke data extraction form, which was then tested on different platforms. After testing, Qualtrics^{XM} was used as it allows versatile extraction of information in various forms (Fisher, 2022). A copy of the data extraction form used can be found in **Appendix 5**. A non-exhaustive list of information extracted included authors, funding source, year of publication, source of the publication, study design, location, period, number of participants, types of adverse perinatal outcomes, outcome incidences, effect estimates, pollutant studied, methods of exposure assessment, spatial resolution of exposure assessment method, and covariates included in the analysis. Data extraction was conducted in duplicate and independently by SL and SS/KM. Disagreements were discussed in regular weekly meetings and resolved by consensus or consultation with a third reviewer (HW/RS/KE).

4.2.6. Quality assessment

Quality assessment was conducted in duplicate and independently by SL and SS/KM. Disagreements were discussed in regular weekly meetings and resolved by consensus or consultation with a third reviewer (HW/RS/KE).

4.2.6.1. WHO Air Quality Guideline Risk of Bias Assessment Instrument

The WHO Air Quality Guidelines Risk of Bias Assessment Instrument is a tool for critiquing studies of the health effects of air pollution. It was developed based on the need to critique these studies and determine the new WHO Air Quality Guidelines (WHO, 2021). It contains six components: confounding, selection bias, exposure assessment, outcome measurement, missing data, and selective reporting. It is not developed for a specific health outcome. Hence, modifications were made to include aspects more relevant to studies of air pollution exposure and PTB. The modifications made are described in the following sections.

4.2.6.2. Determining the critical confounders

The confounding domain classified confounders as “critical” or “potential” confounders. To determine what variables belong to the respective categories, a directed acyclic graph (DAG) was constructed using information from available publications, particularly those that have produced one. Using the list of articles that underwent full-text screening, articles that produced a directed acyclic graph (DAG) were retrieved (Cowell et al., 2021; Kwag et al., 2021; Laurent et al., 2016; Savitz et al., 2019; Siddika et al., 2019). The study and DAGs produced were reviewed, and a new DAG with reference to the preceding ones was produced (**Figure 1**). In this thesis, the critical confounders are defined as strong risk factors of PTB, or confounders that need to be adjusted to close the backdoor paths in the DAG, isolating the association of interest and allowing it to be observed. In **Figure 1**, they are highlighted in red (maternal age, parity, socioeconomic status, season of conception). Potential confounders are those that are less strong or with less consistent evidence, but they can also be considered alongside the critical confounders when necessary. The list of potential confounders considered are smoking, alcohol use, infant sex, year of birth/conception, pre-pregnancy BMI, ethnicity/race, and marital status.

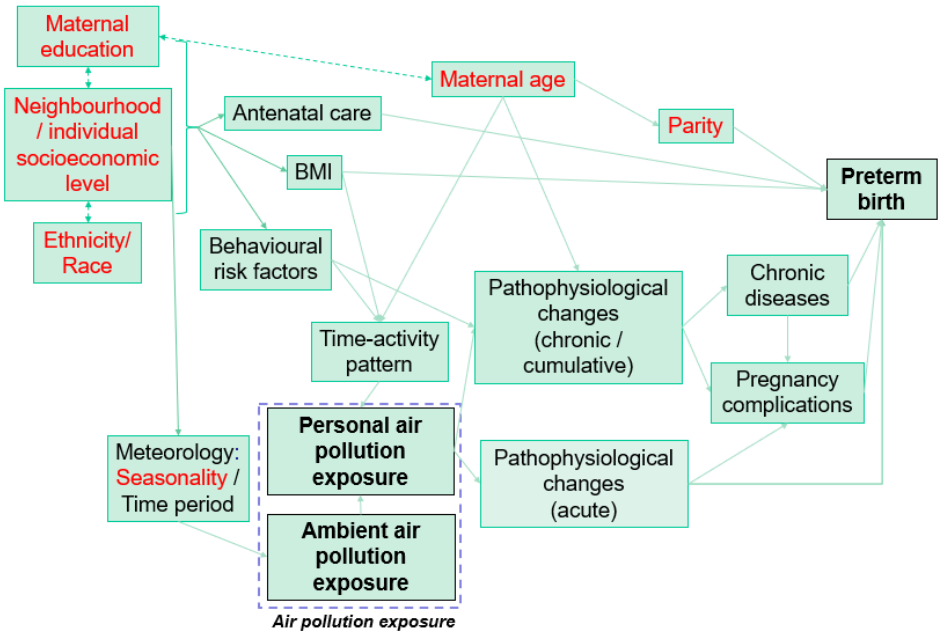


Figure 1. Direct acyclic graph (DAG) produced for the relationship between gestational air pollution exposure and preterm birth.

4.2.6.3. Details of the modifications made

Table 17 shows a summary of the modifications made. The complete version of the modified WHO Risk of Bias Assessment Instrument for Systematic Reviews informing the WHO Global Air Quality Guideline can be found in **Appendix 6**.

If no shortcomings were identified in all domains, the study was classified as having a low risk of bias. Otherwise, the reviewers determined if the shortcoming in a specific domain, for example, outcome measurement, was influential enough to have the study categorised as having a moderate or high risk of bias.

Table 17. Modifications made to the WHO Air Quality Guidelines Risk of Bias Assessment Instrument (continued on next page).

Domain	Modification
Confounding	Two critical confounders were added. (1) Season of conception: Pollutant levels often follow seasonal patterns. The duration of exposure during pregnancy is relatively short, and it is even shorter for pregnancies resulting in PTBs. Therefore, seasonal patterns could heavily determine the pollutant levels during the gestation. (2) Parity: High parity (number of times a woman has given birth) is a strong risk factor of PTB and it correlates with maternal age (see DAG). Hence, parity needs to be adjusted for to avoid its confounding effect on the association. One potential confounder was added. Birth year: Its importance increases as the number of years in the study period increases, as it accounts for annual variations in various factors e.g., pollution levels, particular events that affect air pollution levels or birth outcomes. Not necessary if the study only lasts for a year. Explanations regarding what form of the confounder is most preferable, or in what situation it is acceptable for it to not be adjusted for were added to the tool. It is important as it allows users of the tool to make informed decisions when carrying out the quality assessment. Text to aid users on how to judge validity of measuring of confounding factors was also added.
Selection bias	The name of this domain was changed to study design, to cover more potential issues that may arise in studies of the area. The issues covered in the modified tool were sample size, fixed cohort bias, and reverse causation bias. Explanations were added to aid users' understanding.
Exposure assessment	The original subdomains were modified and renamed into the following: appropriateness of method, consistency (questions largely kept), exposure assignment (spatially and temporally), exposure contrast. Detailed explanations were added to allow users to understand concepts such as maternal mobility, and the preferred granularity of spatial and temporal exposure assignment.

Outcome measurement	The subdomain blinding of outcome measurement was removed, as in most cases information on PTB is obtained via birth certificates or determined by healthcare professionals, which is deemed reliable. Explanation was added to help users of the tool to determine validity of the outcome measurements e.g. last menstrual period or ultrasound dating methods.
Missing data	Content in this domain is largely kept, as in most cases information on PTB is obtained via birth certificates so data is usually relatively comprehensive and suffers less from missing data.
Selective reporting	Content in this domain is largely kept, as the methodological issues stemming from selective reporting is similar to other types of studies.

4.2.7. Addressing gaps in reporting

After quality assessment and data extraction, a list of studies that required additional information or clarification, for instance unclear methodologies, estimates that were only presented graphically, or unknown exposure increment, was created. Using the email addresses available on the respective publications, the corresponding authors of the study were contacted. The list of studies, information required, whether the authors responded, and the information given are shown in **Appendix 13**.

4.2.8. Narrative synthesis

A narrative synthesis was conducted on studies excluded from the meta-analysis and analysed with findings from studies eligible to be included in the meta-analyses but not used. The reasons for excluding each of the studies from the meta-analysis are shown in **Appendix 8**. Studies in the narrative synthesis were organised into various themes with patterns identified, including exposure assessment, environmental factors, maternal factors, outcome assessment, and study design. The findings from the respective themes were then synthesised and discussed, with the aim of providing a holistic understanding of what drives the presence or absence of an association between the pollutants of interest and the risk of PTB.

4.3. Results

4.3.1. Study characteristics

The literature search retrieved 6610 unique records for title/abstract screening, of which 579 articles subsequently underwent full-text screening using the predefined eligibility criteria. The study selection process is shown in **Figure 2**. Good interrater reliability was found for both title/abstract screening (κ = 0.86), and full-text screening (κ = 0.82). After full-text screening and additional restriction due to time-constraints, 205 articles on the wider list of pollutants remained. Among them, 159 studies were about exposure to $PM_{2.5}$, NO_2, or O_3, and hence were included in the narrative synthesis. The characteristics of these studies are shown in **Table 18**.

Few of the 159 studies specified the mode of delivery of the PTBs. Twenty-one and thirteen studies used PTBs that were spontaneously delivered and medically indicated, respectively. Thirteen studies mentioned that the delivery was vaginal, and Eighteen studies mentioned that caesarean sections were performed. The rest did not provide any information. How gestational age was determined was more frequently reported. The last menstrual period (LMP) was most commonly used (n=109), followed by ultrasound (n=53) and best obstetrician estimate (n=34). These methods were used in combination in some studies (n=24).

Most studies have investigated the effect of air pollution exposure over the entire pregnancy (n = 127), followed by trimester 1 (n = 98), trimester 2 (n = 94), trimester 3 (n = 86), and preconception (n =7). Various other exposure windows were also considered (n = 33), including but not limited to last week, first month, or 90 days of pregnancy. $PM_{2.5}$ was the most studied pollutant (n = 115), followed by NO_2 (n = 89) and O_3 (n = 56). The median pollutant concentrations averaged for the entire pregnancy or at specific trimesters were 18.80 μg/m³ (SD: 25.2 μg/m³), 35.57 μg/m³ (SD: 14.67 μg/m³), and 70.76 μg/m³ (SD: 23.10 μg/m³) for $PM_{2.5}$, NO_2, and O_3 respectively. Exposure to pollutants was most commonly estimated using extrapolation from the nearest monitor(s) (n = 88), followed by land-use regression (n = 31) and chemical transport models (n = 30). Various spatial scales were used, ranging from a 50 km radius to a 20 m^2 grid resolution. Studies published more recently were more likely to use hybrid models incorporating various data sources (n = 32), such as satellite remote sensing technologies (n = 21). Logistic regression was the most frequently used statistical method (n = 113), followed by Cox-proportional hazard regression (n = 22). These two groups of studies differed primarily in publication date, with those that used logistic regression dating back to 2000. In contrast, studies using Cox-proportional

hazard regression started to appear in 2016. Studies had typically reported the risk of PTB per increment of pollutant exposure (n = 133), but a considerable number of studies also reported the risk in categorised levels of exposure (n = 18).

Studies included were conducted in various locations, but clear clusters were present. Most of the studies were from the US (n = 58), followed by China (n = 49), Europe (n = 25), Canada (n = 10), and Australia (n = 6). Close to half of the studies from the US were conducted in California (n = 26), with notably fewer studies from other states such as Georgia (n = 4), Ohio (n = 4), Connecticut (n = 3), and Florida (n = 2). Studies from China were mainly in Beijing (n = 6), Shanghai (n = 6), and Guangdong province (n = 9). Studies from the US and China comprise approximately 70% of the studies, and studies from three regions of the world (East Asia, Europe, and North America) comprise more than 90% of the studies included. Only one study took place in Africa. Three studies were conducted in South America, and no studies assessed populations in the Middle East and Southeast Asia. A variety of study designs were also found. The majority of studies were cohort studies (n = 129), and other study designs were case-control (n = 19), cross-sectional (n = 12), and natural experiments (n = 2). The size of the study population varied greatly from 203 to 3972594, with a median of 81886.

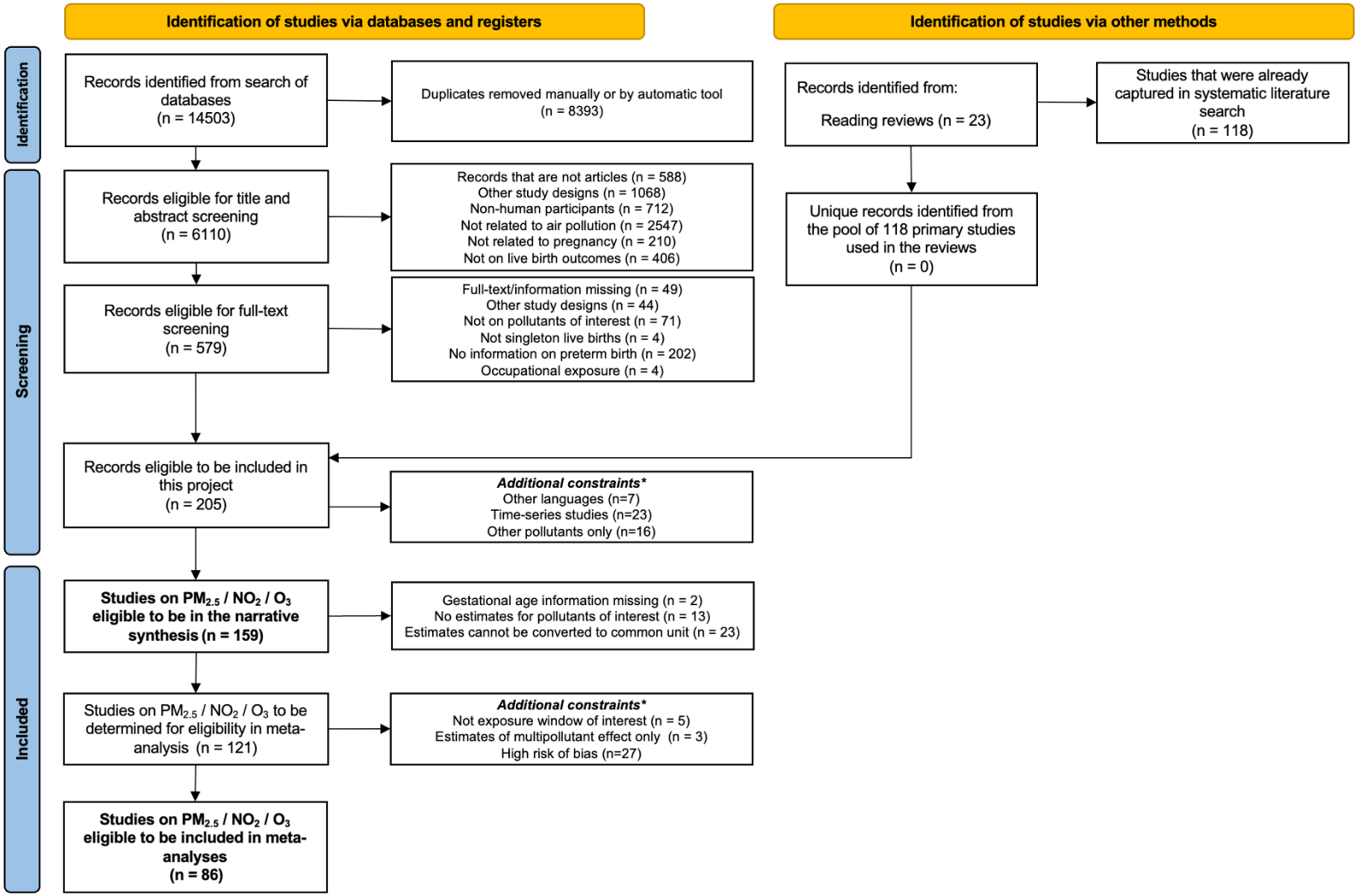


Figure 2. PRISMA diagram showing the study selection process for this systematic review and meta-analysis. Bolded text indicates the stages with number of studies included in the narrative synthesis and the meta-analysis respectively. *Additional constraints are constraints added due to time constraint of this thesis.

Table 18. Characteristics of the primary studies included in this systematic review (studies eligible for meta-analysis highlighted in green).

Study	Location(s) & Period	Study design & births (PTBs)	Exposure assessment	Exposure window(s)	Adjusted covariates	Statistical model	Pollutant(s) & exposure increment(s)	Risk of Bias
Abdo et al., 2019	CO, US 2007-2015	Retrospective cohort 535895 (62436)	***Modelled***: Daily concentrations from a hybrid model at 15 km^2 grid resolution, assigned to ZIP codes of maternal residence at birth.	T1-3 WP	Age, education, smoking, ethnicity, year and month of birth, history of antenatal visit, income, alcohol	Logistic regression	Wildfire smoke $PM_{2.5}$ (1 $\mu g/m^3$)	Mod
Aimaier (亚库甫·艾麦尔) et al., 2013	Beijing, China 2007-2009	Case-control 25872 (unclear)	***Monitor(s)***: Daily concentrations from monitoring network, assigned to maternal residence address at delivery at city level.	First 1 to 8 months	Age, parity, infant sex, history of antenatal visit, alcohol, smoking, maternal conditions, folic acid consumption, meteorology	Logistic regression	PM_{10} / NO_2 / SO_2 (100 $\mu g/m^3$)	High
Alman et al., 2019	CA, US 1999-2006	Case-control 2839 (246)	***Monitor(s)***: Daily concentrations from USEPA Air Quality System (AQS), spatially interpolated using IDW with a 50 km radius, assigned to maternal residence at delivery.	Monthly Weekly	Age, income, ethnicity, parity, maternal origin, smoking, temperature	Bayesian two-level logistic regression	$PM_{2.5}$ (10 $\mu g/m^3$)	Mod
Bachwenkizi et al., 2022	15 African countries[#] 2005-2015	Cross-sectional	***Modelled***: Concentrations from DUACAG at 1 km^2 grid resolution, assigned to maternal residential	WP	Age, education, smoking, infant sex, temperature, history of	Logistic regression	$PM_{2.5}$ (IQR)	Mod

Study	Location(s) & Period	Study design & births (PTBs)	Exposure assessment	Exposure window(s)	Adjusted covariates	Statistical model	Pollutant(s) & exposure increment(s)	Risk of Bias
		131594 (4377)	address at delivery at cluster level.		antenatal visit, cooking fuel source, others*			
Barba-Vasseur et al., 2017	Dijon, Besançon, France 2005-2009	Case-control 1506 (302)	***Modelled***: Monthly concentrations from monitors and ADMS-Urban to produce a 2 m^2 grid resolution, assigned to maternal residence at delivery.	T1-3 WP Last 2 months	Age, smoking, pre-pregnancy BMI, Parity	Conditional logistic regression	NO_2 (10 µg/m^3)	High
Basu et al., 2017	CA, US 2000-2006	Retrospective cohort 231637 (23265)	***Monitor(s)***: Daily concentrations assigned from the closest out of the 7 monitoring sites within 10 km assigned to ZIP codes of maternal residence at birth.	T1-3 WP 1st month Last month	Age, education, ethnicity, infant sex, year of birth, temperature, season of conception, area-level variables	Logistic regression	$PM_{2.5}$ (6.96 µg/m^3)	Mod
Bertin et al., 2015	Brittany, France 2002-2006	Prospective cohort 2509 (83)	***Modelled***: Annual concentrations from LUR model at 100 m^2 grid resolution, assigned to maternal residence at birth.	WP	Education, pre-pregnancy BMI, fish consumption, gestational hypertension, gestational diabetes	Logistic regression	NO_2 (tertiles)	High
Benmarhnia et al., 2017	CA, US 2005-2010	Retrospective cohort 1066783 (86012)	***Monitor(s)***: Averaged annual concentrations from the nearest monitors, assigned to the maternal residence at ZIP code level.	WP	Age, education, parity, infant sex, insurance status, single parenthood	Multilevel logistic regression with random intercepts	$PM_{2.5}$ (1 µg/m^3) NO_2 (1 ppm)	Mod

Study	Location(s) & Period	Study design & births (PTBs)	Exposure assessment	Exposure window(s)	Adjusted covariates	Statistical model	Pollutant(s) & exposure increment(s)	Risk of Bias
Brauer et al., 2008	Vancouver, BC, Canada 1999-2002	Retrospective cohort 70249 (3748)	***Monitor(s)***: Daily concentrations assigned at postal code level by nearest monitor (within 10km) or IDW.	WP	Age, education, smoking, ethnicity, parity, income, infant sex, year of birth, month of birth	Logistic regression	$PM_{2.5}$ / PM_{10} / SO_2 (1 μg/m³), NO_2 / NO (10 μg/m³), CO (100 μg/m³) BC (10^{-5}/m)	Low
Brown et al., 2015	NY (excluding NYC), US 2001-2006	Retrospective cohort 480430 (48885)	***Modelled***: Daily concentrations from USEPA CMAQ and monitor data at 12 km^2 grid resolution, assigned to maternal stress address at birth.	WP	Age, education, smoking, ethnicity, parity, infant sex, Kesser index of prenatal care, pregnancy complications	Cox-PH regression	$PM_{2.5}$ (1 μg/m³), O_3 (1 ppm)	Mod
Cai et al., 2020	China 2015-2016	Retrospective cohort 74109 (6683)	***Modelled***: Monthly concentrations from DUACAG based on AOD, and simulation made by GEOS-Chem CTM, assigned at hospital of delivery.	T1-3 WP	Age, education, pre-pregnancy BMI, gestational hypertension, gestational diabetes, parity, infant sex, method of conception, season of birth, province, temperature, RH	Logistic regression	$PM_{2.5}$ (IQR)	Mod
Cao et al., 2022	Beijing, China 2014-2018	Retrospective cohort 7974 (652)	***Monitor(s)***: Exposure values calculated from daily concentrations calculated using IDW approach,	T1-3 WP	Age, pre-pregnancy BMI, maternal conditions, parity, infant sex, year of	Logistic regression	$PM_{2.5}$ (10 μg/m³, ≥/< 75 μg/m³)	Mod

Study	Location(s) & Period	Study design & births (PTBs)	Exposure assessment	Exposure window(s)	Adjusted covariates	Statistical model	Pollutant(s) & exposure increment(s)	Risk of Bias
			assigned to maternal residence.		conception, season of conception, model of delivery, others*			
Capobussi et al., 2016	Como, Italy 2005-2012	Retrospective cohort 27128 (1516)	***Monitor(s)***: Concentrations from Italian Environmental Protection Agency (APRA), assigned to maternal residence municipality using nearest monitors.	T1-3 WP	Age, education, parity, SES, infant sex, season of birth	Logistic regression	PM_{10} / NO_2 / NO_x / SO_2 / O_3 / CO (Quartile)	Mod
Cassidy-Bushrow et al., 2020	Detroit, MI, US 2008-2010	Retrospective cohort 7961 (891)	***Modelled***: GeoDHOC air pollutant models combined with time series data from the Michigan Air Sampling Network.	T1-3 WP	Age, parity, ethnicity, marital status, clinical site, gestational hypertension, gestational diabetes, poverty quartile	Logistic models	$PM_{2.5}$ / PM_{10} / NO_2 / BTEX (5 μg/m³)	Mod
Chang et al., 2012	NC, US 2001-2005	Retrospective cohort 161078 (14174)	***Modelled***: Weekly concentrations from USEPA AQS at 12km radius, and Fused Air and Deposition Surface data (FSD) at 12 km^2 grid resolution, assigned to maternal street address at delivery.	T1-3 WP Week 1-6	N/A	Discrete time-survival model with logistic link	$PM_{2.5}$ (IQR)	N/A
Chang et al., 2015	Atlanta, GA, US	Retrospective cohort	***Monitor(s)***: Daily concentrations obtained	T1-3	Age, education, infant sex,	Logistic regression	$PM_{2.5}$ (IQR)	Mod

Study	Location(s) & Period	Study design & births (PTBs)	Exposure assessment	Exposure window(s)	Adjusted covariates	Statistical model	Pollutant(s) & exposure increment(s)	Risk of Bias
	1999-2005	175891 (18644)	from 14 air quality monitors within the study region at a 10km radius, using the closest monitor to the geocoded address.	WP First 4 weeks 4-week lag	smoking, temperature, ethnicity, first born, marital status, date and season of conception			
Chen et al., 2018a	Brisbane, Australia 2003-2013	Retrospective cohort 173720 (13994)	***Monitor(s)***: Hourly concentration from 5 fixed monitoring sites in metropolitan area of Brisbane obtained from Queensland Department of Environment and Heritage Protection.	T1-3 WP	Age, infant sex, parity, smoking, maternal conditions, gravidity	Cox-PH regression	$PM_{2.5}$ / NO_2 / SO_2 / O_3 (IQR)	Mod
Chen et al., 2021a	Tianjin, China 2014-2016	Retrospective cohort 10960 (241)	***Modelled***: Daily concentrations obtained from a hybrid exposure at 15 km^2 grid resolution, assigned to the maternal residential address at township level.	PC T1-3 WP	Age, smoking, alcohol use, education, ethnicity, pre-pregnancy BMI, municipality, parity, infant sex, temperature, occupation, gravidity, season of conception, dew point	Cox-PH regression	$PM_{2.5}$ / PM_{10} (5 µg/m³) NO_2 / SO_2 / O_3 (3 µg/m³) CO (0.1 mg/m³)	Low

Study	Location(s) & Period	Study design & births (PTBs)	Exposure assessment	Exposure window(s)	Adjusted covariates	Statistical model	Pollutant(s) & exposure increment(s)	Risk of Bias
Chen et al., 2021b	Shiyan, China 2015-2017	Retrospective cohort 13111 (614)	***Monitor(s)***: Daily concentrations from four the air quality monitoring station. Exposure assigned at city-level.	T1-3 WP	Age, education, infant sex, parity, gravidity	Cox-PH regression	$PM_{2.5}$ / PM_{10} / NO_2 / SO_2 (10 μg/m³)	Mod
Chu et al., 2021	8 Provinces, China 2009-2011	Prospective cohort 5976 (443)	***Modelled***: Monthly $PM_{2.5}$ levels obtained from hybrid model developed by DUACAG at 0.01◦ × 0.01◦ resolution.	T1-3 WP	Age, education, income, alcohol use, smoking, pre-pregnancy BMI, temperature, municipality, paternal age and education, season of delivery, relative humidity	Cox-PH regression	$PM_{2.5}$ (10 μg/m³)	Low
Cossi et al., 2015	CA, US 2000-2006	Retrospective cohort 253704 (29175)	***Monitor(s)***: Daily concentrations from USEPA AQS database spatially interpolated using IDW, assigned to maternal residence at delivery	T2	Age, education, ethnicity, history of antenatal visit, birth weight, Medi-Cal payment for birth expense	Logistic regression	$PM_{2.5}$ / PM_{10} / NO_2 / CO (quartile)	High
Costello et al., 2022	CA, US 2011-2017	Retrospective cohort 2825946 (187235)	***Modelled***: Quarterly mean concentrations derived from CalEnviroScreen using ordinary kriging, assigned to maternal address at	WP	Age, education, ethnicity, smoking, alcohol, parity, pre-pregnancy BMI, infant sex, type of	Log-linear model	$PM_{2.5}$ (IQR)	Mod

Study	Location(s) & Period	Study design & births (PTBs)	Exposure assessment	Exposure window(s)	Adjusted covariates	Statistical model	Pollutant(s) & exposure increment(s)	Risk of Bias
			delivery at census tract level		residence, maternal conditions, drug abuse, insurance, season of birth, interpregnancy interval			
Dėdelė et al., 2011	Kaunas, Lithuania 2008-2009	Prospective cohort 3287 (187)	***Modelled***: Concentration in the exposure periods obtained from AIRVIRO dispersion model, assigned to maternal residence at street level.	T1-3 WP	Education, smoking, parity, family situation, year of birth, maternal condition, previous PTB, stress	Logistic regression	NO_2 (10 μg/m³)	Mod
Dėdelė et al., 2017	Kaunas, Lithuania 2007-2008	Prospective cohort 3292 (235)	***Modelled***: Concentration in the exposure periods obtained from AIRVIRO dispersion model, assigned to maternal residence at street level.	T1-3 WP	Age, education, parity, smoking, alcohol, second hand smoke exposure, ethnicity, year of birth, infant sex, pre-pregnancy BMI, others*	Logistic regression	NO_2 (10 μg/m³)	Mod
DeFranco et al., 2016	OH, US 2007-2010	Retrospective cohort 224921 (19027)	***Monitor(s)***: Daily concentrations from 57 monitors across Ohio from USEPA within a 10km radius, assigned to	T1-3 WP	Age, education, ethnicity, smoking, parity, year of birth, infant sex, history of antenatal visit,	Logistic regression	$PM_{2.5}$ (1 μg/m³)	Mod

Study	Location(s) & Period	Study design & births (PTBs)	Exposure assessment	Exposure window(s)	Adjusted covariates	Statistical model	Pollutant(s) & exposure increment(s)	Risk of Bias
			maternal residential address at delivery.		season of birth, insurance type			
Dibbens and Clements, 2015	Scotland 1994-2008	Retrospective cohort 23086 (1049)	***Modelled***: Concentration from pollution climate model at 1 km^2 grid resolution, assigned to maternal residence at delivery at postcode level.	WP	Age, education, SES, parity, income, year of birth, smoking, crime rate, season of birth, lone parent at birth registration	Logistic regression	PM_{10} / NO_2 / SO_2 (1 μg/m^3)	Low
Dzhambow et al., 2019	Tyrol, Austria 1998-2005	Cross-sectional 573 (81)	***Modelled***: Concentrations from Gaussian propagation model at 100 m^2 grid resolution, assigned to maternal residence at delivery.	WP	Age, education, infant sex, smoking, single mother status, duration of pregnancy, duration of residence before conception, house type	Logistic regression	NO_2 (10 μg/m^3)	High
Estarlich et al., 2016	Asturias, Gipuzkoa, Sabadell, Valencia, Spain 2004-2008	Prospective cohort 2409 (115)	***Modelled***: Daily concentrations from LUR model and monitoring stations, assigned to maternal residential address throughout pregnancy.	T1-3 WP	Age, SES, smoking, type of residence, infant sex, season of delivery	Logistic regression	NO_2 (10 μg/m^3) Benzene (1 μg/m^3)	Low
Fan (樊静洁) et al., 2019	Shenzhen, China	Retrospective cohort	***Monitor(s)***: Daily concentrations from the	T1-3	Age, education, maternal	Logistic regression	$PM_{2.5}$ / PM_{10} / NO_2 / SO_2 / O_3	High

Study	Location(s) & Period	Study design & births (PTBs)	Exposure assessment	Exposure window(s)	Adjusted covariates	Statistical model	Pollutant(s) & exposure increment(s)	Risk of Bias
	2015	400 (200)	Chinese Ministry of Environment Protection.		conditions, previous PTBs, stillbirth, family history of PTB		/ CO (Unclear)	
Fang et al., 2020	Tianjin, China 2014-2016	Prospective cohort 10738 (231)	***Monitor(s)***: Hourly concentrations from 25 monitors within Tianjin City, assigned to maternal residence at delivery.	PC T1-3 Every month	Age, education, second hand smoke exposure, employment, parity, infant sex, temperature, type of residence, dew point, season	Generalised additive distributed lag model	$PM_{2.5}$ (categorised)	Mod
Fang et al., 2022	Beijing, China 2014-2017	Retrospective cohort 24001 (1062)	***Monitor(s)***: Hourly concentrations from a fixed location monitor of Peking University Health Science Centre, assigned at a city level.	T1-3 WP	Age, ethnicity, parity, year of conception, temperature, relative humidity, gravidity, gestational weight gain, season of conception	Logistic regression Cox-PH regression	$PM_{2.5}$ (10 µg/m³)	Mod
Fleischer et al., 2014	Global 2004-2008	Cross-sectional 192900 (13378)	***Modelled***: Estimates derived from global model developed by Dankelaar et al., 2010 at 10 km^2 resolution. A 50 km radius was applied to places without model coverage.	WP	Age, education, infant sex, parity, history of antenatal visits, country-level covariates, seasonality	generalised estimating equation models	$PM_{2.5}$ (10 µg/m³)	Mod

Study	Location(s) & Period	Study design & births (PTBs)	Exposure assessment	Exposure window(s)	Adjusted covariates	Statistical model	Pollutant(s) & exposure increment(s)	Risk of Bias
Gan et al., 2022	Guangzhou, China 2017-2018	Prospective cohort 916 (34)	***Monitor(s)***: Daily concentrations from 11 National Air Quality monitors in Guangzhou city, assigned to maternal residence at delivery.	T1-3	Weight gain during pregnancy	Logistic regression	$PM_{2.5}$ / PM_{10} / NO_2 / SO_2 / O_3 (Unknown)	High
Garshol et al., 2020	Oslo, Akershus, Bergen, Hordaland, Norway 1999-2008	Prospective cohort 16262 (742)	***Modelled***: Concentrations from LUR models developed for ESCAPE, assigned to maternal residence at delivery	T1-3 WP	Age, smoking, pre-pregnancy BMI, gestational diabetes, gestational hypertension	Logistic regression	NO_2 (Categorical)	Mod
Gehring et al., 2011a	Netherlands 1996-1997	Prospective cohort 3853 (165)	***Modelled***: LUR model coupled with air quality monitoring data. Annual averages derived at birth address were adjusted to pregnancy-exposure period.	T1 WP Last month	Maternal origin, study region	Logistic regression	$PM_{2.5}$ / NO_2 / soot (IQR)	Mod
Gehring et al., 2011b	Amsterdam, Netherlands 2003-2004	Prospective cohort 13779 (7541)	***Modelled***: Daily concentration estimated by LUR model for Amsterdam, assigned to maternal residence at delivery.	T1-3 WP	Maternal education	Logistic regression	NO_2 . (Quintile)	Mod
Gehring et al., 2014	Vancouver, BC, Canada 1999-2002	Retrospective cohort 68238 (3654)	***Modelled***: Concentrations from LUR models and monitors, spatially interpolated using IDW with a 50 km radius, assigned to	WP	Age, education, smoking, parity, infant sex, income, year of birth, month of	Logistic regression	$PM_{2.5}$ (1 µg/m³) NO_2 / NO (10 µg/m³) BC (10^{-5} part/m)	Low

Study	Location(s) & Period	Study design & births (PTBs)	Exposure assessment	Exposure window(s)	Adjusted covariates	Statistical model	Pollutant(s) & exposure increment(s)	Risk of Bias
			maternal residential postcode.		birth, first nation status			
Gigoris-Allemand et al., 2017	11 European Countries[#] 1994-2010	Retrospective cohort 71493 (3533)	***Modelled***: LUR model developed which allow estimation of pollution concentration at maternal residence.	T1-3 WP	Age, education, infant sex, smoking, parity, season of conception, temperature, anthropometrics, study centre, atmospheric pressure	Discrete Cox-PH model with logistic link	$PM_{2.5}$ / PM_{coarse} (5 μg/m³), PM_{10} / NO_2 (10 μg/m³), NO_x (20 μg/m³), $PM_{2.5}$ absorbance (10^{-5}/m)	Mod
Gray et al., 2014	NC, US 2002-2006	Retrospective cohort 457642 (40746)	***Modelled***: Daily concentrations from USEPA Hierarchical Bayesian Prediction Model combining AQS and CMAQ output, assigned to maternal residential address at delivery.	WP	Age, education, smoking, parity, infant sex, year of birth, history of antenatal visit, marital status	Logistic regression	$PM_{2.5}$ / O_3 (IQR)	N/A
Gu (顾怡勤) et al., 2019	Shanghai, China 2015-2016	Case-control 29324 (1402)	***Monitor(s):*** Daily concentrations obtained from the National Environmental Air Monitoring System Xuhui Normal University Monitoring Station.	T1-3	Age, education, parity, infant sex, temperature, APGAR score, season of birth, humidity	Logistic regression	$PM_{2.5}$ / NO_2 / SO_2 / O_3 / CO (IQR)	High

Study	Location(s) & Period	Study design & births (PTBs)	Exposure assessment	Exposure window(s)	Adjusted covariates	Statistical model	Pollutant(s) & exposure increment(s)	Risk of Bias
Guo et al., 2018	China 2014	Retrospective cohort 426246 (35261)	***Monitor(s):*** Daily concentrations obtained from National Environmental Monitoring Centre, assigned to maternal residential address at city level.	T1-3 WP	Age, alcohol use, education, pre-pregnancy BMI, smoking, employment, infant sex, gravidity, coastal areas, season of conception	Cox-PH regression	$PM_{2.5}$ (10 μg/m³)	Mod
Guo et al., 2019	Guangdong, China 2014-2017	Retrospective cohort 1098600 (54521)	***Monitor(s):*** Daily concentrations from 54 monitors, assigned to maternal residential address at city level.	T1-3, WP, First 3 months, Last month	Age, ethnicity, smoking, parity, birth month, infant sex, meteorology, birth weight	Generalised estimating equation	NO_2 / SO_2 / CO (IQR)	Mod
Ha et al., 2014	FL, US 2004-2005	Retrospective cohort 423719 (30982)	***Modelled***: Concentration from USEPA Hierarchical Bayesian Prediction Model combining AQS and CMAQ output at 12 km^2 resolution, assigned to maternal residential address.	T1-3 WP	Age, smoking, alcohol use, education, ethnicity, SES, municipality, infant sex, antenatal care, marital status, season of conception	Logistic regression	$PM_{2.5}$ / O_3 (IQR)	Low
Han (韩慧) et al., 2017	Beijing, China 2013-2014	Case-control 384 (166)	***Monitor(s):*** Exposure values calculated from daily concentrations using IDW, assigned to maternal	T1-3 WP	Age, pre-pregnancy BMI, infant sex, antenatal care, maternal weight	Logistic regression	$PM_{2.5}$ (1 μg/m³)	High

Study	Location(s) & Period	Study design & births (PTBs)	Exposure assessment	Exposure window(s)	Adjusted covariates	Statistical model	Pollutant(s) & exposure increment(s)	Risk of Bias
			residence with a 40km radius.		gain during pregnancy, PROM, preeclampsia, eclampsia			
Han et al., 2018	Wuxi, China 2014-2016	Retrospective cohort 6693 (638)	***Monitor(s):*** Daily concentrations obtained from 8 monitors was spatially interpolated using IDW, assigned to maternal residential address at delivery.	T1-3 WP	Age, parity, year of birth, infant sex, maternal conditions, assisted reproductive technology, humidity, mode of delivery	Logistic regression	PM_{10} / O_3 (Quartile)	Low
Hannam et al., 2014	Northwest region, England 2004-2008	Retrospective cohort 203562 (13251)	***Monitor(s):*** Monthly concentrations from 8 monitors in the region, assigned to maternal residence at postcode level.	PC T1-3	Age, parity, smoking, pre-pregnancy BMI, SES, ethnicity, season of birth	Logistic regression	$PM_{2.5}$ / PM_{10} / NO_2 / NO_x / CO (Quartile)	Mod
Hansen et al., 2006	Brisbane, Australia 2000-2003	Retrospective cohort 28200 (1583)	***Monitor(s):*** Hourly concentrations from Queensland Environmental Protection Agency Air Services Unit, assigned to study participants at city level.	T1 Last 90 days	Age, parity, maternal origin, history of antenatal visit, SES, infant sex, marital status, number of previous miscarriages, type of delivery	Logistic regression	PM_{10} / NO_2 / O_3 (IQR)	Mod

Study	Location(s) & Period	Study design & births (PTBs)	Exposure assessment	Exposure window(s)	Adjusted covariates	Statistical model	Pollutant(s) & exposure increment(s)	Risk of Bias
Hao et al., 2016	GA, US 2002-2006	Retrospective cohort 511658 (47321)	***Modelled***: Daily concentrations obtained from CMAQ model at 12 km^2 grid resolution, assigned to maternal census block centroid at delivery.	T1-3 WP	Smoking, education, ethnicity, conception date	Logistic regression (T1-2), Time-to-survival model (T3, WP)	$PM_{2.5}$ / PM_{10} / NO_2 / SO_2 / O_3 / CO (IQR)	Mod
He et al., 2022a	336 cities, China 2010-2015	Retrospective cohort 3723169 (287433)	***Modelled***: Monthly concentrations derived from DUACAG at 0.01◦ × 0.01◦ resolution, assigned to residential address during pregnancy.	T1-3 WP	Age, alcohol, education, pre-pregnancy BMI, smoking, preconception folic acid supplements, employment, parity, infant sex, others	Cox-PH regression	$PM_{2.5}$ (IQR)	Low
He (何进伟) et al., 2022b	Shannxi, China 2018-2019	Case-control 912 (456)	***Monitor(s):*** Daily concentrations from Shannxi Province air quality system (Qingyue Data).	T2 WP	Age, parity, gravidity, ASPGAR score, number of previous caesarean sections, preeclampsia	Logistic regression	$PM_{2.5}$ / PM_{10} / O_3 (unclear)	High
Heft-Neal et al., 2022	CA, US 2006-2011	Panel study 3002014 (219740)	***Modelled***: Daily concentrations from NOAA's Hazard Mapping System at 1 km^2 grid resolution, assigned to	T1-3 WP	Age, education, ethnicity, infant sex, maternal	Linear probability model	Wildfire smoke $PM_{2.5}$	N/A

Study	Location(s) & Period	Study design & births (PTBs)	Exposure assessment	Exposure window(s)	Adjusted covariates	Statistical model	Pollutant(s) & exposure increment(s)	Risk of Bias
			maternal residence at ZIP code level.		origin, age squared			
Huang et al., 2015	Beijing, China 2007-2010	Natural Experiment 50874 (3203)	***Monitor(s):*** Daily concentrations from 28 monitors across the area, assigned to study participants at citywide level.	T1-3	Age, parity, infant sex, type of residence, temperature, precipitation, month of conception	Logistic regression	PM_{10} / NO_2 (10 µg/m³)	Mod
Huang et al., 2018	CA, US 2009-2012	Retrospective cohort 1448600 (99518)	***Modelled***: Quarterly means concentrations from CalEnviroScreen 3.0 were used and assigned at census tract level.	WP	Age, education, ethnicity/race, insurance type	Logistic regression	$PM_{2.5}$ (IQR)	High
Huynh et al., 2006	CA, US 1999-2000	Matched case-control 42692 (10673)	***Monitor(s):*** Daily concentrations obtained from the closest PM monitor from the California Air Resources Board, assigned to maternal residential address.	WP 1st month Last 2 weeks	Age, education, ethnicity, parity, marital status	Conditional logistic regression	$PM_{2.5}$ (10 µg/m³) CO (1ppm)	Mod
Hyder et al., 2014	CT, US 2000-2006	Retrospective cohort 647942 (41868)	***Monitor(s):*** USEPA AQS monitoring data from 50km radius, assigned closest monitor to maternal residence.	T1-3 WP	Age, smoking, education, ethnicity, maternal conditions, parity,	Logistic regression	$PM_{2.5}$ (IQR)	Low

Study	Location(s) & Period	Study design & births (PTBs)	Exposure assessment	Exposure window(s)	Adjusted covariates	Statistical model	Pollutant(s) & exposure increment(s)	Risk of Bias
					infant sex, antenatal care			
Hystad et al., 2014	Vancouver, BC, Canada 1999-2002	Retrospective cohort 64705 (3189)	***Modelled***: Concentration for exposure periods from LUR models developed for the area, assigned to maternal residential postcode address.	WP	Age, education, parity, infant sex, smoking, year of birth, month of birth, first national status, income	Logistic regression	$PM_{2.5}$ (1 μg/m³) NO_2 / NO (10 μg/m³) BC (10^{-5}/m)	Low
Jackson et al., 2009	England & Wales 2000-2001	Retrospective cohort 13131 (unclear)	***Modelled***: Concentration from dispersion model of National Environmental Technology Centre at 1km^2 grid resolution, assigned to maternal residential postcode address at delivery.	WP	Smoking, ethnicity	Bayesian graphical regression	NO_2 / SO_2	N/A
Jalaludin et al., 2007	Sydney, NSW, Australia, 1998-2000	Retrospective cohort 123840 (6011)	***Monitor(s):*** Daily concentrations from 14 monitors averaged across the whole metropolitan area of Sydney, assigned at a city level.	T1, First month, Last 3 months, Last month	Age, smoking, ethnicity, SES, parity, infant sex, antenatal care, season of conception	Logistic regression	$PM_{2.5}$ / PM_{10} (1 μg/m³), NO_2 / SO_2 / O_3 (1 ppb), CO (1 ppm)	Mod
Johnson et al., 2016	NYC, NY, US 2008-2010	Retrospective cohort 258294 (10271)	***Modelled***: Annual concentrations generated from LUR models, temporally adjusted with using 2-week intervals from	T1-2	Age, education, ethnicity, pre-pregnancy BMI, SES, parity, infant sex, year of	Logistic mixed model	$PM_{2.5}$ (10 μg/m³) NO_2 (20ppb)	Low

Study	Location(s) & Period	Study design & births (PTBs)	Exposure assessment	Exposure window(s)	Adjusted covariates	Statistical model	Pollutant(s) & exposure increment(s)	Risk of Bias
			multiple daily regulatory monitors.		conception, others			
Ji et al., 2019	Shanghai, China 2014-2015	Matched case-control 25943 (1172)	***Modelled***: Daily concentrations from LUR model in Shanghai at 100 m^2 grid resolution, assigned to maternal residence at delivery.	T1-3 WP Last month Last week	Age, education, pre-pregnancy BMI, SES, infant sex, paternal age, gravidity, season of conception, season of delivery	Logistic regression (unconditional/ discrete time-to-event)	NO_2 (10 μg/m³)	Low
Kim et al., 2021	South Korea 2015-2018	Retrospective cohort 789595 (240981)	***Modelled***: Monthly concentrations from hybrid models using various data sources, assigned to maternal residence at district level.	WP	Age, income, type of residence	Logistic regression	$PM_{2.5}$ / PM_{10} / NO_2 / SO_2 / CO (quartile)	N/A
Kingsley et al., 2017	Providence, RI, US 2001-2012	Retrospective cohort 61640 (5007)	***Modelled***: Hybrid model consisting of LUR and satellite remote sensing with output of 200 m^2 grid, assigned to maternal residence.	T1-3 WP	Age, smoking, education, ethnicity, parity, health insurance, last menstrual period, neighbourhood SES, marital status	Logistic regression	$PM_{2.5}$ (IQR)	Mod

Study	Location(s) & Period	Study design & births (PTBs)	Exposure assessment	Exposure window(s)	Adjusted covariates	Statistical model	Pollutant(s) & exposure increment(s)	Risk of Bias
Kirwa et al., 2021	PR, US 2011-2018	Prospective cohort 1092 (99)	***Monitor(s):*** Daily concentrations from closest USEPA monitors assigned to municipalities in which mother resided at the time of birth.	WP	Age, education, municipality, parity, infant sex, birth year, antenatal care, season of birth, marital status, population-level variables	Poisson regression	$PM_{2.5}$ (IQR)	High
Kloog et al., 2012	MA, US 2000-2008	Retrospective cohort 634244 (61972)	***Modelled***: Daily concentrations from models using calibrated AOD data at a 10 km^2 grid resolution, assigned to the maternal residence at delivery.	T1-3 WP Last month	Age, smoking, maternal conditions, income, infant sex, antenatal care, others	Logistic mixed regression	$PM_{2.5}$ (10 μg/m³)	Low
Kwag et al., 2021a	Seoul, South Korea 2010-2016	Retrospective cohort 581239 (25840)	***Modelled***: Hourly concentrations from the WRF-SMOKE-CMAQ model, assigned to maternal residential address at district level.	T1-3	Age, education, parity, infant sex, temperature, season of birth, others	Logistic regression	$PM_{2.5}$ (10 μg/m³)	Mod
Kwag et al., 2021b	7 major cities, South Korea 2010-2016	Retrospective cohort 1329991 (61182)	***Modelled***: Hourly concentrations from the WRF-SMOKE-CMAQ model, assigned to maternal residential address at district level.	T1-2	Education, parity, infant sex, year of birth, residential area, season of birth	Logistic regression	$PM_{2.5}$ (percentile)	Mod
Laurent et al., 2016	CA, US	Matched case-control	***Modelled***: Daily concentration from	WP	Age, education, ethnicity, median	Logistic regression	$PM_{2.5}$ / $PM_{0.1}$ / UFP / NO_2 / O_3	Mod

Study	Location(s) & Period	Study design & births (PTBs)	Exposure assessment	Exposure window(s)	Adjusted covariates	Statistical model	Pollutant(s) & exposure increment(s)	Risk of Bias
	2001-2008	3870696 (442314)	California Air Resources Board (CARB) interpolated using kriging, output at 200 m^2 grid resolution.		household income at census block group level		/ CO (IQR)	
Lavigne et al., 2016	Ontario, Canada 2005-2012	Retrospective cohort 818400 (51144)	***Modelled***: Concentrations derived from model combining CTM and AOD data at 1 km^2 grid resolution, assigned to maternal residence over course of pregnancy.	T1-3 WP	Age, smoking, municipality, parity, income, infant sex, year of birth, month of birth, previous caesarean, socioeconomic variables	Logistic regression	$PM_{2.5}$ / NO_2 / O_3 (IQR)	Low
Lavigne et al., 2018	Ontario, Canada 2006-2012	Retrospective cohort 196171 (15378)	***Monitor(s)***: Daily concentrations obtained from ground monitors in Ontario, assigned at maternal residential post code level.	T1-3, WP Last 4 weeks Day of delivery	Age, asthma, smoking, SES, parity, income, infant sex, year of birth, month of birth, others	Logistic regression	$PM_{2.5}$ (IQR)	Mod
Le et al., 2012	Detroit, MI, US 1990-2001	Retrospective cohort 164905 (24954)	***Monitor(s)***: Hourly concentrations from Michigan Department of Environmental Quality, assigned to maternal residential ZIP code within 4 km radius of a monitoring station.	1^{st} month Last month	Age, education, smoking, year of birth, infant sex, ethnicity, history of antenatal visit, season of birth, site of residency	Logistic regression	NO_2 / O_3 (Quartile)	Mod

Study	Location(s) & Period	Study design & births (PTBs)	Exposure assessment	Exposure window(s)	Adjusted covariates	Statistical model	Pollutant(s) & exposure increment(s)	Risk of Bias
Lee et al., 2013	Allegheny, PA, US 1997-2002	Retrospective cohort 34705 (1940)	***Modelled***: Space-time ordinary kriging interpolation was conducted on daily air pollution concentrations to obtain exposure levels, which was assigned at maternal residential ZIP code level.	T1	Age, smoking, ethnicity, parity, year of conception, season of birth	Logistic regression	$PM_{2.5}$ / PM_{10} / O_3 (IQR)	Mod
Leem et al., 2006	Incheon, South Korea 2001-2002	Case-control 52113 (2082)	***Modelled***: Hourly concentrations from 16 monitors surrounding Incheon, assigned to maternal residence at delivery at Dong level using ordinary block kriging.	T1, T3	Age, education, parity, infant sex, season of birth	Logistic regression	PM_{10} / NO_2 / O_3 / CO (Quartile)	Mod
Li et al., 2018	China 2013-2014	Retrospective cohort 1240978 (100433)	***Modelled***: Daily concentrations from hybrid exposure model at 10 km^2 grid resolution, assigned to maternal residential address.	T1-3, WP	Age, smoking, alcohol use, education, pre-pregnancy BMI, employment, infant sex, temperature, others	Cox-PH regression	$PM_{2.5}$ / PM_{10} (10 μg/m³)	Low
Li et al., 2020	Hubei, China 2014	Retrospective cohort 429865 (12809)	***Modelled***: Concentrations from WRF Model validated by China Earth International Exchange stations and air quality monitoring data, output at sub-district level.	WP	Age, education, parity, infant sex, antenatal care, nationality, month of conception, annual income at	Logistic regression	$PM_{2.5}$ (30 μg/m³)	Mod

Study	Location(s) & Period	Study design & births (PTBs)	Exposure assessment	Exposure window(s)	Adjusted covariates	Statistical model	Pollutant(s) & exposure increment(s)	Risk of Bias
					district/country level			
Liang et al., 2019a	9 cities, Guangdong, China 2014-2017	Retrospective cohort 1455026 (64028)	***Monitor(s)***: Exposure assigned at maternal residential district level using daily concentrations from monitors, with a 5 km radius.	T1-3	Age, parity, infant sex, temperature, relative humidity	Cox-PH regression	$PM_{2.5}$ / NO_2 / SO_2 / O_3 (10 μg/m³)	Mod
Liang et al., 2019b	7 cities, Guangdong, China 2015-2017	Retrospective cohort 628439 (29489)	***Monitor(s)***: Daily concentrations from the nearest monitor assigned to maternal residential district at birth.	T1-3, WP Last 2 weeks	Age, parity, infant sex, temperature, relative humidity, gravidity	Cox-PH regression	$PM_{2.5}$ / NO_2 / SO_2 / O_3 (10 μg/m³)	Mod
Liang et al., 2022	Guangdong, China 2014-2017	Retrospective cohort 230902 (17066)	***Monitor(s)***: Daily concentrations from nearest of the 77 monitoring stations in the 21 cities across the province, assigned to the maternal residential district.	T1-3 WP	Age, smoking, alcohol use, pre-pregnancy BMI, parity, infant sex, temperature, others	Cox-PH regression	$PM_{2.5}$ / NO_2 / SO_2 / O_3 (10 μg/m³)	Mod
Lin et al., 2015	Taiwan 2001-2007	Case-control 431120 (86224)	***Monitor(s):*** Monthly concentrations from closest 3 out of the 72 monitors within 25 km across the whole country, assigned to maternal residence at postcode level using IDW.	T1-3	Age, income, infant sex, year and season of conception	Logistic regression	PM_{10} (10 μg/m³), NO_2 / O_3 (10 ppb), SO_2 (1 ppb), CO (100 ppb)	Low

Study	Location(s) & Period	Study design & births (PTBs)	Exposure assessment	Exposure window(s)	Adjusted covariates	Statistical model	Pollutant(s) & exposure increment(s)	Risk of Bias
Liu et al., 2003	Vancouver, BC, Canada 1986-1998	Retrospective cohort 229085 (12141)	***Monitor(s)***: Daily concentrations from monitors at census subdivision level were assigned to maternal residence during pregnancy.	First month Last month	Age, parity, infant sex, birth weight, season of birth	Logistic regression	SO_2 (5ppb) NO_2 / O_3 (10ppb) CO (1ppm)	Mod
Liu et al., 2019a	Guangzhou, China 2016-2017	Nested case-control 4101 (234)	***Modelled***: Daily concentrations from 102 monitors and LUR model, produced at 1 km^2 grid resolution, assigned to maternal residence at delivery.	WP	Unclear	Cox-PH regression	PM_{10} / $PM_{2.5}$ / PM_1 / SO_2 / NO_2 / O_3 (10 μg/m³) CO (100 μg/m³)	High
Liu et al., 2019b	Guangdong, China 2014-2015	Case-control 2751 (687)	***Monitor(s)***: Daily concentrations from 111 national monitoring stations in China at city-level.	T1-3 WP	Age, education, employment, parity, infant sex, registered residence, childbirth time, month of birth	Logistic regression	PM_{10} / $PM_{2.5}$ / SO_2 / NO_2 / O_3 (10 μg/m³) CO (100 μg/m³)	High
Liu et al., 2022b	Shenzhen, Guangzhou, China 2009-2012	Natural experiment 522703 (unclear)	***Monitor(s)***: Daily concentrations from Guangzhou and Shenzhen environmental bureau monitors, assigned to maternal residence at birth at city level.	WP	Age, education, parity, year of conception, season fixed effects, weather variables	Two-stage least square method	PM_{10} / NO_2 / SO_2 (10 μg/m³)	Mod

Study	Location(s) & Period	Study design & births (PTBs)	Exposure assessment	Exposure window(s)	Adjusted covariates	Statistical model	Pollutant(s) & exposure increment(s)	Risk of Bias
Llop et al., 2010	Valencia, Spain 2003-2005	Prospective cohort 785 (47)	***Modelled***: Daily concentrations from LUR model and monitors, assigned to maternal residence and workplace throughout pregnancy, to construct weighted exposure according to time spent in each place.	T1-3 WP	Age, education, employment, origin, smoking, alcohol, type of residence, infant sex, season of conception, caffeine consumption at T1	Logistic regression	NO_2 / Benzene (1 μg/m³)	Mod
Lu et al., 2019	Changsha, Henan, China 2011-2012	Retrospective cohort 3509 (146)	***Monitor(s)***: Daily concentrations obtained from 7 monitors and assessed by IDW, assigned to children's kindergarten.	PC T1-3 WP Last month	Age, employment, SES, infant sex, year of birth, temperature, second hand smoke exposure, season of birth, birth weight, others*	Logistic regression	NO_2 (IQR)	Mod
Maroziene & Grazuleviciene, 2002	Kaunas, Lithuania 1998	Retrospective cohort 3988 (203)	***Monitor(s)***: Daily concentrations from 12 monitors, with one monitor per district, assigned to maternal residence at delivery.	WP	Age, education, smoking, parity, marital status, paternal smoking, season of birth	Logistic regression	NO_2 (10 μg/m³)	Mod
Mekonnen et al., 2021	6 CA counties, US 2007-2011	Retrospective cohort	***Modelled***: Concentrations from Bayesian space-time fusion model integrating with monitoring networks	PC T1-3	Age, smoking, education, ethnicity, year of birth, antenatal	Logistic regression	$PM_{2.5}$ (1 μg/m³)	Low

Study	Location(s) & Period	Study design & births (PTBs)	Exposure assessment	Exposure window(s)	Adjusted covariates	Statistical model	Pollutant(s) & exposure increment(s)	Risk of Bias
		953951 (87495)	and CMAQ estimates at 12 km^2 grid resolution.	WP	care, season of conception, insurance type		O_3 (1 ppb)	
Melody et al., 2019a	Victoria, Australia 2012-2015	Retrospective cohort 3591 (294)	***Modelled***: Daily concentrations of $PM_{2.5}$ from coalmine fires from CTM at 1 km^2 grid resolution, assigned to maternal residence at delivery at Statistical Area Level (SA1).	WP	Age, smoking, SES, infant sex, year of conception, temperature, season of conception, infant aboriginality	Multivariable log-binomial regression	$PM_{2.5}$ (10 µg/m³)	High
Melody et al., 2020	Victoria, Australia 2012-2015	Retrospective cohort 285594 (8558)	***Modelled***: Annual concentrations obtained from LUR models and assigned at geographical unit of statistical area level (SA1).	WP	Age, parity, temperature, smoking, year of conception, SES	Multivariable log-link binomial generalised linear model	$PM_{2.5}$ (IQR)	Mod
Mendola et al., 2016	US 2005-2007	Retrospective cohort 223502 (26144)	***Modelled***: Daily concentrations from a modified CMAQ at 12 km^2 grid resolution, assigned to delivery hospital referral region.	PC WP Last 6 weeks	Age, smoking, alcohol use, ethnicity, pre-pregnancy BMI, parity, study site, insurance status, marital status	Logistic regression	$PM_{2.5}$ / PM_{10} / SO_2 / O_3 / CO (IQR)	Mod
Mendola et al., 2019	Utah, US 2002-2010	Retrospective cohort 50005 (3806)	***Modelled***: Daily concentrations from a modified CMAQ at 12 km^2 grid resolution, assigned to	WP	Age, ethnicity, parity, smoking, alcohol use, asthma, pre-pregnancy BMI,	Poisson regression	$PM_{2.5}$ / PM_{10} / NO_2 / SO_2 / O_3 / CO (IQR)	Mod

Study	Location(s) & Period	Study design & births (PTBs)	Exposure assessment	Exposure window(s)	Adjusted covariates	Statistical model	Pollutant(s) & exposure increment(s)	Risk of Bias
			delivery hospital referral region.		previous PTB, interpregnancy interval, insurance status, marital status			
do Nascimento et al., 2022	São Paulo, Brazil 2010-2016	Cross-sectional 979306 (91173)	***Monitor(s)***: Daily concentrations from Companhnia Ambiental do Estado de São Paulo CETESB, assigned to maternal residence at delivery.	T1-3 WP	Age, education, SES, infant sex, temperature, history of antenatal visit, type of delivery	Logistic regression	PM_{10} / NO_2 / O_3 (10 μg/m³)	Mod
Nealy et al., 2018	Atlanta, Georgia, US 2002-2006	Retrospective cohort 267801 (16087)	***Modelled***: Daily concentrations from CMAQ bias-corrected by monitors at 12 km^2 grid resolution, assigned to maternal residence at delivery at census block level.	T1-2	Education, ethnicity, SES, residential county, poverty level, date of conception	Logistic regression	$PM_{2.5}$ / CO / NO_x / EC / OC (IQR)	High
Olsson et al., 2012	Stockholm, Sweden 1988-1995	Retrospective cohort 120251 (6149)	***Monitor(s)***: Daily concentrations from monitors of the City of Stockholm Environment and Health Administration, assigned to mothers at citywide level.	T1-2 Last week	Age, parity, infant sex, smoking, relative humidity, seasonal variations, long-term trend	Logistic regression	NO_2 / O_3 (10 μg/m³)	Mod
Olsson et al., 2013	Stockholm, Sweden	Retrospective cohort	***Monitor(s)***: Daily concentrations obtained from monitors stations,	T1	Age, education, parity, asthma, area of origin,	Logistic regression	NO_x / O_3 (10 μg/m³)	Mod

Study	Location(s) & Period	Study design & births (PTBs)	Exposure assessment	Exposure window(s)	Adjusted covariates	Statistical model	Pollutant(s) & exposure increment(s)	Risk of Bias
	1997-2006	120755 (5291)	assigned to mothers at citywide level.		year of conception, season of conception			
Ottone et al., 2020	48 munipalities, Northern Italy 2012-2014	Retrospective cohort 23708 (1311)	***Modelled***: Daily concentrations linked to maternal residence for entire pregnancy estimated from LUR and yearly concentration map from NINFA-PESCO chain model at 1 km^2 grid resolution.	WP	Age, smoking, education, parity, infant sex, area of origin	Logistic regression	$PM_{2.5}$ (1 $\mu g/m^3$)	Low
Padula et al., 2014	CA, US 2000-2006	Retrospective cohort 263204 (2938)	***Modelled***: Daily concentrations from USEPA AQS database were spatially interpolated using IDW with maximum radius of 50km, assigned to maternal residence at delivery.	T1-3 WP Last 6 weeks Last month	Age, education, ethnicity, history of antenatal visit, birth weight, Medi-Cal payment of birth costs	Logistic regression	$PM_{2.5}$ / PM_{10} / NO_2 / CO (Quartile)	High
Padula et al., 2018	CA, US 2009-2012	Retrospective cohort 53843 (4560)	***Modelled***: Annual concentrations from CalEnviroScreen 2.0 spatially interpolated using kriging from nearest monitors, assigned to	WP	Age, education, ethnicity, payment for delivery	Logistic regression	$PM_{2.5}$ / O_3 (percentile)	High

Study	Location(s) & Period	Study design & births (PTBs)	Exposure assessment	Exposure window(s)	Adjusted covariates	Statistical model	Pollutant(s) & exposure increment(s)	Risk of Bias
			maternal residence at census tract level.					
Padula et al., 2019	CA, US 2000-2006	Retrospective cohort 252205 (21225)	***Monitor(s)***: Daily concentrations from USEPA AQS database, interpolated using IDW, assigned to maternal residence at delivery at street level.	T1-3 WP Last 6 weeks	Age, education, ethnicity, history of antenatal visit, insurance type	Logistic regression	PM_{10} / $PM_{2.5}$ / NO_2 / O_3 (Quartile)	High
Panasevich et al., 2016	Oslo, Akershus, Bergen, Hordaland, Norway 1999-2008	Prospective cohort 17533 (733)	***Modelled***: Daily concentrations from LUR model developed for ESCAPE and Norwegian Institute for Air research Database, assigned to maternal residence at delivery.	WP	Age, education, infant sex, smoking, parity, anthropometrics, marital status, birth season	Logistic regression	NO_2	Low
Pereira et al., 2014a	CT, US 2000-2006	Retrospective cohort 61688 (1808)	***Monitor(s)***: Daily concentrations from the closest USEPA monitor within 40km radius of maternal residence.	T1-3 WP	Age, smoking, parity	Conditional logistic regression	$PM_{2.5}$ (IQR)	Mod
Pereira et al., 2014b	Perth, Australia 1997-2007	Retrospective cohort 31567 (2304)	***Monitor(s)***: Daily concentrations from the closest monitor of Western Australian Department of Environment Conservation within 40 km radius of maternal residence.	T1-3 WP	Age, smoking, parity, unmeasured temporal factors	Logistic regression	$PM_{2.5}$ (1 μg/m³)	Mod

Study	Location(s) & Period	Study design & births (PTBs)	Exposure assessment	Exposure window(s)	Adjusted covariates	Statistical model	Pollutant(s) & exposure increment(s)	Risk of Bias
Pereira et al., 2014c	CT, US 2000-2006	Retrospective cohort 48208 (2586)	***Monitor(s)***: Weekly concentration from closest monitors, assigned to maternal residence within 40 km radius.	T1-3 WP	Age, smoking, parity	Logistic regression	$PM_{2.5}$ sources (IQR)	Mod
Pereira et al., 2016	Rochester, New York State, US 2004-2012	Retrospective cohort 7121 (668)	***Monitor(s)***: Weekly concentrations from the closest USEPA monitor assigned to maternal residence during pregnancy.	T1-3 WP	Age, smoking, parity, physical activity, prepregnancy weight	Conditional logistic regression	$PM_{2.5}$ (1 μg/m³)	Mod
Poirier et al., 2015	Halifax, Nova Scotia, Canada 2008-2012	Retrospective cohort 14752	***Modelled***: Concentrations from LUR model based on data from Air Health Science Division in Health Canada, assigned to maternal residence at postcode level.	WP	Age, smoking, SES, parity, neighbourhood income	Logistic regression	PM_1 / $PM_{2.5}$ / PM_{10} / NO_2 / SO_2 / Benzene / Toluene (IQR)	Mod
Qian et al., 2016	Wuhan, China 2011-2013	Prospective cohort 95911 (4308)	***Monitor(s)***: Daily concentrations from the closest monitor, assigned to the maternal residential community at birth.	T1-3 WP	Age, education, employment, parity, infant sex, temperature, season of conception, gravidity	Logistic regression	$PM_{2.5}$ (5 μg/m³)	Mod
Rappazzo et al., 2014	PA, NJ, OH, US 2000-2005	Retrospective cohort	***Modelled***: Daily concentrations from bias-corrected CMAQ model with monitor network data,	T1-3 WP	Age, education, ethnicity, marital status, ozone	Modified Poisson regression	$PM_{2.5}$ (1 μg/m³)	N/A

Study	Location(s) & Period	Study design & births (PTBs)	Exposure assessment	Exposure window(s)	Adjusted covariates	Statistical model	Pollutant(s) & exposure increment(s)	Risk of Bias
		17815217 (142151)	assigned to maternal residence at delivery.					
Rappazzo et al., 2015	PA, NJ, OH, US 2000-2005	Retrospective cohort 1771255 (140987)	***Modelled***: Daily concentrations from CMAQ model bias-corrected using monitors, assigned to maternal residence at delivery.	T1-3 WP Weekly	Age, smoking, education, ethnicity, temperature, marital status, season of conception, ozone	Modified Poisson regression	EC (0.25 μg/m³) NO_3 / OC / SO_4 (1 μg/m³)	Mod
Rappazzo et al., 2016	PA, NJ, OH, US 2000-2005	Retrospective cohort 1588595 (252219)	***Modelled***: Daily concentrations from CMAQ model bias-corrected using monitors, assigned to maternal residence at delivery.	Weekly	Age, education, ethnicity, marital status, ozone	Linear risk regression	$PM_{2.5}$ (1 μg/m³)	N/A
Riddell et al., 2021	Oakland and San Jose, CA, US 2013-2015	Retrospective cohort 8223 (759)	***Monitor(s)***: Daily concentrations from measurements of Google vehicles, assigned to maternal residential address at delivery.	WP	Age, education, income, SES, parity, season of conception, city of residence, insurance type	Logistic regression	NO_2 / UFP (percentile)	Mod
Ritz et al., 2007	Los Angeles county, CA, US	Nested case-control 58316 (unclear)	***Monitor(s)*:** Daily concentrations from closest monitor from SCAQMD, assigned to maternal residential ZIP code at delivery.	T1 WP Last 6 weeks	Age, education, ethnicity, parity, season of birth	Logistic regression	$PM_{2.5}$ (10 μg/m³) NO_2 / O_3 (1 pphm) CO (1 ppm)	Mod

Study	Location(s) & Period	Study design & births (PTBs)	Exposure assessment	Exposure window(s)	Adjusted covariates	Statistical model	Pollutant(s) & exposure increment(s)	Risk of Bias
Ruan (阮海燕) et al., 2010	Panyu, Guangzhou, China 2006-2008	Case-control 9648 (4924)	***Monitor(s)*:** Daily concentrations from one monitor at Panyu district were used and assigned to mothers at a city level.	T1-3	Age, education, infant sex, parity	Logistic regression	PM_{10} / NO_2 / SO_2 (Unclear)	High
Saldiva et al., 2018	São Paulo, Brazil 2012-2013	Case-control 348337 (40755)	**Sampling:** Passive sample tubes used for 7-10 days, and average concentrations were attributed to closest maternal residence at delivery.	WP	Age, type of pregnancy, type of delivery	Logistic regression	NO_2 / O_3 (Quartile)	High
Salihu et al., 2012	Hillsborough, FL, US 2000-2007	Retrospective cohort 103961 (9549)	***Monitor(s)*:** Daily concentrations from 14 monitoring stations, assigned to maternal residential ZIP code using nearest monitor approach.	WP	Age, education, smoking, parity, infant sex, year of birth, marital status, maternal conditions, adequacy of prenatal care	WP	PM_{10} / $PM_{2.5}$ (Dichotomised)	Mod
Savitz et al., 2019	New York City, NY, US 2008-2010	Retrospective cohort 257956 (18996)	***Modelled***: Biweekly concentrations from New York Community Air Survey with information from LUR model and USEPA monitors, assigned at maternal residence at delivery.	T2	Age, education, ethnicity, pre-pregnancy BMI, SES, parity, year of conception, Medicaid status, hospital	Logistic regression	$PM_{2.5}$ (10 μg/m³) NO_2 (10 ppb)	Mod

Study	Location(s) & Period	Study design & births (PTBs)	Exposure assessment	Exposure window(s)	Adjusted covariates	Statistical model	Pollutant(s) & exposure increment(s)	Risk of Bias
Schifano et al., 2016	Rome, Italy Barcelona, Spain 2001-2012	Retrospective cohort 27255 (1266)	***Monitor(s)*:** Daily concentrations from monitors of Lazio Environmental Protection Agency or network of the Catalan Government, assigned to maternal residence at citywide level.	WP	Age, education, moth, year, nationality of mother, maternal conditions	Cox-PH regression	PM_{10} / NO_2 / O_3 (IQR)	Mod
Seabrook et al., 2019	Ontario, Canada 2009-2014	Retrospective cohort 25263 (1894)	***Modelled***: Concentration of O_3 from Global Environmental Multi-Scale Modelling Air Quality and Chemistry Model at 10 km^2 grid resolution, and SO_2 from the Ozone Monitoring Instrument satellite data at 20 km^2 grid resolution, assigned to maternal residence at delivery at 6-digit postal code level.	WP	Age, education, smoking, infant sex, SES, pre-pregnancy BMI, family situation, alcohol, physical environmental variable, population density, maternal conditions	Logistic regression	SO_2 / O_3 (unclear)	Mod
Sheridan et al., 2019	CA, US 2005-2010	Retrospective cohort 2293218 (187275)	***Monitor(s)*:** $PM_{2.5}$ estimate for each zip code was interpolated IDW using measurements from the 3 nearest monitors within 20 km of the zip code.	T1-3 WP	Age, education, parity, infant sex, ethnicity/race, year of conception, Medicaid insurance status, season of conception	Cox-PH regression	$PM_{2.5}$ (10 µg/m³)	Mod

Study	Location(s) & Period	Study design & births (PTBs)	Exposure assessment	Exposure window(s)	Adjusted covariates	Statistical model	Pollutant(s) & exposure increment(s)	Risk of Bias
Shi et al., 2021	Zhejiang, Huangzhou, China 2017-2020	Retrospective cohort 439 (93)	***Modelled***: Daily concentration from China High Air Pollutants (CHAP) datasets at 1 km^2 grid resolution, and national monitors of China National Environmental Monitoring Center, assigned to maternal residence at delivery.	WP	Age, smoking, education, pre-pregnancy BMI, employment, type of residence, temperature, relative humidity	Cox-PH regression	PM_{10} / $PM_{2.5}$ / NO_2 / SO_2 / O_3 / CO (IQR)	High
Siddika et al., 2019	Espoo, Finland 1984-1900	Prospective cohort 2453 (195)	***Modelled***: Daily concentrations from the system for integrated modelling of atmospheric compositions (SILAM) at 0.1° x 0.1° resolution, assigned to maternal residence at delivery.	WP	Age, smoking, SES, infant sex, single parenthood, season of birth	Poisson regression	PM_{10} / $PM_{2.5}$ (10 µg/m^3) NO_2 / SO_2 / O_3 / CO (10 ppb)	Low
Siddika et al., 2020	Espoo, Finland 1984-1900	Prospective cohort 2427 (178)	***Modelled***: Daily concentrations from SILAM at 0.1° x 0.1° resolution, assigned to maternal residence at delivery.	T1-3	Age, SES, smoking, second hand smoke exposure, infant sex, season of birth, single parenthood	Poisson regression	PM_{10} / $PM_{2.5}$ / NO_2 / SO_2 / O_3 / CO (Quartile)	Mod
Smith et al., 2020	London, United Kingdom 2006-2010	Retrospective cohort 578382 (33712)	***Modelled***: Concentrations from the KCL urban model using ADMS v4 and road source model v2.3 at 20 m^2	T1-2	Age, ethnicity, infant sex, year of birth, birth registration type, tobacco	Logistic regression	PM_{10} / $PM_{2.5}$ / NO_2 / O_3 (IQR)	Low

Study	Location(s) & Period	Study design & births (PTBs)	Exposure assessment	Exposure window(s)	Adjusted covariates	Statistical model	Pollutant(s) & exposure increment(s)	Risk of Bias
			grid resolution, assigned to maternal address.		expenditure (COA-level), Carstairs quintile, season of conception, random intercept for MSOA			
Stieb et al., 2016a	Canada 1999-2008	Retrospective cohort 2966705 (184765)	***Modelled***: Monthly concentrations from a LUR model that incorporates monitoring and satellite-derived estimates, assigned to maternal post code.	T1-3 WP	Age, maternal origin, municipality, parity, infant sex, year of birth, marital status, paternal age, area-level variables	Logistic regression	$PM_{2.5}$ / NO_2 (IQR)	Low
Stieb et al., 2016b	Canada 1999-2008	Retrospective cohort 779040 (49850)	***Modelled***: Monthly concentrations derived from a North American LUR model with monitoring and satellite-derived estimates, assigned to maternal 6-character postal code.	WP	Age, maternal origin, municipality, parity, infant sex, year of birth, marital status, paternal age, area-level variables	Logistic regression	$PM_{2.5}$ / NO_2 (IQR)	Low
Sun et al., 2019	Zhejiang, China 2013-2017	Prospective cohort 6275 (372)	***Monitor(s)*:** Daily concentrations obtained from the closest among the 47 ground-based monitoring stations,	T1-3 WP	Age, pre-pregnancy BMI, employment status, parity, infant sex,	Logistic survival model	PM_{10} / $PM_{2.5}$ / NO_2 / SO_2 / O_3 / CO (10 µg/m³)	Mod

Study	Location(s) & Period	Study design & births (PTBs)	Exposure assessment	Exposure window(s)	Adjusted covariates	Statistical model	Pollutant(s) & exposure increment(s)	Risk of Bias
			assigned to maternal residential address at delivery.		gravidity time, occupation, gestational hypertension, other maternal conditions			
Sun et al., 2020	CA, US 2005-2013	Retrospective cohort 1967300 (127039)	***Monitor(s)***:Daily concentrations obtained from USEPA AQS, interpolated using IDW with a 20 km radius around the maternal residence ZIP code centroids.	WP	Age, education, ethnicity, parity, infant sex, insurance, season	Cox-PH regression	PM_{10} / $PM_{2.5}$ / NO_2 / O_3	N/A
Symanski et al., 2014	Harris, TX, US 2005-2007	Retrospective cohort 177816 (18489)	***Monitor(s)***: Daily concentrations calculated from USEPA AQS monitors, assigned to maternal residential county at delivery.	Every 4 weeks	Age, education, parity, pre-pregnancy BMI, ethnicity, season, others*	Logistic regression	$PM_{2.5}$ (10 μg/m³)	Mod
Symanski et al., 2016	Houston, Harris, TX, US 2005-2007	Retrospective cohort 152214 (14278)	***Monitor(s)***: Daily concentrations calculated from USEPA AQS monitors, assigned to maternal residential county at delivery.	Every 4 weeks	Age, education, smoking, ethnicity, adjacent pregnancy periods, season of birth	Logistic regression	O_3 (10 ppb)	Mod
Tapia et al., 2020	Lima, Peru 2012-2016	Retrospective cohort	***Model(s)***: Daily concentrations obtained from satellite	T1-3 WP	Age, pre-pregnancy BMI, SES, parity,	Logistic regression	$PM_{2.5}$ (IQR)	Mod

Study	Location(s) & Period	Study design & births (PTBs)	Exposure assessment	Exposure window(s)	Adjusted covariates	Statistical model	Pollutant(s) & exposure increment(s)	Risk of Bias
		123034 (8897)	measurements and chemical transport model data at 1km^2 grid resolution, assigned to district of maternal residence.		antenatal care, preeclampsia, gestational diabetes mellitus			
Trasande et al., 2013	US 1997-2006	Cross-sectional 222359 (75978)	***Monitor(s)***: Monthly concentrations of pollutants obtained from USEPA AIRS database, assigned to address of hospital of delivery with a 10-mile radius.	Month of birth	Ethnicity, infant sex, primary expected payer, median income quartile for patient ZIP code, hospital region, admission month, teaching hospital status	Logistic regression	$PM_{2.5}$ (1 μg/m³) NO_2 (1 ppb)	Mod
Tu & Tu, 2018	Atlanta, GA, US 2000	Cross-sectional 116112 (11445)	***Modelled***: Daily concentrations obtained from USEPA Fused Air Quality Predictions Surface (AQPS), assigned to maternal residential address at census tract level.	WP	Age, smoking, alcohol use, education, ethnicity, parity, income, infant sex, antenatal care, marital status, urban land use	Global logistic regression	$PM_{2.5}$ (1 μg/m³) O_3 (1 ppb)	Mod
van den Hooven et al., 2012	Rotterdam, Netherlands 2001-2005	Prospective cohort 7772 (412)	***Modelled***: Trimester-specific and WP concentrations obtained from monitors and dispersion modelling,	WP	Age, education, ethnicity, pre-pregnancy BMI, smoking, alcohol use, infant sex, folic acid	Logistic regression	PM_{10} / NO_2 (1 μg/m³)	Low

Study	Location(s) & Period	Study design & births (PTBs)	Exposure assessment	Exposure window(s)	Adjusted covariates	Statistical model	Pollutant(s) & exposure increment(s)	Risk of Bias
			assigned to maternal residence at delivery.		consumption, anthropometrics, noise exposure			
Villeneuve et al., 2022	Canada 1998-2006	Retrospective cohort 2227819 (139495)	***Modelled***: Monthly concentrations obtained from a North America LUR model, developed using monitoring stations and satellites, assigned at maternal address at postcode level.	WP	Age, marital status, maternal place of birth and residence, maternal province of residence, season of birth, neighbourhood SES, visible minority	Logistic regression	$PM_{2.5}$ / NO_2 (IQR)	Low
Wang et al., 2018	Guanzhou, Guangdong, China 2015-2017	Retrospective cohort 469975 (25879)	***Monitor(s)***: Daily concentrations obtained from monitors and were assigned at maternal residential address at district level.	T1-3 WP	Age, parity, infant sex, year of birth, temperature, season of conception, delivery mode, district, maternal conditions	Cox-PH regression	PM_{10} / $PM_{2.5}$ NO_2 / SO_2 / O_3 (IQR)	Mod
Wang et al., 2020	Guangzhou, Guangdong, China 2015-2017	Retrospective cohort 215509 (11109)	***Monitor(s)***: Daily concentrations obtained from State Environmental Protection Administration monitors, assigned at maternal residential district level.	WP	Age, parity, month of birth, temperature, relative humidity, residency district, heatwave exposure, gestational $PM_{2.5}$ at 20-28th weeks	Cox-PH regression	$PM_{2.5}$ (10 μg/m³)	Mod

Study	Location(s) & Period	Study design & births (PTBs)	Exposure assessment	Exposure window(s)	Adjusted covariates	Statistical model	Pollutant(s) & exposure increment(s)	Risk of Bias
					(general PTB only)			
Wang et al., 2022	Jinan, Shandong, China 2018-2019	Prospective cohort 424 (17)	***Monitor(s)***: Hourly concentrations obtained from Tianqiao station, averaged to everyday level and assigned to maternal hospital of delivery.	WP	Age, education, parity, infant sex, pre-pregnancy BMI, second-hand smoke exposure, paternal alcohol use, menstruation status, gravidity, age at menarche	Logistic regression	PM_{10} / $PM_{2.5}$ NO_2 / SO_2 / O_3 / CO (IQR)	High
Warren et al., 2012	Texas, US 2002-2004	Retrospective cohort 32170 (unclear)	***Modelled***: Daily concentrations obtained from USEPA AQS and FSD, estimated by CMAQ at 12 km^2 and 36 km^2 grid resolution, assigned to maternal residence at delivery.	WP	Age, education, ethnicity, infant sex, paternal ethnicity, paternal age, paternal education	Bayesian probit regression	$PM_{2.5}$ (10 μg/m^3) O_3 (1 ppm)	N/A
Weber et al., 2019	San Joaquin, CA, US 2000-2006	Retrospective cohort 2552205 (21225)	***Monitor(s)***: Daily concentrations from USEPA AQS and 4 monitors, spatially interpolated using IDW with 50 km radius, assigned to maternal residential street address.	T1-3 WP Last 6 weeks	Age, education, ethnicity, payment of birth expenses, insurance type, prenatal care	Logistic regression	PM_{10} / $PM_{2.5}$ NO_2 / CO (Quartile)	High

Study	Location(s) & Period	Study design & births (PTBs)	Exposure assessment	Exposure window(s)	Adjusted covariates	Statistical model	Pollutant(s) & exposure increment(s)	Risk of Bias
Wilhelm et al., 2003	Los Angeles, CA, US 1994-1996	Case-control 34588 (13464)	***Monitor(s)***: Annual background concentrations were obtained from SCAQMD, assigned to maternal residence using nearest monitor approach.	WP	Age, education, ethnicity, infant sex, parity, history of antenatal visit, interval since previous livebirth, previous LBW/PTB, season of birth	Logistic regression	PM_{10} (10 µg/m³) NO_2 / O_3 (1 pphm) CO (1 ppm)	Mod
Wilhelm & Ritz, 2005	Los Angeles county, CA, US 1994-2000	Case-control 106483 (9268)	***Monitor(s)***: Daily concentrations obtained from SCAQMD, assigned at maternal zip code level.	T1-2 1st month Last 6 weeks	Age, education, ethnicity, parity, infant sex, infant sex, antenatal care, interval since previous childbirth, previous PTBs, season of birth	Logistic regression	$PM_{2.5}$ (10 µg/m³)	Mod
Wilhelm et al., 2011	Los Angeles, CA, US 2004-2008	Matched case-control 11203 (10134)	***Monitor(s)***: Daily concentrations from MATES monitoring stations, assigned to maternal residential address at a 5-mile radius.	WP	Age, education, ethnicity, parity, marital status	Conditional logistic regression	$PM_{2.5}$ (5 µg/m³)	Low
Williams et al., 2021	US 2002-2008	Retrospective cohort 204891 (19888)	***Modelled***: Hourly concentrations from modified CMAQ model	T1-3 WP	Age, smoking, alcohol use, ethnicity, pre-pregnancy BMI, maternal	Binary poisson regression	$PM_{2.5}$ (IQR)	High

Study	Location(s) & Period	Study design & births (PTBs)	Exposure assessment	Exposure window(s)	Adjusted covariates	Statistical model	Pollutant(s) & exposure increment(s)	Risk of Bias
			assigned at hospital referral regions.		conditions, health insurance, marital status, area-level factors			
Wing et al., 2022	Los Angeles, CA, US 2008-2016	Retrospective cohort 174186 (15134)	***Modelled***: Monthly concentrations from LUR model and CARB routine monitor network, assigned to maternal residence at delivery.	WP	Age, education, ethnicity, SES, smoking, aircraft-origin UFP	Logistic regression	NO_2 (Quartile)	High
Wu et al., 2009	Los Angeles, Orange, CA, US 1997-2006	Retrospective cohort 81186 (6712)	***Modelled***: Monthly concentrations from a modified CALINE4 dispersion model, assigned to maternal residential address.	T1-3 WP	Age, ethnicity, prenatal care insurance, poverty, season of conception, pyelonephritis	Logistic regression	$PM_{2.5}$ (10 μg/m³)	Low
Wu et al., 2011	Los Angeles and Orange, CA, US 1997-2006	Retrospective cohort 81186 (3928)	***Monitor(s)***: Monthly concentrations from CARB, assigned using nearest monitoring stations to residential location.	T1-3 WP	Age, ethnicity, pyelonephritis, prenatal care insurance type, poverty, season of conception	Conditional logistic regression	PM_{10} / $PM_{2.5}$ / NO_2 / O_3 / CO (IQR)	Mod
Wu et al., 2016	CA, US 2001-2008	Matched case-control 3972594 (356078)	***Modelled***: Hourly concentrations spatially interpolated from CARB using kriging with output at 200 m^2 grid resolution, assigned to maternal residence.	WP	Age, education, ethnicity, median household income at census block group level	Conditional logistic regression	$PM_{2.5}$ / $PM_{0.1}$ / NO_2 / O_3 / CO / UFP (IQR)	Low

Study	Location(s) & Period	Study design & births (PTBs)	Exposure assessment	Exposure window(s)	Adjusted covariates	Statistical model	Pollutant(s) & exposure increment(s)	Risk of Bias
Xiao et al., 2018	Shanghai, China 2011-2014	Retrospective cohort 132783 (5855)	***Modelled***: Monthly concentrations derived from CMAQ model with AOD data at 1 km^2 grid resolution, assigned to maternal residential address.	T1-3 WP	Age, education, parity, infant sex, temperature, paternal education level, birth location type, season of conception, long-term temporal trend, average surface temperature	Logistic regression (T1-2), Time-to-survival model (T3, WP)	$PM_{2.5}$ (10 μg/m^3)	Low
Xu (徐丽珍) et al., 2008	Taiyuan, Shanxi, China 2005-2007	Retrospective cohort 31145 (1092)	***Monitor(s)***: Daily concentrations from Taiyuan Environmental Protection Monitoring Stations.	First 3 months Last 3 months	Age, education, smoking, alcohol	Logistic regression	PM_{10} / NO_2 / SO_2 (categorical)	High
Yang (杨敏娟) et al., 2016	Shanghai, China 2013	Retrospective cohort 38083 (1782)	***Monitor(s)***: Daily concentrations from monitoring network, assigned to maternal residence address at delivery at district level.	T1-3 WP	Age, education, infant sex, parity, paternal household registration, season of delivery, maternal cultural background, congenital anomaly	Logistic regression	PM_{10} / NO_2 / SO_2 (unclear)	Mod

Study	Location(s) & Period	Study design & births (PTBs)	Exposure assessment	Exposure window(s)	Adjusted covariates	Statistical model	Pollutant(s) & exposure increment(s)	Risk of Bias
Yang et al., 2020b	Guangdong, Guangzhou, China 2014-2017	Retrospective cohort 1455026 (64028)	***Monitor(s)***: Concentrations from monitors were assigned to maternal residence address at delivery at district level.	T1-3 WP Last 4 weeks	Age, temperature, infant sex, parity, relative humidity, gravidity	Cox-PH regression	O_3 (10 μg/m³)	Mod
Ye et al., 2018	Taizhou, China 2013-2016	Retrospective cohort 24246 (1501)	***Monitor(s)***: Exposure values calculated from daily concentrations using IDW approach, assigned to maternal residential address.	T1-3 WP	Age, education, parity, infant sex, year of birth, season of birth	Logistic regression	PM_{10} / $PM_{2.5}$ / NO_2 (IQR)	Mod
Yuan et al., 2020	Shanghai, China 2013-2016	Prospective cohort 3692 (168)	***Modelled***: Weekly concentrations retrieved from gaps filled AOD and other data sources at 1 km^2 grid resolution, assigned at maternal residential address.	T1-3 WP	Age, alcohol use, education, pre-pregnancy BMI, smoking, parity, infant sex, temperature, humidity, cooking fuel type, use of ventilator, birth season	Cox-PH regression	$PM_{2.5}$ (10 μg/m³)	Mod
Zhang et al., 2020	Wuhan, Hubei, China 2013-2015	Retrospective cohort 2101 (273)	***Modelled***: LUR method with kriging interpolation. Spatial scale of 1 km^2 grid. Exposure assigned using geocoded maternal addresses.	T1-3 WP	Age, education, parity, infant sex, delivery mode, gravidity, season of conception, gestational diabetes,	Logistic regression	$PM_{2.5}$ (10 μg/m³)	Low

Study	Location(s) & Period	Study design & births (PTBs)	Exposure assessment	Exposure window(s)	Adjusted covariates	Statistical model	Pollutant(s) & exposure increment(s)	Risk of Bias
					gestational hypertension			
Zhang et al., 2021b	Beijing, China 2017-2018	Prospective cohort 203 (15)	***Modelled***: Daily concentrations obtained using MAIAC aerosol optical depth at a 1 km^2 grid resolution, at the maternal residential address.	T1-3 WP	Age, pre-pregnancy BMI, gestational hypertension, gestational diabetes, parity, infant sex, temperature, relative humidity, season of conception, gravidity, mode of delivery	Logistic regression	$PM_{2.5}$ (10 μg/m³)	Mod
Zhao et al., 2020a	Hohhot, Inner Mongolia, China 2015-2016	Case-control 69945 (5842)	***Monitor(s)***: Daily concentrations from the 8 monitoring stations in the region, assigned to maternal residence at delivery using the nearest monitor approach.	T1-3 WP	Age, birth weight, season of conception	Logistic regression	PM_{10} / $PM_{2.5}$ / NO_2 / SO_2 / O_3 / CO (IQR)	Mod
Zhao (赵姗) et al., 2020b	Hebei, China 2015-2016	Retrospective cohort 91756 (5190)	***Monitor(s)*** Concentrations from automatic air quality monitoring and release system of Hebei Province.	T1-3 WP	Age, education, parity, infant sex, history of antenatal visit, maternal conditions, hospital, season of delivery,	Logistic regression	$PM_{2.5}$ (unclear)	High

Study	Location(s) & Period	Study design & births (PTBs)	Exposure assessment	Exposure window(s)	Adjusted covariates	Statistical model	Pollutant(s) & exposure increment(s)	Risk of Bias
					season of pregnancy			
Zhou et al., 2022a	Henan, China 2014-2016	Retrospective cohort 697316 (15224)	***Monitor(s)***: Daily concentrations calculated using IDW approach, assigned at county level.	T1-3 WP	Age, smoking, alcohol use, education, pre-pregnancy BMI, parity, infant sex, temperature, season of conception, relative humidity	Logistic regression	PM_{10} / $PM_{2.5}$ / NO_2 / SO_2 / O_3 (10 μg/m³) CO (1 mg/m³)	Mod
Zhou et al., 2022b	Chongqing, China 2015-2020	Retrospective cohort 572116 (33669)	***Monitor(s)***: Daily concentration obtained from nearest monitor with 10 km radius from China National Environmental Monitoring Stations at address level.	T1-3 WP	Age, temperature, humidity, paternal age, birth weight	General additive model	PM_{10} / $PM_{2.5}$ / NO_2 / SO_2 / O_3 / CO (SD)	Mod
Zhu (朱娟) et al., 2011	Guangdong, China 2009	Case-control 857 (86)	***Monitor(s)***: Daily concentrations obtained from monitors, assigned to maternal residence at district level.	PC T1-3 WP	Age, education, income, gestational sexual activity, maternal conditions, others*	Logistic regression	PM_{10} / NO_2 / SO_2 (Quartile)	High
Zou et al., 2021	Shanghai, China 2011-2012	Retrospective cohort 2527 (107)	***Monitor(s)***: Daily concentrations from Shanghai Environmental Monitoring Centre assigned	T1-3 WP	Age, smoking, infant sex, asthma, employment, second-hand	Logistic regression	PM_{10} / NO_2 / SO_2 (IQR)	High

Study	Location(s) & Period	Study design & births (PTBs)	Exposure assessment	Exposure window(s)	Adjusted covariates	Statistical model	Pollutant(s) & exposure increment(s)	Risk of Bias
			to maternal residence at district level.		smoke exposure, maternal conditions, others*			

Row in green: studies eligible to be included in the meta-analysis
Row in grey: studies not eligible to be included in the meta-analysis
List of countries

- Backwenkizi et al. (2022): Angola, Benin, Burundi, Cameroon, Chad, Ethiopia, Guinea, Mali, Nigeria, Malawi, South Africa, Tanzania, Uganda, Zambia, Zimbabwe.
- Gigoris-Allemand et al. (2017): Denmark, England, France, Greece, Hungary, Italy, Lithuania, Netherlands, Norway, Spain, Sweden

Others*

- Backwenkizi et al. (2022): wealth quantile, place of residence, ozone, relative humidity, annual government expenditure on health services
- Ji et al. (2019): anthropometrics, placental pathologies, pathological uterus
- Lu et al. (2019): wind speed, cockroaches, perinatal pets, water damage, living near traffic
- Pereira et al. (2016): Illegal drug use during pregnancy, depression, unmeasured temporal factors
- Symanski et al. (2014): receipt of WIC services, trimester where prenatal care initiated, paternal support, type of payment for delivery
- Zhu (朱娟) et al., 2011: previous preterm birth or stillbirth, congenital abnormalities, family history of stillbirth or preterm birth
- Zou et al. (2021): ownership of residence, home dampness, renovation of residence, distance to main road

Abbreviations for US/Canadian states

BC: British Columbia	GA: Georgia	OH: Ohio
CA: California	MA: Massachusetts	PA: Pennsylvania
CO: Colorado	MI: Michigan	PR: Puerto Rico
CT: Connecticut	NC: North Carolina	RI: Rhodes Island
FL: Florida	NY: New York	

4.3.2. Quality assessment

More than half of the studies were assigned a moderate risk of bias (n = 92) (**Table 18**). Similar numbers of studies were assigned a high risk of bias (n=28) and a low risk of bias (n = 30). Shortcomings in the domains of exposure assessment and confounder adjustment were the main causes of studies being assigned a moderate risk of bias. Studies that were assigned a high risk of bias had shortcomings in the two aforementioned domains as well as the study design domain, usually due to the sample size being very small, for instance, less than 3,000, but did not provide a justification for the sample size. Studies were also assigned high risk of bias when key information was missing from the publication, such as the confounders adjusted for, spatial resolution of the model used for assessing exposure, concentration distribution of the pollutants, and the unit (ppb or $\mu g/m^3$) of pollutant(s). All studies in Chinese were assigned a high risk of bias, mostly because of missing information and the rest due to issues in exposure assessment.

Most studies had not adjusted for all critical confounders specified in the quality assessment framework (n = 110). However, in most cases, at least one critical confounder was adjusted for. Maternal age was the most common critical confounder to be adjusted for (n = 140), followed by socioeconomic status (n = 121), parity (n = 98), and season of conception (n = 86). For the potential confounders of interest, infant sex was most frequently adjusted for (n = 90), followed by smoking (n = 75), ethnicity (n = 64), year of birth (n = 37), pre-pregnancy BMI (n = 30), alcohol consumption (n = 22), and marital status (n = 22).

4.3.3. Narrative synthesis and interpretation of findings

Table 19 shows the number of studies and those that found an increase in the risk of PTB (or its subtypes) in each pollutant-exposure period-outcome relationship. The percentage of studies finding an increase in the risk of PTB associated with trimester 1-3 and whole pregnancy exposure for all pollutants ranged from 21.9% (Trimester 1 – O_3 exposure) to 50.6% (Whole pregnancy – $PM_{2.5}$ exposure).

Fewer studies have studied the effect of air pollution exposure on the risk of PTB subgroups, and a few studies showed a more consistent increase in risk than the others. For instance, Chen et al. (2021b) was the only study that found an increase in risk of PTB, moderate-to-late PTB, and very PTB associated with preconception $PM_{2.5}$ exposure, with no study showing an increase in risk for NO_2 or O_3 across all exposure windows and types of PTB. It was also the only study that found an increase in the risk of very PTB associated with trimester 2 O_3 exposure, and risk of extremely PTB associated with trimester 1 and 3 NO_2

exposure. Other studies that have frequently found an increase in risk across various pollutant-exposure window-PTB subtype relationships include Fang et al. (2022) (n = 6), Weber et al. (2019) (n = 6), Zhou et al. (2022a) (n = 6), and Liang et al. (2022) (n = 5).

Table 19. The distribution of studies and those that have reported an increase in risk of PTB or subgroups of PTB using ORs/RRs/HRs in each pollutant-exposure period-outcome relationship, among those that were eligible to be included in the narrative synthesis.

Pollutant	Type*	Exposure window*				
		PC (n = 7)	T1 (n = 98)	T2 (n = 94)	T3 (n = 86)	WP (n = 127)
$PM_{2.5}$ (n = 115)	PTB (n = 104)	1 (5)	20 (68)	22 (67)	23 (59)	44 (87)
	mlPTB (n = 9)	1 (1)	5 (6)	6 (7)	4 (6)	7 (9)
	vPTB (n = 12)	1 (1)	6 (9)	5 (10)	4 (9)	9 (10)
	exPTB (n = 8)	0 (1)	2 (5)	1 (6)	1 (5)	3 (6)
NO_2 (n = 89)	PTB (n = 82)	0 (4)	12 (48)	11 (45)	12 (42)	20 (62)
	mlPTB (n = 6)	0 (1)	2 (3)	2 (4)	1 (3)	3 (5)
	vPTB (n = 7)	0 (1)	3 (5)	2 (6)	2 (5)	4 (6)
	exPTB (n = 4)	0 (1)	1 (2)	0 (3)	1 (2)	3 (4)
O_3 (n = 56)	PTB (n = 54)	0 (3)	7 (32)	8 (27)	6 (25)	10 (37)
	mlPTB (n = 4)	0 (1)	0 (2)	0 (2)	0 (2)	0 (3)
	vPTB (n = 4)	0 (1)	0 (2)	1 (2)	0 (2)	0 (3)
	exPTB (n = 3)	0 (1)	0 (1)	0 (1)	0 (1)	0 (2)

Numbers in brackets are number of studies in that particular pollutant-exposure period-outcome relationship. *Those that used other exposure windows and PTB types were not included here, as not all the estimates for these exposure windows and PTB types were extracted. PC: preconception, T1: trimester 1, T2: trimester 2, T3: trimester 3, WP: whole pregnancy.

4.3.3.1. Exposure assessment

4.3.3.1.1. Did the method of exposure assessment influence the risk of PTB observed?

Six studies compared the risk estimates derived from different exposure assessment methods (Brauer et al., 2008; Hannam et al., 2014; Hyder et al., 2014; Liang et al., 2022; Stieb et al., 2016b; Xiao et al., 2018). Five compared monitor-based approaches to modelled approaches, for instance, LUR models (Brauer et al., 2008; Hannam et al., 2014; Hyder et al., 2014; Stieb et al., 2016b; Xiao et al., 2018), and one compared nearest monitors and IDW (Liang et al., 2022). Those that provided continuous estimates are presented in **Figure 3**. Negligible differences were found between the estimates derived from the nearest monitor

and IDW (Liang et al., 2022). Similarly, minimal differences were observed when monitor-based vs model-based approaches for estimating $PM_{2.5}$ exposures were compared (Brauer et al. 2008; Hannam et al., 2014, Hyder et al., 2014; Xiao et al., 2018). Xiao et al. (2018) found that satellite-based modelling produced higher and statistically significant risk estimates compared to monitors, which conform with the notion that monitor-based estimates are prone to greater measurement error and were often attenuated. However, this is not found in Hyder et al. (2014) which did the same comparison, which could be due to the much higher concentrations available in Xiao et al. (2018) compared to Hyder et al. (2014) which allows for greater exposure contrast.

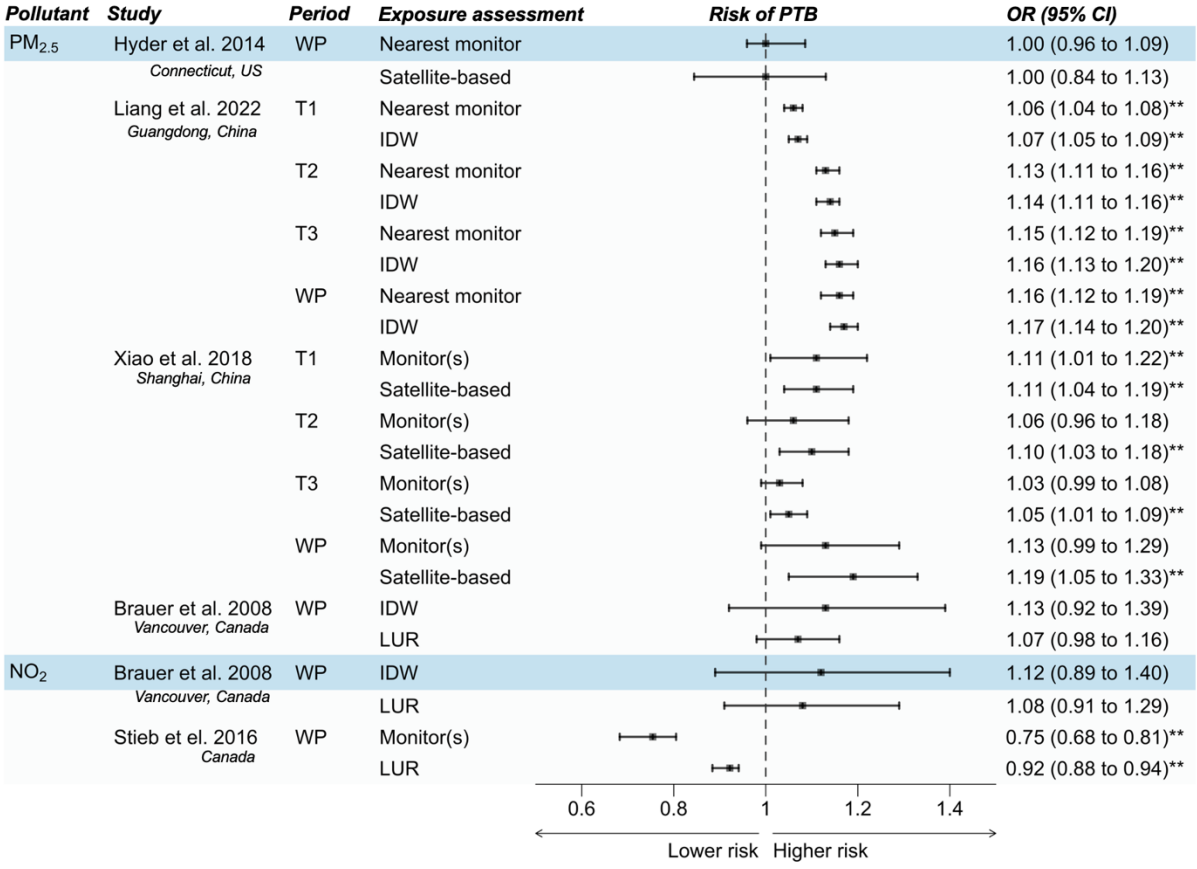


Figure 3. Primary studies that have provided estimates from both non-modelled and modelled approaches. The estimates are in odds ratios (OR) and in 10 μg/m³ increment. Estimates from Hannam et al. (2014) were not presented in this figure as they were categorical.

On the contrary, Stieb et al. (2016) found a statistically significant decrease in the risk of PTB from NO_2 exposure estimated using monitor(s) and a temporally adjusted LUR model. The two estimates were statistically different, with no overlaps in 95% CI. Stieb et al. (2016) had the largest sample size among the six studies and relatively low NO_2 levels. Therefore,

the findings are inconclusive and have two implications: (1) There were no differences in the estimates derived using monitor-based versus modelled approaches, (2) there were differences in the estimates, but it can only be detected with a sufficiently large sample size, or the differences were more pronounced for pollutants with greater spatial contrast in concentration, e.g. NO_2.

4.3.3.1.2. Did accounting for residential mobility or time-activity pattern produce estimates for the risk of PTB that were statistically different from those that did not?

Most studies did not account for maternal mobility during pregnancy. Six studies incorporated the changes in address during pregnancy, or conducted a stratified analysis according to whether an address change occurred or the amount of time spent at home (Estarlich et al., 2016; Gehring et al., 2011b; He et al., 2022a; Lavigne et al., 2016; Llop et al., 2010; Panasevich et al., 2016). **Figure 4** shows the estimates from two studies that were derived using continuous exposure.

Gehring et al. (2011b) compared those who had moved during pregnancy and those who did not, and Panasevich et al. (2016) provided the estimate for non-movers compared to the main estimate, and both found no significant differences. Whether maternal mobility influences the risk estimate produced might depend on the distance the mothers have moved from the original address. This was not provided by the two studies, but another study of successive pregnancies that had explored this found that half of its population had moved a median of 4.9 km. However, the nearest monitor did not change for 95% of mothers, so there was limited change in the estimate produced (Pereira et al., 2016).

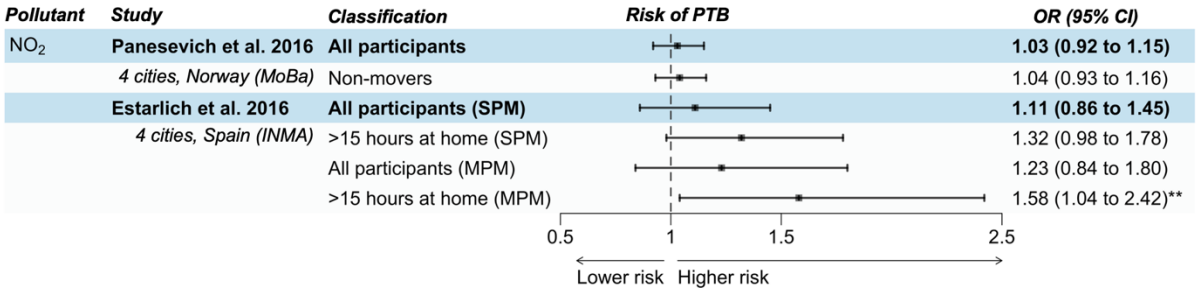


Figure 4. Primary studies that have provided estimates for continuous exposure for participants that have changed address or spent more than a certain amount of time at home daily during pregnancy. Odds ratios (ORs) presented are in 10 μg/m³ increment. The location of the study is shown below its name, with the name of the cohort in brackets. INMA: INfancia y Medio Ambiente – Childhood and Environment project. MoBa: Norwegian Mother and Child Cohort Study. SPM: Single pollutant model. MPM: two-pollutant model (NO_2 – Benzene)

Two studies that used the Spanish INMA cohort have compared the results derived using the entire study population and those that have spent more than a certain amount of time at home during pregnancy (Estarlich et al., 2016; Llop et al., 2010). Llop et al. (2010) found no change in the risk of PTB for those who spent ≥ 16 hours at home. Estarlich et al. (2016) utilised a more extended time period and greater sample size. They observed that mothers who spent more than 15 hours on average at home per day during pregnancy have a higher risk, which became statistically significant when two-pollutant models (NO2 – Benzene) were used.

Findings from the studies retrieved suggest that residential mobility may have a limited impact on the risk of PTB observed. However, the impact could increase if most of the study population moved a great distance and caused changes in the monitor/grid of the model assigned to the address. On the other hand, the time-activity pattern of mothers could be influential to the risk of PTB observed. However, the impact could only be detected with a greater sample size and adjustment of other pollutants.

4.3.3.2. Environmental factors

4.3.3.2.1. Did the risk of PTB remain robust after adjustment of other pollutants of interest?

Thirty-four studies in the narrative synthesis have used co/multipollutant models (MPM) to disentangle the effects of different air pollutants. After excluding studies that have only provided MPM results graphically or only on $PM_{2.5}$ and its sources and constituents (**Appendix 9**), there were 18 studies that compared results from single-pollutant models (SPM) and MPM for whole pregnancy $PM_{2.5}$, NO_2, and O_3 exposure. Key information and results from their analyses are shown in **Appendix 10**. Among the 18 studies, 9 had provided estimates of continuous exposure for $PM_{2.5}$ and NO_2 in SPM and MPMs, where the two pollutants were weakly to moderately correlated (**Figure 5**).

For $PM_{2.5}$ exposure, the addition of NO_2 and/or O_3 exposure(s) did not result in notable changes in the estimates across all exposure periods in most studies, although there were some small increases/decreases in the magnitude of the estimates. A statistically significant increase in risk of $PM_{2.5}$ exposure was found by Zhou et al. (2022b) in trimester 3 but was no longer present when other pollutants were adjusted for. Similarly, minimal changes were found after adjustment for other pollutants except for Zhou et al. (2022b) for O_3 exposure, where a statistically significant increase in risk of PTB was found after adjustment for $PM_{2.5}$, PM_{10}, SO_2, NO_2, and CO.

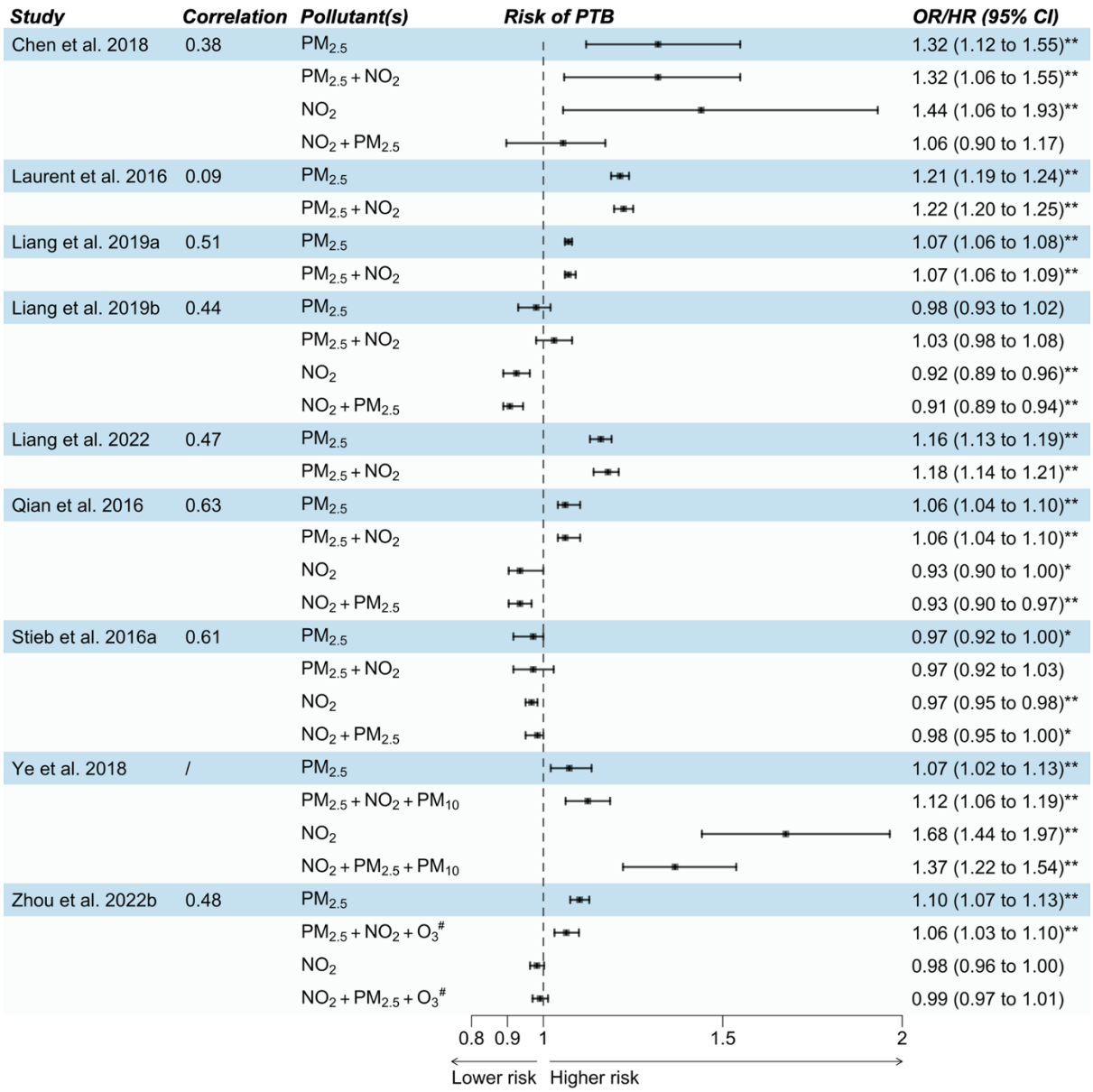


Figure 5. Estimates from studies that have compared results from SPM and MPM with $PM_{2.5}$ and NO_2 exposure. The estimates were converted to per 10μg/m^3 increment. Values presented in the Correlation column were the Spearman coefficient given for the correlation between NO_2 and $PM_{2.5}$ concentrations in that study. #: PM_{10}, SO_2, and CO were also adjusted.

For NO_2 exposure, most studies also did not find notable changes with adjustment for other pollutants. However, two studies found a noticeable attenuation of the effect estimate after $PM_{2.5}$ exposure was adjusted for (Chen et al., 2018; Ye et al., 2018), and for one study the effect was no longer statistically significant (Chen et al., 2018). Smith et al. (2020) found a statistically significant decrease in risk of PTB for trimester 1 NO_2 exposure that was no longer present after O_3 exposure was adjusted for.

Overall, adjustment for exposure to other pollutants has a minimal impact on the risk of PTB observed for $PM_{2.5}$ exposure, suggesting that the estimates derived from SPM likely did not overestimate the effect, or were not confounded by pollutants such as NO_2 and O_3. There was a relatively low correlation between $PM_{2.5}$ and NO_2 exposure in these studies, suggesting that this observation may not be generalisable to settings where they highly correlate with each other. $PM_{2.5}$ and O_3 concentrations ranged from moderately negative to strongly positive. However, in a study where they strongly correlated with each other, the results from SPM of $PM_{2.5}$ or O_3 exposure remained robust in MPM (Laurent et al., 2016). The attenuation of the estimates derived for NO_2 exposure upon adjustment for $PM_{2.5}$ (and PM_{10}) in two settings with similar levels of correlation between $PM_{2.5}$ and NO_2 concentrations suggests that it is possible that the effect of NO_2 exposure observed is partly driven by $PM_{2.5}$. However, it is likely that other factors driving such attenuation in these two studies were present, as this observation was not replicated in other studies retrieved.

4.3.3.2.2. Was the risk of PTB associated with $PM_{2.5}$ exposure generated from adverse events different from that of ambient sources?

Three studies have explored the impact on $PM_{2.5}$ exposure from adverse events, of which two were on wildfire smoke (Abdo et al., 2019; Heft-Neal et al., 2022), and one on coalmine fires (Melody et al., 2019a).

Abdo et al. (2019) found a statistically significant increase in risk of PTB associated with per 10μg/m^3 increase in trimester 2 (OR: 3.45 [2.32-5.15]) and whole pregnancy (OR: 2.08 [1.17-3.64]) wildfire $PM_{2.5}$ exposure, which were much stronger than risk found by studies of ambient $PM_{2.5}$ exposure (refer to Chapter 3 Table 1). The finding is supported by Heft-Neal et al. (2022), which found that additional days of exposure to any wildfire smoke is positively associated with the risk of PTB except in trimester 1. Both studies were conducted in the US where wildfire smoke is becoming a major source of $PM_{2.5}$, with evidence suggesting that it could be more harmful to the human body due to different compositions (Aguilera et al., 2021). Wildfires produce a large range of hazardous pollutant, including high levels of polycyclic aromatic hydrocarbons (PAHs) and volatile organic compounds (VOCs) which has been shown to be adversely linked to PTB (Cassidy-Bushrow et al., 2020; Freije et al., 2022; Miller et al., 2020). It is possible that the increased risk of PTB observed is partly attributable to the combined effect of $PM_{2.5}$ and other hazardous pollutants released, or other factors not considered in these studies such as stress.

Melody et al. (2020) did not find an association between $PM_{2.5}$ attributable to coal mine fires and the risk of PTB. This association is not frequently explored in the literature, but coal

mine fires exposure had shown to be linked with other adverse health outcomes in children, such as reduced lung functions and respiratory symptoms (Shao et al., 2019; Willis et al., 2020).

4.3.3.2.3. Was the risk of PTB associated with $PM_{2.5}$ exposure driven by certain sources or constituents?

16 studies explored if the risk of PTB observed was driven by certain sources or constituents. 11 of these studies reported a positive association (Basu et al., 2017; Cai et al., 2020; Fang et al., 2022; Hao et al., 2016; He et al., 2022; Kingsley et al., 2017; Laurent et al., 2016; Nealy et al., 2018; Ottone et al., 2020; Rappazzo et al., 2015; Wilhelm et al., 2011), where 5 of them did not (Hystad et al., 2014; Gehring et al., 2011a; Pereira et al., 2014c; Riddell et al., 2021; Smith et al., 2020).

Evidence regarding the risk of PTB associated with $PM_{2.5}$ sources was mainly on traffic (**Figure 6**) (Laurent et al., 2016; Ottone et al., 2020; Smith et al, 2020; Wilhelm et al., 2011). Whole pregnancy exposure to traffic-related diesel $PM_{2.5}$ emissions were found to be associated with higher risk of PTB in 3 studies (Laurent et al., 2016; Ottone et al., 2020; Wilhelm et al., 2011), and the risk could be much higher with no overlap in 95%CI than that of total ambient $PM_{2.5}$ exposure (Laurent et al., 2016). However, these findings were not replicated by Smith et al. (2020) that analysed the effect of exposure to traffic exhaust $PM_{2.5}$ in trimester 1 and 2. There were studies suggesting a greater risk of PTB associated with $PM_{2.5}$ generated from oil combustion (Ottone et al., 2020), biomass burning (Wilhelm et al., 2011), cooking (Laurent et al., 2016), and soil dust (Cai et al., 2020).

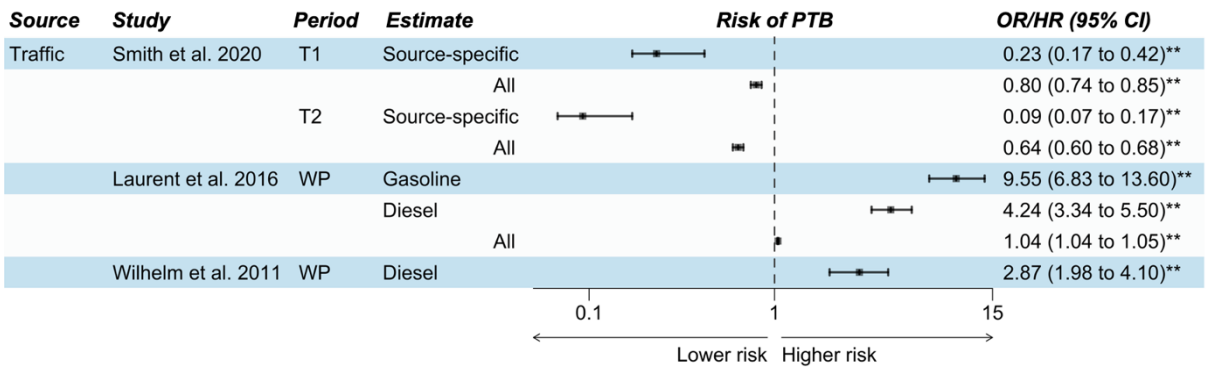


Figure 6. Results from studies that have provided estimates for the risk of PTB associated with continuous $PM_{2.5}$ exposure from traffic. The estimates were all converted to per 10μg/m³ increment. "All" refers to the estimate for total ambient $PM_{2.5}$ exposure. Wilhelm et al. (2011) did not provide an estimate for total ambient $PM_{2.5}$ exposure. Ottone et al. (2020) only provided estimates of categorical exposures and thus was excluded.

Most studies have suggested a positive association between whole pregnancy exposure to $PM_{2.5}$ constituents such as OC, EC, sulphates, nitrates, and ammonium, and the risk of PTB (**Figure 7**). The risk found, particularly for EC and ammonium, are much higher than that of ambient $PM_{2.5}$ exposure. The findings for OC and EC are consistent with that shown in **Figure 6**, as traffic generates a significant proportion of OC and EC in ambient $PM_{2.5}$ (Keuken & Brink, 2009). There are also studies showing positive associations with a variety of metals (Basu et al., 2017; Laurent et al., 2016). While there are still much fewer studies than those of ambient $PM_{2.5}$ exposure, this body of literature is growing, and a systematic review and/or meta-analysis can potentially be conducted to synthesise the effect shown by these studies. However, it could be challenging to interpret the findings from these studies, as the components often highly correlate with each other, thus it is unclear if a certain component is driving most of the effects observed. For instance, PAH is a constituent of $PM_{2.5}$ that was not investigated in the studies included here but has been shown to be linked with an increased risk of PTB (Freije et al., 2022; Jiang et al., 2022).

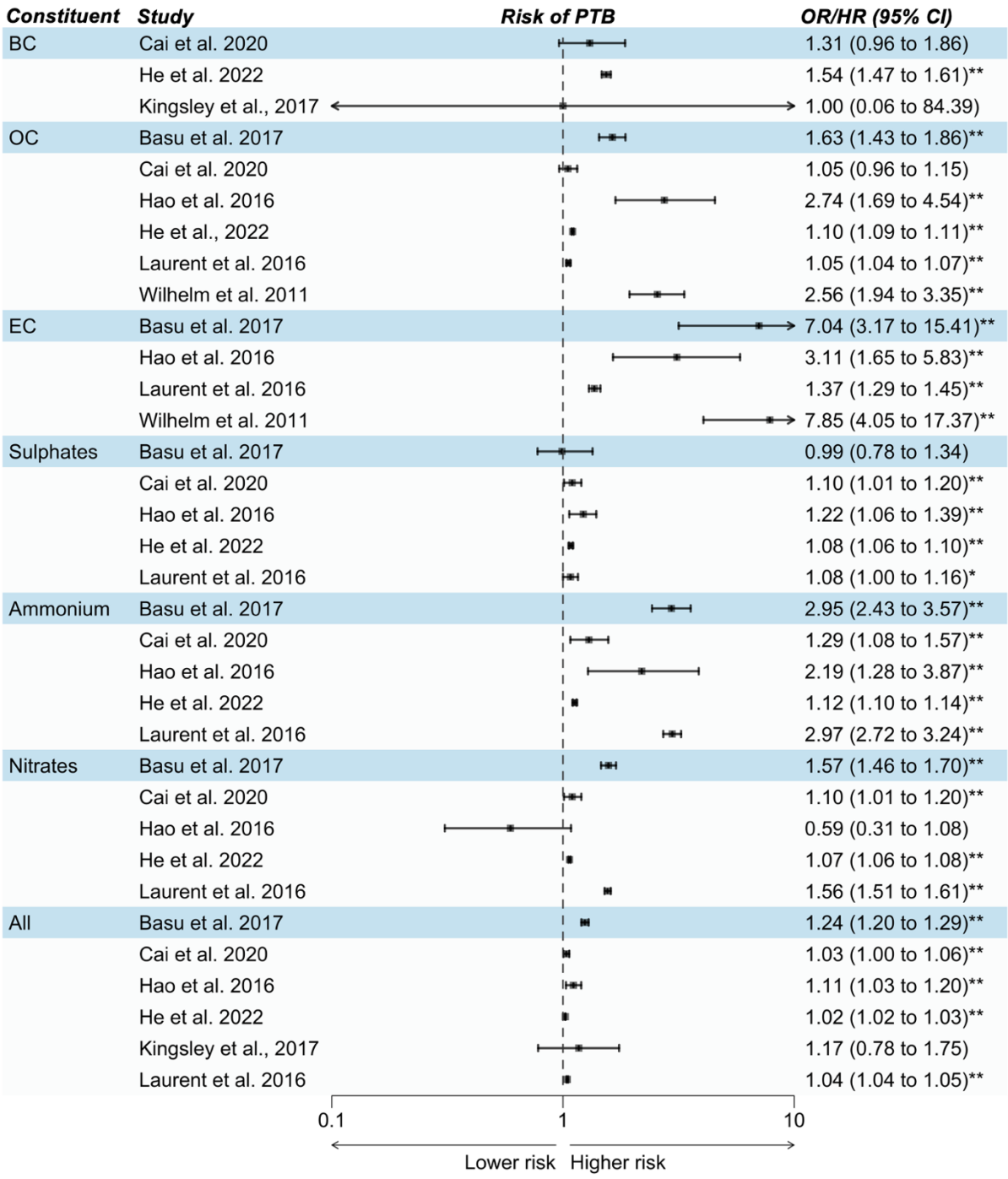


Figure 7. The studies that have provided estimates for whole pregnancy exposure to $PM_{2.5}$ components and the risk of PTB based on continuous exposure. The estimates were all converted to per 10μg/m^3 increment. Fang et al. (2022), Nealy et al. (2018), Ottone et al. (2020) provided trimester-specific estimates, and Rappazzo et al. (2015) provided estimates per gestational week, thus they were not shown here. Basu et al. (2017) and Laurent et al. (2016) provided estimates for metals such as titanium (Ti) and potassium (K), but the list of metals analysed was extensive and thus not included here. "All" refers to the estimate for exposure to total ambient $PM_{2.5}$ by the studies included. Wilhelm et al. (2011) did not provide one and thus was not shown.

4.3.3.2.4. Did temperature modify the risk of PTB associated with air pollution exposure

Five studies have explored if temperature modified the risk of PTB observed in their study population (Alman et al., 2019; Chen et al., 2018; Kwag et al., 2021a; Sun et al., 2020; Wang et al., 2020). Their stratification for temperature or definition of heatwaves are shown in **Table 20**. Alman et al. (2019) found no association between $PM_{2.5}$ exposure and the risk of PTB, and a broad dichotomisation at 19.2°C did not alter the findings. Chen et al. (2018) stratified their analysis at lower temperature levels (0 to 12.73°C), and found that the increased risk of PTB associated with $PM_{2.5}$ exposure only persisted in the stratum of 0 to 5.41°C, which was similar to the finding of Kwag et al. (2021a) that found an increased risk of PTB only in the stratum of -14.5 to 3.4°C.

Two of the five studies explored if heatwaves modified the association between air pollution and the risk of PTB using excess risk due to interaction (RERI) (Sun et al., 2020; Wang et al., 2020). Both found synergistic effects of heatwave and $PM_{2.5}$ exposure for less intense heatwaves (daily mean temperature ≥33.6°C for 2-4 consecutive days) but not more intense heatwaves. Furthermore, one of them found synergistic effects of more intense heatwave and O_3 exposure (daily mean temperature ≥36.7°C for 2-4 consecutive days) (Sun et al., 2020).

Both low or high ambient temperature have been found to be associated with PTB, suggesting that the relationship is potentially U-shaped (Chersich et al., 2020; Ruan et al., 2023). Findings from the five studies retrieved suggest that temperature modifies the risk of PTB associated with $PM_{2.5}$ exposure in similar ways. On the other hand, it is more uncertain if temperature plays a role in the relationship between NO_2 and O_3 exposure and PTB, and more studies are needed to ascertain its role.

Table 20. Studies that have studied how temperature or heatwaves modify the association between air pollution exposure and the risk of PTB.

Study	Stratification for temperature or heatwaves
Alman et al. 2019	Above or below the 75th percentile of the mean temperature during pregnancy (66^{o}F = 19.2^{o}C).
Chen et al. 2018a	Tertiles of mean temperature during pregnancy, where 1st tertile (0 to 5.41^{o}C) is "low" temperature, 2nd tertile (5.42 to 6.58^{o}C) is "moderate temperature, 3rd tertile (6.59 to 12.73^{o}C) is "high" temperature.
Kwag et al. 2021a	Quartiles of mean temperature of each trimester (4th quartile: -14.5 to 3.4^{o}C, 3rd quartile: 3.4 to 14.5^{o}C, 2nd quartile: 14.5 to 22.8^{o}C, 1st quartile: 22.8 to 31.8^{o}C.
Wang et al. 2020	Two types: (1) Daily mean temperature ≥ 33^{o}C for 2, 3, or 4 consecutive days, (2) Daily mean temperature ≥ 34.6^{o}C (75th percentile, 35.7^{o}C (90th percentile), 36.4^{o}C (95th percentile), or 37^{o}C (98th percentile), for 2, 3, or 4 consecutive days.
Sun et al. 2020	Daily mean temperature ≥ 31.2^{o}C (88.1^{o}F) (75th percentile, 33.6^{o}C (92.5^{o}F) (90th percentile), 35.1^{o}C (95.1^{o}F) (95th percentile), or 36.7^{o}C (98.1^{o}F) (98th percentile), for 2, 3, or 4 consecutive days.

4.3.3.2.5. Did season of birth or conception modify the risk of PTB associated with air pollution exposure?

Six studies have explored if season of birth or conception is a modifier for the association between air pollution exposure and the risk of PTB (Jalaludin et al., 2007; Olsson et al., 2012; Siddika et al., 2019; Siddika et al., 2020; Weber et al., 2019; Zhao et al., 2020a). Their findings are shown in **Figure 8.** Siddika et al. (2019) stratified the births into those that were born during warm seasons (spring and summer) and cold seasons (autumn and winter) and found no significant differences in the risk of PTB for $PM_{2.5}$, NO_2, and O_3.

Jalaludin et al. (2007) conducted a study in Sydney, Australia and found that births conceived during autumn or winter generally had a higher risk of PTB associated with $PM_{2.5}$ and NO_2 exposure, compared to births conceived during summer, but the risk estimates were not always statistically significant. The seasonal pattern was opposite of what was found in Zhao et al. (2020a) which was conducted in Inner Mongolia, China, and found a

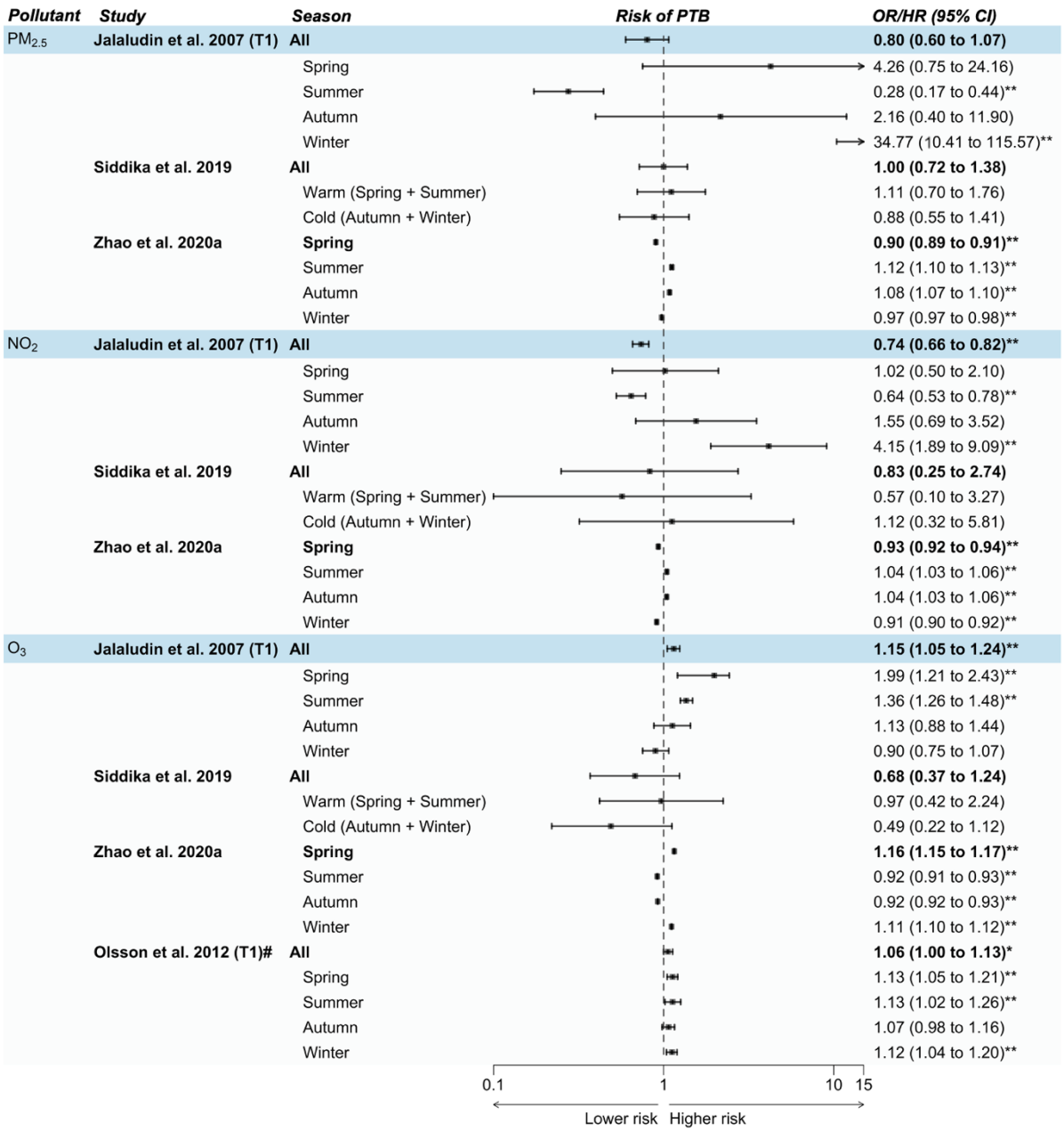


Figure 8. Studies that have conducted a stratified analysis by season of birth or conception for air pollution exposure and the risk of PTB, based on continuous exposure. The estimates were all converted to per 10 μg/m$_3$ increment. Siddika et al. (2020) and Weber et al. (2019) used categorised exposures. "All" refers to the estimate without stratification by season of birth or conception. Zhao et al. (2020a) did not provide an unstratified estimate and thus it is not shown here. #: Estimates from Olsson et al. (2012) are adjusted for NO_2 exposure. T1: Estimate used is based on trimester 1 exposure.

greater risk of PTB in summer and autumn. For O_3 exposure, the seasonal pattern exhibited by Jalaludin et al. (2007) and Zhao et al. (2020a) were the reverse of NO_2 and $PM_{2.5}$ exposure. The findings could be due to the negative correlation between concentrations of

NO_2 and O_3. Olsson et al. (2012) presented season-specific estimates for O_3 exposure adjusted for NO_2 and found no seasonal patterns.

It is possible that the differences in findings are driven by the different seasonal pattern of pollutants in climates with distinct characteristics. For instance, using the Köppen climate classification system, Sydney and Hohhot (Inner Mongolia) are classified as humid subtropical climate (Cfa) and cold arid steppe climate (BSk) respectively (Beck et al., 2018). Sydney and Hohhot have distinct temperature patterns over the year. Moreover, Cfa generally has higher ground O_3 levels and lower $PM_{2.5}$ levels in summer, whereas BSk typically has higher $PM_{2.5}$ levels in summer and lower O_3 levels overall (Morshed et al., 2024), which can explain some of the findings shown **Figure 8**.

4.3.3.2.6. Did noise exposure play any role in the relationship between air pollution exposure and the risk of PTB?

Four studies have explored the joint impact of noise and air pollution exposure and the risk of PTB, alongside the impact of air pollution or noise exposure alone (Barba-Vasseur et al., 2017; Dzhambov et al., 2019; Gehring et al., 2014; Smith et al., 2020). The studies were conducted in different locations (France, London, Tyrol, Vancouver). No significant difference was found between the estimates derived from the impact of air pollution alone and the combined effect of noise and air pollution. However, Smith et al. (2020) found that air pollution may be a negative confounder between the effect of noise exposure and the risk of PTB. Results from these studies suggest that noise exposure play a limited role in the relationship between air pollution exposure and the risk of PTB. Yet, the finding is far from conclusive as it may not be applicable to settings with much higher median residential noise levels (Hammer et al., 2014).

4.3.3.3. Other factors

4.3.3.3.1. Were mothers with certain conditions/demography of greater susceptibility to the risk of PTB associated with air pollution exposure?

There were nine studies that have explored if mothers with certain conditions are of greater risk of delivering a PTB associated with air pollution exposure during pregnancy (Garshol et al., 2020; Gan et al., 2022; Lavigne et al., 2016; Lavigne et al., 2018; Liang et al., 2022; Mendola et al., 2016; Padula et al., 2019; Weber et al., 2019; Zhang et al., 2021b). A variety of conditions which were shown to be linked with air pollution exposure or the risk of PTB

were included, such as asthma and gestational diabetes. Some of their findings are presented in **Figure 9**.

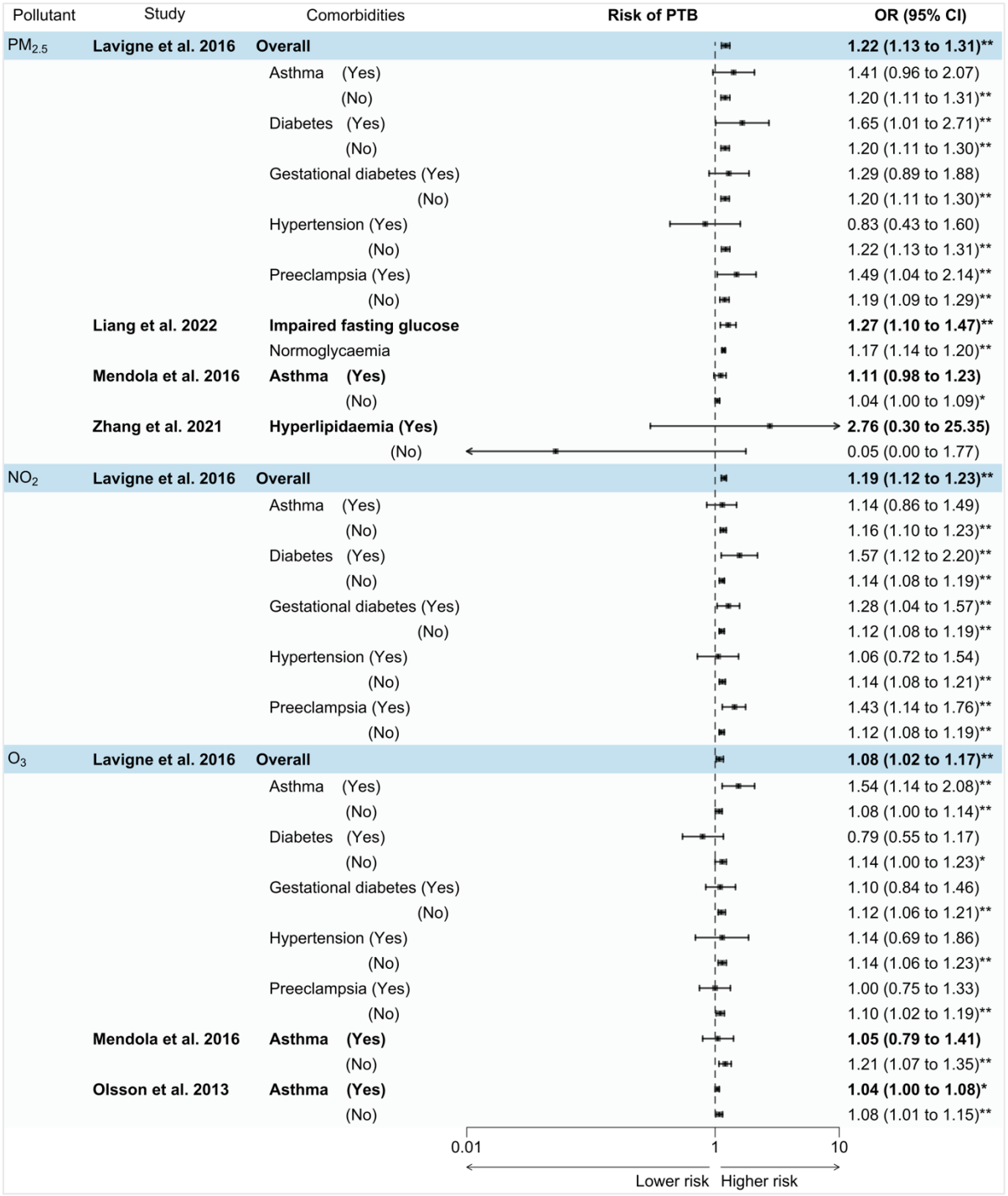


Figure 9. Studies that have conducted a stratified analysis by presence or absence of certain conditions for the risk of PTB associated with air pollution exposure. The estimates were all converted to per 10μg/m^3 increment. Garshol et al. (2020), Padula et al. (2019), and Weber et al. (2019) presented estimates derived from categorised levels of exposure. Gan et al. (2022) did not provide the increment of exposure. Lavigne et al. (2018) and Liu et al. (2019b) did not provide stratified estimates. These studies were therefore not shown here. Overall: unstratified estimates.

Three studies that conducted stratified analysis by asthmatic status found a higher risk of PTB associated with $PM_{2.5}$ exposure among asthmatic mothers, but the findings were less consistent for NO_2 and O_3 (Mendola et al., 2016; Lavigne et al., 2016; Olsson et al., 2013). Mothers with conditions related to elevation of blood glucoses, for instance impaired fasting glucose, type 2 diabetes mellitus, and gestational diabetes, were found to be at greater risk of PTB associated with both $PM_{2.5}$ and NO_2 exposure (Lavigne et al., 2016; Liang et al., 2022a). A similar pattern was observed for mothers with inflammatory conditions, such as hyperlipidaemia and inflammatory bowel diseases (Garshol et al., 2020; Zhang et al., 2021b). However, a group of studies found that mothers with comorbidities had a lower risk of PTB associated with air pollution exposure (Lavigne et al., 2016; Padula et al., 2019; Weber et al., 2019). The apparent lower risk found did not imply that a protective effect was present, as the estimates in the respective stratum were not statistically different from each other (**Figure 9**). Hence, it is inconclusive if mothers with these conditions are of greater or lower risk of PTB associated with air pollution exposure.

Lavigne et al. (2018) created an oxidative potential metric using glutathione (GSH) and found no increase in risk of PTB associated with different quartiles of $PM_{2.5}$ exposure when mothers were stratified by glutathione (GSH) oxidative potential (Lavigne et al., 2018). Gut microbiota taxa may also play a role in the association, with one study finding *parabacteroides* explaining around half of the correlation between $PM_{2.5}$ exposure and risk of PTB (Gan et al., 2022). It is worth noting that these as well as some of the conditions mentioned in previous paragraphs could be mediators in the association between air pollution and the risk of PTB.

Five studies conducted in the US have investigated if the risk of PTB associated with air pollution exposure was different among mothers of different ethnicities, and three of them found greater risk for non-Hispanic black mothers compared to non-Hispanic whites, and one found an elevated risk for Hispanic women as well (Benmarhnia et al., 2017; Gray et al., 2014; Le et al., 2012; Pereira et al., 2014a; Riddell et al., 2021). Two studies stratified their study population by markers of socioeconomic status, such as years of maternal education (Gray et al., 2014; do Nascimento et al., 2022). Both studies found that studies of lower socioeconomic status are of greater risk of PTB associated with air pollution exposure.

4.3.3.3.2. *Did the method of gestational dating affect the risk of PTB presented?*

Three studies have compared estimates derived from PTB defined by different gestational dating methods (Nealy et al., 2018; Panasevich et al., 2016; Rappazzo et al., 2016). Two studies found that more births were being classified as PTB when last menstrual period (LMP) was used compared to clinical estimation (CE) (Nealy et al., 2018; Rappazzo et al., 2016). However, Nealy et al. (2018) found no significant differences in the estimates for the risk of PTB derived using LMP or CE for trimester 1 and 2 $PM_{2.5}$, NO_x, EC, OC, or CO exposure. The finding was broadly consistent with Panesevich et al. (2016), which found minimal differences in the risk of PTB derived using both LMP and CE, or LMP only. Rappazzo et al. (2016) calculated the risk differences in the risk of PTB associated with $PM_{2.5}$ exposure per gestational week and 1 $\mu g/m^3$ increase in $PM_{2.5}$ exposure, and compared the results from populations that defined gestational age that used LMP and/or CE. They found a greater divergence in the risk differences as pregnancy advanced. However, it is impossible to know if the risk estimates and the 95%CI used for deriving the risk differences were statistically different from each other (authors no longer had access to the data). Hence, it is unlikely for the choice of gestational dating method to have a strong influence on the risk of PTB observed associated with air pollution exposure.

4.3.3.3.3. *Was the risk of PTB greater among mothers that underwent certain methods of conception or delivery?*

Eleven studies have investigated if the risk of PTB associated with air pollution exposure was stronger for mothers with certain trigger of delivery (spontaneous or medically indicated), type of delivery (vaginal or caesarean section), or method of conception (Cai et al., 2020; Han et al., 2018; Johnson et al., 2016; Olsson et al., 2013; Kingsley et al., 2017; Kirwa et al., 2021; Padula et al., 2018; Padula et al., 2019; Savitz et al., 2019; Weber et al., 2019; Williams et al., 2021). Some of their findings are shown in **Figure 10**.

Seven of the studies provided separate risk estimates for different triggers of delivery (Johnson et al., 2016; Kingsley et al., 2017; Padula et al., 2018; Padula et al., 2019; Savitz et al., 2019; Weber et al., 2019; Williams et al., 2021). All of them found no significant differences between the risk estimates derived using the all births, or restricted to spontaneous or medically-indicated deliveries. Three studies have explored if excluding births from caesarean section would lead to substantially different risk estimate for PTB associated with air pollution exposure, and the exclusion resulted in negligible differences in the risk estimate (Han et al., 2018; Kirwa et al., 2021; Olsson et al., 2013). Two studies conducted a stratified analysis by mode of conception (Cai et al., 2020; Han et al., 2018).

Han et al. (2018) found no significant differences in the risk of PTB associated with O_3 exposure, whereas Cai et al. (2020) found that the statistically significant increase in the risk of PTB associated with $PM_{2.5}$ exposure derived using the whole study population dissipated in those that were naturally conceived, but was strengthened among those that used assisted reproductive therapies (ART). Findings from these studies suggest that the risk of PTB associated with air pollution exposure did not differ significantly among mothers with certain triggers of delivery, type of delivery, and method of conception.

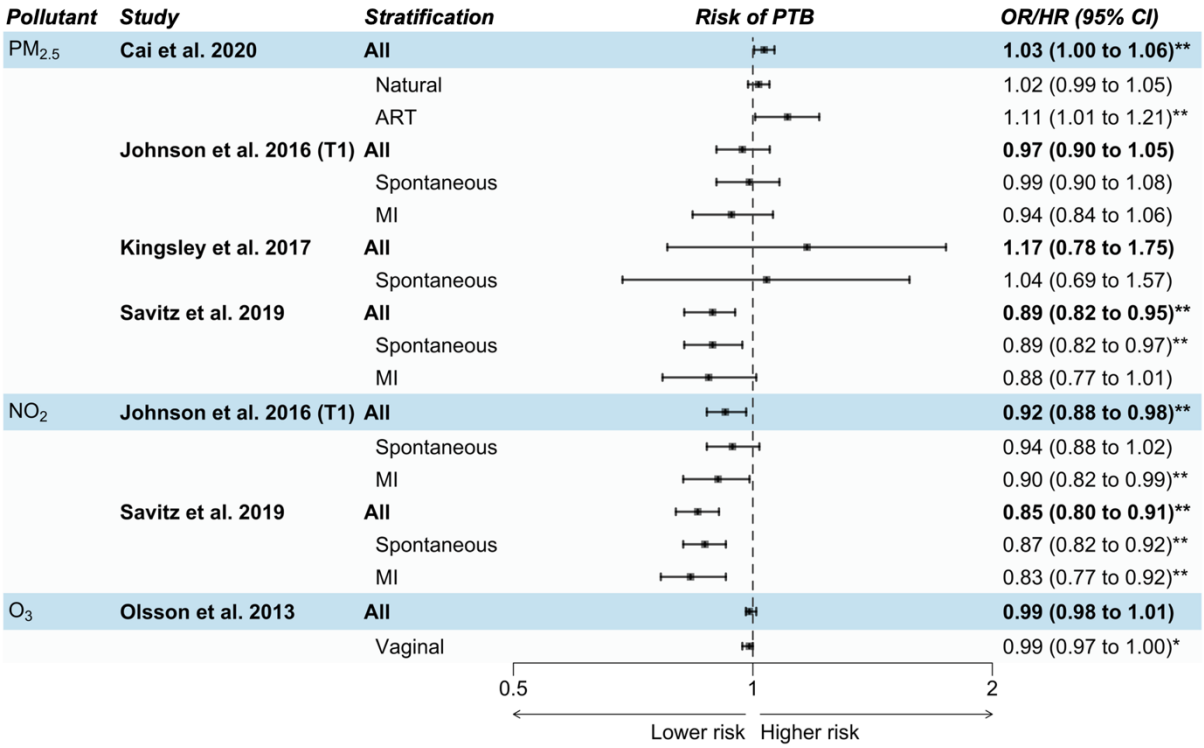


Figure 10. Studies that have conducted a stratified analysis by trigger of delivery, type of delivery, and method of conception. The estimates were all converted to per 10 μg/m³ increment. Han et al. (2018), Padula et al. (2019), Weber et al. (2019), and Williams et al. (2021) derived estimates from categorised levels of exposure which cannot be transformed to a common unit. Padula et al. (2018) did not show the results of the stratified analysis. These studies were therefore not shown here. "All": estimate using all births. ART: assisted reproductive therapies.

4.3.3.4. Studies that used different designs or methods

4.3.3.4.1. Categorical exposures

Some studies had provided risk estimates for PTB associated with air pollution exposure that apply to pollution levels at a certain range, such as quartiles or percentiles. By comparing the risk of PTB for those at higher levels to lower levels of exposure, it helps detect potential non-linear relationships and threshold levels. Their findings can offer qualitative support to

the narrative synthesis, but are hard to incorporate into a meta-analysis, as it is difficult to adjust the estimates derived to a common scale, for instance per 10 μg/m^3 increase, without making substantial assumptions about the distribution of the pollution levels.

There were fourteen studies compared the risk of PTB for mothers exposed to the pollution levels at the highest level, such as the 4th quartile, to the other levels (Alman et al., 2019; Bertin et al., 2015; Capobussi et al., 2016; Cossi et al., 2015; Gehring et al., 2011b; Han et al., 2018; Hannam et al., 2014; Kwag et al., 2021b; Leem et al., 2006; Padula et al., 2014; Padula et al., 2019; Riddell et al., 2021; Siddika et al., 2020; Weber et al., 2019), with 2 studies further stratified the population according to diabetic or hypertension status (findings discussed in **Figure 9**) (Padula et al., 2019; Weber et al., 2019). Three studies on O_3 exposure found no difference in the risk of PTB at higher exposure categories (Capobussi et al., 2015; Han et al., 2018; Siddika et al., 2020). There were 3 studies that found no difference in the risk of PTB associated with $PM_{2.5}$ exposure (Alman et al., 2019; Hannam et al., 2014; Kwag et al., 2021b), and 5 studies found no difference in the risk of PTB for NO_2 exposure (Capobussi et al., 2015; Gehring et al., 2011b; Hannam et al., 2014; Leem et al., 2006; Riddell et al., 2021). However, one study found increased in risk of PTB associated with NO_2 exposure at the 2nd and 3rd tertile among mothers living in urban areas (Bertin et al., 2015), and one finding of greater risk for trimester 3 NO_2 and $PM_{2.5}$ exposures as well (Siddika et al., 2020). Studies that have investigated the risk for subgroups of PTB found a more consistent increase in risk for $PM_{2.5}$, but less for NO_2 and O_3 exposure (Cossi et al., 2015; Padula et al., 2019; Weber et al., 2019).

Salihu et al. (2012) have dichotomised the study population to those that were exposed to above or below/equal to the median of $PM_{2.5}$ concentration (11.28 μg/m³) and found no differences in the risk. One study compared the risk for mothers who have experienced different numbers of days with $PM_{2.5}$ exposure above specific levels in China. They found statistically significant increases in risk of PTB across most of the gestation for $PM_{2.5}$ exposure above 200 μg/m³ (Fang et al., 2020).

4.3.3.4.2. Successive pregnancies

Five studies have compared successive pregnancies in the same women to control for various maternal risk factors discussed previously to better disentangle the effect of air pollution exposure (Mendola et al., 2019; Pereira et al., 2014a; Pereira et al., 2014b; Pereira et al., 2014c; Pereira et al., 2016). 4 of the studies that were conducted in the US found an increased risk of PTB in pregnancies with elevated $PM_{2.5}$ exposures (Mendola et al., 2019; Pereira et al., 2014a; Pereira et al., 2016), but the increase in risk dissipated when $PM_{2.5}$

exposures was subcategorised into different anthropogenic sources (Pereira et al., 2014c). Pereira et al. (2014b), a study that was conducted in Australia, also showed no change in the risk of PTB when pregnancies of higher or lower $PM_{2.5}$ levels were compared.

4.3.3.4.3. Bayesian or natural experiment approaches with low compatibility with logistic/Cox-PH regression in a meta-analysis

Statistical approaches used in primary studies as well as meta-analyses can be frequentist or Bayesian in nature. Therefore, if studies that used both frequentist and Bayesian approaches were included in the same meta-analysis, it poses difficulty in interpretating the findings, Frequentist approaches, such as logistic regression, are based on fixed parameters of interest, whereas for Bayesian frameworks the posterior distributions of the parameter of interest are informed by priors.

Four studies used Bayesian analytical methods to investigate the risk of PTB associated with air pollution exposure (Chang et al., 2012; Chang et al., 2015; Jackson et al., 2009; Warren et al., 2012). 3 of them found an increase in risk of PTB associated with trimesters 1 and 2 (Chang et al., 2015; Warren et al., 2012), and whole pregnancy $PM_{2.5}$ exposure (Chang et al., 2012), and one found no association across the entire pregnancy for NO_2 exposure (Jackson et al., 2009).

Three of these studies attempted to identify the critical window (month or week) of exposure during pregnancy, as Bayesian framework allows time-varying exposure in pregnancy to be considered (Chang et al., 2012; Chang et al., 2015; Warren et al., 2012). Chang et al. (2012) found no critical window of exposure, but Chang et al. (2015) found the strongest risk of PTB associated with week 30 of $PM_{2.5}$ exposure. Warren et al. (2012) found the strongest risk of PTB associated with around week 15 of $PM_{2.5}$ exposure using USEPA AQS but not FSD data, and highest risks of PTB associated with O_3 exposure in the first 5 weeks of pregnancy. Their findings will be discussed alongside studies that have used frequentist approaches (logistic regression and Cox-proportional hazard regression) in the next section.

Two studies used events leading to significant air quality improvement to study the impact of air pollution exposure on the risk of PTB using natural experiments, which allows adjustment of unmeasured spatiotemporal factors (Huang et al., 2015; Liu et al., 2022b). Huang et al. (2015) found no association between NO_2 exposure and the risk of PTB using pollution changes as a result of measures implemented for the 2008 Beijing Olympics, but Liu et al. (2022) found a small increase in the risk of PTB associated with NO_2 exposure using the pollution changes during the 2010 Asian Games in Guangzhou. A US study that used

geographically-weighted and global logistic regression to generate ORs and 95% CIs for each birth found that the risk of PTB is primarily driven by births from mothers of lower socioeconomic status in rural areas (Tu & Tu, 2018), which was partly consistent with the findings in **section 4.3.3.3.1**.

4.3.3.4.4. Monthly or weekly risks of PTB

28 studies have provided weekly or monthly estimates for the risk of PTB associated with air pollution exposure, as it helps to identify the potential critical window(s) of exposure. 17 of these studies were excluded from the meta-analyses due to various reasons (details in **Appendix 8**). The majority used logistic regression but there were 3 studies that used Bayesian approaches (Chang et al., 2012; Chang et al., 2015; Warren et al., 2012), and 2 studies that used Cox-proportional hazard regression (Wang et al., 2018; Yang et al., 2020).

17 studies provided monthly and weekly estimates for the risk of PTB associated with $PM_{2.5}$ exposure, and among them 6 studies found no associations with any window of exposure (Alman et al., 2019; Chang et al., 2012; Kloog et al., 2012; Lavigne et al., 2018; Trasande et al., 2013; Weber et al., 2019). 9 studies found an increase in risk of PTB and its subgroups towards the end of pregnancy, starting from week 25 to 30 (Chang et al., 2015; Mendola et al., 2016; Wang et al., 2018), to the last 3 to 1 month before delivery (Jalaludin et al., 2007; Liu et al., 2019b; Padula et al., 2014; Padula et al., 2019; Symanski et al., 2014; Wilhelm & Ritz, 2005). 3 studies found a greater risk of PTB associated with exposure to $PM_{2.5}$ at the first 1 to 2 months of pregnancy (Huynh et al., 2006; Liu et al., 2019b; Symanski et al., 2014), and one study found the highest risk of PTB for exposures at week 15 (Warren et al., 2012).

16 studies provided monthly and weekly estimates for NO_2 exposure, and more than half of them found no associations with any month(s) or week(s) of exposure (Guo et al., 2019; Hansen et al., 2006; Jalaludin et al., 2007; Le et al., 2012; Liu et al., 2003; Padula et al., 2019; Weber et al., 2019; Aimaier et al., 2013; Xu et al., 2008). Similar to $PM_{2.5}$ exposure, most studies suggest that the critical window is at the last month or week of pregnancy (Ji et al., 2019; Liu et al., 2019b; Olsson et al., 2012; Padula et al., 2014; Trasande et al., 2013), with one study suggesting week 25 (Wang et al., 2018), and two studies suggesting that the first and second month of pregnancy (Liu et al., 2019b; Lu et al., 2019).

12 studies provided monthly and weekly estimates for O_3 exposure, and 7 of them found no associations with any month(s) or week(s) of exposure (Hansen et al., 2006; Jalaludin et al., 2007; Liu et al., 2003; Liu et al., 2019b; Olsson et al., 2012; Padula et al., 2019; Trasande et

al., 2013). 4 studies suggested that a critical window at the last month or week before delivery (Mendola et al., 2016; Symanski et al., 2016; Wang et al., 2018; Yang et al., 2020), with one study suggesting the first 5 weeks (Warren et al., 2012).

The findings from these studies were relatively consistent and indicated that $PM_{2.5}$, NO_2, and O_3 exposure at the last months or weeks of pregnancy were likely associated with greater risk of PTB. There is emerging evidence on the first month or weeks of pregnancy, however, there might be greater degrees of exposure misclassification present as the exact date of conception is unavailable most of the time.

4.4. Discussion

4.4.1. Summary of findings

Various systematic reviews and/or meta-analyses had synthesised the evidence on the relationship between gestational air pollution exposure and the risk of PTB (**Table 7**). However, these previous reviews primarily focused on findings derived from studies employing traditional logistic regression approaches. There was less focus on some of the findings from the literature that could offer valuable insights, for instance, how some factors such as co-exposures or comorbidities may affect the presence or magnitude of the association of interest. Moreover, findings from studies that cannot be incorporated into a meta-analysis, such as those from Bayesian approaches were often overlooked. The work conducted in this chapter overcame these limitations and synthesised the findings, which are summarised in **Table 21**.

Table 21. Summary of findings from narrative synthesis (contnued on next page).

<table>
<tr><th rowspan="2">Aspects</th><th colspan="3">Pollutant</th></tr>
<tr><th>$PM_{2.5}$</th><th>NO_2</th><th>O_3</th></tr>
<tr><td>Exposure assessment</td><td colspan="3">Limited differences in findings detected, but differences may be detected with sufficiently large sample size or pollutants with greater spatial contrast e.g. NO_2</td></tr>
<tr><td>Residential mobility/ Time-activity pattern</td><td colspan="3">Residential mobility likely had little impact on the risk observed. Time-activity pattern could be influential in studies with larger sample size and adjustment of other pollutants</td></tr>
<tr><td>Adjustment for other pollutants</td><td>Likely have minimal impact on the risk observed</td><td>The risk observed could be partly driven by $PM_{2.5}$</td><td>Inconclusive</td></tr>
<tr><td>$PM_{2.5}$ from adverse events</td><td>Emerging evidence for wildfire exposure, but the risk observed could be driven by other hazardous pollutants or factors such as stress</td><td colspan="2" rowspan="2"></td></tr>
<tr><td>Sources or constituents of $PM_{2.5}$</td><td>Emerging evidence for constituents but challenges remain in interpretating the findings due to high correlations between constituents.</td></tr>
<tr><td>Temperature</td><td>Likely modifier of the association</td><td colspan="2">Inconsistent findings</td></tr>
<tr><td>Season of birth/conception</td><td colspan="3">Differences in risk observed could be attributable to pattern of air pollutants in different climates</td></tr>
<tr><td>Noise</td><td colspan="3">Findings from the evidence available to date showed negligible impact on the risk of PTB observed</td></tr>
</table>

Maternal conditions	Inconclusive findings. The differences in risk observed were not statistically different from each other. Suggestive evidence for greater susceptibility for US non-Hispanic blacks, and mothers of lower SES.	
Gestational age dating method	Few studies available. Unlikely to have a strong influence on the risk observed.	
Types/methods of conception or delivery	Few studies available. Risk observed were not statistically different from each other.	
Studies of categorical exposure	More consistent findings associated with the risk of PTB subgroups	Less consistent findings
Studies comparing successive pregnancies	Findings from the few studies were suggestive of an association with $PM_{2.5}$ exposure	Inconclusive
Studies using other methods		
Risk of PTB by gestational months/weeks	Suggestive of an association at later stages (more consistent) and early stages (less consistent) of pregnancy. Meta-analyses could be conducted for windows more frequently used to determine the presence and magnitude of the association.	

4.4.2. Interpretation and implication of findings

A qualitative comparison can be made with the preceding SR/MAs found in Chapter 3 (**Table 7**). A more direct comparison to findings of preceding SR/MAs was conducted in the following chapter (Chapter 5), where the association between air pollution exposure and the risk of PTB was meta-analysed.

It is clear that $PM_{2.5}$ has been studied most extensively, followed by NO_2 and O_3, when it comes to association with the risk of PTB. Hence, it remains possible that when the evidence base for NO_2 and O_3 exposure reach a breadth and depth similar to that of $PM_{2.5}$, their relationship with the risk of PTB can be better elucidated. There was generally a higher proportion of studies reporting a positive association at trimester 3 and whole pregnancy exposure (**Table 19**). This is consistent with the findings from numerous SR/MAs that suggest an increase in risk associated with either trimester 3 or whole pregnancy $PM_{2.5}$ or NO_2 exposure (Lamicchane et al., 2015; Ghosh et al., 2021; Ju et al., 2023; Yu et al., 2022; Zhu et al., 2015), as well as studies that compared successive pregnancies to control for maternal factors (Pereira et al., 2014a; Pereira et al., 2014c; Pereira et al., 2016; Mendola et al., 2016). Findings from studies of the risk of PTB at particular gestational weeks or months were in line with the observation, showing a plethora of studies suggesting a greater risk of PTB at later stages of pregnancy for $PM_{2.5}$, NO_2, and O_3 exposure.

Much fewer studies were available for the association between air pollution exposure and the risk of subcategories of PTB. However, the proportion of studies reporting a positive association for a particular pollutant-exposure period-outcome combination was higher than that of the equivalent combinations for PTB. One of the preceding SR/MAs found an increased risk of moderate-to-late PTB and very PTB associated with $PM_{2.5}$ exposure (Yu et al., 2022) compared with PTB overall. Findings from studies of categorised exposures were more consistent for the risk of subcategories of PTB associated with $PM_{2.5}$ exposure as well (Cossi et al., 2015; Padula et al., 2019; Weber et al., 2019). It is possible that a threshold effect is present for $PM_{2.5}$ exposure, where the risk of PTB associated with it becomes much stronger once it reaches a certain level.

The presence of an effect of $PM_{2.5}$ exposure is supported by the studies of wildfire exposures (Abdo et al., 2019; Heft-Neal et al., 2022). However, more studies are needed to ascertain the effect of wildfire $PM_{2.5}$ and provide a more comprehensive comparison with the effect of ambient $PM_{2.5}$ exposure, as they could have differences in toxicity and composition, such as having higher levels of PAHs and carbonaceous compounds. Findings from the wider literature suggested a link between PAH exposure and the risk of PTB (Chapter 1 **section 1.3.2.3**). The effect of carbonaceous compounds, such as OC, EC, as well as other constituents, for instance ammonium, were shown to be linked to increased risk of PTB. In addition, the risk of PTB associated with $PM_{2.5}$ exposure appeared to be robust after adjustment of NO_2 and O_3, suggesting low risk of confounding by other pollutants.

All in all, the information considered so far from preceding SR/MAs and this narrative synthesis of relevant studies, whose results would not be well-captured by the meta-analyses conducted in Chapter 5, suggested that an association could be found between $PM_{2.5}$ exposure and the risk of PTB and its subtypes, but less so for NO_2 and O_3 exposure. The conclusion was preliminary and ahead of considering the results of the updated meta-analyses reported in Chapter 5. Findings from the wider literature not included in the narrative synthesis, such as those of air pollution exposure and reductions in gestational age (Huang et al., 2017; Qiu et al., 2020), were brought in and discussed alongside the findings of Chapter 5.

This narrative synthesis found that conditions such as extreme temperature and climatic pattern of pollutants, likely interact with the effect of air pollution exposure on the risk of PTB. The risk of PTB associated with air pollution exposure increased under low ambient temperature as well as heatwaves (Chen et al., 2018; Sun et al., 2020; Wang et al., 2020). However, as there is linkage between heatwaves and high O_3 levels, therefore it is possible that at very intense heatwaves, the risk of PTB is predominantly driven by O_3 exposure,

leading to the synergistic effect found for heatwaves and O_3 exposure at higher but not lower heatwave intensities (Sun et al., 2020). Climatic patterns of pollutants were likely what lead to difference in risk of PTB associated with air pollution exposure across seasons. It was shown to be attenuated once pollutants that negatively correlate with each other were adjusted (Olsson et al., 2012), demonstrating using estimates adjusted for other exposures is likely important in future studies.

4.4.2.1. Exposure assessment, residential mobility, and time-activity pattern

Findings from studies that compared different exposure assessment methodologies, residential mobility, or took time activity pattern into account showed that their influence on the risk of PTB associated with air pollution exposure observed is likely small. However, it does not imply that they should not be considered. In the case of PTB where the risk is likely to be small, whether a study took the aforementioned factors into account could result in a slight shift of the risk estimate and its treatment of uncertainty, leading it to be interpreted as showing an increase in risk to having null effect. It is thus important to assess exposure using information and models from multiple sources and accurate address history for studies covering large geographical areas, as the participants' entire profile of exposure is often built entirely based on their addresses. This is supported by findings from studies showing that the differences in risk estimates may depend on the distance between the old and updated addresses (Hodgson et al., 2015; Lupo et al., 2010; Warren et al., 2018). Moreover, ambient air pollutants can be generated in indoor environments, such as home and workplace (Tran et al., 2020). It is uncertain if mothers who spent more time indoors had greater risk of PTB associated with air pollution exposure (Estarlich et al., 2016; Llop et al., 2010), and no study retrieved has measured both ambient and indoor air pollution. Therefore, the risk estimates derived could differ once indoor air pollution exposure was considered.

4.4.2.2. Maternal and clinical factors

Studies included in this narrative synthesis suggest that maternal or clinical factors had limited impact on the risk of PTB associated with air pollution exposure, but there were findings suggesting greater susceptibility for those that are non-Hispanic blacks in the US and mothers of lower socioeconomic status (Benmarhnia et al., 2017; Gray et al., 2014). The finding for ethnicity was consistent with a review of racial disparity and pregnancy outcomes which found greater risk of PTB for non-Hispanic blacks compared to non-Hispanic whites

(Dzekem et al., 2024). The finding for socioeconomic status is similar to those of other health outcomes, where those with lower socioeconomic status were often exposed to higher levels of air pollution, and the adverse effects were likely stronger for those of lower socioeconomic status and living in urban areas (Boing et al., 2022; van den Brekel et al., 2024).

4.4.2.3. Quality assessment

The results of the quality assessment conducted in the present review were very different from those of the preceding publications. Most studies were assigned a moderate risk of bias, whereas they were mainly categorised as high quality in previous works. The discrepancy in finding can be partly explained by the use of conventional frameworks in previous studies. For instance, the Newcastle-Ottawa Scale (NOS) that cannot provide a comprehensive assessment of study quality was used in 13 previous systematic reviews or meta-analyses (SR/MAs) (Chapter 3 **section 3.3.3.2**). Three SR/MAs provided a more comprehensive quality assessment (HEI, 2022; Ghosh et al., 2021; Rappazzo et al., 2021). However, methodological differences in the framework, and judgements made regarding what the important aspects to consider were, potentially still resulted in differences in findings. For instance, the work by Ghosh et al. (2021) has residential mobility as one of the major aspects to be considered when assessing study quality. On the other hand, it played a less important role in the exposure assessment domain in the quality assessment framework used by the present review, and the decision is in line with the findings from the narrative synthesis.

In addition to the present work, Ghosh et al. (2021) and Yu et al. (2022) investigated the extent of variable adjustment when assessing the risk of bias of primary studies. Yu et al. (2022) had 12 variables categorised as “important confounders”, some of which overlap with the “critical confounders” used in the present work (maternal age, parity). In Ghosh et al. (2021), sufficient adjustment was defined as having 4 variables (infant sex, socioeconomic status, weight gain during pregnancy, smoking) adjusted in the analysis. Otherwise, studies were deemed to have insufficient confounder adjustment. These approaches gave no consideration to the relative importance of each confounder, and the impact is elaborated below.

Using infant sex as an example, a study could have treated it as an effect modifier and not a confounder, but have adjusted for other potentially more influential factors such as maternal age (Cossi et al., 2015). Around 40% of the studies that were quality assessed did not adjust for infant sex, so adopting this approach could move a substantial number of studies to a category of higher risk of bias. Smoking is a strong confounder of the association, but some

studies may not have adjusted for it as it was conducted in regions where smoking is exceptionally rare for pregnant women, as it would no longer act as a confounder. In this case, the study would be unnecessarily penalised. Hence, the present work has done a more comprehensive analysis and categorisation of the confounders adjusted for, and how it affects the quantitative summary risk estimates is presented in Chapter 5 **section 5.3.7.1**.

4.4.3. Strength and limitations

This systematic review was rigorously conducted, from the use of a protocol, validation of search strings and results, to a bespoke quality assessment framework for the studies of the area developed using information on important confounders. Moreover, the narrative synthesis used findings from individual studies that were often given less attention in previous systematic reviews and meta-analyses, and identified gaps in research which could explain the inconsistencies in the findings. Both Chinese and English literature were also screened to overcome language limitations, as previous systematic reviews and meta-analyses used only evidence from English language studies.

Several limitations are also present. The quality assessment framework used was modified from the WHO Air Quality Guidelines Risk of Bias Assessment Instrument. The modification ensured that it could comprehensively assess the quality of the studies retrieved. Nonetheless, the modified framework is not pilot-tested, and thus, its validity cannot be ensured; however, modifications to this framework have been made in previous systematic reviews and meta-analyses of traffic-related air pollution and health outcomes, and thus, its versatility has been demonstrated (Chen and Hoek, 2020; HEI, 2022). All Chinese studies were also assigned a high risk of bias. Still, as Chinese journals do not necessarily follow the format and requirements of English language journals, some key information was likely missing due to journalistic requirements instead of reporting issues.

4.4.4. Interim conclusion

Following the recommendations in Chapter 3, a systematic review was conducted with a narrative synthesis of findings from studies that were not eligible to be included in a meta-analysis. Findings from the studies included in the narrative synthesis suggest that there could be an increase in risk of PTB associated with air pollution, in particular $PM_{2.5}$ exposure. If a true association is present, then there is evidence showing that the association may be modified by factors such as extreme temperatures and climate-related pattern of air pollutants, but more research is needed for some other factors and for NO_2 and O_3 exposure.

In the next chapter, meta-analyses were conducted to derive quantitative summary risk estimates regarding the association between air pollution exposure and the risk of PTB, as well as its subgroups. The results from the meta-analyses were interpreted in conjunction with findings from this chapter, to allow a better overview of the evidence in the area to be made.

Chapter 5. Meta-analysis

5.1. Introduction

The findings from **Chapter 3** and **Chapter 4** suggests that a positive association could be present between air pollution exposure and the risk of PTB, with whole pregnancy $PM_{2.5}$ exposure being the most plausible. In this chapter, meta-analyses were conducted to determine the presence, and if present, the magnitude and direction, of the quantitative relationship between the pollutants of interest ($PM_{2.5}$, NO_2, O_3) and the outcome (PTB and its subcategories) for specific exposure time periods (preconception, trimesters 1, 2, 3, and whole pregnancy). The methods for conducting the meta-analyses are first outlined, then the results of the main meta-analyses are presented. Afterwards, sensitivity analyses were conducted to investigate the potential drivers of heterogeneity, where the main meta-analyses were further stratified by location, risk of bias (RoB), concentration levels, exposure assessment methodology and source of data. The results of the main and sensitivity analyses are discussed in conjunction with the findings from the preceding chapters, as well as the Bradford Hill Criteria, to determine if a causal relationship is likely to be present.

5.2. Methods

5.2.1. Conversion of effect estimates

Effect estimates from studies were in the form of odds ratio (OR), risk ratio (RR) and hazard ratios (HR). OR is the ratio of the odds of an outcome in one group versus another group. Similarly, RR is the ratio of the risk of an outcome in one group versus another group (Dettori et al., 2021). HR measures how often an outcome occurs, if it has not already occurred, in one group versus another group over time (Spruance et al., 2004). It is a time-to-event estimate measuring instantaneous risk, which is different from the comparison of odds by OR. The summary effect estimates from the meta-analyses were calculated as ORs, as it was what most studies had used (n = 113/159, **section 4.3.1**). Estimates from studies reporting RRs were converted to ORs (Grant, 2014; Zhang & Yu, 1998). The formula for converting RRs to ORs is as follows:

$$OR = \frac{RR\ (1 - P_0)}{1 - (RR * P_0)}$$

where P_0 equals to the baseline risk (number of PTBs divided by total study population).

In this thesis, it was decided to conduct separate meta-analyses for estimates of HRs due to two reasons. First, HRs approximate RRs when the frequency of the outcome is relatively uncommon (<10%). Therefore, it is common for meta-analyses to directly use HRs as RRs. However, as previously described, HRs are more precisely "relative rates" instead of relative risks and thus should be interpreted differently, despite common practices for primary studies to interpret HRs as RR (Sutradhar & Austin, 2018). Second, HRs (or the use of Cox-PH regression) are potentially preferable in observational studies of gestational air pollution exposure and birth outcomes, as Cox-PH regression accounts for the differences in exposure duration between preterm and term births and thus overcomes the issue of fixed cohort bias, considers seasonal exposure patterns, and are easier than ORs to interpret clinically. However, there is no direct method to convert ORs to HRs. Moreover, as found in **section 4.3.1** as well as Pereira (2022), studies on air pollution exposure and the risk of PTB that have used HRs were published more recently. By conducting separate meta-analyses for studies reporting HRs and ORs, one can potentially understand if studies that differ in the statistical model used and time published have derived different results.

All estimates were transformed to a per 10 μg/m^3 increment to facilitate comparison. To achieve this, estimates were first converted to logarithmic odds/hazards using natural logarithms, then divided by the exposure increment used to be transformed to per 1 μg/m^3 or 1 ppb/ppm/parts per hundred million (pphm), depending on how the estimates were originally presented in the primary studies (1 ppb = 0.1 pphm = 0.001 ppm). If the estimates were presented per 1 ppb/ppm/pphm, they were converted to per 1 μg/m^3 using the conversion factors stated in **Table 22** (National Environmental Technology Centre, 2003). After the conversion, the estimates were multiplied by 10 to be transformed to per 10 μg/m^3.

Table 22. Conversion factors used for converting between parts per billion (ppb) to micrograms per cubic metres (μg/m^3), at 20°C and 1013 mb (National Environmental Technology Centre, 2003).

Pollutant	Conversion factor
NO_2	1 ppb = 1.9125 μg/m^3
O_3	1 ppb = 1.9957 μg/m^3

5.2.2. Measurement metric of O_3

The concentration of O_3 assigned to a study population could be very different depending on the measurement metric (averaging time) used (details in **section 1.1.3**). In this meta-analysis, studies providing estimates for O_3 were analysed separately according to the measurement metric, as conversion to a common metric could introduce considerable uncertainty (Anderson & Bell, 2010). The measurement metric used in each study included in the meta-analysis is shown in **Appendix 11**.

5.2.3. Selection of effect estimates

When a study has provided multiple estimates for the same PTB type, the effect estimate that was shown as the main result, or not restricted to subgroups e.g. mode of delivery, was selected. Several studies have only provided estimates in subgroups (Liang et al., 2019; Liang et al., 2022a; Mendola et al., 2016; Williams et al., 2021). For instance, Mendola et al. (2016) provided estimates for exposure to $PM_{2.5}$ and the risk of PTB for mothers with or without asthma. The two categories (non-asthmatic and asthmatic) would together constitute the whole study population. Therefore, fixed-effect meta-analysis, which assumes that estimates included in the meta-analysis have the same true effect size derived from a homogeneous population, was conducted to obtain an estimate for the whole study population. The studies with estimates in subgroups, the types of subgroups analysed, and the summary estimate derived from fixed-effect meta-analysis are available in **Appendix 12**.

There were fourteen studies that have categorised pollutant exposure into different levels instead of having estimates for exposure on a continuous scale (**section 4.3.3.4.1**). Methods for converting estimates from categorical exposures have been explored, such as estimating the median and quartile distribution of pollutants by assuming a normal distribution or taking the mean of the exposure categories as the typical exposure to be used for pooling (Berlin et al., 1993). However, distributions of pollutants are usually right-skewed (examples include Ji et al., 2019; Melody et al., 2020; Stieb et al., 2016a). Hence, after careful consideration these studies were excluded from the meta-analysis.

There were studies that required additional information from authors to determine whether they can be used in the meta-analysis. **Appendix 13** shows the list of studies, the information required, if the authors have responded, and the information provided. Studies assigned a high risk of bias were excluded from the main meta-analyses. However, they were added to the pool of studies during the sensitivity analysis to assess their influence on the summary estimates.

5.2.4. Analysing study dependencies and selection of studies for inclusion in meta-analysis

Study dependencies occur when studies using identical or overlapping populations in overlapping time periods are included in the same meta-analysis. When dependent studies were included in the same meta-analysis, it violates the assumption of independent observations and results in a summary estimate biased towards locations, time periods, or populations that were studied more often, as well as an inflation of the sample size which would lead to an underestimation of uncertainty shown by narrower CIs (Higgins, 2022).

Estimates from studies eligible to be meta-analysed were first sorted by the statistical method used, pollutant, respective exposure window(s), and type(s) of PTBs investigated (**Table 23**). Then, how study dependencies were identified and the rationale for study selection for each statistical method-pollutant-exposure period-outcome combination is shown in **Figure 11.** The studies were ranked according to risk of bias (RoB), prioritising studies with lower RoB, then sample size, prioritising studies with larger sample size. It was decided that the assessment of study dependencies should be centred around studies with lower risk of bias, as they were likely more rigorously conducted and more informative.

For studies within the same pollutant-exposure period-outcome-estimate combination, the studies were first sorted by the countries in which they were conducted. If the studies were from different countries, they were deemed independent. Regional studies covering multiple countries were also considered. If a nationwide study was completely embedded within a regional study—using the same population/cohort and overlapping time periods—the national study was excluded. For studies conducted within the same country, temporal dependencies were assessed first. Studies with no overlap in time periods were considered independent. For those with overlapping periods, they were sorted by risk of bias (overall level, then confounder adjustment) and sample size. The studies were first divided into two groups: nationwide studies (covering more than two states) and state-level studies.

Among the nationwide studies, the highest-ranked study was selected first, followed by the next one that was not temporally dependent on the previously selected studies, in order of ranking. One nationwide study was retained because the nationwide studies retrieved were hospital-based rather than using birth registries, resulting in generally small population overlap compared to registry-based studies, which were usually at state-level. However, nationwide studies can provide additional information from areas not covered by other statewide or substate level studies, making the inclusion of one beneficial.

State-level studies were then further divided into statewide studies and substate level studies (county/city/district level). In each group, the highest-ranked study was selected first, followed by others that were not dependent on the previously selected studies, in order of ranking. The selected studies from both groups were then compared to check for temporal overlap exceeding 50%. If there was no significant overlap, both studies were retained. If there was an overlap, the area covered and the rankings of the statewide and substate level studies were compared. If the within-state study covered areas not included in the statewide study and was ranked higher, it was kept; otherwise, it was removed.

There is a preference for retaining statewide studies that provide larger spatial coverage of areas not addressed by studies focusing on a single city, county, or district. However, if a within-state study covers areas not included in the statewide study and is ranked higher, it may offer valuable insights due to better confounder adjustment or a larger sample size, thus providing more comprehensive coverage of a refined region. In such cases, both studies would be retained. A similar rationale applies to keeping both a national study and a statewide study, as the actual overlap in populations would be limited unless the exact same population or data source was used.

Table 23. The statistical method-pollutant-exposure period-outcome combinations of interest.

Pollutant	Statistical method	Exposure period	Outcome			
$PM_{2.5}$	Odds ratios	Preconception	PTB	Moderate-to-late PTB	Very PTB	Extremely PTB
		Trimester 1				
		Trimester 2				
		Trimester 3				
		Whole pregnancy				
	Hazard ratios	Preconception				
		Trimester 1				
		Trimester 2				
		Trimester 3				
		Whole pregnancy				
NO_2	Odds ratios	Preconception				
		Trimester 1				
		Trimester 2				
		Trimester 3				
		Whole pregnancy				
	Hazard ratios	Preconception				
		Trimester 1				
		Trimester 2				
		Trimester 3				
		Whole pregnancy				
O_3	Odds ratios	Preconception				
		Trimester 1				
		Trimester 2				
		Trimester 3				
		Whole pregnancy				
	Hazard ratios	Preconception				
		Trimester 1				
		Trimester 2				
		Trimester 3				
		Whole pregnancy				

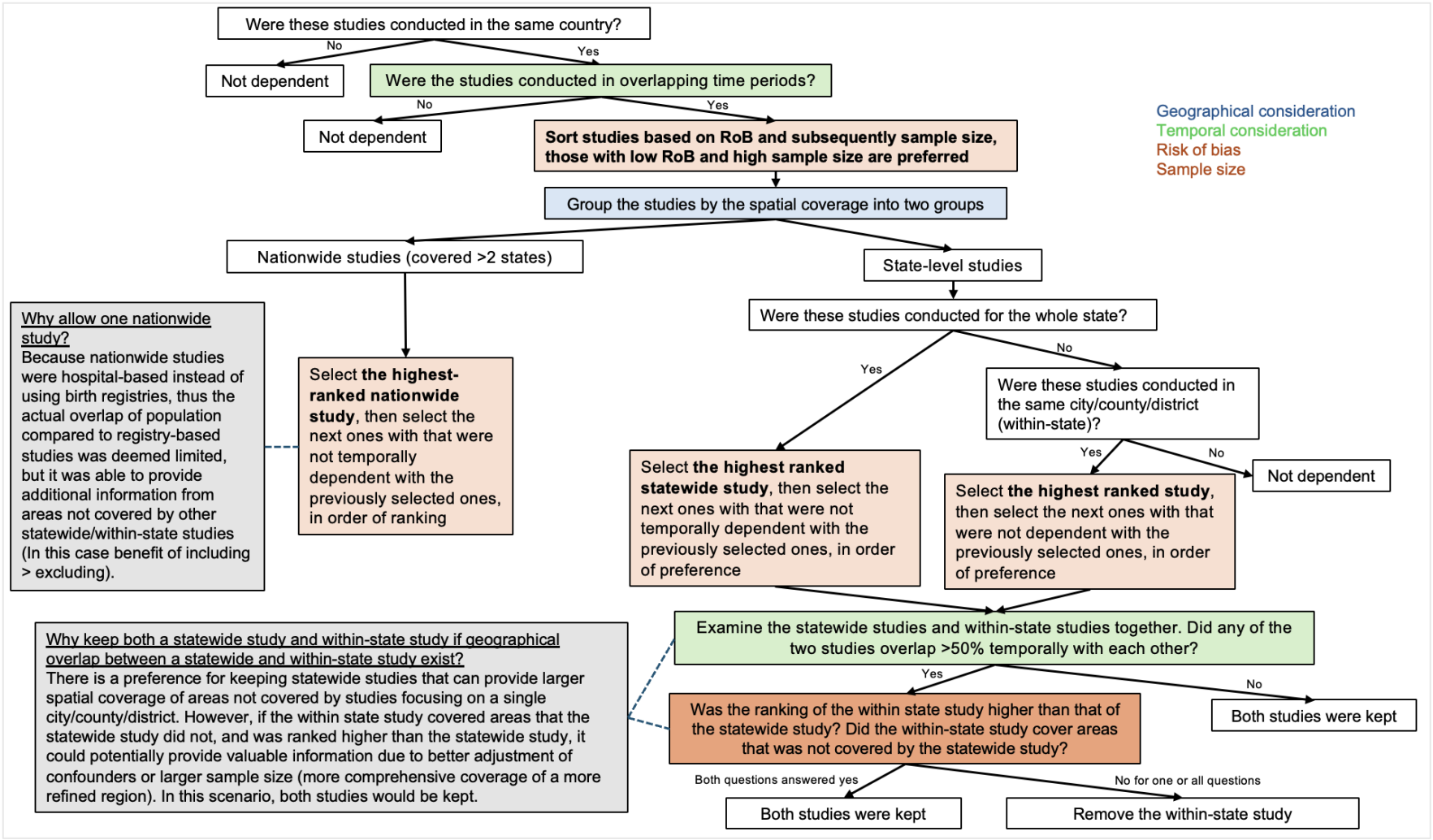


Figure 11. A flowchart showing how study dependencies between studies were identified, and how a study(ies) was selected to be included in the meta-analysis.

5.2.4.1. Sorting study dependencies: an example using studies in California

Studies in the US that have derived an odds ratio (OR) for whole pregnancy $PM_{2.5}$ exposure and the risk of PTB are shown in **Figure 12**.

1. Were these studies conducted in the same country? Yes
2. Were the studies conducted in overlapping time period? Yes
3. Group studies by spatial coverage
    a. Nationwide studies: There was one nationwide study (Mendola et al., 2016), which was kept as it was the only nationwide study available.
    b. Statewide studies: There were 6 statewide studies in California (Mekonnen et al., 2021; Basu et al., 2017; Benmarhnia et al., 2017; Huynh et al., 2006; Wu et al., 2016; Laurent et al., 2016, in order of preference). Mekonnen et al. (2021) was ranked highest thus was selected, followed by Basu et al. (2017) which did not temporally overlap with Mekonnen et al. (2021).
    c. Within-state studies: There were 3 within-state studies in California (Wu et al., 2009; Wu et al., 2011; Wilhelm et al., 2011, in order of preference). Wu et al. (2009) was selected as it was ranked highest, and the other two studies were not included due to complete temporal overlap with Wu et al. (2009)
4. Did any of the selected statewide studies and within state studies overlap >50% temporally with each other? Yes, Wu et al. (2009) and Basu et al. (2017)
5. Was the ranking of the within state study higher than that of the statewide study? And, did the within-state study cover areas that was not covered by the statewide study?
    a. Yes for both questions. Wu et al. (2009) covered LA and Orange County, and Basu et al. (2017) were able to offer information on 6 additional counties in addition to LA (Excluding orange), hence despite the >50% temporal overlap, the the extent of spatial overlap was small (in the perspective of Basu et al. (2017). However, this means that 50% of Wu et al. (2009)'s area was covered by Basu et al. (2017). As it was ranked higher than Basu et al. (2017), in this case it has merits by being more informative (lower RoB and give detail for Orange County), and thus there was benefits in keeping both studies.
6. Conclusion: 4 studies were kept (Basu et al., 2017; Mekonnen et al., 2021; Mendola et al., 2016; Wu et al.,2009)

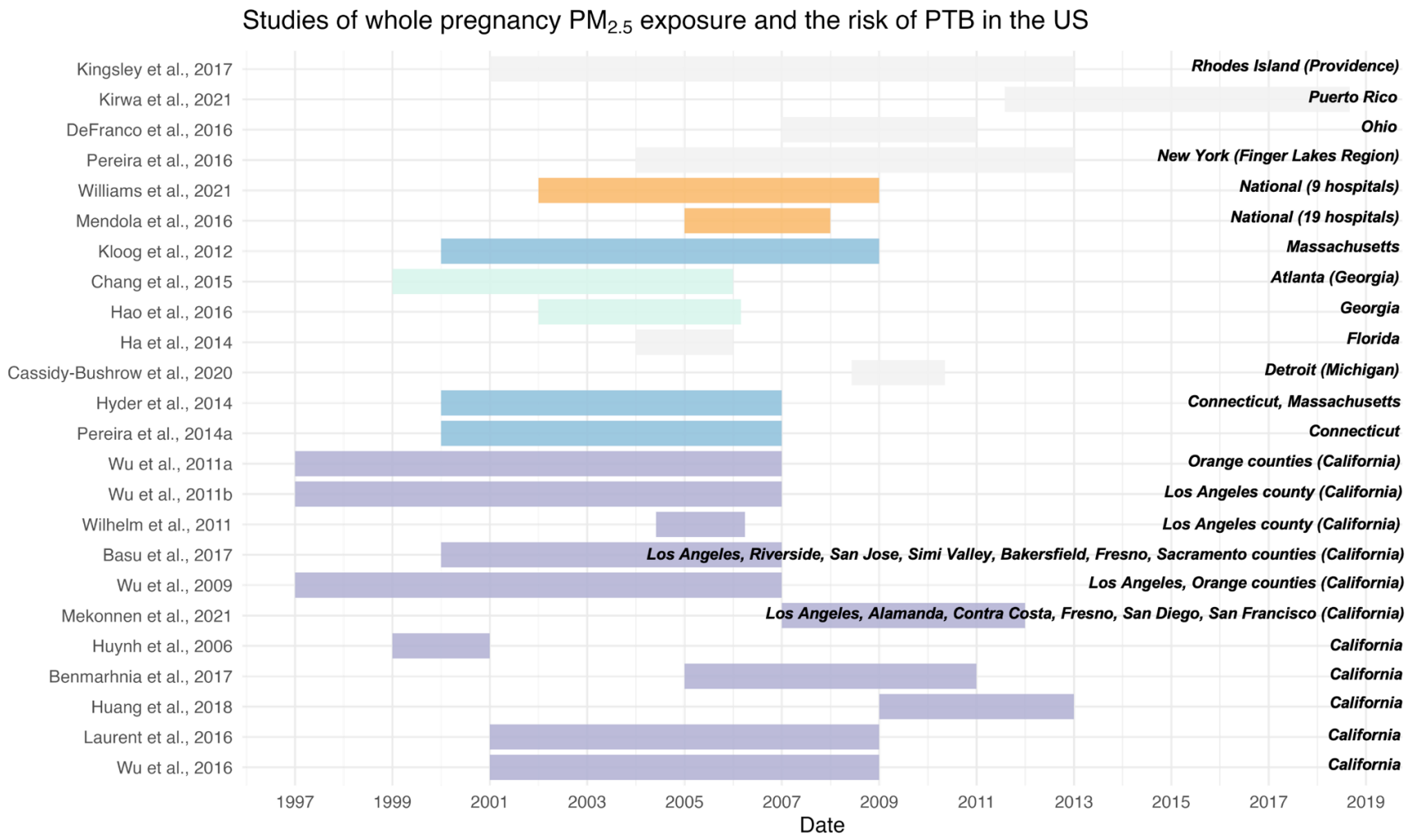


Figure 12. The location, start and end date of studies reporting ORs in the US for whole pregnancy $PM_{2.5}$ exposure and risk of PTB. Studies with potential dependencies are highlighted in purple (California), blue (Connecticut and/or Massachusetts), green (Georgia), or orange (nationwide), and those without are highlighted in grey. Wu et al. 2011a and 2011b were from the same study (Wu et al., 2011) that had provided estimates for Los Angeles (Wu et al., 2011a) and Orange (Wu et al., 2011b) counties.

5.2.5. Meta-analytical method of choice

Meta-analysis can be conducted using fixed-effect or random-effect models. The use of fixed-effect models was covered in **section 5.2.3**. Random-effect models assume that there is a distribution of the true effect size, and variation in observed effect sizes are due to both sampling error and between-study heterogeneity (Higgins et al., 2009). The studies included in the meta-analyses were expected to be heterogenous due to variations in the populations, health systems, and study designs used. Hence, random-effects meta-analyses were conducted for all exposure-outcome pairs, apart from the meta-analyses of estimates derived from subgroups of populations in a study (details in **Appendix 12**).

There are different ways to estimate the pooled effect from heterogeneous studies, for example, the DerSimonian Laird method, the Sidik-Jonkman method, or the Restricted Maximum-Likelihood method (DerSimonian & Laird, 1986; Tanriver-Ayder et al., 2021). The DerSimonian Laird method is adopted by most preceding publications and has shown that between-study heterogeneity was very high (Stieb et al., 2012; Zhu et al., 2015; Ju et al., 2023; Yu et al., 2022). The Sidik-Jonkman method is shown to produce more robust meta-analysis results when studies were expected to be heterogeneous and when number of studies available are small (which could be relevant in sensitivity analyses) (IntHout et al., 2014). Hence, meta-analyses were run using both the Sidik-Jonkman and the DerSimonian Laird method separately, and the results were very similar (results not shown), hence the DerSimonian Laird method was used to calculate the pooled effect (InHout et al., 2016).

The meta-analyses conducted were performed on studies per exposure window-PTB type (**Table 23**) pair using R version 4.3.1. with package “meta”, if more than two studies per type of effect estimate (OR/HR) were available (Balduzzi et al., 2019; R Core Team, 2023). Between-study heterogeneity was quantified using I^2 and τ^2 statistics, which estimates the variation across the studies used in the meta-analysis that is not due to chance, and between-study variance in the effect sizes of the studies respectively (Thorlund et al., 2012).

Prediction intervals (PI), which show the range in which estimates from future studies would lie, are derived from τ^2 statistics and are also included in the forest plots produced (IntHout et al., 2016). Forest plots with scaled effect estimates, pooled summary estimates, the respective 95% confidence intervals (CIs), PIs, I^2 and τ^2 statistics were produced.

5.2.6. Leave-one-out analysis

Leave-one-out analysis was conducted to explore the influence of particular studies on the summary estimate and test the robustness of the results. It was not conducted for all

pollutant-exposure period-outcome combinations due to time-constraints and the amount of material that would be generated. Hence, it was conducted only for PTB (but not its subgroups) when there were fewer than 10 studies in the meta-analysis, with meta-analyses including studies that were weighted more than 20%, or containing primary studies with estimates showing no overlap with the 95% CI of the summary estimate.

5.2.7. Investigating drivers of heterogeneity

Heterogeneity found in meta-analyses can originate from various aspects of the studies included, ranging from pollutant composition, healthcare systems, to analytical methods used, or research groups. Based on *a priori* decisions, this thesis focused on heterogeneity originating from risk of bias, location, exposure assessment methodology, source of data, and concentration of pollutants. Sensitivity analyses were conducted if there were more than two studies available for a statistical method-pollutant-exposure period-outcome combination. Stratified analyses were conducted to explore if any of these factors was a significant contributor to the heterogeneity observed. **Figure 13** shows the stratified analyses conducted.

5.2.7.1. Risk of bias

Risk of bias (or quality of studies) can be a source of heterogeneity as studies of different risk of bias may have systematic differences in their effect sizes due to factors such as extent in controlling for bias, comprehensiveness in exposure assessment methodologies, and sample size. To investigate its impact, three sensitivity analyses were conducted:

1. Studies assigned a high risk of bias were excluded from the main meta-analyses conducted. Here, studies assigned a high risk of bias were added back to the studies included in the main meta-analyses, to understand their influence on the summary estimate.
2. Stratified analyses were conducted, where the studies in the meta-analyses conducted in (1) were stratified according to their risk of bias level (Low/Moderate/High).
3. For studies in the “Low” or “Moderate” risk of bias level, they were further stratified into studies that have adjusted for all critical confounders, or those that did not.

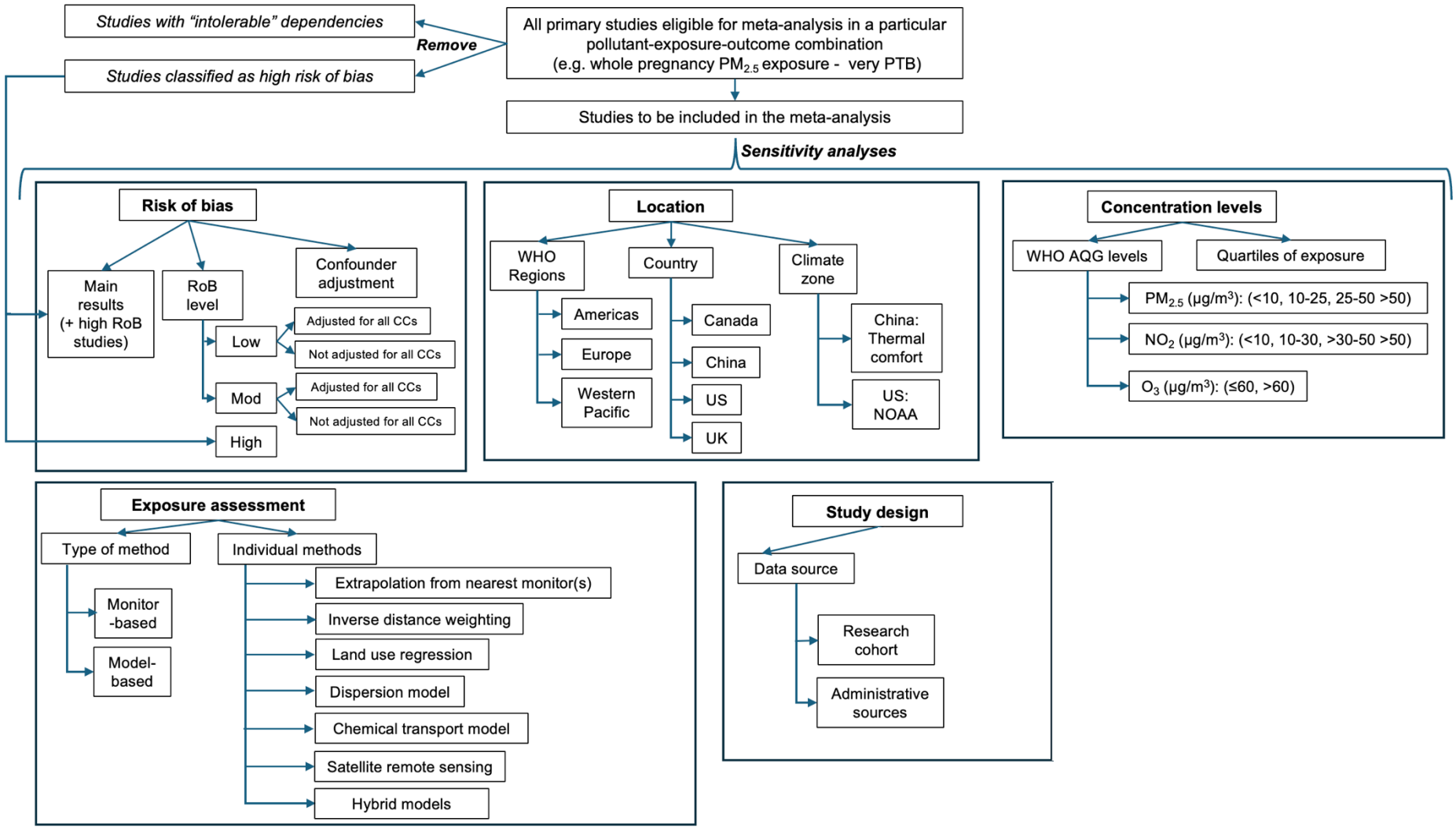


Figure 13. A flowchart showing the process of sensitivity analysis conducted for the meta-analysis for each main result, if more than two studies were available.

The first and second sensitivity analyses aimed to show if risk of bias had influenced the results shown in the primary studies and thus the pooled effect size. The third stratification served to explore if studies with more complete adjustment of confounders have reported different results compared to those that did not. If risk of bias or confounder adjustment were found to be a substantial contributor of heterogeneity, then a more cautious interpretation of the main results would be needed, and it calls for more future research that is more robust.

5.2.7.2. Location

Location can be a source of heterogeneity as it is a proxy of a variety of differences between studies included in a meta-analysis, which could range from population characteristics, healthcare systems or policies, to pattern, composition, or concentrations of pollutants, as well as culture or resources available. Three sensitivity analyses were conducted to explore if location is a contributor of heterogeneity observed in the main meta-analyses:

1. Estimates from studies were stratified by WHO regions, to understand if broad regional heterogeneity is present (WHO, 2024).
2. Estimates from studies were stratified according to the country in which the study took place.
3. If more than three estimates were available after country-level stratification, then a further stratification by climate region within the country was carried out. Results from section 4.3.1. suggest that there would only be sufficient studies for this sensitivity analysis to be conducted in the US or China. US studies were stratified according to the National Oceanic and Atmospheric Administration (NOAA) climate region (NOAA, 2024), which divided the US into nine climate consistent regions. It is well-used in climate region analysis, and preceding work has found regional heterogeneity in particulate matter constituents between these regions (Achilleos et al., 2017). For China, stratification was conducted using the classification by thermal comfort from a recent publication, which divided China into five climate regions with different weather patterns (severe cold, cold, hot summer warm winter, hot summer cold winter, and mild) (Yang et al., 2023).

The three sensitivity analyses described aimed to understand the extent to which location, at a broad, country level, or climate regions level, accounted for the heterogeneity observed. If it was found to be a substantial contributor, then it calls for meta-analyses to be conducted according to location, and caution should be exercised when interpreting the results from the main meta-analyses.

5.2.7.3. Exposure assessment methodology

Exposure assessment methodology can contribute to the heterogeneity observed in meta-analyses due to underlying variability in the methodologies, and thus precision and accuracy of different exposure assessment methods. It could be a proxy of geographical or temporal variations, as some exposure assessment methods could be preferred by researchers in certain locations or time periods. As a result, there could be systematic differences in the effect estimates derived using different exposure assessment methods. Two sensitivity analyses were conducted:

1. Estimates from primary studies were stratified according to two categories of exposure assessment methodology: methods that were predominantly monitor-based (extrapolation from nearest monitors, inverse distance weighting), and methods that were model-based (land-use regression, dispersion modelling, chemical transport models, satellite remote sensing, or hybrid models that used a combination of these). There were no studies that used methods that did not fit in either category.
2. Estimates from primary studies were stratified by the respective exposure assessment method (extrapolation from nearest monitors, inverse distance weighting, land-use regression, dispersion modelling, chemical transport models, satellite remote sensing, or hybrid models).

The first sensitivity analysis explored if heterogeneity could originate from greater measurement error from relatively crude exposure assessment methods that were predominantly monitor-based, compared to more sophisticated model-based approaches that could provide more accurate exposure assessment. The second sensitivity analysis was done to understand the extent to which the variety of exposure assessment methods used has contributed to the overall heterogeneity found. If exposure assessment method contributed to the heterogeneity observed, then it potentially emphasises the importance of robust and comprehensive exposure assessment methods in future studies.

5.2.7.4. Concentration levels

Concentration levels of pollutants can be a potential source of heterogeneity as studies included in a meta-analysis may have used populations exposed to drastically different levels of air pollutants, which has several implications. First, the extent of the health impacts could differ at different pollutant levels, and the effect at higher exposure levels could be more easily detected due to greater exposure contrast. Second, populations exposed to high

levels of exposure may differ in characteristics and hence susceptibility to the outcome. Two sensitivity analyses were conducted.

1. Studies were stratified according to WHO Air Quality Guidelines (AQG) target levels (WHO, 2021). The respective cut-offs are shown in the following sections.
2. After deriving the overall distribution of median (or mean, if median value was unavailable) level of exposure of the primary studies included in the meta-analysis, the studies were stratified into the respective quartiles of median values.

If concentration level is a substantial contributor to the heterogeneity observed, then caution is advised when interpreting the main meta-analyses results, as it could be influenced by studies of a particular range of concentration. There was one instance where the concentrations were only presented graphically (Giorgis-Allemand et al., 2017). An image analysis tool was used to extract the value presented (Rohatgi, 2024).

5.2.7.4.1. $PM_{2.5}$

The WHO AQG annual average $PM_{2.5}$ is 5 μg/m^3. Initial data exploration showed that most studies were not conducted in settings with annual average concentrations of $PM_{2.5}$ less than 5 μg/m^3. The estimates were stratified into four groups: (1) <10 μg/m^3 (WHO interim target 4), (2) 10-25 μg/m^3 (WHO interim target 2), (3) 25-50 μg/m^3 (WHO interim target 2 for 24-hour exposure), (4) >50 μg/m^3.

5.2.7.4.2. NO_2

The WHO AQG annual average NO_2 is 10 μg/m^3. The estimates were stratified into four groups: (1) <10 μg/m^3 (WHO AQG annual target), (2) 10-30 μg/m^3 (WHO interim target 2), (3) 30-50 μg/m^3, (4) >50 μg/m^3.

5.2.7.4.3. O_3

The WHO AQG annual average O_3 is 60 μg/m^3. As there were far fewer estimates for O_3, a dichotomous division of ≤60 μg/m^3 and >60 μg/m^3 was used.

5.2.7.5. Source of data

Data source can be a source of heterogeneity, as they could be proxies of availability of information on covariates, population coverage, bias potential, or sample size. Studies were

stratified by source of data used, namely administrative data and research cohorts. Studies that used information such as birth registries, electronical medical records, or government statistical agencies were classified as using administrative data sources. Studies that used information from prospective cohorts that recruited individuals and obtained information through methods such as questionnaires, survey, and clinical assessment, sometimes with administrative data were classified as using research cohorts.

If data source was a substantial contributor to the heterogeneity observed, then future studies would need to consider how data source may impact their results, and for impact assessment, if a concentration-response function from certain data sources is preferable over one derived from all studies.

5.2.8. Publication bias

Egger's test and funnel plots were used to examine potential publication/small study bias where there were more than ten studies included in the meta-analysis (Egger et al., 1997; Sterne et al., 2011).

5.2.9. Interpreting the results of meta-analysis

Figure 14 shows how the results of the meta-analyses conducted were interpreted. The following questions were asked:

1. Did the summary estimate (and its 95%CI) show an increase/decrease in risk? Yes was chosen when the 95%CI of the summary estimate did not overlap with 1. No was chosen when the 95% CI of the summary estimate overlapped with 1. If answered yes, then the questions 2 to 4 were asked. If answered no, then only question 4 was asked.
2. What is the magnitude of the increase/decrease in risk?
3. What is the certainty in the magnitude of the association?
4. How robust is the conclusion to the addition of future studies?

Figure 14 also shows the wording that was used to describe the magnitude, certainty, and robustness of the association found, to ensure consistency in the interpretation of the findings.

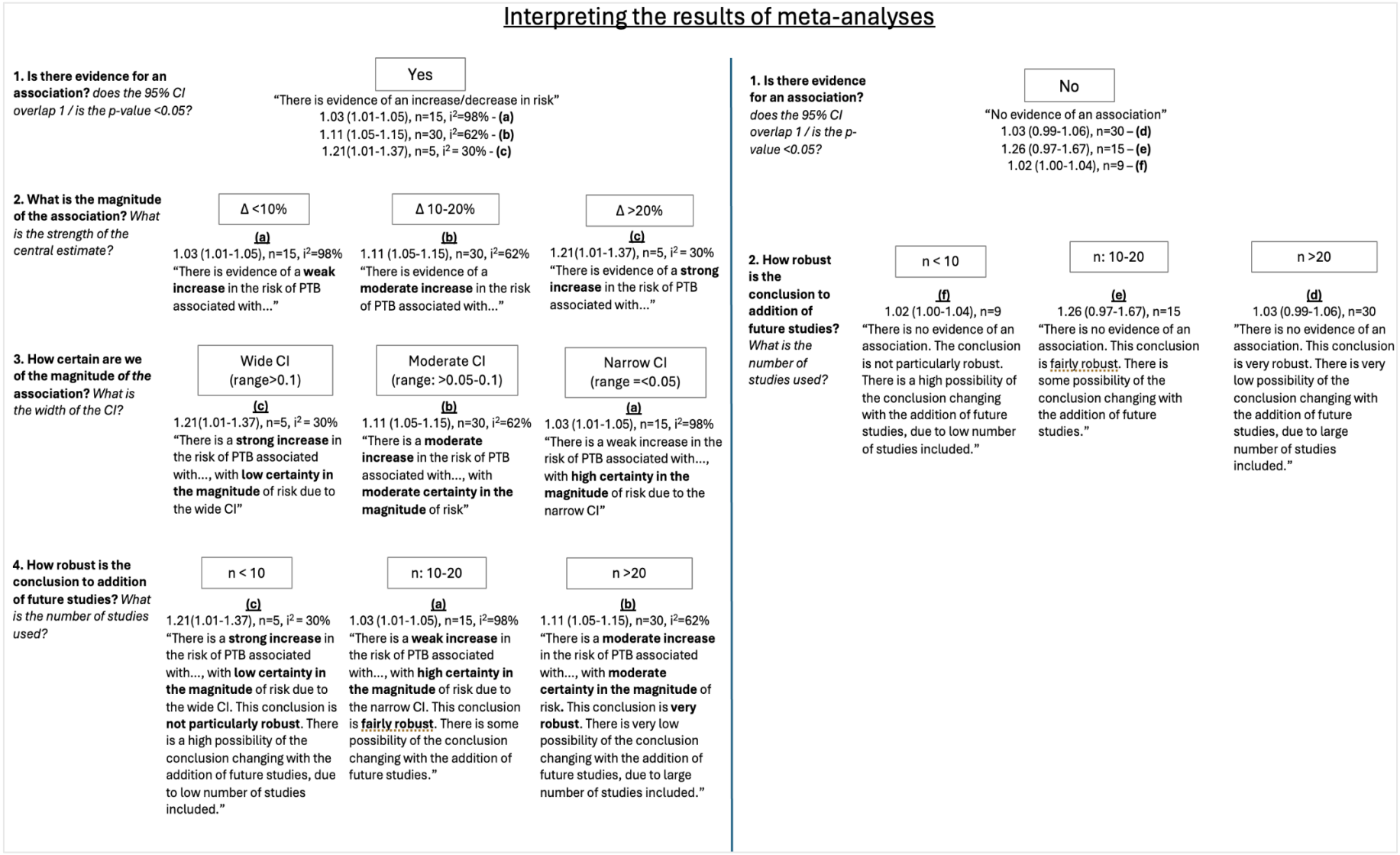


Figure 14. The interpretation of the result of a meta-analysis regarding the presence, magnitude, and robustness of the association.

5.2.10. Interpreting the results of the stratified analysis

Based on the results of the stratified analyses conducted, a factor is determined to be not a driver, a partial driver, or a driver of the high heterogeneity observed using the following questions:

1. Did a reduction in heterogeneity occur in any strata? Yes (≥ 10%) / No (<10%)
2. If yes, what was the strength of the reduction? Weak (10-20%) / Moderate (>20-40%) / Strong (>40%)
3. How often did the reduction occur in other strata? 1 stratum / some but not all strata / all strata
4. How often did the reduction occur in other exposure windows? Only 1 window / some but not all windows / all windows

Table 24 shows the definitions for not being a driver, being a partial driver, or a full driver of the heterogeneity observed are as follows:

Table 24. The definitions for a factor not being a driver, being a partial driver, or a driver of the heterogeneity observed (continued on next page).

Category	Definition
Not a driver	Reduction in heterogeneity was not observed at all, or a weak reduction in the heterogeneity was observed in only 1 stratum and 1 exposure window.
Partial driver	Reduction in heterogeneity observed in at least 1 but not all strata, in more than 1 or all exposure windows. It implied that a portion of the heterogeneity observed in the main result can be partially explained by the factor. However, heterogeneity could be driven by other factors in other strata and other exposure windows.
Driver	Reduction in heterogeneity observed in all strata in at least 1 exposure window - 1 exposure window: full driver in one window - >1 exposure window: full driver for some windows - All windows: full driver of the association It implied that a portion of the heterogeneity observed in the main result can potentially be fully explained by the factor. However, it could be a proxy of other unmeasured factors that contributed to the

	heterogeneity, and it may not be the ***only*** driver of the heterogeneity observed.

A Q-test for subgroup differences was also conducted. However, as the number of studies available in each stratum was usually low (<10), the results of the Q-test were not very reliable. Low certainty was particularly flagged when there were fewer than five studies available in each stratum (or when there was only one stratum available).

5.2.11. Causality discussion based on the Bradford Hill criteria

The Bradford Hill Criteria are a set of nine principles outlined by Sir Austin Bradford Hill that can be used to determine if an environmental risk factor causes a disease (Hill, 1965). The nine criteria are as follows: strength, consistency, specificity, temporality, biological gradient, plausibility, coherence, experiment, and analogy. Their definitions are shown in **Table 4**

These criteria were used as guidance in discussing whether there is sufficient evidence that any observed increased risk of PTB or its subgroup associated with $PM_{2.5}$, NO_2, O_3 exposure, could be causal.

5.3. Results

The results of sorting study dependencies are shown in **Appendix 14** for each of the exposure-outcome relationships that can be meta-analysed. 33 authors were contacted for further information or clarification, and 18 of them replied (**Appendix 13**). This has resulted in additional estimates being available across various exposure-outcome pairs in Chapter 5. Result of the leave-one-out analyses conducted is shown in **Appendix 18**.

5.3.1. Preconception exposure

Insufficient studies (<3) were available for all pollutants and exposure windows for preconception exposure after sorting for study dependencies. Hence, meta-analysis was not conducted.

5.3.2. Effect of $PM_{2.5}$ exposure

5.3.2.1. Trimester 1 $PM_{2.5}$ exposure (OR)

There were 41 studies that gave estimates for this exposure-outcome pair. Most of them were conducted in either China or the US. 13 estimates (from 12 studies) were removed due to intolerable dependencies or being assigned a high RoB. 28 studies were included in the meta-analysis.

There was no evidence of an association between trimester 1 $PM_{2.5}$ exposure and the risk of PTB (OR = 1.01, 95% CI: 0.99 -1.03) (**Figure 15**). This is expected as most primary studies reported estimates overlapping the null, with one study reported a strong increase in risk (Pereira et al., 2016). The conclusion is very robust with low possibility of the conclusion changing with the addition of future studies, due to the large number of studies included. High heterogeneity was observed (I^2 = 87%, τ^2 = 0.0016), with PI overlapping the null (0.93 - 1.10).

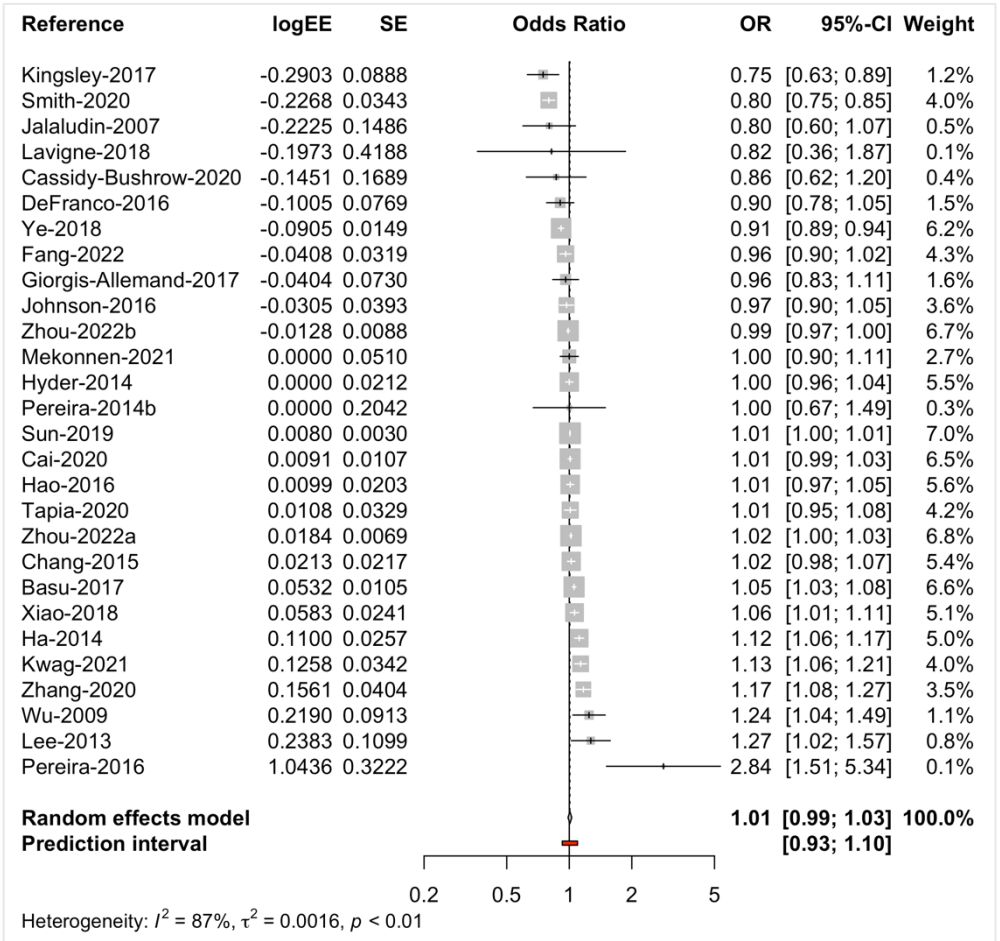


Figure 15. Forest plot for the meta-analysis of the association between trimester 1 $PM_{2.5}$ exposure and the risk of PTB from studies using ORs. The individual effect sizes and 95%CIs, log of the effect sizes (logEE), standard error (SE), summary estimate (in diamond shape), and its 95%CI, prediction intervals (PI), and I^2 statistics are presented.

5.3.2.2. Trimester 2 $PM_{2.5}$ exposure (OR)

There were 40 studies that gave estimates (n = 41) for this exposure-outcome pair. Most of them were conducted in China or the US. 12 studies (13 estimates) were removed due to intolerable dependencies or being assigned a high RoB. 27 studies were included in the meta-analysis.

There was no evidence of an association between trimester 2 $PM_{2.5}$ exposure and the risk of PTB (OR = 1.03, 95% CI: 1.00 - 1.06) (**Figure 16**). High heterogeneity was observed (I^2 = 96%, τ^2 = 0.0043), with PI overlapping the null (0.90 - 1.18). The conclusion is very robust with low possibility of the conclusion changing with the addition of future studies, due to the large number of studies included.

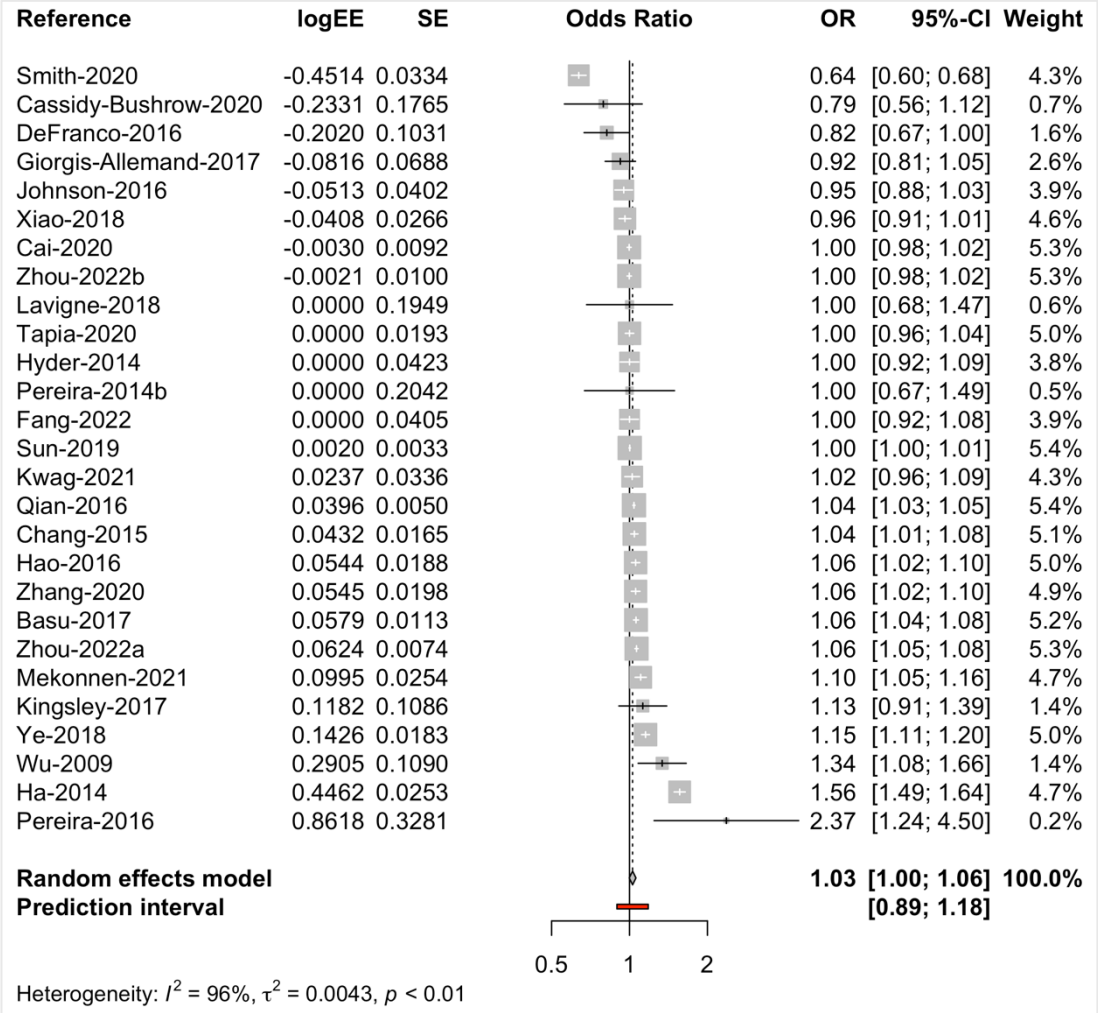


Figure 16. Forest plot for the meta-analysis of the association between trimester 2 $PM_{2.5}$ exposure and the risk of PTB from studies using ORs. The individual effect sizes and 95%CIs, log of the effect sizes (logEE), standard error (SE), summary estimate (in diamond shape), and its 95%CI, prediction intervals (PI), and I^2 statistics are presented.

5.3.2.3. Trimester 3 $PM_{2.5}$ exposure (OR)

There were 30 studies that gave estimates for this exposure-outcome pair. Most of them were conducted in China or the US. 8 studies were removed due to intolerable dependencies or being assigned a high RoB. 22 studies/estimates were used in the meta-analysis.

There was evidence of a weak increase in the risk of PTB associated with a per 10 μg/m^3 increase in trimester 3 $PM_{2.5}$ exposure (OR = 1.05, 95% CI: 1.02 - 1.07), with high certainty in the magnitude due to the narrow CI. The conclusion is very robust with low possibility of the conclusion changing with the addition of future studies, due to the large number of studies included (**Figure 17**). High heterogeneity was observed (I^2 = 95%, τ^2 = 0.0028), with PI overlapping the null (0.93 - 1.17).

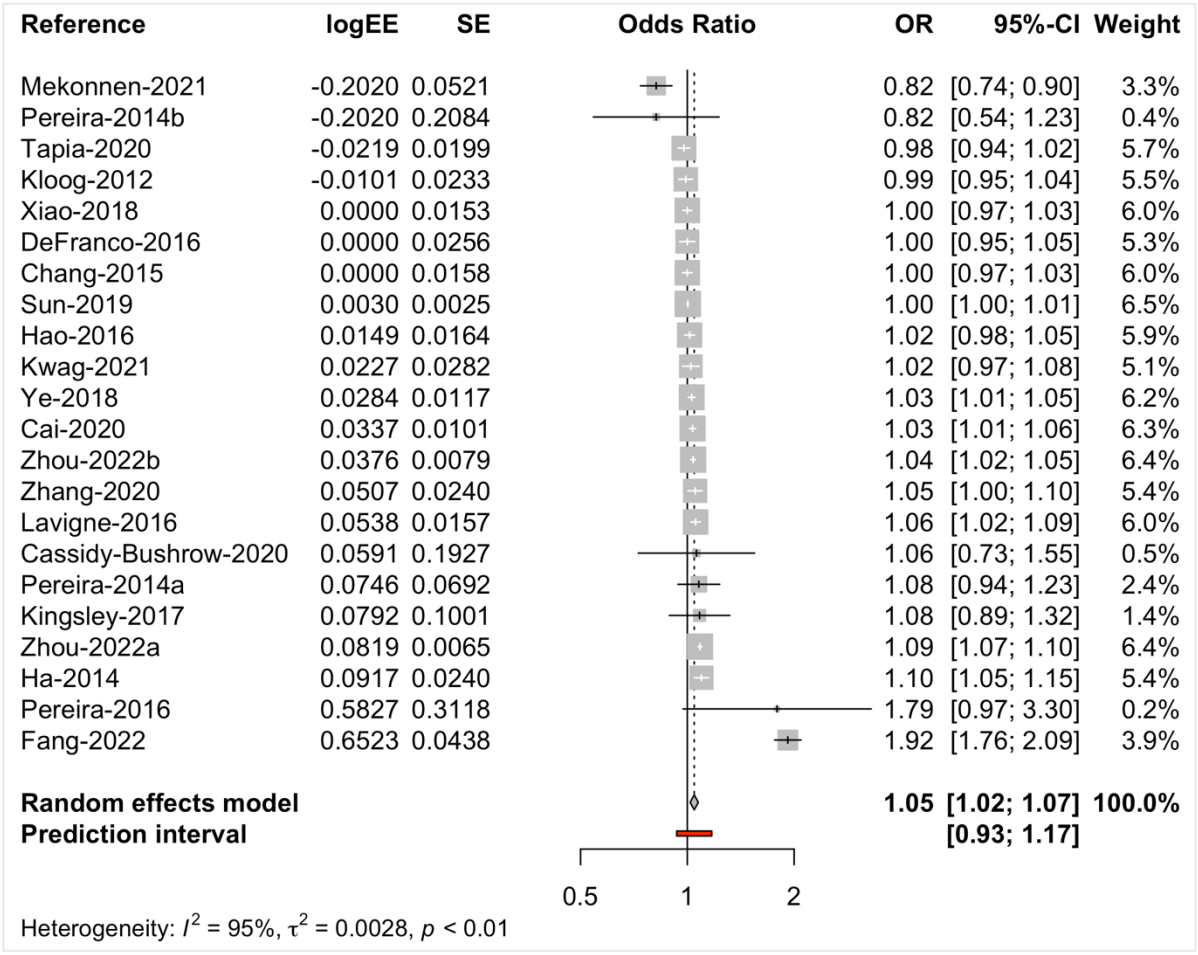


Figure 17. Forest plot for the meta-analysis of the association between trimester 3 $PM_{2.5}$ exposure and the risk of PTB from studies using ORs. The individual effect sizes and 95%CIs, log of the effect sizes (logEE), standard error (SE), summary estimate (in diamond shape), and its 95%CI, prediction intervals (PI), and I^2 statistics are presented.

5.3.2.4. Whole pregnancy $PM_{2.5}$ exposure (OR)

There were 51 studies with estimates (n = 52) for this exposure-outcome pair. Most studies were conducted in the US or China. 18 studies were removed due to intolerable dependencies or high risk of bias. 34 studies were used in the meta-analysis.

There was a moderate increase in the risk of PTB associated with a per 10 µg/m^3 increase in whole pregnancy exposure (OR = 1.10, 95% CI: 1.06 - 1.14) (**Figure 18)**, with moderate certainty in the magnitude of risk. The conclusion is very robust with low possibility of the conclusion changing with the addition of future studies, due to the large number of studies included. High heterogeneity was observed (I^2 = 92%, τ^2 = 0.0066), with PI overlapping the null (0.93 - 1.30).

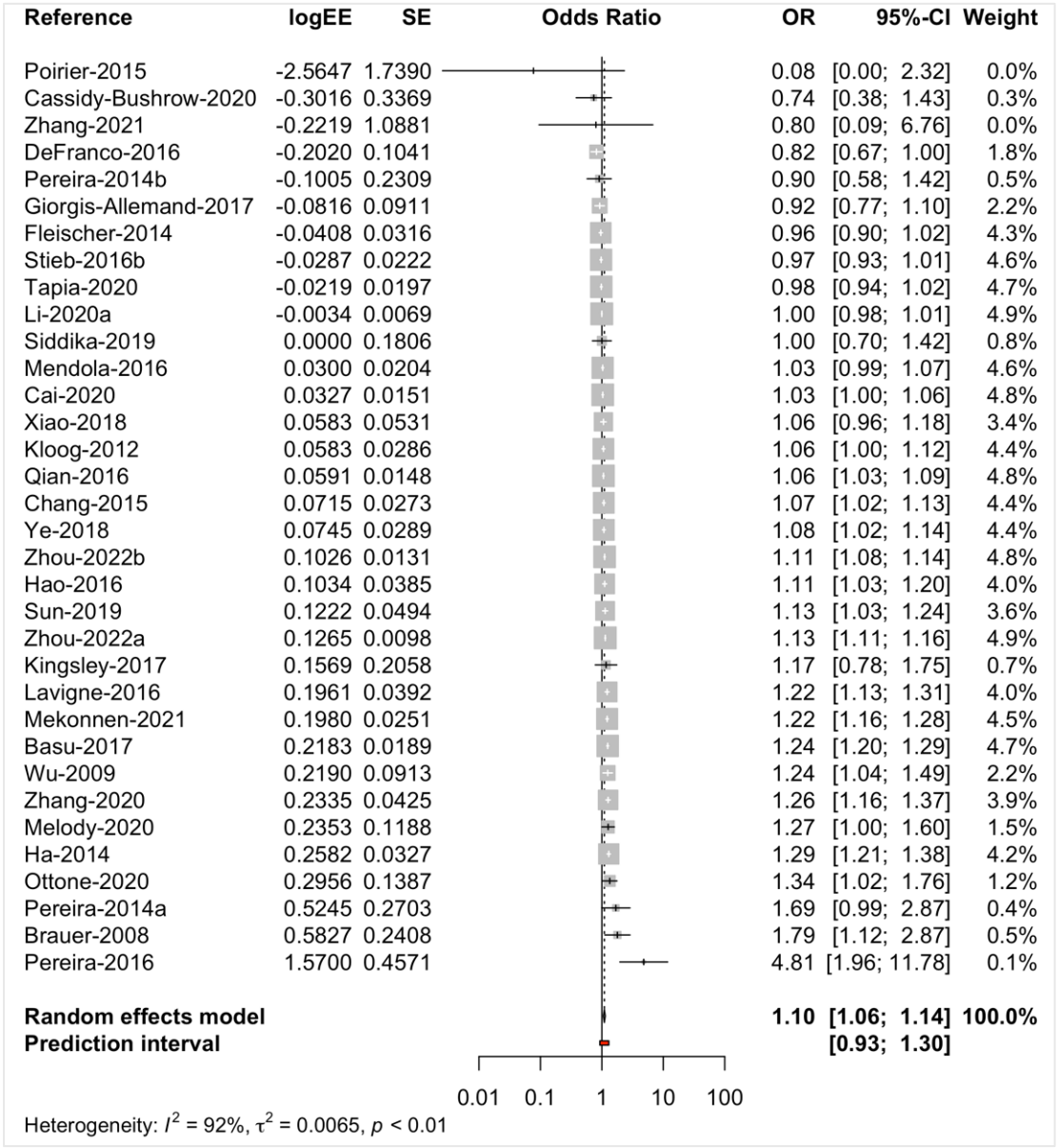


Figure 18. Forest plot for the meta-analysis of the association between whole pregnancy $PM_{2.5}$ exposure and the risk of PTB from studies using ORs. The individual effect sizes and 95%CIs, log of the effect sizes (logEE), standard error (SE), summary estimate (in diamond shape), and its 95%CI, prediction intervals (PI), and I^2 statistics are presented.

5.3.2.5. Trimester 1 $PM_{2.5}$ exposure (HR)

There were 13 studies that gave estimates for this exposure-outcome pair. Most studies took place in China. After removing 5 Chinese studies with intolerable dependencies, 8 studies/estimates were used in the meta-analysis. No studies were assigned a high risk of bias.

There was no evidence of an association between trimester 1 $PM_{2.5}$ exposure and the risk of PTB (HR = 1.02, 95% CI: 0.99 - 1.04) (**Figure 19**). This was expected as most primary studies reported estimates overlapping the null. This conclusion is not particularly robust. There is a high possibility of the conclusion changing with the addition of future studies, due to low number of studies included.

There were 2 studies that reported a statistically significant increase in risk, with one reporting a strong increase in risk (Chen et al., 2022; He et al., 2022). High heterogeneity was observed (I^2 = 97%, τ^2 = 0.0005), with PI overlapping the null (0.96 - 1.08). Leave-one-out analysis showed minimal influence of the primary studies to the summary estimate, indicating the robustness of the pooled results (**Appendix 18**).

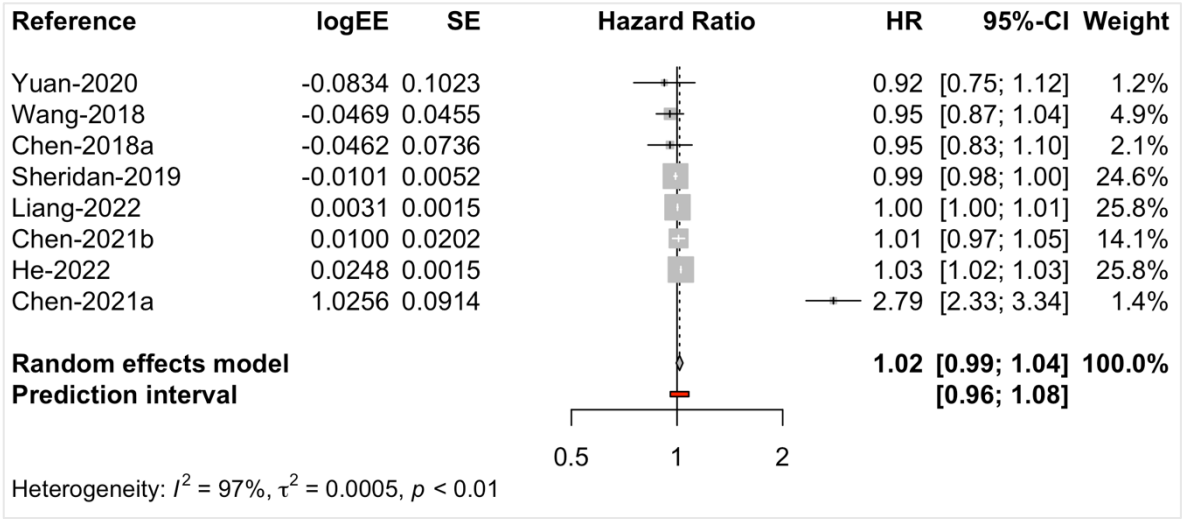


Figure 19. Forest plot for the meta-analysis of the association between trimester 1 $PM_{2.5}$ exposure and the risk of PTB from studies using HRs. The individual effect sizes and 95%CIs, log of the effect sizes (logEE), standard error (SE), summary estimate (in diamond shape), and its 95%CI, prediction intervals (PI), and I^2 statistics are presented.

5.3.2.6. Trimester 2 $PM_{2.5}$ exposure (HR)

There were 13 studies that gave estimates for this exposure-outcome pair. Most of them were conducted in China. After removing 5 studies with intolerable dependencies, 8 studies were used in the meta-analysis. No studies were assigned a high risk of bias (Appendix 13).

There was no evidence of an association between trimester 2 $PM_{2.5}$ exposure and the risk of PTB (HR = 1.01, 95% CI: 1.00 - 1.03, per 10 µg/m^3 increase) (**Figure 20**). High heterogeneity was observed (I^2 = 86%, τ^2 = <0.0001), with PI overlapping the null (0.99 - 1.04). This conclusion is not particularly robust. There is high possibility of the conclusion changing with the addition of future studies, due to low number of studies included. Leave-one-out analysis showed minimal influence of the primary studies to the summary estimate, indicating the robustness of the pooled results.

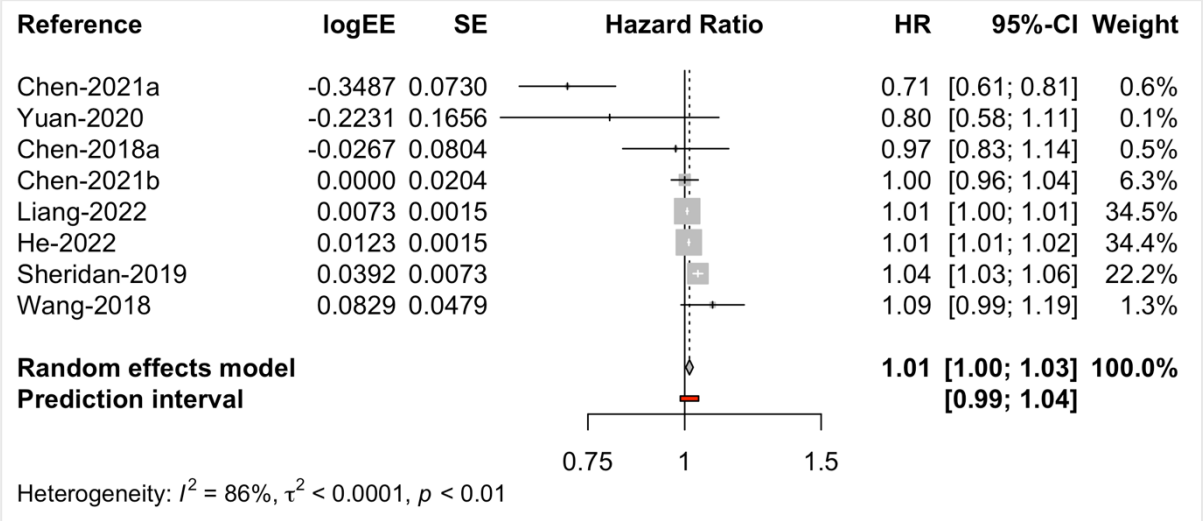


Figure 20. Forest plot for the meta-analysis of the association between trimester 2 $PM_{2.5}$ exposure and the risk of PTB from studies using HRs. The individual effect sizes and 95%CIs, log of the effect sizes (logEE), standard error (SE), summary estimate (in diamond shape), and its 95%CI, prediction intervals (PI), and I^2 statistics are presented.

5.3.2.7. Trimester 3 $PM_{2.5}$ exposure (HR)

There were 15 studies that gave estimates for this exposure-outcome pair. Most of them were conducted in China. After removing 5 studies with intolerable dependencies, 10 studies were meta-analysed. No studies were assigned a high risk of bias (Appendix 13).

There was evidence of a weak increase in the risk of PTB associated with a per 10 µg/m^3 increase in trimester 3 $PM_{2.5}$ exposure (OR = 1.05, 95% CI: 1.03 - 1.08, per 10 µg/m^3 increase), with high certainty in the magnitude due to the narrow CI (**Figure 21**). The conclusion is not particularly robust. There is high possibility of the conclusion changing with the addition of future studies, due to low number of studies included. High heterogeneity was observed (I^2 = 97%, τ^2 = 0.0006), with PI overlapping the null (0.99 - 1.12).

Leave-one-out analysis showed the removal of He et al. (2022) or Liang et al. (2022) strengthened the summary estimate (OR = 1.11, 95% CI: 1.04 - 1.18, when either study removed), suggesting some sensitivity to these 2 studies.

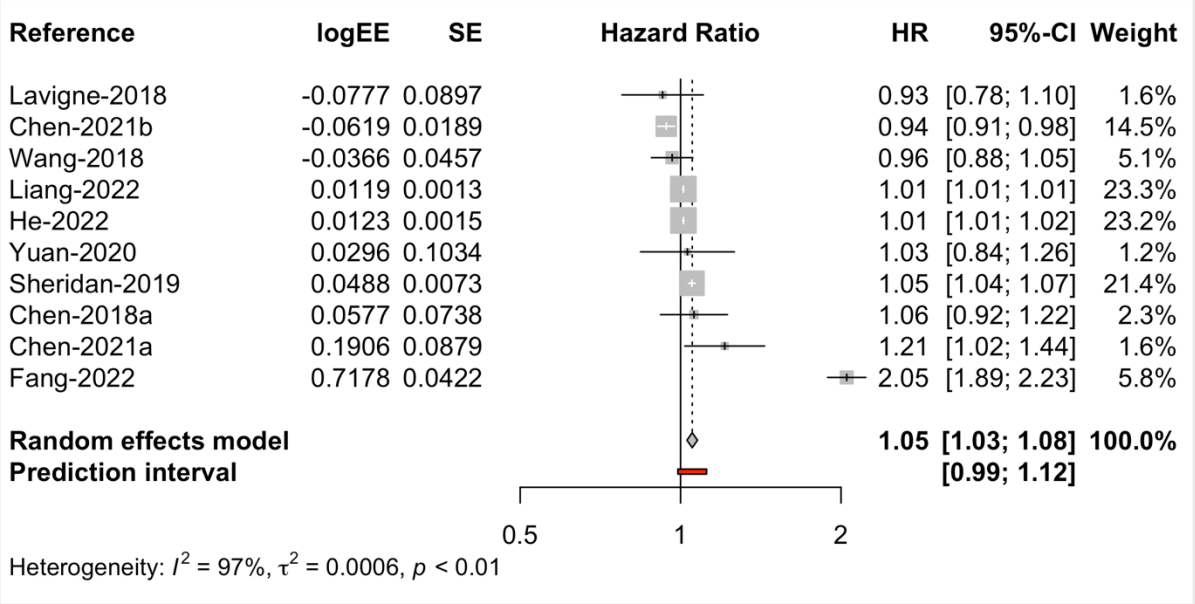


Figure 21. Forest plot for the meta-analysis of the association between trimester 3 $PM_{2.5}$ exposure and the risk of PTB from studies using HRs. The individual effect sizes and 95%CIs, log of the effect sizes (logEE), standard error (SE), summary estimate (in diamond shape), and its 95%CI, prediction intervals (PI), and I^2 statistics are presented.

5.3.2.8. Whole pregnancy $PM_{2.5}$ exposure (HR)

There were 15 studies that gave estimates for this exposure-outcome pair. Most of them were conducted in China. After removing 6 studies with intolerable dependencies, 9 studies were available to be meta-analysed. No studies were assigned a high risk of bias (Appendix 13).

There was evidence of a weak increase in the risk of PTB associated with a per 10 μg/m^3 increase in whole pregnancy $PM_{2.5}$ exposure (HR = 1.05, 95% CI: 1.03 - 1.07, per 10 μg/m^3 increase) (**Figure 22**), with high certainty in the magnitude of the risk due to the narrow CI. High heterogeneity was observed (I^2 = 95%, τ^2 = 0.0003), with PI touching the null (1.00 - 1.10). The conclusion is not particularly robust, with high possibility of the conclusion changing with the addition of future studies, due to low number of studies included.

Leave-one-out analysis showed the removal of He et al. (2022) or Liang et al. (2022) increased the summary estimate to 1.14 (1.04 - 1.24) and 1.14 (1.05 - 1.24) respectively, suggesting some sensitivity to these 2 studies.

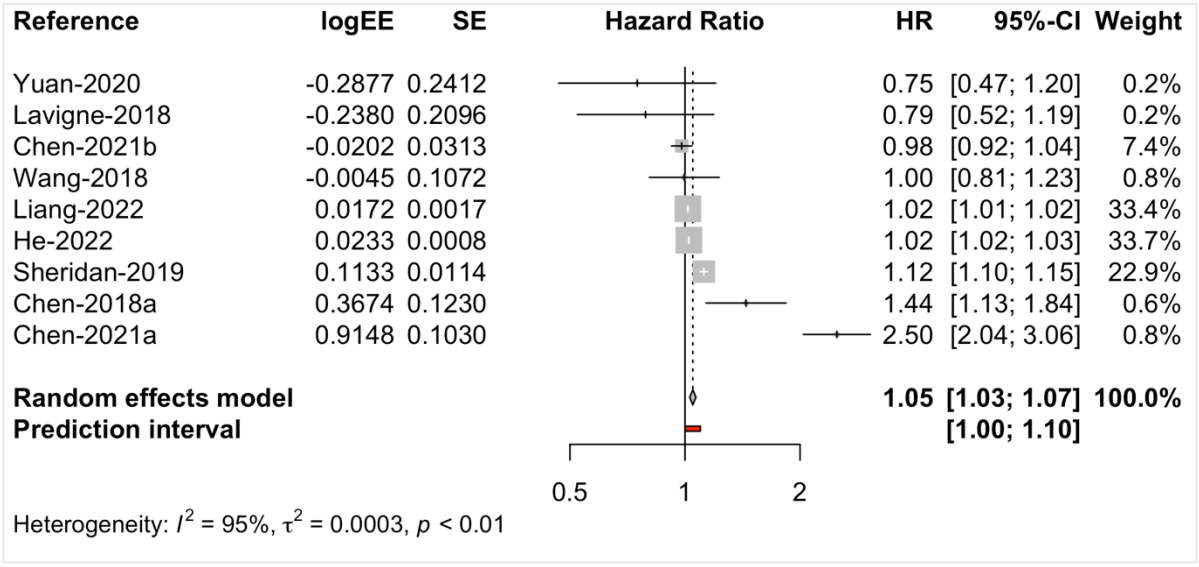


Figure 22. Forest plot for the meta-analysis of the association between whole pregnancy $PM_{2.5}$ exposure and the risk of PTB from studies using HRs. The individual effect sizes and 95%CIs, log of the effect sizes (logEE), standard error (SE), summary estimate (in diamond shape), and its 95%CI, prediction intervals (PI), and I^2 statistics are presented.

5.3.3. Effect of NO_2 exposure

5.3.3.1. Trimester 1 NO_2 exposure (OR)

There were 29 studies that gave estimates (n = 30) for this exposure-outcome pair. They were conducted in a variety of locations, including the US, China, Europe, Australia, Brazil, and Taiwan. 9 estimates were removed due to intolerable dependencies. 21 estimates were used in the meta-analysis.

There was no evidence of an association between trimester 1 NO_2 exposure and the risk of PTB (OR = 1.03, 95% CI: 0.96 - 1.09, per 10 μg/m^3 increase) (**Figure 23**). The conclusion is very robust with low possibility of the conclusion changing with the addition of future studies, due to the large number of studies included. High heterogeneity was observed (I^2 = 99%, τ^2 = 0.0186), with PI overlapping the null (0.76 - 1.38).

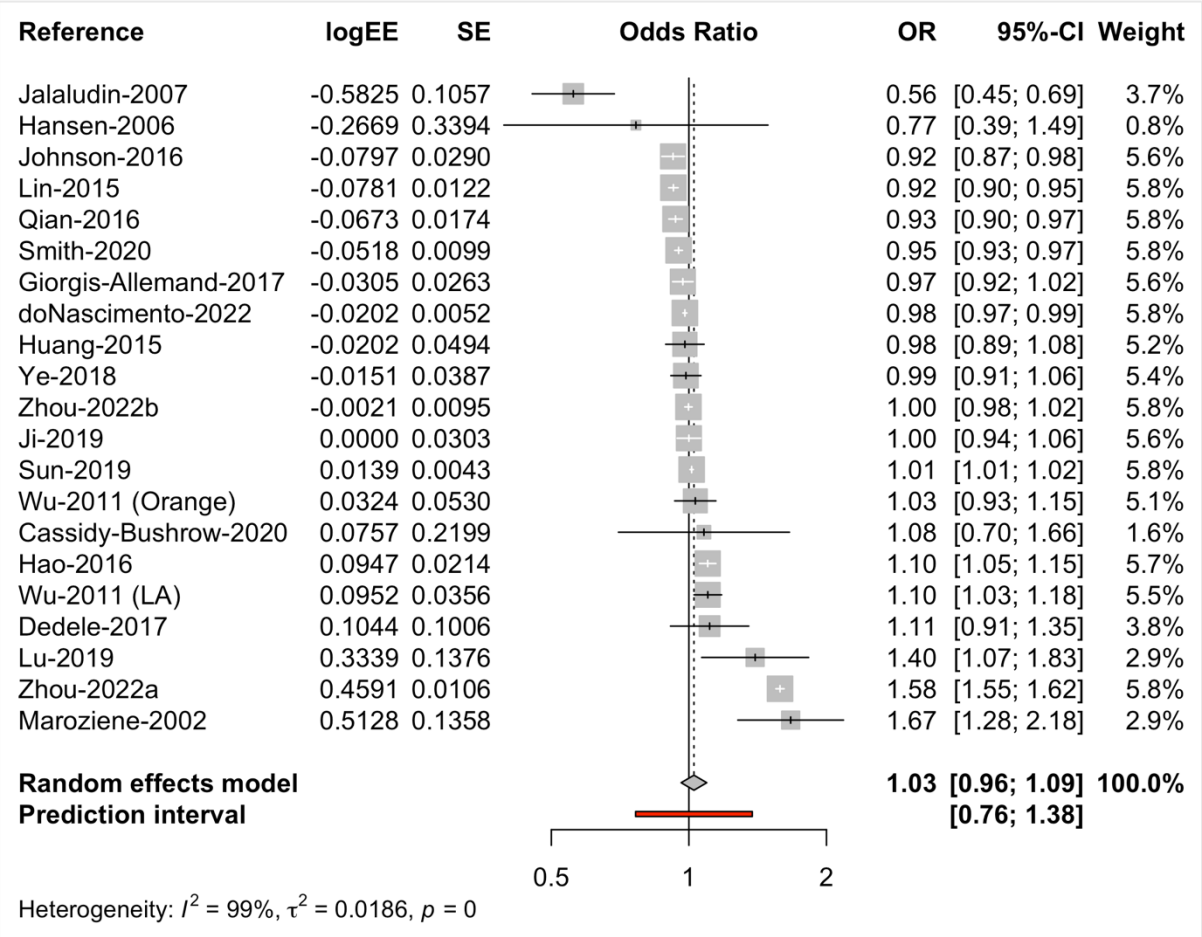


Figure 23. Forest plot for the meta-analysis of the association between trimester 1 NO_2 exposure and the risk of PTB from studies using ORs. The individual effect sizes and 95%CIs, log of the effect sizes (logEE), standard error (SE), summary estimate (in diamond shape), and its 95%CI, prediction intervals (PI), and I^2 statistics are presented.

5.3.3.2. Trimester 2 NO_2 exposure (OR)

There were 25 studies that gave estimates (n = 26) for this exposure-outcome pair. They were conducted in a variety of locations. 7 estimates were removed due to intolerable dependencies or the study being assigned a high RoB. 19 estimates (from 18 studies) were used in the meta-analysis.

There was no evidence of an association between trimester 2 NO_2 exposure and the risk of PTB (OR = 1.03, 95% CI: 0.96 - 1.11, per 10 μg/m^3 increase) (**Figure 24**). The conclusion is fairly robust, with some possibility of the conclusion changing with the addition of future studies. High heterogeneity was observed (I^2 = 99%, τ^2 = 0.0237), with PI overlapping the null (0.74 - 1.44).

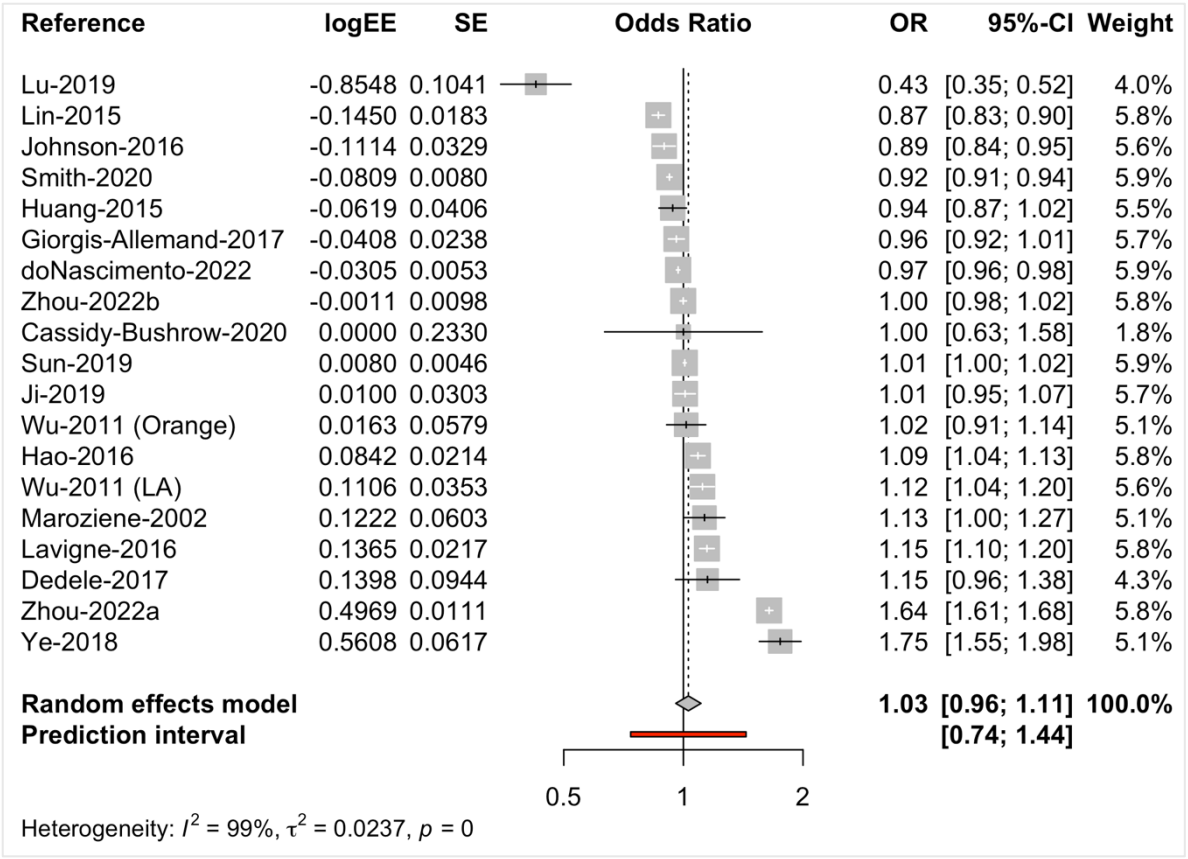


Figure 24. Forest plot for the meta-analysis of the association between trimester 2 NO_2 exposure and the risk of PTB from studies using ORs. The individual effect sizes and 95%CIs, log of the effect sizes (logEE), standard error (SE), summary estimate (in diamond shape), and its 95%CI, prediction intervals (PI), and I^2 statistics are presented.

5.3.3.3. Trimester 3 NO_2 exposure (OR)

There were 21 studies that gave estimates for this exposure-outcome pair. They were conducted in a variety of locations. 5 studies were removed due to intolerable dependencies or high RoB. 16 studies were used in the meta-analysis.

There was evidence of moderate increase in the risk of PTB associated with a per 10 μg/m^3 increase in trimester 3 NO_2 exposure (OR = 1.20, 95%CI: 1.07 - 1.34, per 10 μg/m^3 increase) (**Figure 25**), with low certainty in the magnitude of risk due to the wide CI. The conclusion is fairly robust, with some possibility of the conclusion changing with the addition of future studies. High heterogeneity was observed (I^2 = 99%, τ^2 = 0.0481), with PI overlapping the null (0.74 - 1.95). In the leave-one-out analysis, most primary studies showed minimal influence to the summary estimate except when Zhou et al. (2022a) or Fang et al. (2022) (OR: 1.13 [1.01 - 1.27]) was removed, where the OR decreased in strength to 1.13 (95%CI: 1.07 - 1.19) and 1.13 (95%CI: 1.01 - 1.27) respectively (**Appendix 18**).

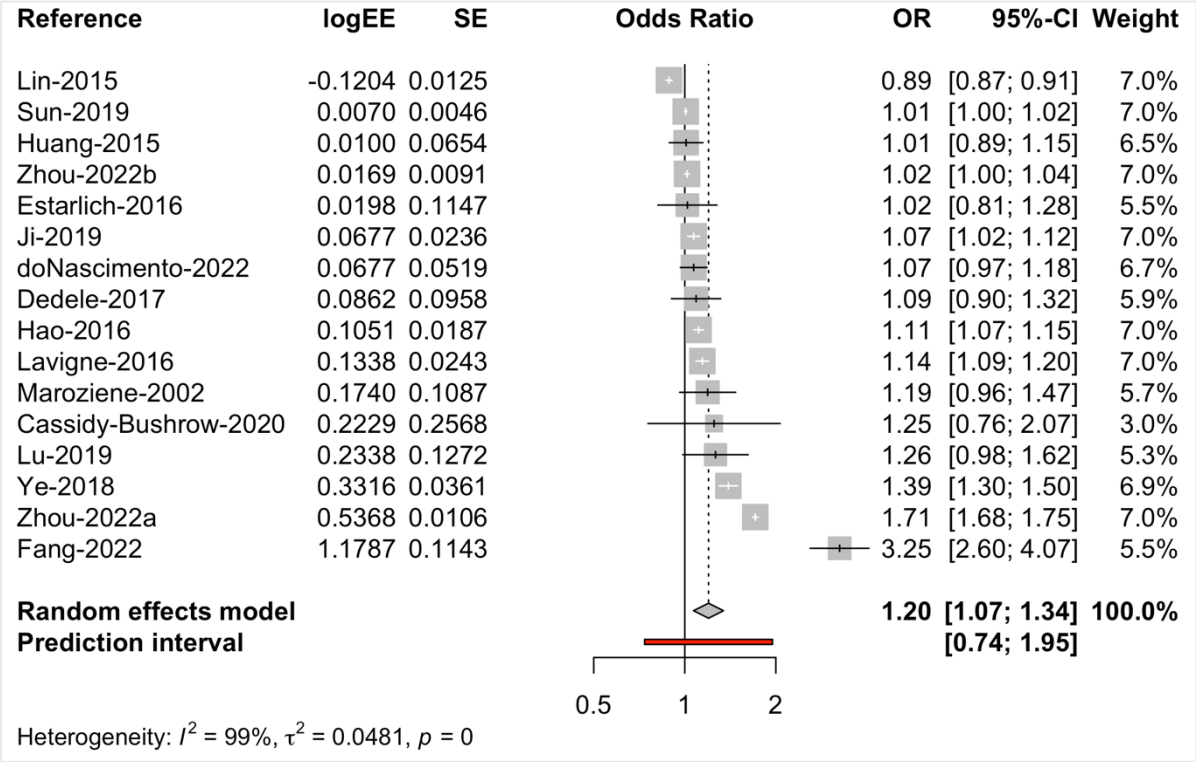


Figure 25. Forest plot for the meta-analysis of the association between trimester 3 NO_2 exposure and the risk of PTB from studies using ORs. The individual effect sizes and 95%CIs, log of the effect sizes (logEE), standard error (SE), summary estimate (in diamond shape), and its 95%CI, prediction intervals (PI), and I^2 statistics are presented.

5.3.3.4. Whole pregnancy NO_2 exposure (OR)

There were 33 studies that gave estimates (n = 34) for this exposure-outcome pair. They were conducted in a variety of locations. 10 studies were removed due to intolerable dependencies or being assigned a high RoB. 23 studies (24 estimates) were used in the main meta-analysis.

There was no evidence of an association between whole pregnancy NO_2 exposure and the risk of PTB (OR = 1.10, 95%CI: 0.99 - 1.23, per 10 μg/m^3 increase) (**Figure 26**), with low certainty in the magnitude of risk due to the wide CI. The conclusion is very robust with low possibility of the conclusion changing with the addition of future studies, due to the large number of studies included. High heterogeneity was observed (I^2 = 99%, τ^2 = 0.0612), with PI overlapping the null (0.65 - 1.86).

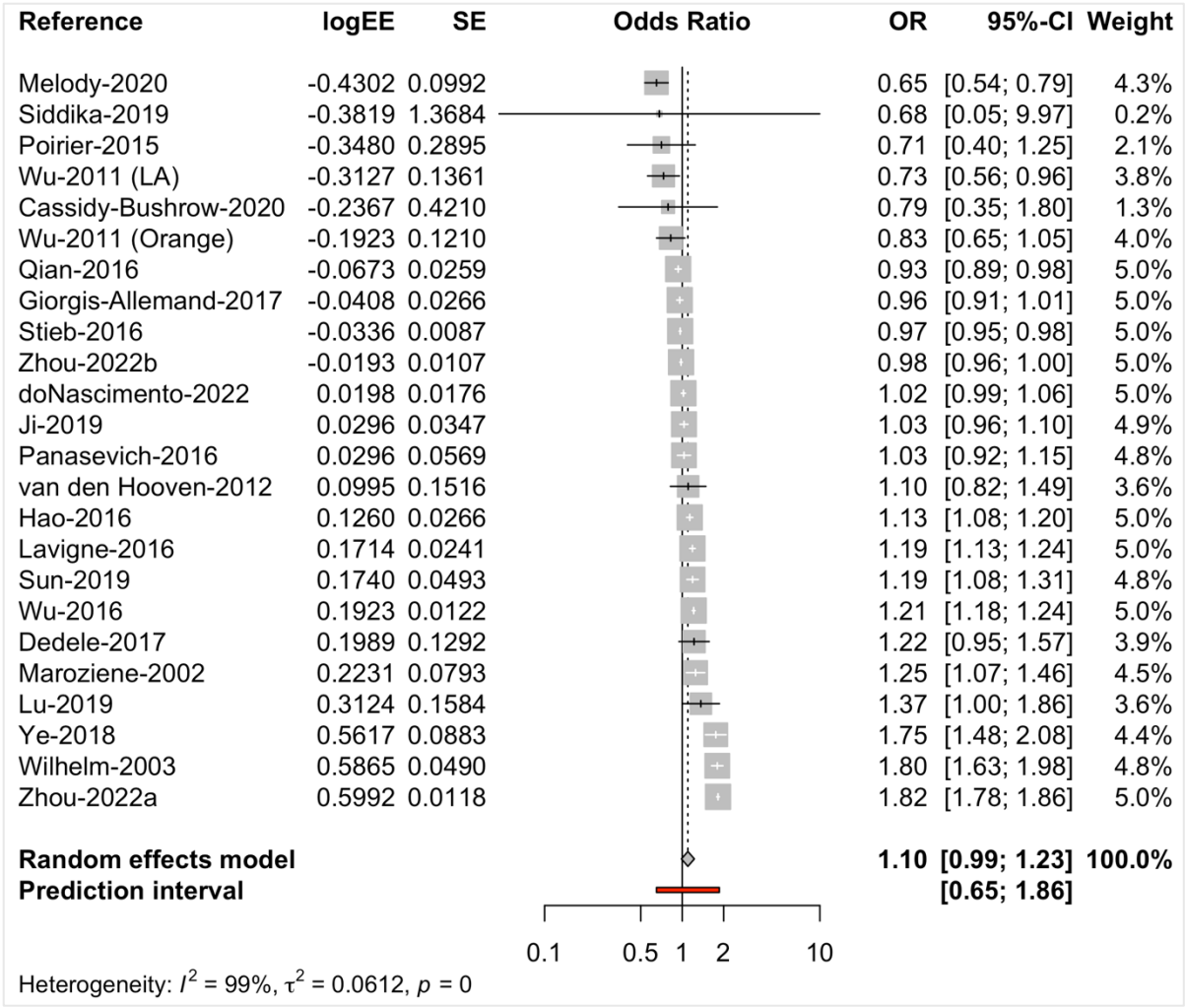


Figure 26. Forest plot for the meta-analysis of the association between whole pregnancy NO_2 exposure and the risk of PTB from studies using ORs. The individual effect sizes and 95%CIs, log of the effect sizes (logEE), standard error (SE), summary estimate (in diamond shape), and its 95%CI, prediction intervals (PI), and I^2 statistics are presented.

5.3.3.5. Trimester 1 NO_2 exposure (HR)

There were 5 studies that gave estimates for this exposure-outcome pair. Apart from 1 that took place in Australia, all the other studies were conducted in China. After removing 1 study with intolerable dependencies, 4 estimates remained and were available to be meta-analysed. No studies were assigned a high risk of bias.

There was evidence of a strong increase in the risk of PTB associated with a per 10 μg/m^3 increase in trimester 1 NO_2 exposure (HR = 1.62, 95%CI: 1.03 - 2.55, per 10 μg/m^3 increase) (**Figure 27**), with low certainty in the magnitude of risk due to the very wide CI. High heterogeneity was observed (I^2 = 98%, τ^2 = 0.1897), with PI overlapping the null (0.19 - 13.54). The conclusion is not particularly robust, with high possibility of the conclusion

changing with the addition of future studies, due to low number of studies included. The wide 95% CI was driven by Chen et al. (2021b) which found a very strong increase in risk.

In the leave-one-out analysis, the statistically significant increase in risk was no longer present when Chen et al. (2021b) (OR: 1.89 [0.47 - 7.56]) or Liang et al. (2019b) (OR: 1.88 [0.46 - 7.66]) were removed. The magnitude of the increase in risk became much smaller when Chen et al. (2021a) was removed (HR: 1.04, 95% CI: 1.03 - 1.04), indicating some sensitivity to these studies.

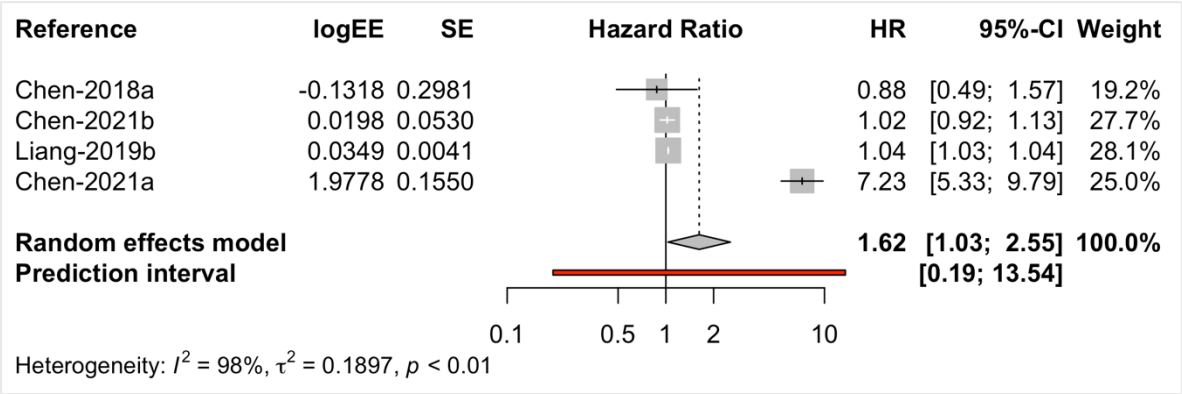


Figure 27. Forest plot for the meta-analysis of the association between trimester 1 NO_2 exposure and the risk of PTB from studies using HRs. The individual effect sizes and 95%CIs, log of the effect sizes (logEE), standard error (SE), summary estimate (in diamond shape), and its 95%CI, prediction intervals (PI), and I^2 statistics are presented.

5.3.3.6. Trimester 2 NO_2 exposure (HR)

There were 5 studies that gave estimates for this exposure-outcome pair. Apart from 1 that took place in Australia, all the other studies were conducted in China. After removing 1 study with intolerable dependencies, 4 studies were available to be meta-analysed. No studies were assigned a high risk of bias.

There was no evidence of an association between trimester 2 NO_2 exposure and the risk of PTB, with low certainty in the magnitude of risk due to the very wide CI (HR = 0.91, 95% CI: 0.71 - 1.17, per 10 μg/m^3 increase) (**Figure 28**). High heterogeneity was observed (I^2 = 94%, τ^2 = 0.0496), with PI overlapping the null (0.30 - 2.74). The conclusion is not particularly robust, with high possibility of the conclusion changing with the addition of future studies, due to low number of studies included.

Leave-one-out analysis revealed that removal of any single primary study had minimal impact on the summary estimate, indicating the robustness of the pooled results.

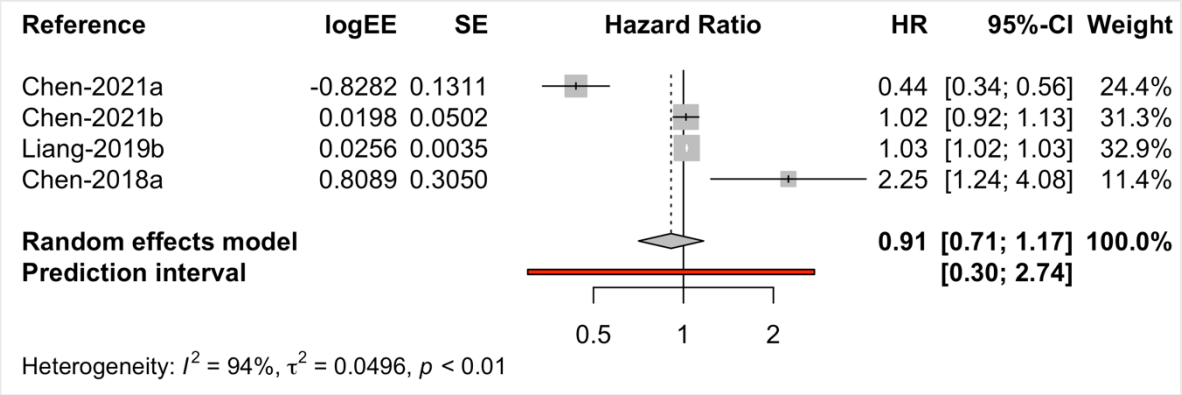


Figure 28. Forest plot for the meta-analysis of the association between trimester 2 NO_2 exposure and the risk of PTB from studies using HRs. The individual effect sizes and 95%CIs, log of the effect sizes (logEE), standard error (SE), summary estimate (in diamond shape), and its 95%CI, prediction intervals (PI), and I^2 statistics are presented.

5.3.3.7. Trimester 3 NO_2 exposure (HR)

There were 6 studies that gave estimates for this exposure-outcome pair. Apart from 1 that took place in Australia, all the other studies were conducted in China. After removing 1 study with intolerable dependencies, 5 studies remained in the main meta-analysis. No studies were assigned a high risk of bias.

There was evidence of a strong increase in the risk of PTB associated with trimester 3 NO_2 exposure (HR = 1.57, 95% CI: 1.10 - 2.23, per 10 μg/m^3 increase) (**Figure 29**), with low certainty in the magnitude of the risk due to the wide CI. The conclusion is not particularly robust, with high possibility of the conclusion changing with the addition of future studies, due to low number of studies included. High heterogeneity was observed (I^2 = 97%, τ^2 = 0.01409), with PI overlapping the null (0.42 - 5.90). The wide 95% CI was driven by Chen et al. (2018a) and Fang et al. (2022) which found very strong increases in risk.

In the leave-one-out analysis, the statistically significant increase in risk persisted when Chen et al. (2018a) was removed, but not when each of the other primary studies were removed (**Appendix 18**).

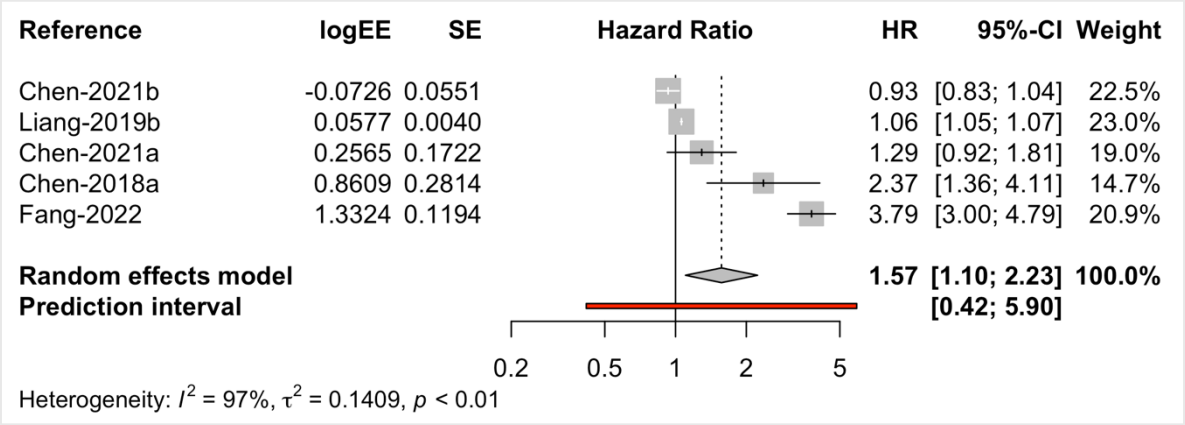


Figure 29. Forest plot for the meta-analysis of the association between trimester 3 NO_2 exposure and the risk of PTB from studies using HRs. The individual effect sizes and 95%CIs, log of the effect sizes (logEE), standard error (SE), summary estimate (in diamond shape), and its 95%CI, prediction intervals (PI), and I^2 statistics are presented.

5.3.3.8. Whole pregnancy NO_2 exposure (HR)

There were 7 studies gave estimates (n = 8) for this exposure-outcome pair. They were conducted in a variety of location. After removing 1 study with intolerable dependencies, 6 studies (7 estimates) were used in the main meta-analysis. No studies were assigned a high risk of bias.

There was no evidence of an association between whole pregnancy NO_2 exposure and the risk of PTB (HR = 1.00, 95% CI: 0.99 -1.01, per 10 μg/m^3 increase) (**Figure 30**), with high certainty in the magnitude of risk due to the narrow CI. High heterogeneity was observed (I^2 = 89%, τ^2 = 0.0001), with PI overlapping the null (0.97 - 1.04). The conclusion is not particularly robust, with high possibility of the conclusion changing with the addition of future studies, due to low number of studies included. The finding could be driven by the four studies reporting null findings as they dominated by weight in the meta-analysis.

Leave-one-out analysis showed minimal influence of the primary studies to the summary estimate, indicating the robustness of the pooled results.

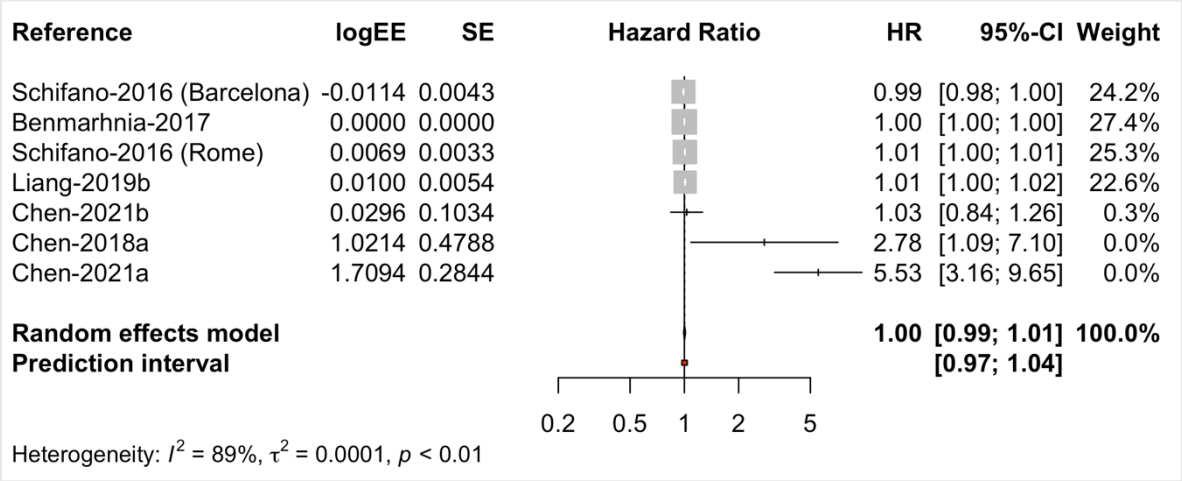


Figure 30. Forest plot for the meta-analysis of the association between whole pregnancy NO_2 exposure and the risk of PTB from studies using HRs. The individual effect sizes and 95%CIs, log of the effect sizes (logEE), standard error (SE), summary estimate (in diamond shape), and its 95%CI, prediction intervals (PI), and I^2 statistics are presented.

5.3.4. Effect of O_3 exposure

The following meta-analyses were based on studies that have used 8-hour maximum or 8-hour average (10 am to 6 pm) O_3 levels, as there were insufficient studies available in all exposure windows of interest for other measurement metrics (**Appendix 10**).

5.3.4.1. Trimester 1 O_3 exposure (OR)

There were 13 studies that gave estimates (n = 14) for this exposure-outcome pair. They took place in a variety of locations. 3 studies were removed due to intolerable dependencies or being assigned a high risk of bias. Hence, 10 studies (11 estimates) were used for the meta-analysis.

There was no evidence of an association between trimester 1 O_3 exposure and the risk of PTB (OR = 1.04, 95% CI: 1.00 - 1.07, per 10 μg/m^3 increase) (**Figure 31**). High heterogeneity was observed (I^2 = 97%, τ^2 = 0.0023), with PI overlapping the null (0.93 - 1.16). The conclusion is fairly robust, with some possibility of the conclusion changing with the addition of future studies.

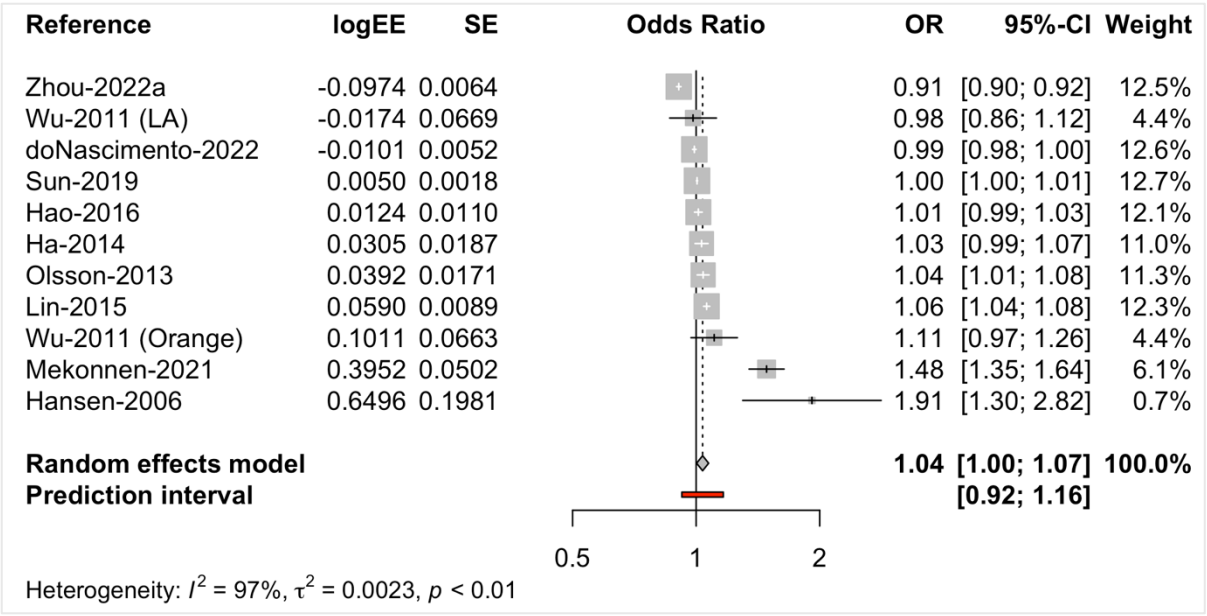


Figure 31. Forest plot for the meta-analysis of the association between trimester 1 O_3 exposure and the risk of PTB from studies using ORs. The individual effect sizes and 95%CIs, log of the effect sizes (logEE), standard error (SE), summary estimate (in diamond shape), and its 95%CI, prediction intervals (PI), and I^2 statistics are presented.

5.3.4.2. Trimester 2 O_3 exposure (OR)

There were 12 studies that gave estimates (n = 13) for this exposure-outcome pair. They took place in a variety of locations. 3 studies were removed due to intolerable dependencies or being assigned a high RoB. Hence, 9 studies (10 estimates) were used in the meta-analysis.

There was no evidence of an association between trimester 2 O_3 exposure and the risk of PTB (OR = 1.01, 95%CI: 0.98 - 1.05, per 10 μg/m^3 increase) (**Figure 32**). High heterogeneity was observed (I^2 = 99%, τ^2 = 0.0028), with PI overlapping the null (0.89 - 1.16). The conclusion is fairly robust, with some possibility of the conclusion changing with the addition of future studies

In the leave-one-out analysis, most primary studies had a minimal impact on the summary estimate except Zhou et al. (2022a), where an increase in risk was found when it was removed (OR: 1.04 [1.01 - 1.07]), indicating some sensitivity of the results to the study.

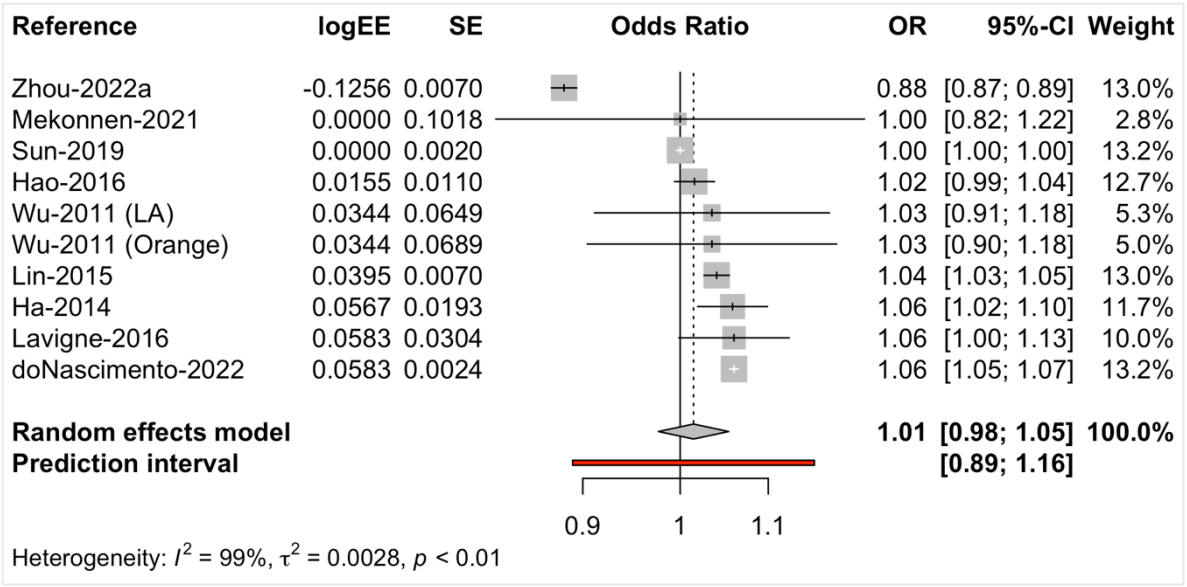


Figure 32. Forest plot for the meta-analysis of the association between trimester 2 O_3 exposure and the risk of PTB from studies using ORs. The individual effect sizes and 95%CIs, log of the effect sizes (logEE), standard error (SE), summary estimate (in diamond shape), and its 95%CI, prediction intervals (PI), and I^2 statistics are presented.

5.3.4.3. Trimester 3 O_3 exposure (OR)

There were 12 studies that gave estimates for this exposure-outcome relationship. They were conducted in a variety of locations. 3 studies were removed due to intolerable dependencies or high RoB. 9 studies were used in the meta-analysis.

There was no evidence of an association between trimester 3 O_3 exposure and the risk of PTB (OR = 1.02, 95% CI: 0.94 - 1.10, per 10 μg/m^3 increase) (**Figure 33**). High heterogeneity was observed (I^2 = 99%, τ^2 = 0.0143), with PI overlapping null (0.75 - 1.37). There is high possibility of the conclusion changing with the addition of future studies, due to the low number of studies included

Leave-one-out analysis showed minimal influence of the primary studies to the summary estimate, indicating robustness of the pooled results.

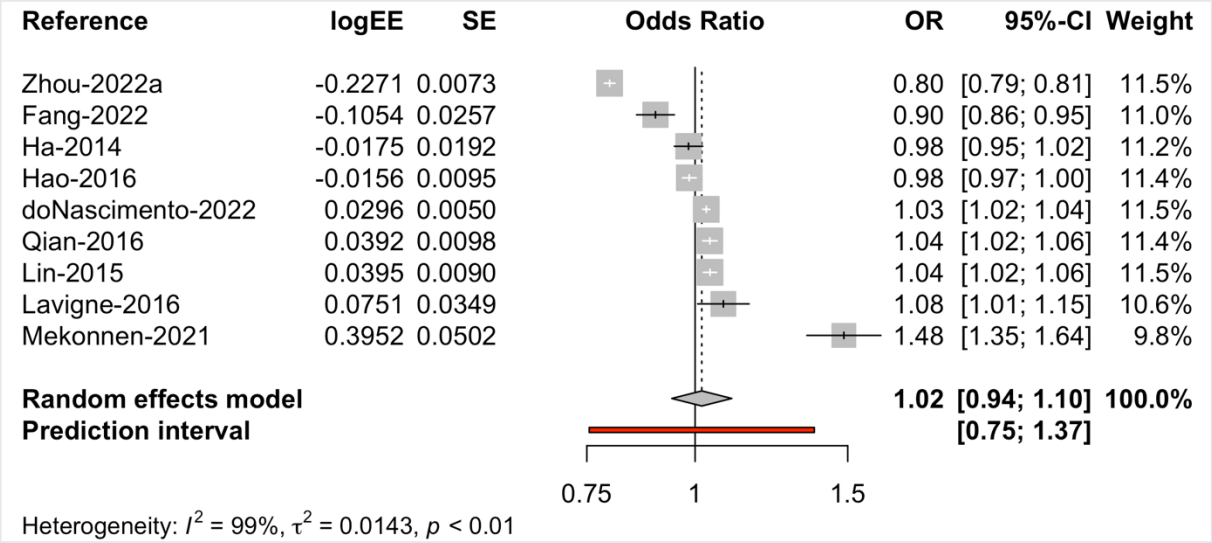


Figure 33. Forest plot for the meta-analysis of the association between trimester 3 O_3 exposure and the risk of PTB from studies using ORs. The individual effect sizes and 95%CIs, log of the effect sizes (logEE), standard error (SE), summary estimate (in diamond shape), and its 95%CI, prediction intervals (PI), and I^2 statistics are presented.

5.3.4.4. Whole pregnancy O_3 exposure (OR)

There were 13 studies that gave estimates (n = 14) for this exposure-outcome relationship. The studies took place in various locations. 3 studies with intolerable dependencies or high RoB were removed. Hence, 10 studies (11 estimates) were used in the main meta-analysis.

There was no evidence of an association between whole pregnancy O_3 exposure and the risk of PTB (OR = 1.04, 95%CI: 0.93 - 1.17, per 10 µg/m^3 increase) (**Figure 34**). High heterogeneity was observed (I^2 = 99%, τ^2 = 0.0351), with PI overlapping the null (0.67 - 1.62). This conclusion is fairly robust, with some possibility of the conclusion changing with the addition of future studies.

Leave-one-out analysis showed minimal influence of most primary studies except for doNascimento et al. (2022), where an increase in risk was found when it was removed (OR: 1.09 [1.05 - 1.13]).

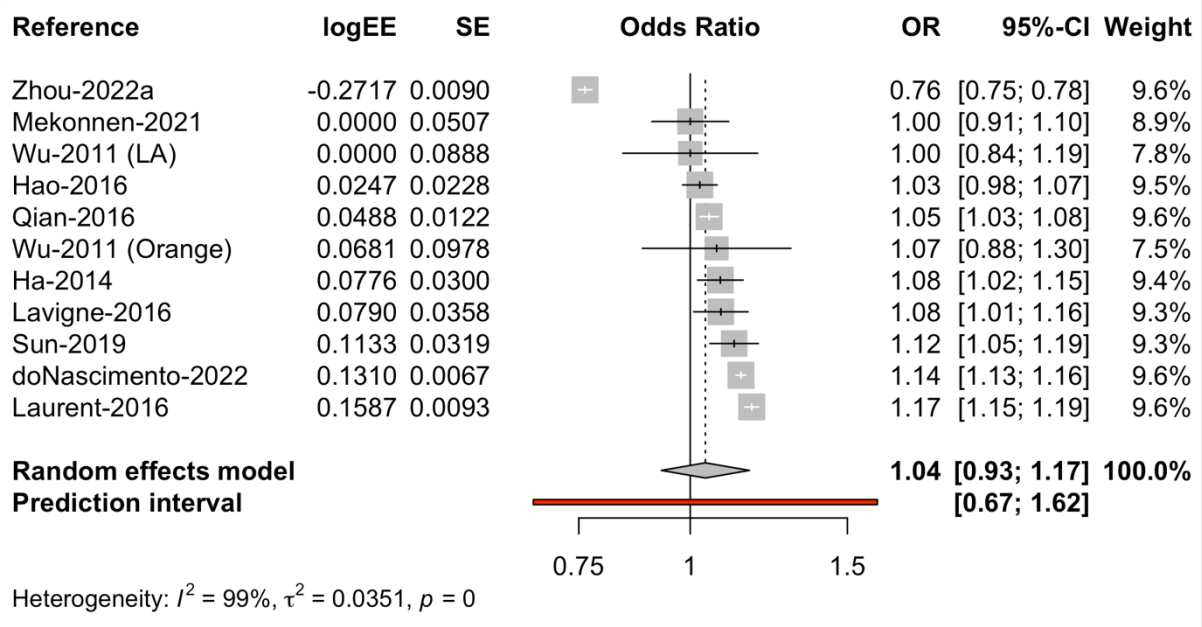


Figure 34. Forest plot for the meta-analysis of the association between whole pregnancy O_3 exposure and the risk of PTB from studies using ORs. The individual effect sizes and 95%CIs, log of the effect sizes (logEE), standard error (SE), summary estimate (in diamond shape), and its 95%CI, prediction intervals (PI), and I^2 statistics are presented.

5.3.4.5. Trimester 1 O_3 exposure (HR)

There were 3 studies that gave estimates for this exposure-outcome relationship. Two of them were conducted in China, and the remaining one was conducted in Australia. No study was excluded due to intolerable dependencies or high risk of bias.

There was no evidence of an association between trimester 1 O_3 exposure and the risk of PTB (HR = 0.76, 95% CI: 0.37 - 1.58, per 10 μg/m^3 increase) (**Figure 35**). The conclusion is not particularly robust. There is a high possibility of the conclusion changing with the addition of future studies, due to low number of studies included. High heterogeneity was observed (I^2 = 99%, τ^2 = 0.4069), with the PI being extremely wide, which could be because the 95% CIs of the three studies included in the meta-analysis do not overlap at all, and they indicate a strong decrease, small increase, and large increase in risk respectively.

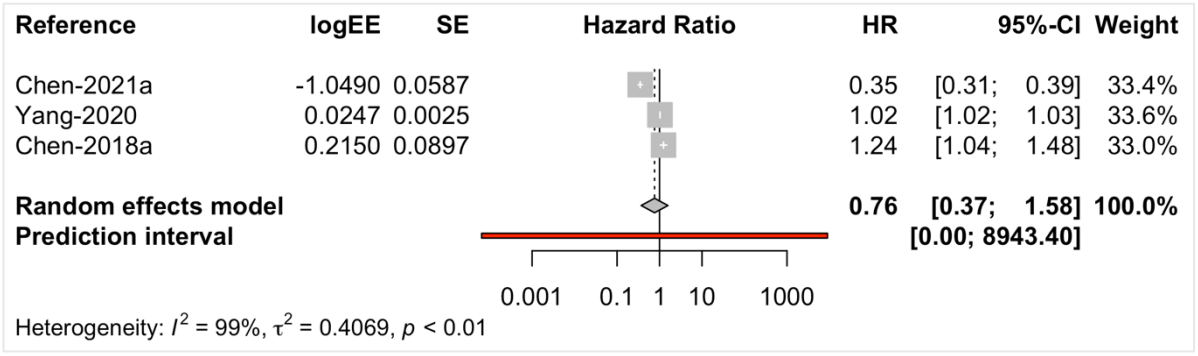


Figure 35. Forest plot for the meta-analysis of the association between trimester 1 O_3 exposure and the risk of PTB from studies using HRs. The individual effect sizes and 95%CIs, log of the effect sizes (logEE), standard error (SE), summary estimate (in diamond shape), and its 95%CI, prediction intervals (PI), and I^2 statistics are presented. HK: Knapp-Hartung adjustment

5.3.4.6. Trimester 2 O_3 exposure (HR)

There were 3 studies that gave estimates for this exposure-outcome relationship. Two of them were conducted in China, and the remaining one was conducted in Australia. No study was excluded due to intolerable dependencies or high risk of bias.

There was no evidence of an association between trimester 2 O_3 exposure and the risk of PTB (HR = 1.26, 95% CI: 0.97 - 1.63, per 10 μg/m^3 increase) (**Figure 36**). High heterogeneity was observed (I^2 = 93%, τ^2 = 0.0473), with the PI being very wide (0.05 - 31.77). Despite all 3 studies showing a statistically significant increase in risk, there was one study with much smaller increase in risk compared to the other two. This indicates that a large uncertainty remains for this exposure-outcome relationship, which likely resulted in a summary estimate overlapping the null, wide PI as well as high heterogeneity shown by I^2. Therefore, the conclusion is not particularly robust. There is high possibility of the conclusion changing with the addition of future studies, due to low number of studies included.

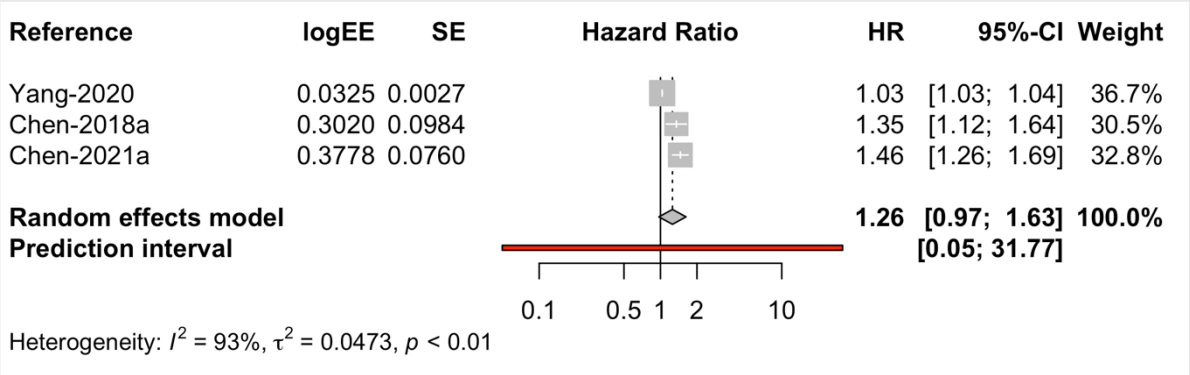


Figure 36. Forest plot for the meta-analysis of the association between trimester 2 O_3 exposure and the risk of PTB from studies using HRs. The individual effect sizes and 95%CIs, log of the effect sizes (logEE), standard error (SE), summary estimate (in diamond shape), and its 95%CI, prediction intervals (PI), and I^2 statistics are presented.

5.3.4.7. Trimester 3 O_3 exposure (HR)

There were 3 studies that gave estimates for this exposure-outcome pair. Two of them took place in China, and one took place in Australia. No studies were removed due to intolerable dependencies or high risk of bias. Hence, 3 studies were available for the meta-analysis.

There was no evidence of an association between trimester 3 O_3 exposure and the risk of PTB (HR = 1.19, 95% CI: 0.63 - 2.27, per 10 μg/m^3 increase) (**Figure 37**). High heterogeneity was observed (I^2 = 99%, τ^2 = 0.3202), with PI overlapping the null (0.00 - 4889.83). The conclusion is not particularly robust. There is high possibility of the conclusion changing with the addition of future studies, due to low number of studies included.

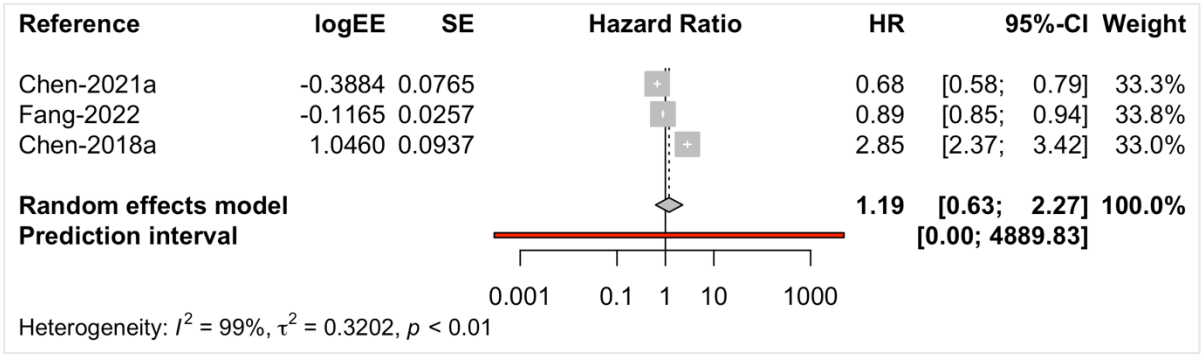


Figure 37. Forest plot for the meta-analysis of the association between trimester 3 O_3 exposure and the risk of PTB from studies using HRs. The individual effect sizes and 95%CIs, log of the effect sizes (logEE), standard error (SE), summary estimate (in diamond shape), and its 95%CI, prediction intervals (PI), and I^2 statistics are presented.

5.3.4.8. Whole pregnancy O_3 exposure (HR)

Four studies gave estimates for this exposure-outcome pair. They took place in China, Australia, Spain, and Italy. No study was removed due to study dependencies. There were no studies with a high risk of bias.

There was no evidence of an association between whole pregnancy O_3 exposure and the risk of PTB (OR = 1.01, 95% CI: 0.97 - 1.05, per 10 µg/m^3 increase) (**Figure 38**). High heterogeneity was observed (I^2 = 98%, τ^2 = 0.0009), with a wide PI overlapping the null (0.86 - 1.18). The two estimates from Schifano et al. (2016) for Rome and Barcelona constituted more than 95% of the total weight in the meta-analysis. The conclusion is not particularly robust. There is a high possibility of the conclusion changing with the addition of future studies, due to low number of studies included.

Leave-one-out analysis showed minimal influence of the primary studies to the summary estimate, indicating the robustness of the pooled results. However, the 95% CI became much wider when the one of the two estimates from Schifano et al. (2016) was removed (OR: 1.06 [0.33 - 3.41]).

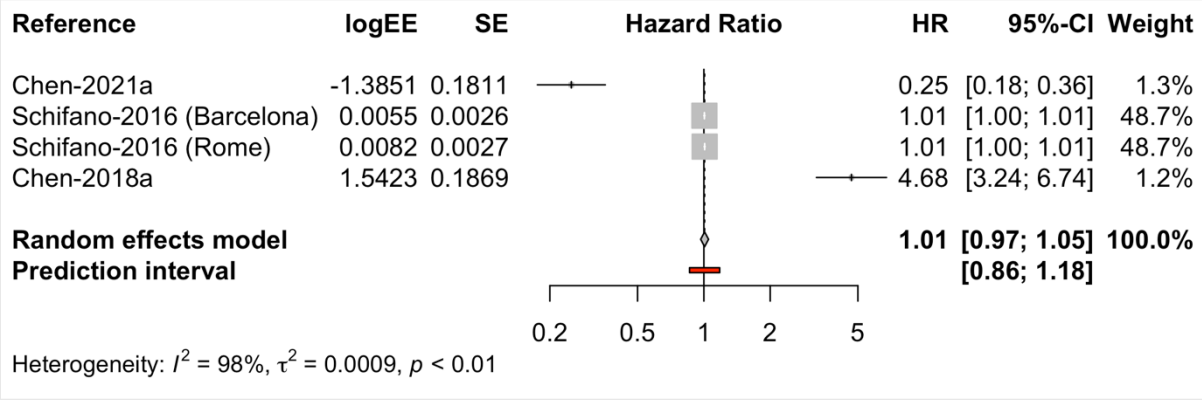


Figure 38. Forest plot for the meta-analysis of the association between whole pregnancy O_3 exposure and the risk of PTB from studies using HRs. The individual effect sizes and 95%CIs, log of the effect sizes (logEE), standard error (SE), summary estimate (in diamond shape), and its 95%CI, prediction intervals (PI), and I^2 statistics are presented.

5.3.5. Summary of the effects of $PM_{2.5}$, NO_2, O_3 exposures on the risk of PTB

The results of the main meta-analyses conducted are shown in **Figure 39**. For $PM_{2.5}$ exposure, an increase in risk was found for trimester 3 and whole pregnancy in the main meta-analyses among studies that used ORs, and the findings persisted in studies that used HRs. For NO_2 exposure, an increase in risk was found for trimester 3 exposure for studies using ORs and HRs. In addition, an increase in risk was found for trimester 1 exposure, but only among studies that used HRs. O_3 exposure was not associated with the risk of PTB in any exposure period for studies that used ORs or HRs.

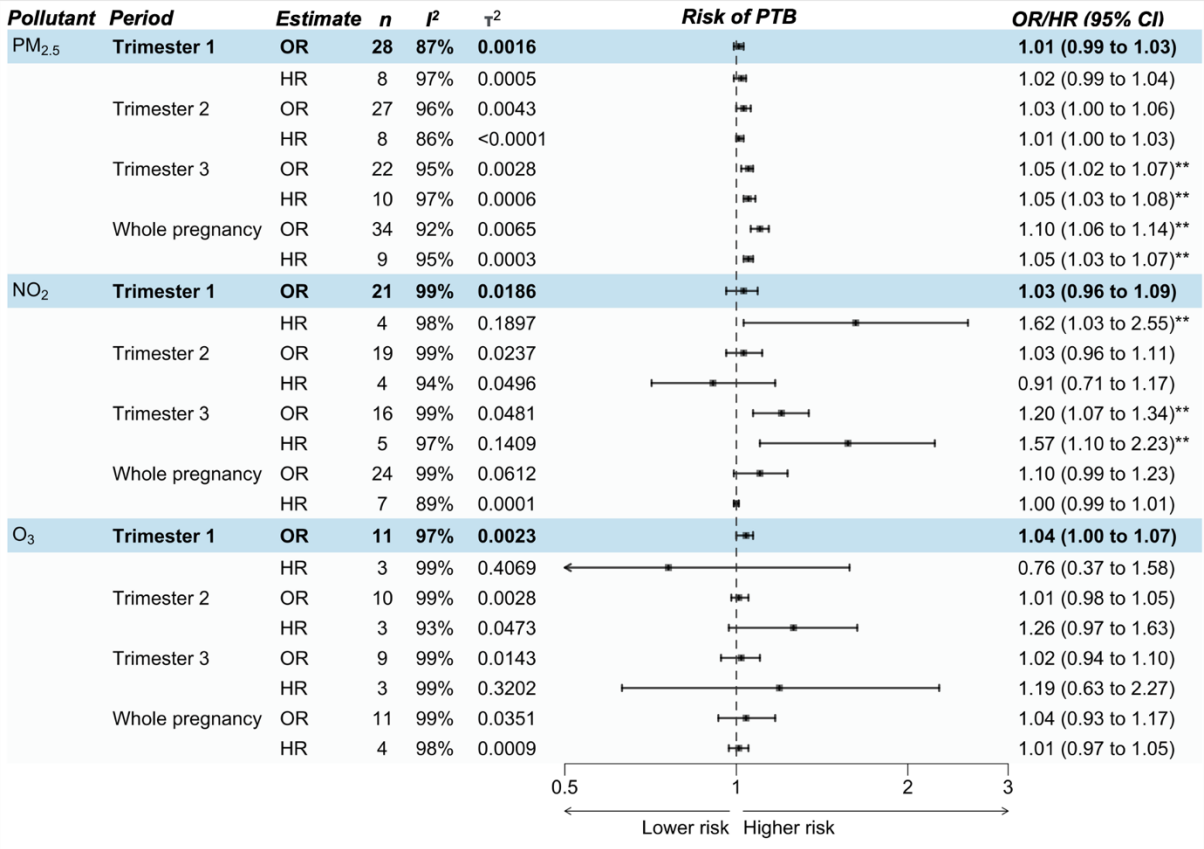


Figure 39. Forest plot for the summary estimates of the main meta-analyses. The summary effect sizes (per 10 μg/m³ increase in exposure) and 95%CIs, I^2 statistics, and number of estimates (n) are presented.

5.3.6. The effects of $PM_{2.5}$, NO_2, O_3 exposures on PTB subgroups

Subgroups of PTBs were moderate-to-late PTB (mlPTB, 33 to 36 gestational weeks), very PTB (vPTB, 28 to 32 gestational weeks), and extremely PTB (exPTB, <28 gestational weeks). Meta-analyses were not conducted for the effect of O_3 exposure due to there being fewer than 3 primary studies. For NO_2 exposure, meta-analyses were only conducted for

vPTB for HR studies, as it was the only exposure-outcome pair with enough studies available. Summaries of the results for $PM_{2.5}$ and NO_2 exposure are shown in **Figure 40** and **Figure 41**. The individual forest plots are available in **Appendix 15 to 17**.

For $PM_{2.5}$ exposure, the statistically significant increase in risk associated with whole pregnancy exposure for studies reporting ORs and HRs was strengthened for PTB subgroups. However, the statistically significant increase in risk associated with trimester 3 exposure only persisted for mlPTB among studies that reported ORs. The risk dissipated for other exposure-outcome pairs. In trimester 1 and 2, there were no associations found between $PM_{2.5}$ exposure and PTB. However, an increase in risk was found for trimester 1 exposure and vPTB among studies that reported HRs. Similarly, an increase in risk of vPTB was associated with trimester 2 exposure for studies using ORs, as well as for mlPTB for studies that used HRs and ORs.

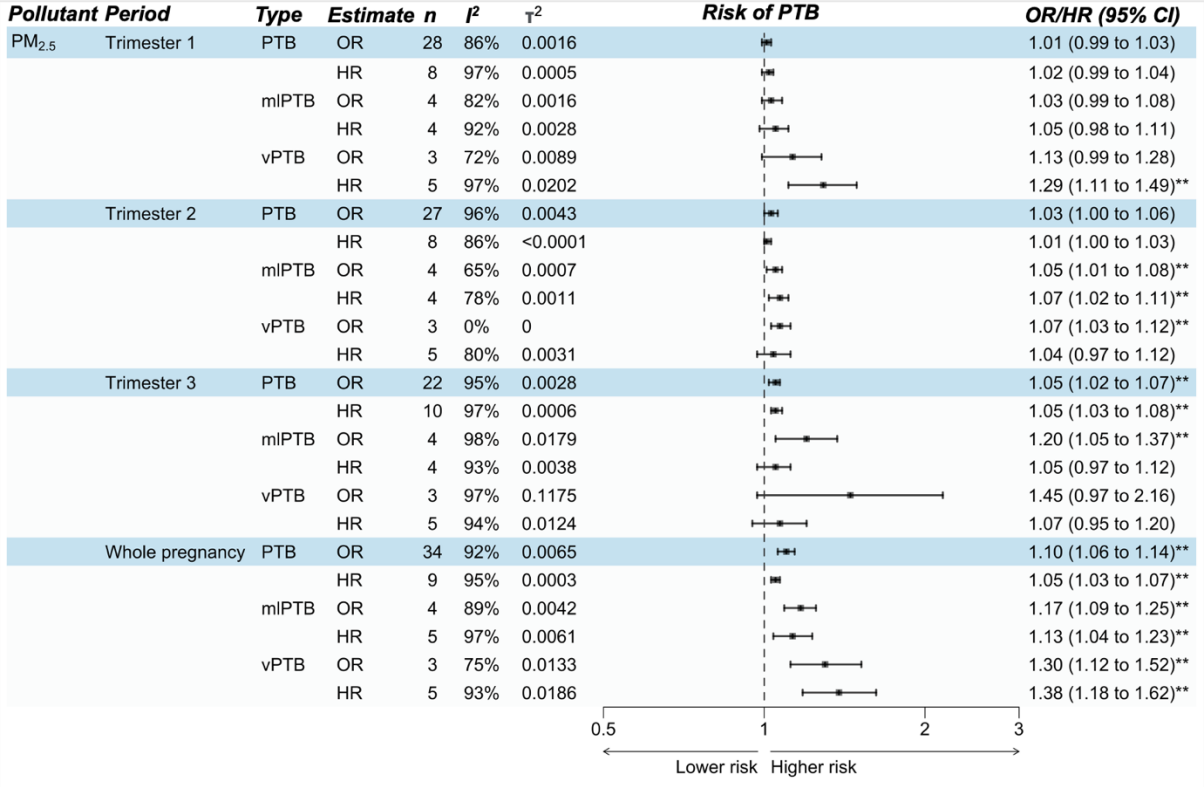


Figure 40. Forest plot for the summary estimates of the meta-analyses conducted for the association between $PM_{2.5}$ exposure at trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP) and PTB, mlPTB, and vPTB. The summary effect sizes (per 10 µg/m³ increase in exposure) and 95%CIs, I^2 statistics, and number of estimates (n) are presented.

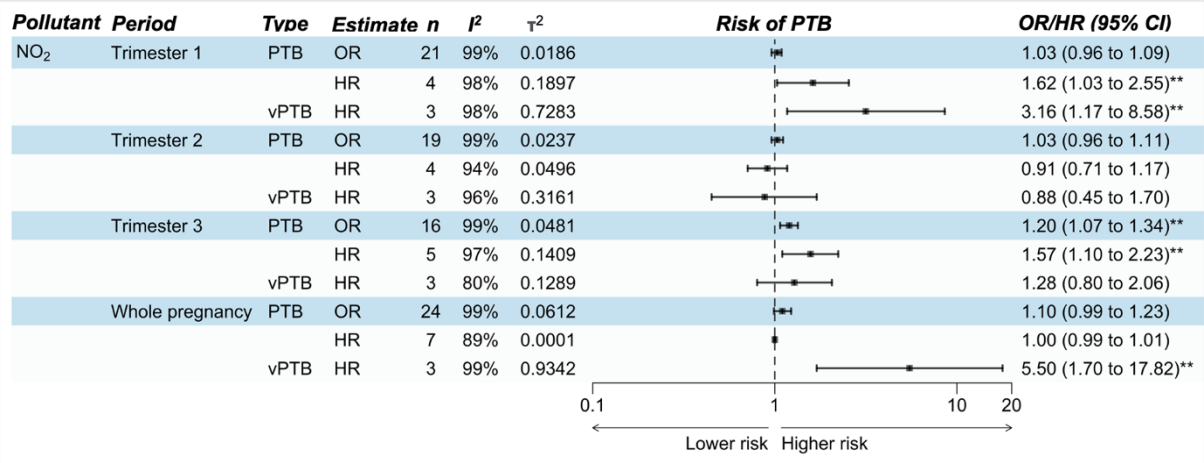


Figure 41. Forest plot for the summary estimates of the meta-analyses conducted for the association between NO_2 exposure at trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP) and PTB and vPTB. The summary effect sizes (per 10 µg/m³ increase in exposure) and 95%CIs, I^2 statistics, and number of estimates (n) are presented.

For NO_2 exposure, the increase in risk found for trimester 1 exposure (HR analysis) was strengthened for vPTB. However, the statistically significant increase in risk found for trimester 3 exposure dissipated in the vPTB analysis of HR studies. No association was found for the effect of whole pregnancy exposure for PTB, but a strong increase in risk of vPTB was found in HR studies.

5.3.7. Investigating the drivers of heterogeneity

5.3.7.1. Risk of bias

5.3.7.1.1. Stratified analysis by risk of bias for $PM_{2.5}$ exposure (ORs)

A summary of the results is shown in **Figure 42**, with p-values obtained from the Q-test shown in **Table 25**. No significant subgroup differences were found between risk of bias levels or extent of confounder adjustment. Including studies of a high risk of bias did not notably change the main results. RoB did not drive the high heterogeneity observed in the main meta-analysis (without studies assigned a high risk of bias). However, meta-analyses of trimester 1 and 2 exposure using only studies assigned a high risk of bias have resulted in a strong decrease in heterogeneity (I^2 = 0%).

The increase in risk of trimester 3 exposure dissipated among studies assigned a low risk of bias (OR: 0.99 [0.92 - 1.06]). There was no evidence suggesting that confounder adjustment level was a driver of the heterogeneity observed, as limited reduction in heterogeneity were present in all strata except two.

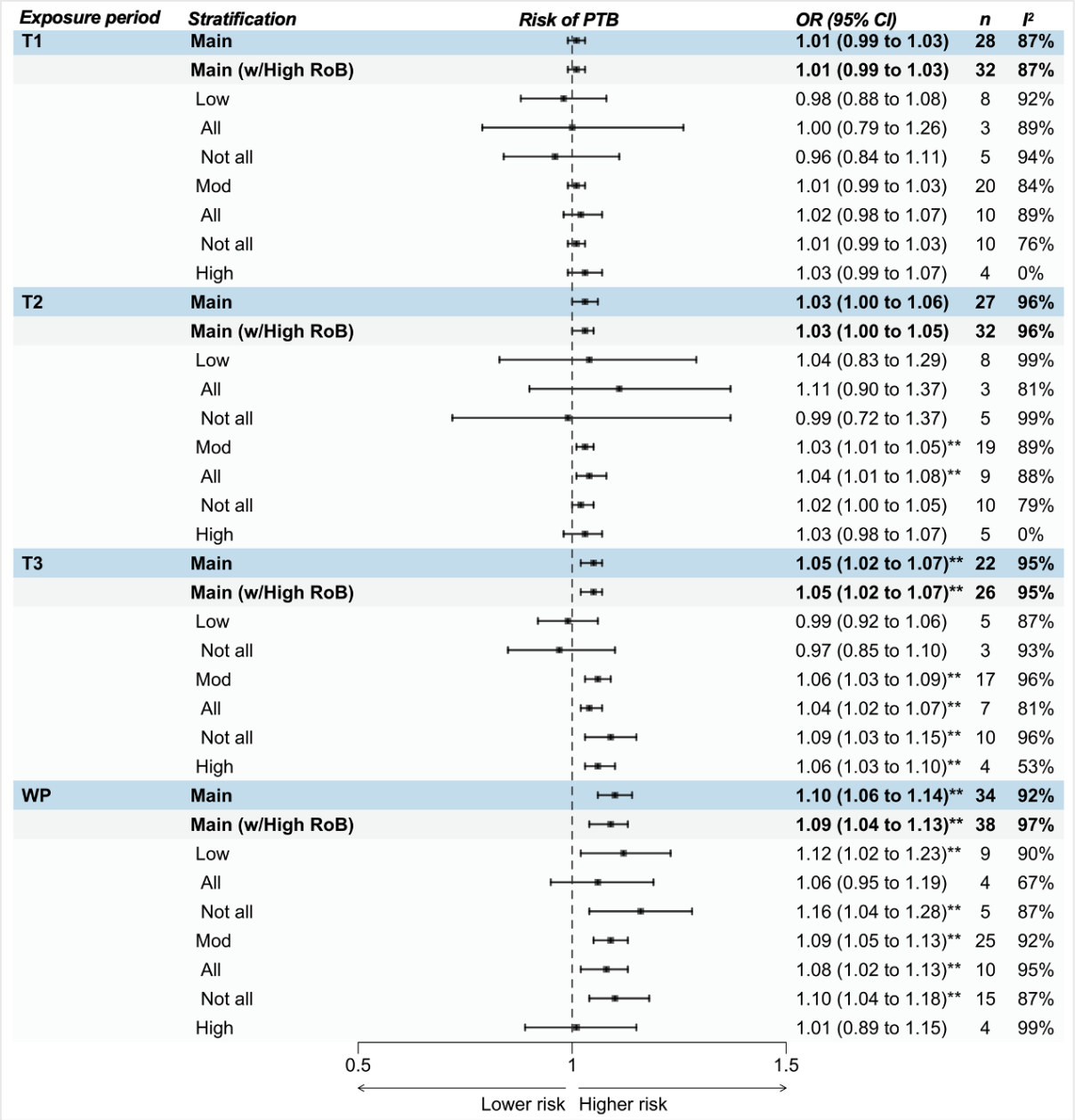


Figure 42. Forest plot for the summary estimates of the meta-analyses conducted for the association between $PM_{2.5}$ exposure at trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP), and the risk of PTB from studies using ORs, stratified by the risk of bias (RoB) levels of the studies used (Low, Moderate, High), and a further stratification by the extent of confounder adjustment (All, Not All). The summary effect sizes (per 10 μg/m^3 increase in exposure) and 95%CIs, I^2 statistics, and number of estimates (n) are presented. The forest plots which show the individual studies included in the sensitivity analyses are available in Appendix 22.

Table 25. P-values obtained from Q-test of subgroup differences for stratification by risk of bias and confounder adjustment for meta-analyses of studies of $PM_{2.5}$ exposure reporting ORs.

Stratification	T1	T2	T3	WP
Risk of bias	0.27	0.84	0.05	0.50
Level of confounder adjustment – studies of low RoB	0.76	0.64	0.58	0.34
Level of confounder adjustment – studies of moderate RoB	0.80	0.26	0.20	0.53

5.3.7.1.2. Stratified analysis by risk of bias for $PM_{2.5}$ exposure (HRs)

A summary of the results is shown in **Figure 43**, with p-values obtained from the Q-test shown in **Table 26**. No significant subgroup differences were found. The results suggested that risk of bias or confounder adjustment were not drivers of the high heterogeneity observed, as no decreases in I^2 were observed. However, there is low certainty in the finding as there were only enough studies to be meta-analysed for the strata of moderate risk of bias, and those that did not adjust for all critical confounders among studies assigned a moderate risk of bias.

Table 26. P-values obtained from Q-test of subgroup differences for stratification by risk of bias and confounder adjustment for meta-analyses of studies of $PM_{2.5}$ exposure reporting HRs.

Stratification	T1	T2	T3	WP
Risk of bias	0.84	0.76	0.33	0.30
Level of confounder adjustment – studies of moderate RoB	0.13	0.56	0.31	0.09

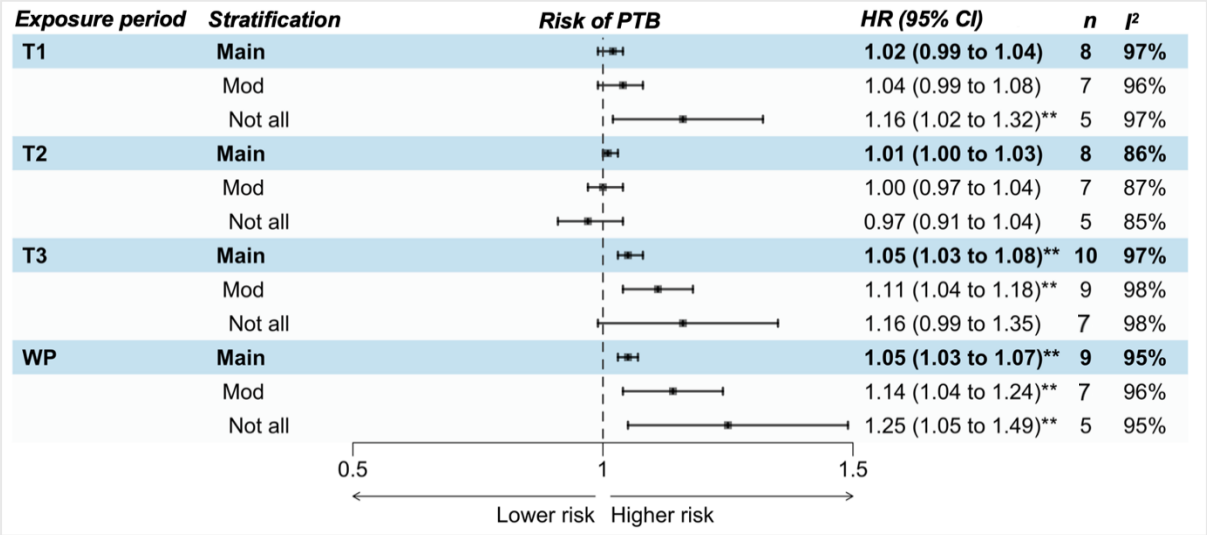


Figure 43. Forest plot for the summary estimates of the meta-analyses conducted for the association between $PM_{2.5}$ exposure at trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP), and the risk of PTB from studies using HRs, stratified by the risk of bias (RoB) levels (Low, Moderate, High) of the studies used, and a further stratification by the extent of confounder adjustment (All, Not all). The summary effect sizes (per 10 µg/m³ increase in exposure) and 95%CIs, I^2 statistics, and number of estimates (n) are presented. The forest plots which show the individual studies included in the sensitivity analyses are available in Appendix 22. No studies were assigned a high risk of bias.

5.3.7.1.3. Stratified analysis by risk of bias for NO_2 exposure (ORs)

A summary of the results is shown in **Figure 44**, with p-values obtained from the Q-test shown in **Figure 27**. The results suggested that inclusion of studies assigned a high RoB did not notably change the results, as there was minimal impact to the summary estimates of the main result, and limited decreases in I^2 were observed.

The RoB level was suggested to be a partial driver of the heterogeneity observed. Meta-analyses using only studies assigned a high RoB in all exposure windows have resulted in a strong decrease in heterogeneity (I^2 = 0%). In trimester 1 and 2, studies assigned a low RoB showed a decrease in risk which was distinct from that of moderate or high RoB, with 95% CIs minimally overlapped with that of studies of moderate as well as high RoB, and the main meta-analyses which that found no effect. Strong decrease in heterogeneity was also found for trimester 1 and whole pregnancy exposure (I^2 decreased <40%) for studies of low RoB. No significant subgroup differences were found by the Q-test.

The level of confounder adjustment was suggested to be a partial driver of the heterogeneity observed. Further decrease in heterogeneity shown by I^2 was found for low RoB studies without adjustment for all critical confounders, but this decrease was not found for studies of

moderate RoB, despite a significant subgroup difference being found for trimester 2 for confounder adjustment among studies assigned a moderate RoB.

Table 27. P-values obtained from Q-test of subgroup differences for stratification by risk of bias and confounder adjustment for meta-analyses of studies of NO_2 exposure reporting ORs. *: P-value <0.05. /: 3 or less studies available for the stratum.

Stratification	T1	T2	T3	WP
Risk of bias	0.19	0.11	0.11	0.40
Level of Confounder adjustment – studies of Low RoB	0.06	0.06	/	0.75
Level of confounder adjustment – studies Moderate RoB	0.83	<0.01*	0.32	0.19

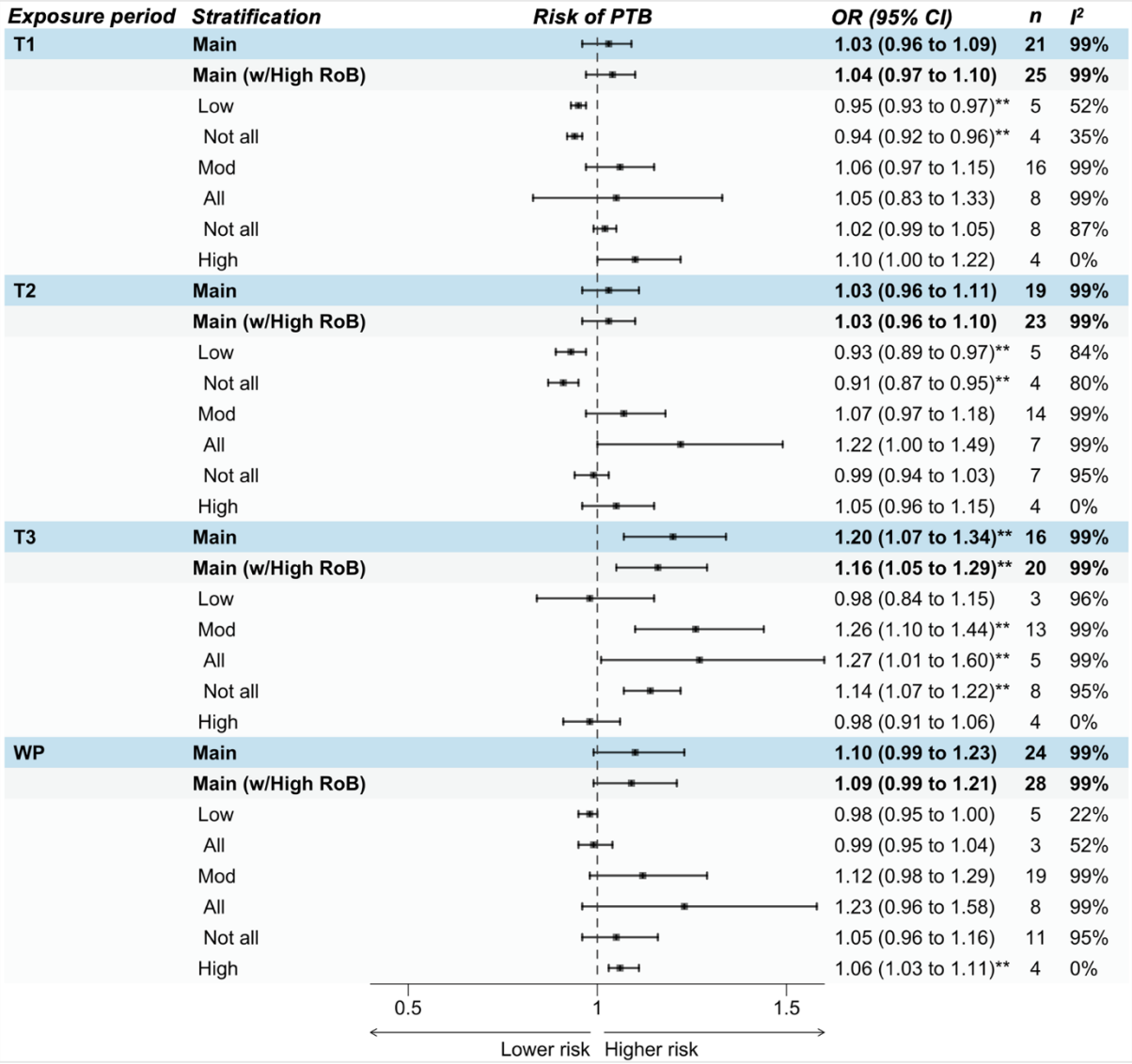


Figure 44. Forest plot for the summary estimates of the meta-analyses conducted for the association between NO_2 exposure at trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP), and the risk of PTB from studies using ORs, stratified by the risk of bias (RoB) levels (Low, Moderate, High) of the studies used, and a further stratification by the extent of confounder adjustment (All, Not all). The summary effect sizes (per 10 µg/m^3 increase in exposure) and 95%CIs, I^2 statistics, and number of estimates (n) are presented. The forest plots which show the individual studies included in the sensitivity analyses are available in Appendix 22.

5.3.7.1.4. Stratified analysis by risk of bias for NO_2 exposure (HRs)

All studies were assigned a moderate risk of bias and did not adjust for all critical confounders. Hence a stratified analysis was not conducted.

5.3.7.1.5. Stratified analysis by risk of bias for O_3 exposure (ORs)

A summary of the results is shown in **Figure 45**, with p-values obtained from the Q-test shown in **Table 28**. Inclusion of studies assigned a high risk of bias (RoB) did not produce notable changes to the results. RoB or levels of confounder adjustment were not drivers of the high heterogeneity observed, as minimal impact to the summary estimates and I^2 were observed. The only exception was for the meta-analysis of studies assigned a low RoB for trimester 2 exposure, where an increase in risk and a strong decrease in heterogeneity was observed (I^2 = 0%), suggesting that risk of bias level can partially explain the heterogeneity in that exposure window but not the others. Compared to the null findings in the main meta-analyses, an increase in risk was also found for studies of low RoB for trimester 1 exposure (OR: 1.16 [1.04 - 1.29]), and studies of moderate RoB without adjustment for all critical confounders (OR: 1.12 [1.07 - 1.17]).

Table 28. P-values obtained from Q-test of subgroup differences for stratification by risk of bias and confounder adjustment for meta-analyses of studies of O_3 exposure reporting ORs. *: P-value <0.05.

Stratification	T1	T2	T3	WP
Risk of bias	<0.01*	0.39	0.08	0.99
Level of confounder adjustment – studies assigned a moderate RoB	0.24	0.15	0.92	0.15

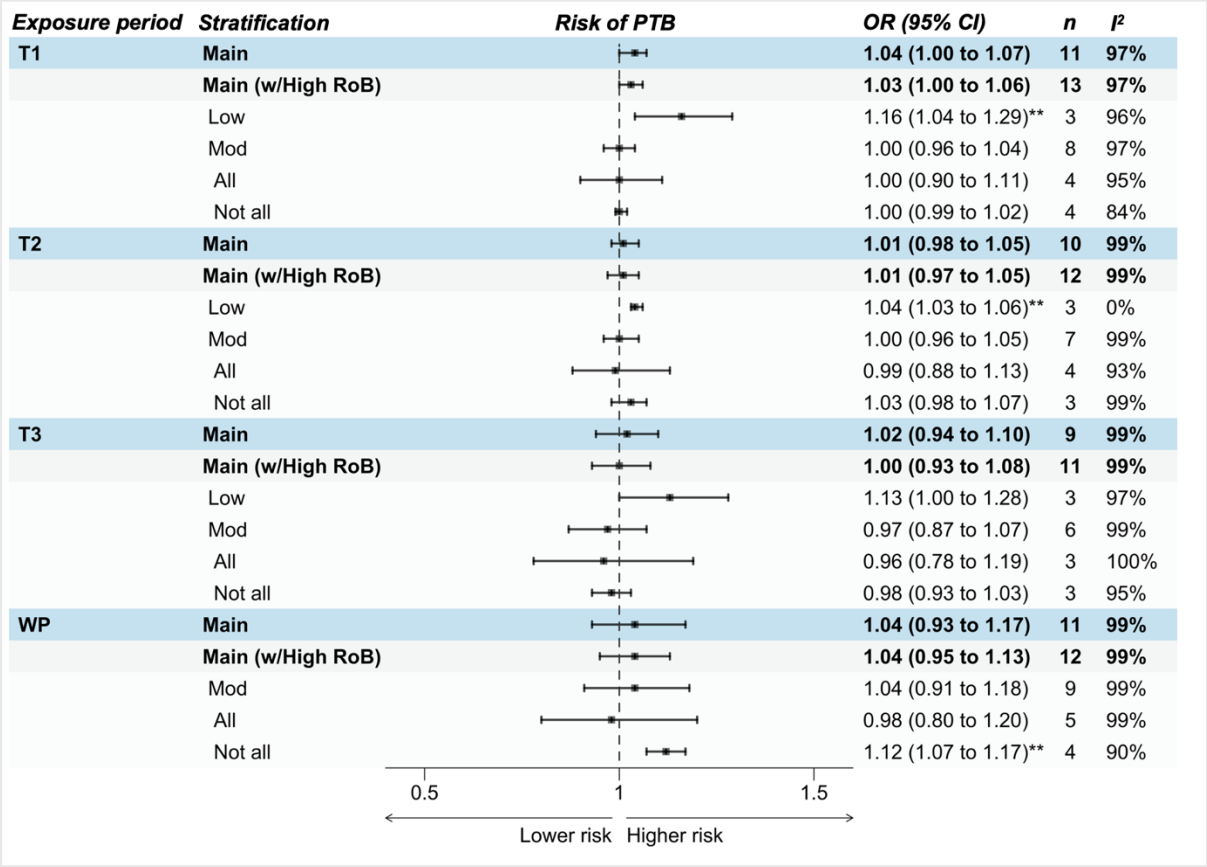


Figure 45. Forest plot for the summary estimates of the meta-analyses conducted for the association between O_3 exposure at trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP), and the risk of PTB from studies using ORs, stratified by the risk of bias (RoB) levels (Low, Moderate, High) of the studies used, and a further stratification by the extent of confounder adjustment (All, Not all). The summary effect sizes (per 10 µg/m^3 increase in exposure) and 95%CIs, I^2 statistics, and number of estimates (n) are presented. The forest plots which show the individual studies included in the sensitivity analyses are available in Appendix 22.

5.3.7.1.6. Stratified analysis by risk of bias for O_3 exposure (HRs)

All studies were assigned a moderate risk of bias and did not adjust for all critical confounders. Hence a stratified analysis was not conducted.

5.3.7.2. Location

As most studies reporting HRs are from China, country-level analysis was not conducted for countries apart from China for studies reporting HRs.

5.3.7.2.1. Stratified analysis by location for $PM_{2.5}$ exposure (ORs)

A summary of the results is shown in **Figure 46**, with p-values obtained from the Q-test shown in **Table 29**. The findings from the main results were robust to the stratification conducted in most cases. The results suggested that climate region could be a partial driver of the high heterogeneity observed in US studies, and WHO regions for the high heterogeneity observed in the main meta-analysis. There was no evidence suggesting that country or climate region (among Chinese studies) drove the high heterogeneity observed.

A moderate to strong reduction in heterogeneity shown by I^2 was observed in Europe and the western region of the US. The statistical significance of the increase in risk for whole pregnancy exposure dissipated when only European studies were used, despite the summary estimate being of similar magnitude, which could be due to the meta-analysis being based on fewer studies (n = 4) resulting in wider CIs. Some significant subgroup differences were detected.

For the western part of the US, an increase in risk was found for trimester 2 exposure which contrasted the null findings in the main meta-analysis, showing a relatively precise range of effect despite being based on few studies (OR: 1.10 [1.03 - 1.16], n = 3). The increase in risk found for whole pregnancy exposure in the main meta-analysis was no longer present when the analysis was restricted to studies in the western region of the US.

Table 29. P-values obtained from Q-test of subgroup differences for stratification by location for studies reporting ORs for the association between $PM_{2.5}$ exposure and the risk of PTB. *: P-value <0.05

Stratification	T1	T2	T3	WP
WHO region	<0.01*	<0.01*	0.02*	0.18
Country	<0.01*	<0.01*	0.11	0.26
US climate regions	0.21	0.23	0.02*	0.04*
China climate regions	0.71	0.75	<0.01*	0.58

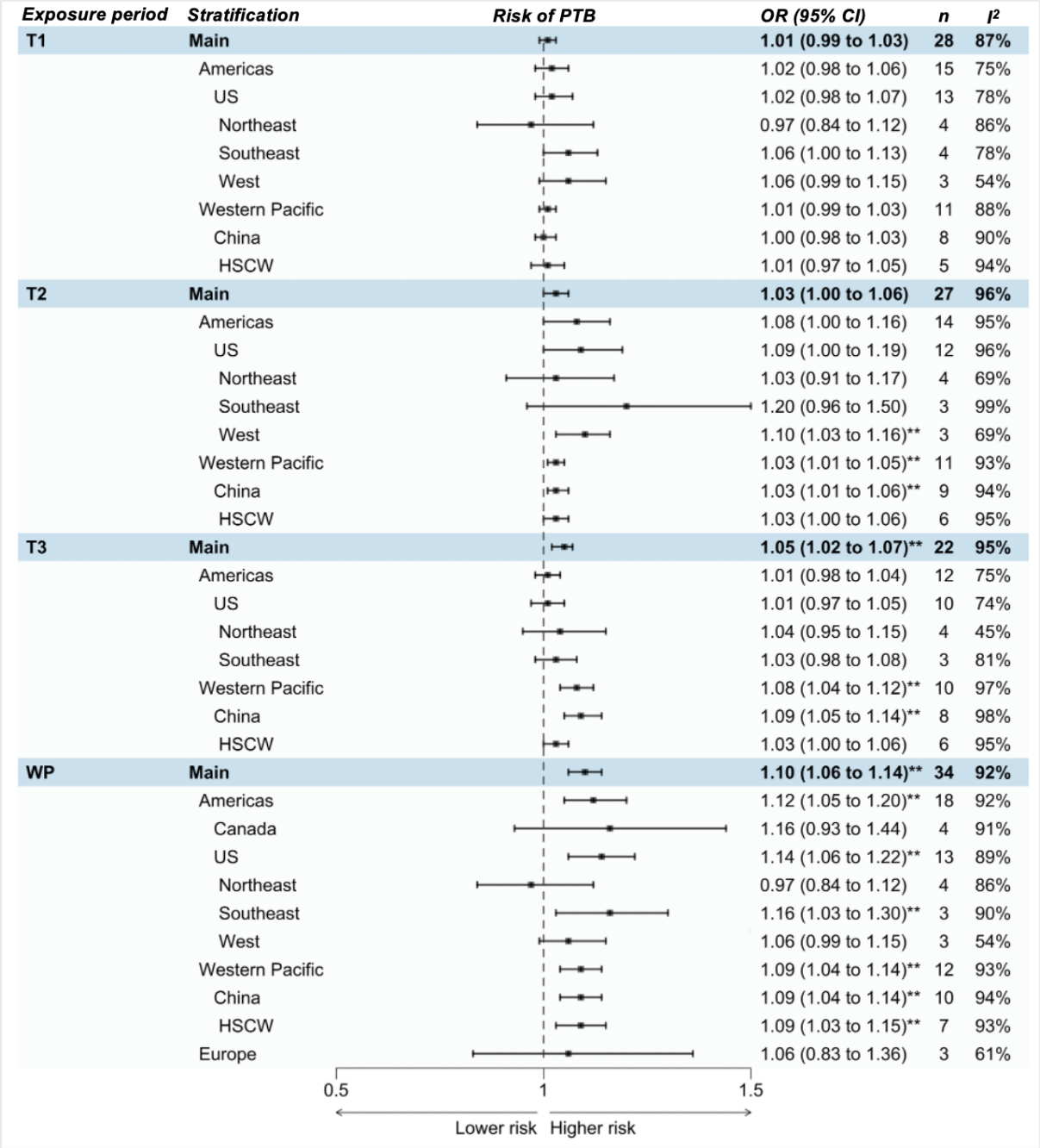


Figure 46. Forest plot for the summary estimates of the meta-analyses conducted for the association between $PM_{2.5}$ exposure at trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP), and the risk of PTB from studies using ORs, stratified in the levels of WHO region (Americas, Europe, Western Pacific), country (China, US, Canada), and climatic regions within the China and the US. The summary effect sizes (per 10 μg/m³ increase in exposure) and 95%CIs, I^2 statistics, and number of estimates (n) are presented. The forest plots which show the individual studies included in the sensitivity analyses are available in Appendix 22.

5.3.7.2.2. *Stratified analysis by location for $PM_{2.5}$ exposure (HRs)*

A summary of this sensitivity analysis is shown in **Figure 47**, with p-values obtained from Q-test shown in **Table 30**. The findings from the main results were robust to the stratification conducted in most cases. The results suggested that climate region was a driver of the heterogeneity found among Chinese studies. The summary estimates remained relatively stable when the analysis used only studies from the WHO region of Western Pacific or China across all exposure periods. A strong decrease in heterogeneity was found when only studies from the hot summer region in China was used except for trimester 3, where the increase in risk found in the main meta-analysis dissipated. Significant subgroup differences were found from the Q-test for Chinese climate regions. However, the number of studies used was low and no other climate regions had enough studies to corroborate the findings, indicating large uncertainties in the findings.

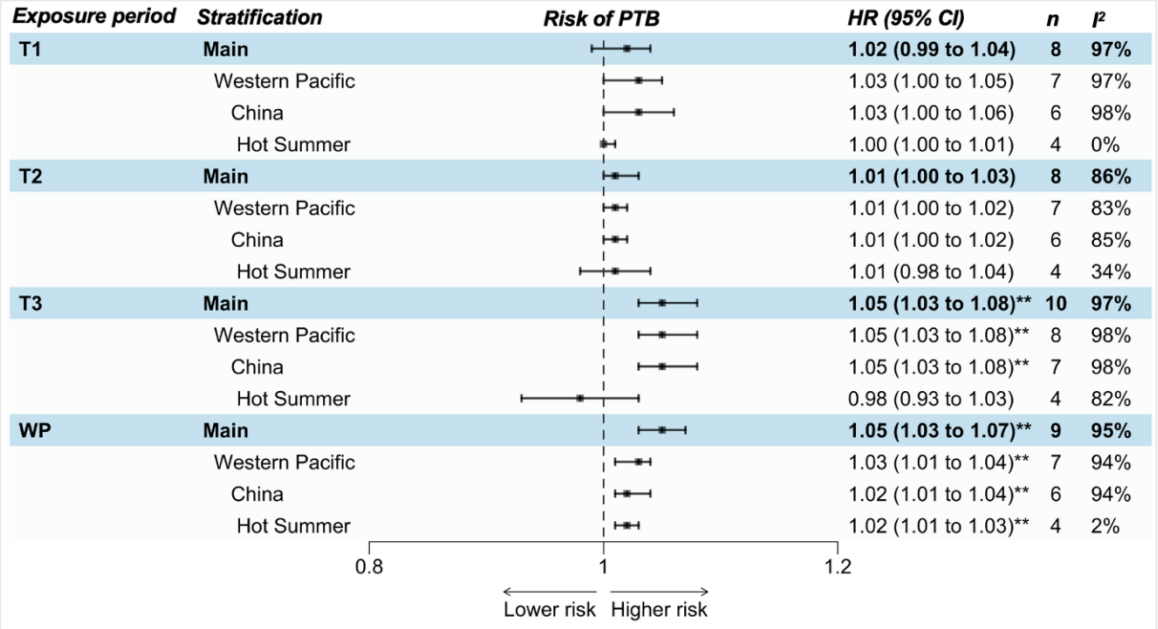


Figure 47. Forest plot for the summary estimates of the meta-analyses of the association between trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP) $PM_{2.5}$ exposure and the risk of PTB from studies using HRs, stratified by WHO region (Americas, Europe, Western Pacific), country (China, US, Canada), and climatic regions within the China and the US. The summary effect sizes (per 10 µg/m³ increase in exposure) and 95%CIs, I^2 statistics, and number of estimates (n) are presented. The forest plots which show the individual studies included in the sensitivity analyses are available in Appendix 22.

Table 30. P-values from Q-test of subgroup differences for stratification by location for studies reporting HRs for the association between $PM_{2.5}$ exposure and the risk of PTB. *: P-value <0.05

Stratification	T1	T2	T3	WP
WHO region	0.18	0.01*	0.65	<0.01*
Country	0.26	0.03*	0.59	<0.01*
China climate regions	<0.01*	<0.01*	<0.01*	<0.01*

5.3.7.2.3. Stratified analysis by location for NO_2 exposure (ORs)

A summary of the results is shown in **Figure 48**, with p-values obtained from Q-test shown in . The results suggested that WHO region could be a partial driver of the heterogeneity observed. There was no evidence for an influence of country or climate regions within China, despite some significant subgroup differences being found. The findings from the main results were robust to the stratification conducted, apart from an increase in risk of trimester 2 exposure found for the “hot summer, cold winter” (HSCW) region of China, where no association was found in the main meta-analysis.

The increase in risk found for trimester 3 exposure remained in all strata, although some changes in the magnitude of the summary estimate were observed. A strong decrease in heterogeneity was observed when only studies in the Americas were used (I^2= 0%), and a decrease of smaller magnitude was observed for European studies for trimester 1 and whole pregnancy exposure.

Table 31. P-values from Q-test of subgroup differences for stratification by location for studies reporting ORs for the association between NO_2 exposure and the risk of PTB. *: P-value <0.05.

Stratification	T1	T2	T3	WP
WHO region	0.85	0.99	0.42	0.65
Country	0.25	0.99	0.76	0.80
China climate regions	<0.01*	0.13	<0.01*	0.04*

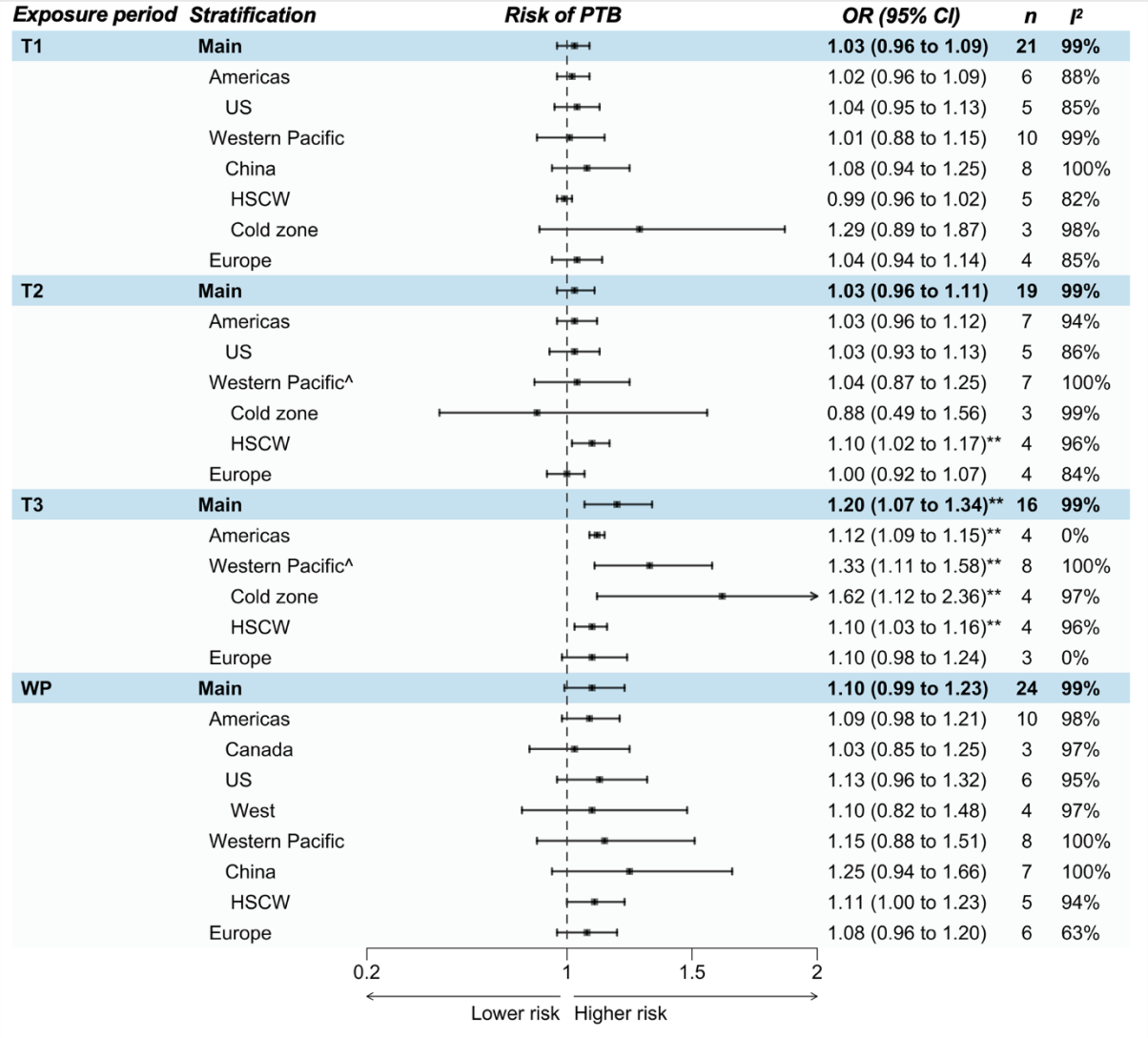


Figure 48. Forest plot for the summary estimates of the meta-analyses conducted for the association between NO_2 exposure at trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP), and the risk of PTB from studies using ORs, stratified in the levels of WHO region (Americas, Europe, Western Pacific), country (China, US, Canada), and climatic regions within the China and the US. The summary effect sizes (per 10 μg/m^3 increase in exposure) and 95%CIs, I^2 statistics, and number of estimates (n) are presented. The forest plots which show the individual studies included in the sensitivity analyses are available in Appendix 22. ^All the studies in the WHO region of Western Pacific were conducted in China for Trimester 2 and 3.

5.3.7.2.4. Stratified analysis by location for NO_2 exposure (HRs)

A summary of the results is shown in **Figure 49**, with p-values obtained from Q-test shown in

Table 32. The findings from the main results were robust to the stratification conducted. Due to the few studies available, only a subgroup analysis of studies that took place in China was conducted. The results suggested that location was not a driver of the high heterogeneity observed in the main results here., but there was low certainty in the findings due to the few studies available.

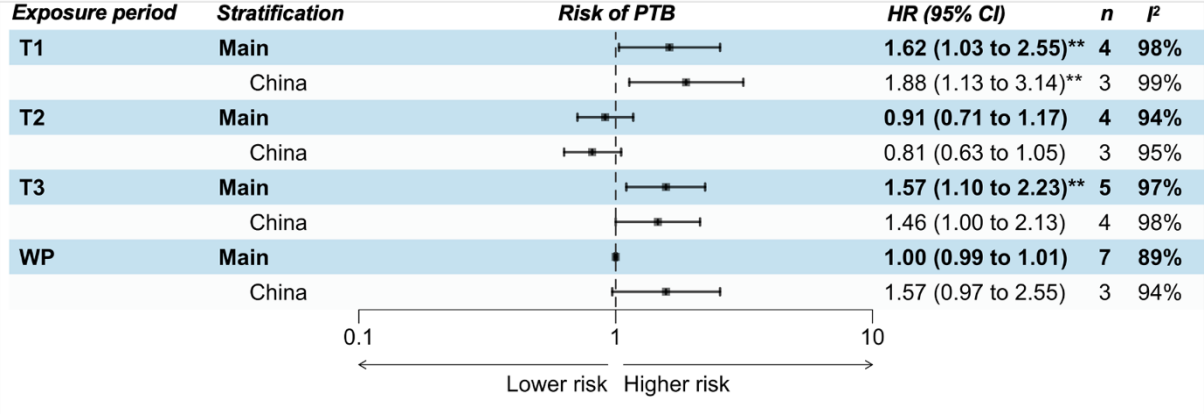


Figure 49. Forest plot for the summary estimates of the meta-analyses conducted for the association between NO_2 exposure at trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP), and the risk of PTB from studies using HRs, stratified according to WHO region of Western Pacific, China, and the hot summer region within China. The summary effect sizes (per 10 µg/m³ increase in exposure) and 95%CIs, I^2 statistics, and number of estimates (n) are presented. The forest plots which show the individual studies included in the sensitivity analyses are available in Appendix 22.

Table 32. P-values obtained from Q-test of subgroup differences for stratification by location for meta-analyses of studies of NO_2 exposure reporting HRs. *: P-value <0.05

Stratification	T1	T2	T3	WP
Country	0.20	0.01*	0.34	0.49

5.3.7.2.5. Stratified analysis by location for O_3 exposure (ORs)

A summary of the results is shown in **Figure 50**, with p-values obtained from the Q-test shown in **Table 33**. The results suggested that location could be a partial driver of the heterogeneity observed, as a moderate decrease in heterogeneity was found for trimester 2 exposure when studies in the Americas were used, and a further decrease occurred when only studies in the US or western region of the US were used (I^2= 0%). However, heterogeneity remained high for the Western Pacific stratum across all exposure periods. Studies in the US showed an increase in risk across all exposure periods, contrasting the null findings in the main meta-analyses.

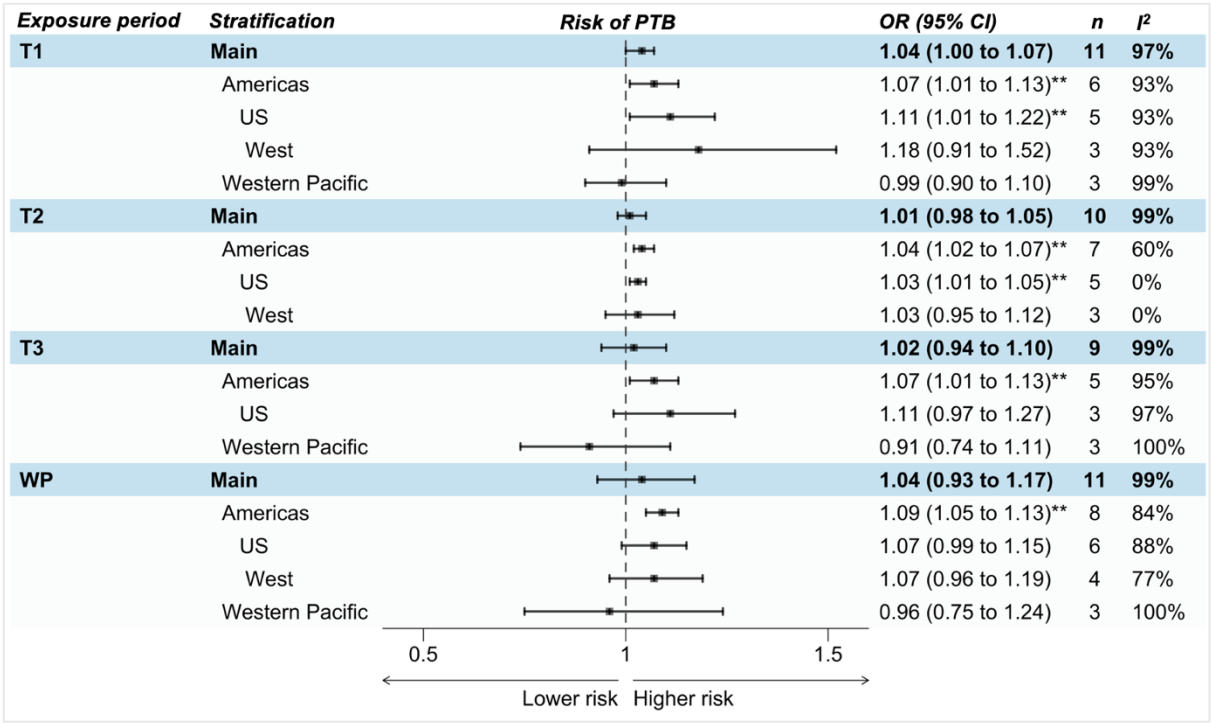


Figure 50. Forest plot for the summary estimates of the meta-analyses conducted for the association between O_3 exposure at trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP), and the risk of PTB from studies using ORs, stratified in the levels of WHO region (Americas, Europe, Western Pacific), country (China, US, Canada), and climatic regions within the China and the US. The summary effect sizes (per 10 µg/m^3 increase in exposure) and 95%CIs, I^2 statistics, and number of estimates (n) are presented. The forest plots which show the individual studies included in the sensitivity analyses are available in Appendix 22.

Table 33. P-values obtained from Q-test of subgroup differences for stratification by location for meta-analyses of studies of O_3 exposure reporting ORs. *: P-value <0.05

Stratification	T1	T2	T3	WP
WHO region	0.36	0.07	0.04	0.23
Country	0.02	0.65	0.61	0.83
US Climate regions	0.04	0.32	<0.01*	0.67

5.3.7.2.6. Stratified analysis by location for O_3 exposure (HRs)

There were fewer than three studies in all stratifications, hence a stratified analysis was not conducted here.

5.3.7.3. Exposure assessment methodology

5.3.7.3.1. Stratified analysis by exposure assessment methodology for $PM_{2.5}$ exposure (ORs)

A summary of the results is shown in **Table 34**, with p-values obtained from the Q-test shown in **Figure 51**. The results suggested that exposure assessment method could be a partial driver of the heterogeneity observed, as various degrees of reduction in heterogeneity were observed when the meta-analyses were restricted to some methods of exposure assessment but not the others. The increase in risk found for trimester 3 and whole pregnancy exposure dissipated among studies that used model-based exposure assessment methods, except for those that used hybrid exposure models in whole pregnancy exposure.

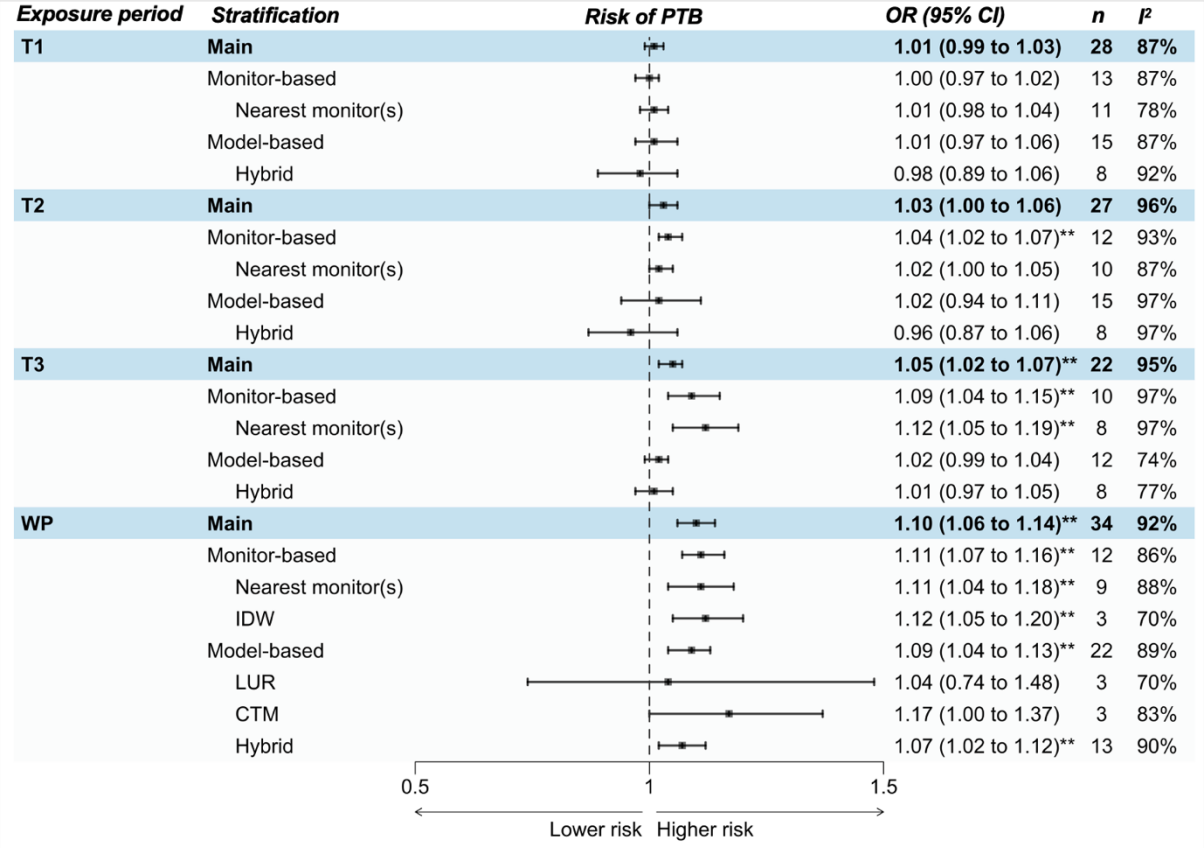


Figure 51. Forest plot for the summary estimates of the meta-analyses conducted for the association between $PM_{2.5}$ exposure at trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP), and the risk of PTB from studies using ORs, stratified by different exposure assessment methods. The summary effect sizes (per 10 µg/m^3 increase in exposure) and 95%CIs, I^2 statistics, and number of estimates (n) are presented. The forest plots which show the individual studies included in the sensitivity analyses are available in Appendix 22. IDW: Inverse distance weighting; LUR: Land-use regression; CTM: Chemical transport models.

Table 34. P-values obtained from Q-test of subgroup differences for stratification by exposure assessment for meta-analyses of studies of $PM_{2.5}$ exposure reporting ORs. *: P-value <0.05

Stratification	T1	T2	T3	WP
Model-based vs Monitor-based	0.04*	0.57	0.01*	0.45
Each method	0.21	<0.01*	0.11	0.51

5.3.7.3.2. Stratified analysis by exposure assessment methodology for $PM_{2.5}$ exposure (HRs)

A summary of the results is shown in **Figure 52**, with p-values obtained from the Q-test shown in **Figure 35**. The results suggested that the method of exposure assessment could be a partial driver of the high heterogeneity observed. The strata with studies that used nearest monitor(s) all have reduced heterogeneity except in trimester 3. On the other hand, the strata with studies of model-based exposure assessment methods have a strong reduction in heterogeneity (I^2 = 52%) in trimester 3 only. There was limited change in the summary estimate with 95% CIs heavily overlapping with each other. There was low certainty in the finding as few studies were available in most strata.

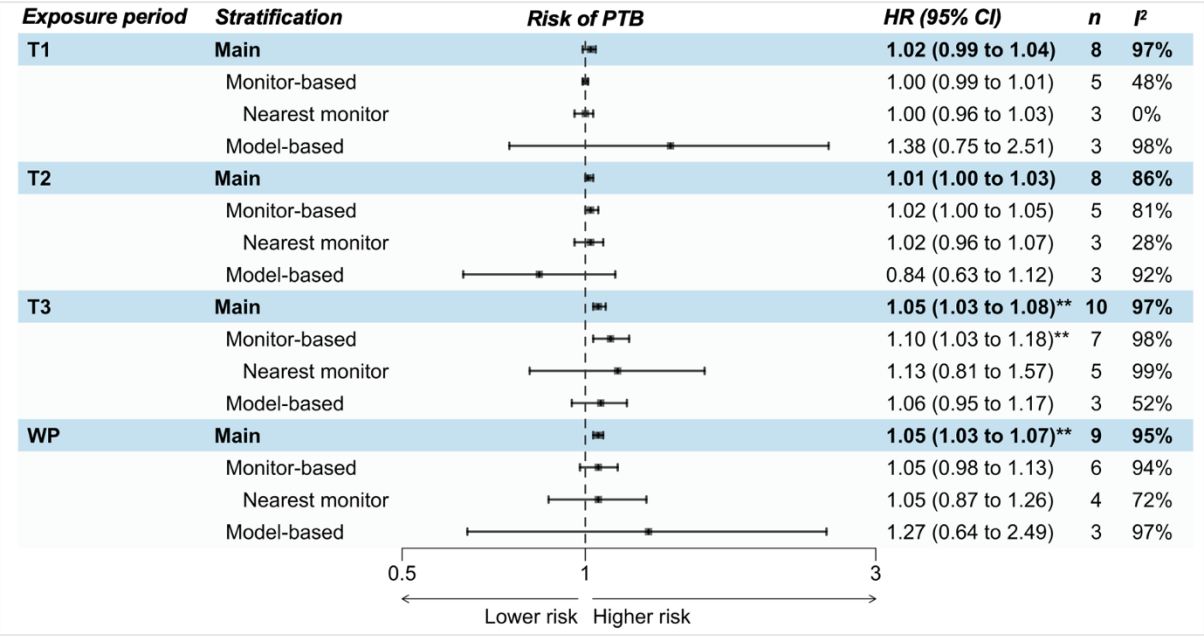


Figure 52. Forest plot for the summary estimates of the meta-analyses conducted for the association between $PM_{2.5}$ exposure at trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP), and the risk of PTB from studies using HRs, stratified by different exposure assessment methods. The summary effect sizes (per 10 µg/m3 increase in exposure) and 95%CIs, I^2 statistics, and number of estimates (n) are presented. The forest plots which show the individual studies included in the sensitivity analyses are available in Appendix 22.

Table 35. P-values obtained from Q-test of subgroup differences for stratification by exposure assessment for meta-analyses of studies of $PM_{2.5}$ exposure reporting HRs.

Stratification	T1	T2	T3	WP
Model-based vs Monitor-based	<0.01*	0.06	0.61	0.02
Each method	/	/	/	/

5.3.7.3.3. Stratified analysis by exposure assessment methodology for NO_2 exposure (ORs)

A summary of the results is shown in **Figure 53**, with p-values obtained from the Q-test shown in **Table 36**. The results suggested that exposure assessment method could be a partial driver of the high heterogeneity observed. The strongest reduction in heterogeneity was found for model-based methods in trimester 3 (I^2 = 0%). The increase in risk of trimester 3 exposure remained in all strata except for studies that used inverse distance weighting (IDW), which could be attributable to having fewer studies in the strata resulting in a wide CI. For whole pregnancy exposure, an increase in risk was found among studies that used monitor-based approaches but not model-based approaches. The increase in risk from studies that used nearest monitors(s) (OR: 1.16 [1.04 - 1.30], n = 7) was distinct with no overlaps in 95%CI with studies that used IDW (OR: 1.75 [1.58 - 1.95], n = 3). On the other hand, a decrease in risk was suggested by studies that used LUR (OR: 0.89 [0.80-0.99], n = 7).

Table 36. P-values obtained from Q-test of subgroup differences for stratification by exposure assessment for meta-analyses of studies of NO_2 exposure reporting ORs. *: P-value <0.05

Stratification	T1	T2	T3	WP
Model-based vs Monitor-based	0.62	0.82	0.30	<0.01*
Each method	0.33	0.85	0.94	<0.01*

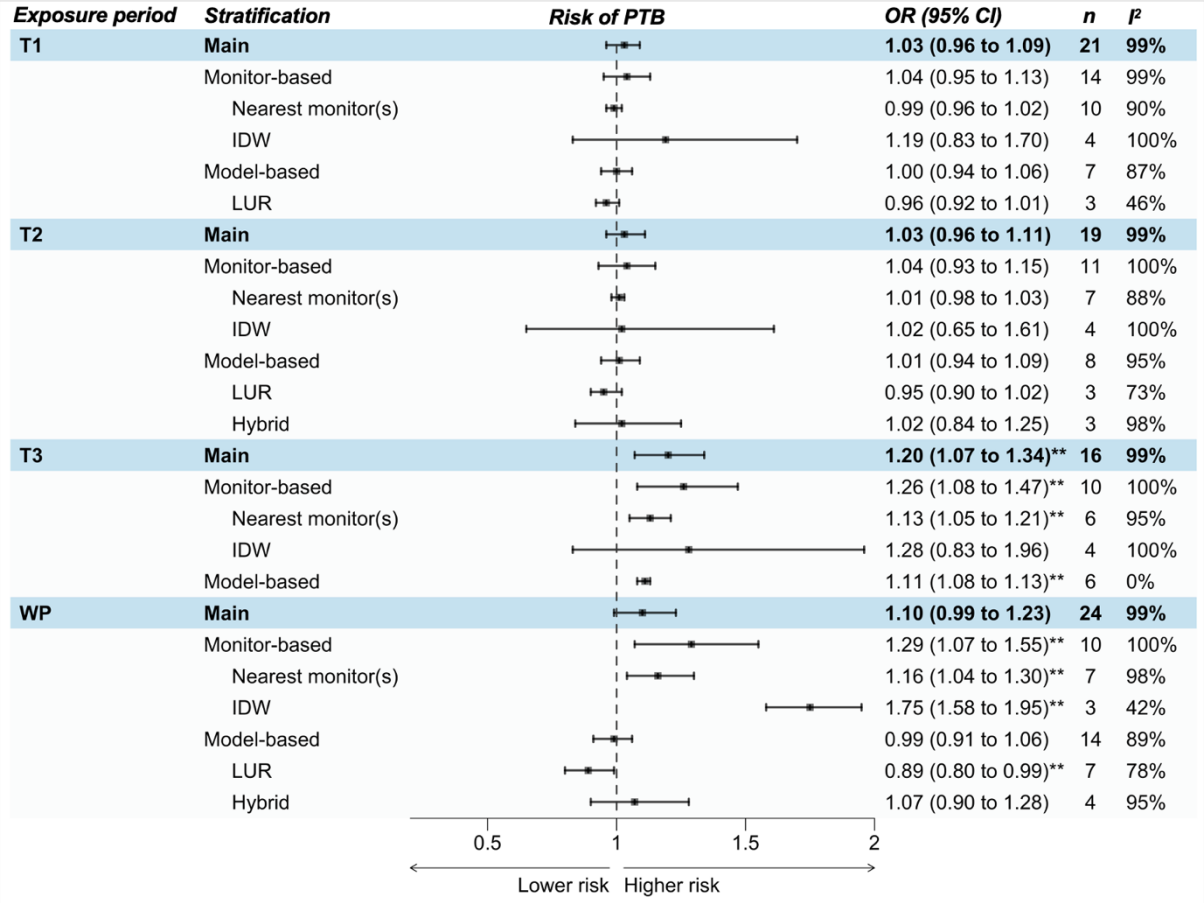


Figure 53. Forest plot for the summary estimates of the meta-analyses conducted for the association between NO_2 exposure at trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP), and the risk of PTB from studies using ORs, stratified by different exposure assessment methods. The summary effect sizes (per 10 µg/m^3 increase in exposure) and 95%CIs, I^2 statistics, and number of estimates (n) are presented. The forest plots which show the individual studies included in the sensitivity analyses are available in Appendix 22. IDW: Inverse distance weighting; LUR: Land-use regression; CTM: Chemical transport models.

5.3.7.3.4. Stratified analysis by exposure assessment methodology for NO_2 exposure (HRs)

A summary of the results is shown in **Figure 54**. The results suggested that exposure assessment method could be a partial driver of the high heterogeneity found, with heterogeneity reduced in all exposure periods except trimester 3 among studies that used nearest monitor(s). Large uncertainty was present for the finding, as there were few studies available, and subgroup analysis was only conducted for studies using extrapolation from nearest monitor(s).

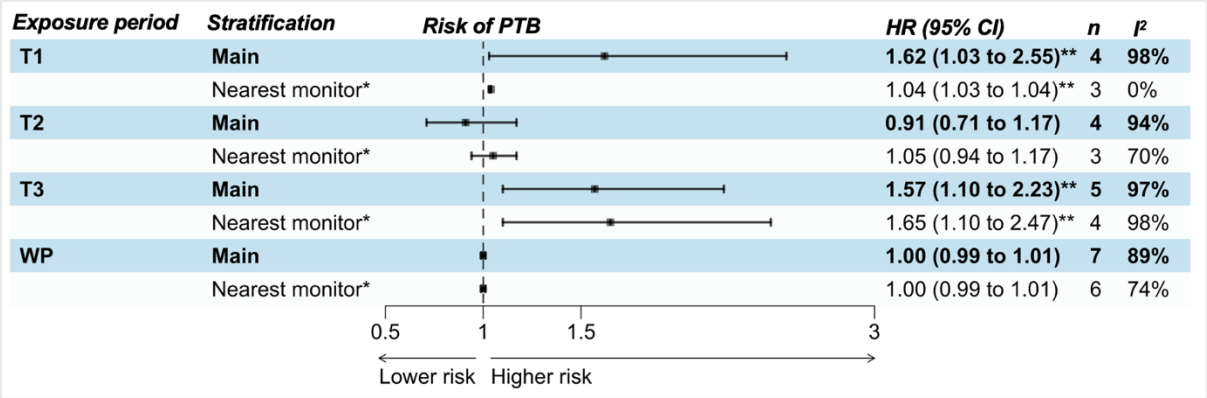


Figure 54. Forest plot for the summary estimates of the meta-analyses conducted for the association between NO_2 exposure at trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP), and the risk of PTB from studies using HRs, with restriction to studies that have extrapolated exposure from nearest monitors only. The summary effect sizes (per 10 µg/m³ increase in exposure) and 95%CIs, I^2 statistics, and number of estimates (n) are presented. The forest plots which show the individual studies included in the sensitivity analyses are available in Appendix 22.

5.3.7.3.5. Stratified analysis by exposure assessment methodology for O_3 exposure (ORs)

A summary of the results is shown in **Figure 55**, with p-values obtained from the Q-test shown in **Table 37**. The summary estimates remained relatively stable across most strata. The results suggested that the method of exposure assessment could be a partial driver of the heterogeneity found, with a strong decrease in heterogeneity found in studies that used model-based exposure assessment methods for trimester 2 and whole pregnancy exposure. An increase in risk was found for trimester 1 and 2 exposures among studies that used model-based approaches, and for whole pregnancy exposure among studies that used nearest monitor(s), which contrasted the null findings in the main meta-analyses. However, large uncertainty is present for the finding as few studies were used.

Table 37. P-values obtained from Q-test of subgroup differences for stratification by exposure assessment for meta-analyses of studies of O_3 exposure reporting ORs. *: P-value <0.05

Stratification	T1	T2	T3	WP
Modelling vs. Not	0.02*	0.35	0.10	0.96
Each method	<0.01*	0.35	0.06	0.01*

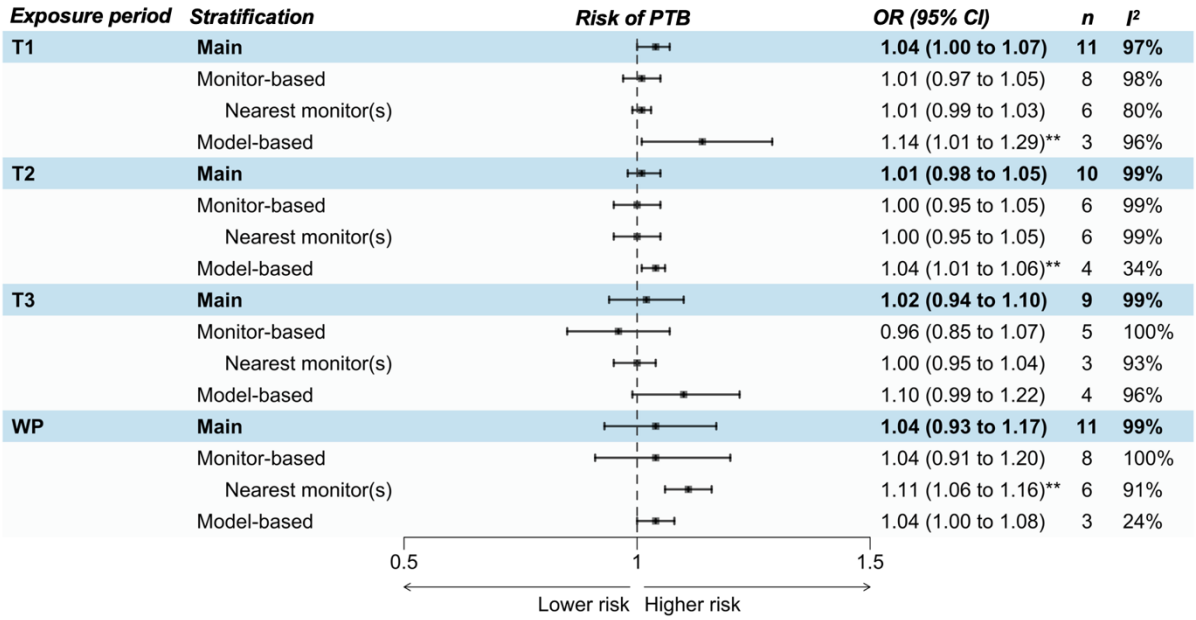


Figure 55. Forest plot for the summary estimates of the meta-analyses conducted for the association between O_3 exposure at trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP), and the risk of PTB from studies using ORs, stratified by different exposure assessment methods. The summary effect sizes (per 10 µg/m^3 increase in exposure) and 95%CIs, I^2 statistics, and number of estimates (n) are presented. The forest plots which show the individual studies included in the sensitivity analyses are available in Appendix 22.

5.3.7.3.6. Stratified analysis by exposure assessment methodology for O_3 exposure (HRs)

The sensitivity analysis was only conducted for whole pregnancy exposure as there were not enough studies in the other exposure periods (**Figure 56**). There was no evidence that exposure assessment method contributed to the heterogeneity observed.

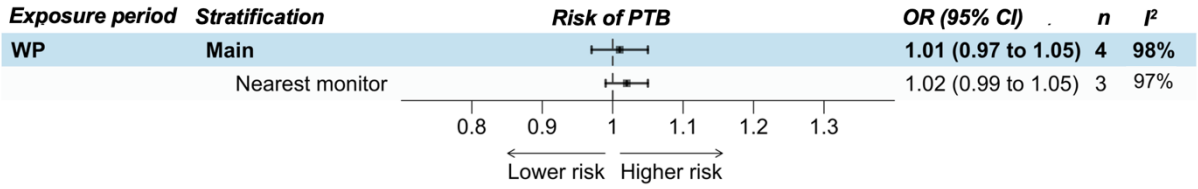


Figure 56. Forest plot for the summary estimates of the meta-analyses conducted for the association between O_3 exposure at whole pregnancy (WP) and the risk of PTB from studies using HRs, with restriction to studies that have extrapolated exposure from nearest monitors only. The summary effect sizes (per 10 µg/m^3 increase in exposure) and 95%CIs, I^2 statistics, and number of estimates (n) are presented. The forest plots which show the individual studies included in the sensitivity analyses are available in Appendix 22.

5.3.7.4. Concentration levels

5.3.7.4.1. Stratified analysis by concentration levels for $PM_{2.5}$ exposure (ORs)

A summary of the results is shown in **Figure 57**, with p-values obtained from the Q-test shown in **Table 38**. The results suggested that concentration levels could be a partial driver of the high heterogeneity observed. The summary estimate was relatively robust across most strata, except for studies in the <10 μg/m^3 stratum for trimester 2 exposure, where an increase in risk was found (OR: 1.30 [1.07 - 1.58]) that was distinct and did not overlap with the null finding of the main meta-analysis (OR: 1.03 [1.00 - 1.06]). Similarly, the increase in risk in trimester 3 exposure found for studies in the >25 to 50 μg/m^3 stratum showed no heterogeneity (OR: 1.04 [1.02 - 1.05], I^2 = 0%, n = 3) and did not overlap with that of >50 μg/m^3 stratum (OR: 1.12 [1.06 - 1.19]). These findings suggest that concentration could be an effect modifier but not the sole driver of heterogeneity.

The increase in risk found for trimester 3 exposure dissipated in strata with levels <25 μg/m^3, whereas it only dissipated in the stratum for 10 to 25 μg/m^3 for whole pregnancy exposure. The results of the meta-analysis by stratification of AQG values are broadly similar to that of the quartile distribution of $PM_{2.5}$ in the primary studies used (**Appendix 21**).

Table 38. P-values obtained from Q-test of subgroup differences for stratification by pollution levels for meta-analyses of studies of $PM_{2.5}$ exposure reporting ORs. *: P-value <0.05

Stratification	T1	T2	T3	WP
AQG	0.32	<0.01*	0.03*	0.61
Quartile	0.20	<0.01*	<0.01*	0.64

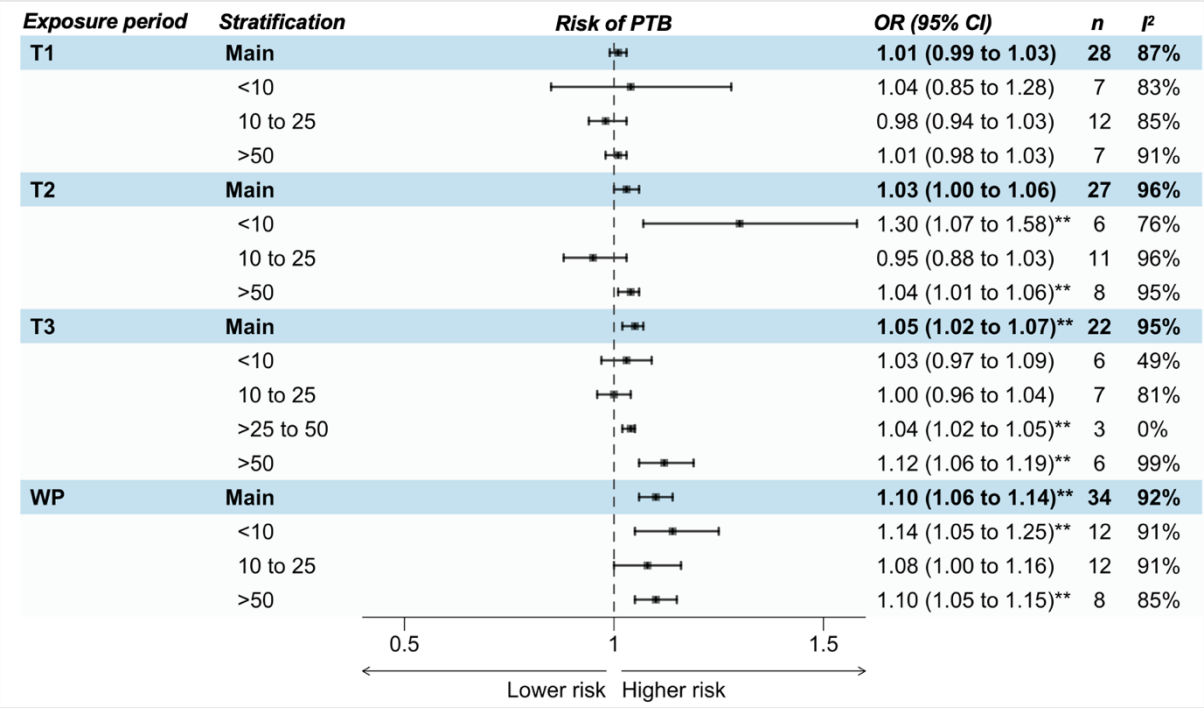


Figure 57. Forest plot for the summary estimates of the meta-analyses conducted for the association between $PM_{2.5}$ exposure at trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP), and the risk of PTB from studies using ORs, stratified by brackets of $PM_{2.5}$ set using the WHO Air Quality Guideline levels. The summary effect sizes (per 10 µg/m³ increase in exposure) and 95%CIs, I^2 statistics, and number of estimates (n) are presented. The forest plots which show the individual studies included in the sensitivity analyses are available in Appendix 22. The levels of $PM_{2.5}$ strata are expressed in µg/m³.

5.3.7.4.2. Stratified analysis by concentration levels for $PM_{2.5}$ exposure (HRs)

A summary of the results is shown in **Figure 58**, with p-values obtained from the Q-test shown in **Table 39**. The results suggested that concentration levels could be a partial driver of the heterogeneity observed, but uncertainties were present due to the few studies available in the strata. In the strata of >25 to 50 µg/m^3 across all exposure periods, heterogeneity was reduced by varying degrees and to 0% in trimester 1 and whole pregnancy exposure. The increase in risk found for trimester 3 exposure dissipated in all strata, whereas it remained in all strata for whole pregnancy exposure. Stratification by quartiles of distribution was not conducted as only the 3rd quartile contained enough (>2) studies to be meta-analysed. There was low certainty in the findings as relatively few studies available.

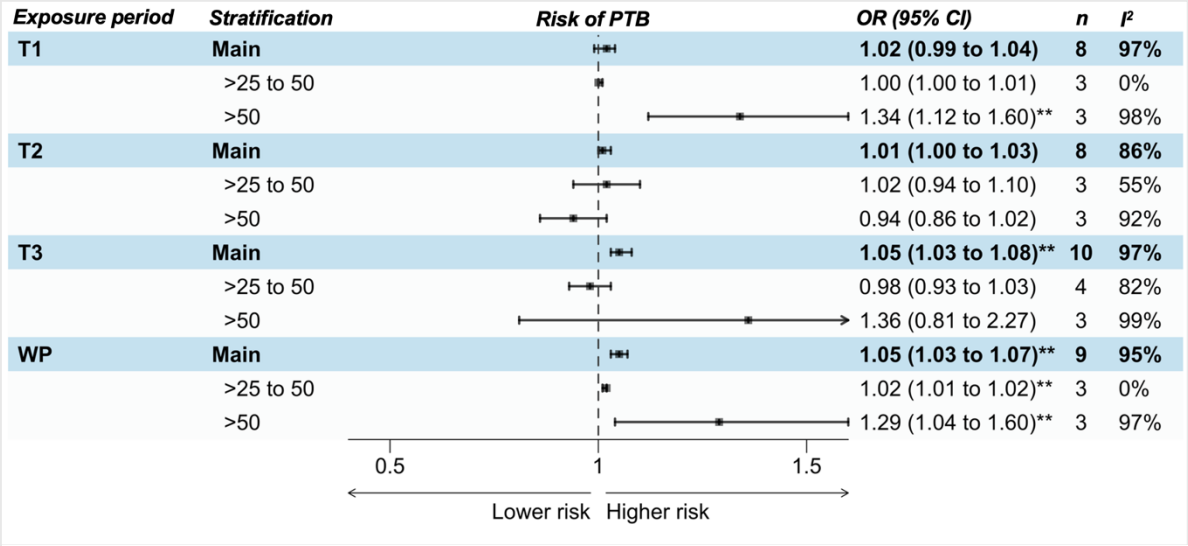


Figure 58. Forest plot for the summary estimates of the meta-analyses conducted for the association between $PM_{2.5}$ exposure at trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP), and the risk of PTB from studies using HRs, stratified by brackets of $PM_{2.5}$ set using the WHO Air Quality Guideline levels. The summary effect sizes (per 10 µg/m³ increase in exposure) and 95%CIs, I^2 statistics, and number of estimates (n) are presented. The forest plots which show the individual studies included in the sensitivity analyses are available in Appendix 22. The levels of $PM_{2.5}$ strata are expressed in µg/m³.

Table 39. P-values obtained from Q-test of subgroup differences for stratification by pollution levels for meta-analyses of studies of $PM_{2.5}$ exposure reporting HRs. *: P-value <0.05

Stratification	T1	T2	T3	WP
AQG	0.04	0.51	0.15	0.36
Quartile	<0.01*	<0.01*	<0.01*	0.47

5.3.7.4.3. Stratified analysis by concentration levels for NO_2 exposure (ORs)

A summary of the results is shown in **Figure 59**, with p-values obtained from the Q-test shown in **Table 40**. The results suggested that concentration level could be a partial driver of the heterogeneity observed, with more pronounced reduction in heterogeneity only found in the 10 to 30 µg/m³ stratum of trimester 1 and 3 exposure. The increase in risk found for trimester 3 exposure dissipated in the >30 to 50 µg/m³ stratum, but remained in all other strata. The results of the meta-analysis by stratification of AQG values were identical to that of the quartile distribution of NO_2 in the primary studies used and thus are not shown.

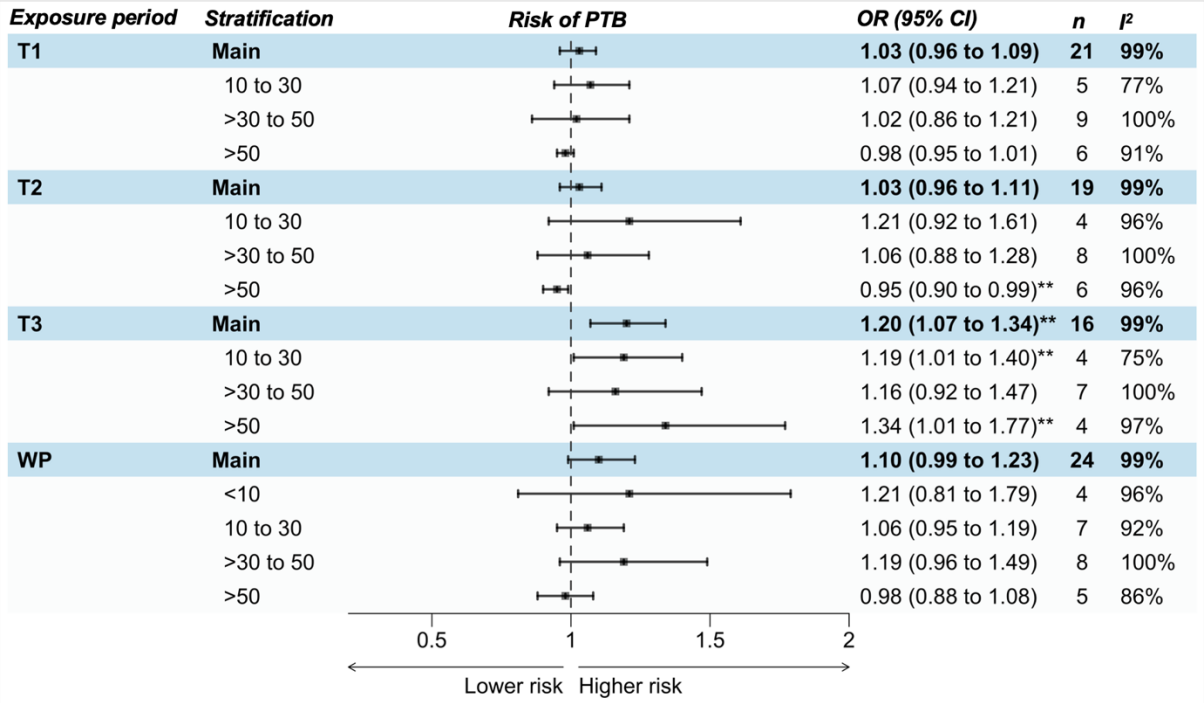


Figure 59. Forest plot for the summary estimates of the meta-analyses conducted for the association between NO_2 exposure at trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP), and the risk of PTB from studies using ORs, stratified by brackets of NO_2 set using the WHO Air Quality Guideline levels. The summary effect sizes (per 10 µg/m³ increase in exposure) and 95%CIs, I^2 statistics, and number of estimates (n) are presented. The forest plots which show the individual studies included in the sensitivity analyses are available in Appendix 22. The levels of NO_2 strata are expressed in µg/m³.

Table 40. P-values obtained from Q-test of subgroup differences for stratification by pollution levels for meta-analyses of studies of NO_2 exposure reporting ORs.

Stratification	T1	T2	T3	WP
AQG	0.79	0.05	0.88	0.23
Quartile	<0.01*	<0.01*	0.88	0.10

5.3.7.4.4. Stratified analysis by concentration levels for NO_2 exposure (HRs)

Only one stratum contains enough (>2) studies to be meta-analysed, hence a stratified analysis was not conducted here.

5.3.7.4.5. Stratified analysis by concentration levels for O_3 exposure (ORs)

A summary of the results is shown in **Figure 60** (AQG) and **Figure 61** (Quartile), with p-values obtained from the Q-test shown in **Table 41**. The results suggested that concentration levels could be a partial driver of heterogeneity observed. Minimal changes to the summary estimate were observed in most strata. The strongest decrease in heterogeneity was in the stratum of 60 or less µg/m³ for trimester 2 and whole pregnancy exposure, and the stratum

for exposure in the first and second quartile for trimester 2 and 1 exposure respectively. There were no changes in the findings when the studies were stratified by AQG. However, increase in risk was found for trimester 1, 2, and whole pregnancy exposure at the first and second quartile after stratification.

Table 41. P-values obtained from Q-test of subgroup differences for stratification by pollution levels for meta-analyses of studies of O_3 exposure reporting ORs. *: P-value <0.05

Stratification	T1	T2	T3	WP
AQG	<0.01*	0.50	0.01*	0.91
Quartile	<0.01*	0.36	<0.01*	0.06

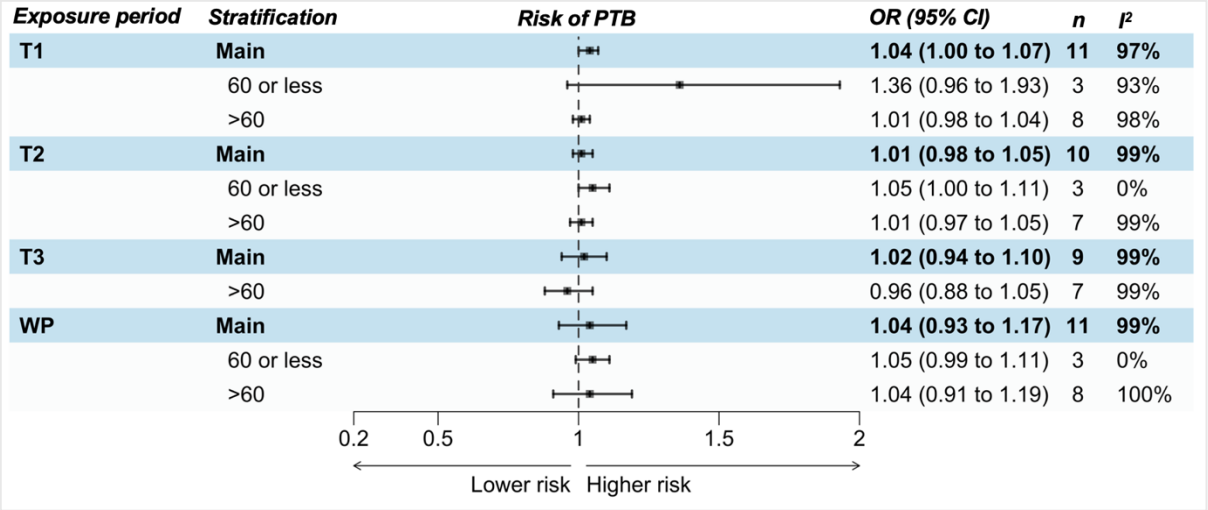


Figure 60. Forest plot for the summary estimates of the meta-analyses conducted for the association between O_3 exposure at trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP), and the risk of PTB from studies using ORs, stratified by brackets of O_3 set using the WHO Air Quality Guideline levels. The summary effect sizes (per 10 µg/m^3 increase in exposure) and 95%CIs, I^2 statistics, and number of estimates (n) are presented. The forest plots which show the individual studies included in the sensitivity analyses are available in Appendix 22. The levels of O_3 strata are expressed in µg/m^3.

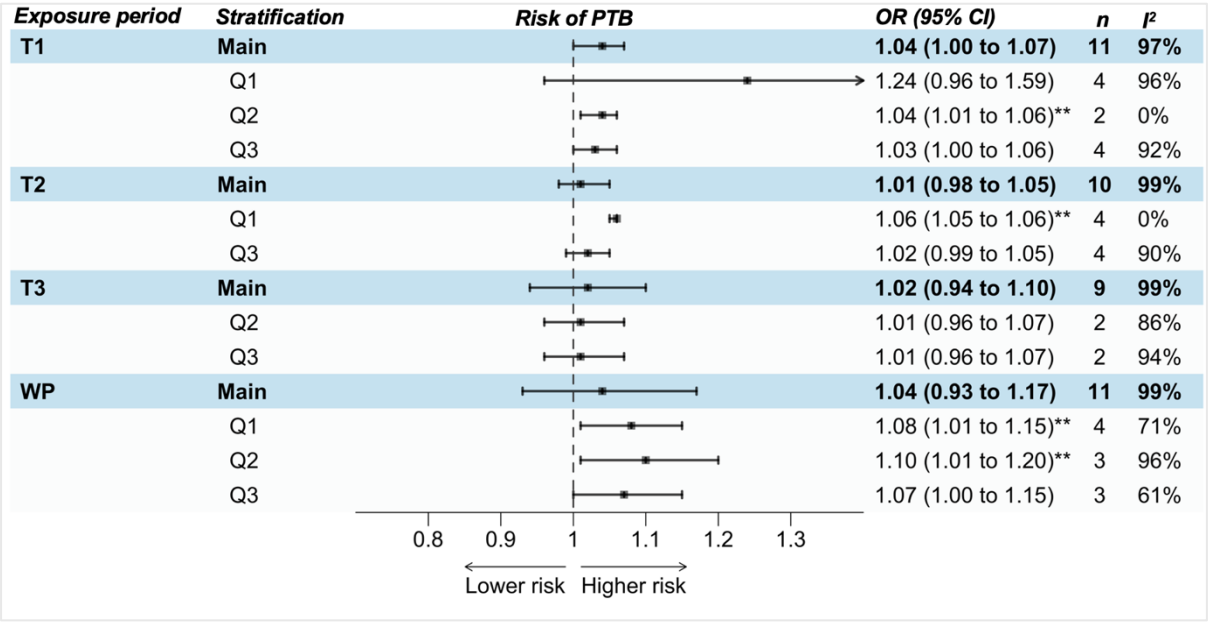


Figure 61. Forest plot for the summary estimates of the meta-analyses conducted for the association between O_3 exposure at trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP), and the risk of PTB from studies using ORs, stratified by the quartiles of the distribution of O_3 levels in the primary studies used in the meta-analyses. The summary effect sizes (per 10 µg/m^3 increase in exposure) and 95%CIs, I^2 statistics, and number of estimates (n) are presented. The forest plots which show the individual studies included in the sensitivity analyses are available in Appendix 22. The levels of O_3 strata are expressed in µg/m^3.

5.3.7.4.6. Stratified analysis by concentration levels for O_3 exposure (HRs)

Only one stratum contains enough (>2) studies to be meta-analysed, hence a stratified analysis was not conducted here.

5.3.7.5. Data source

5.3.7.5.1. Stratified analysis by data source for $PM_{2.5}$ exposure (ORs)

A summary of the results is shown in **Figure 62**, with p-values obtained from the Q-test shown in **Table 42**. No significant subgroup differences were found. Notable reductions in heterogeneity were found for whole pregnancy exposure, restricted to studies that have used research cohorts, which the increase in risk was dissipated and I^2 was reduced to 18%, suggesting that source of data potentially explained some of the heterogeneity in whole pregnancy exposure. Data source could be a partial driver of the high heterogeneity observed but there were not enough studies that used research cohorts in exposure windows rather than whole pregnancy.

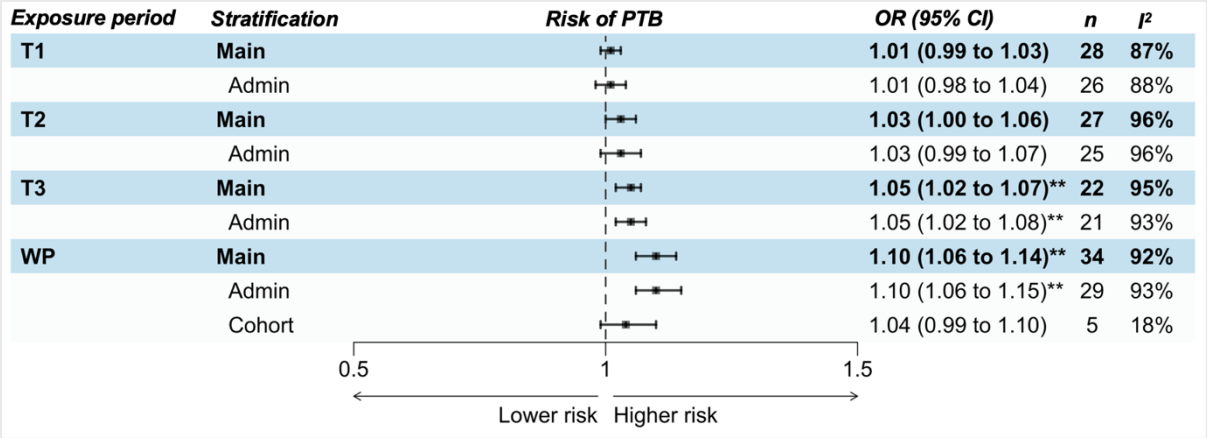


Figure 62. Forest plot for the summary estimates of the meta-analyses conducted for the association between $PM_{2.5}$ exposure at trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP), and the risk of PTB from studies using ORs, stratified by the source of data of the study: Administrative (Admin) or Research Cohort (Cohort). The summary effect sizes (per 10 µg/m³ increase in exposure) and 95%CIs, I^2 statistics, and number of estimates (n) are presented. The forest plots which show the individual studies included in the sensitivity analyses are available in Appendix 22.

Table 42. P-values obtained from Q-test of subgroup differences for stratification by data source for meta-analyses of studies of $PM_{2.5}$ exposure reporting ORs. *: P-value <0.05

Stratification	T1	T2	T3	WP
Data source	0.78	0.37	0.47	0.30

5.3.7.5.2. Stratified analysis by data source for $PM_{2.5}$ exposure (HRs)

A summary of the results is shown in **Figure 63**, with p-values obtained from the Q-test shown in **Table 43**. The results suggested that data source was not a driver of the high heterogeneity observed, as there were minimal changes to the magnitude and direction of summary estimates with minimal reduction in heterogeneity.

Table 43. P-values obtained from Q-test of subgroup differences for stratification by data source for meta-analyses of studies of $PM_{2.5}$ exposure reporting HRs. *: P-value <0.05

Stratification	T1	T2	T3	WP
Data source	<0.01*	<0.01*	0.29	<0.01*

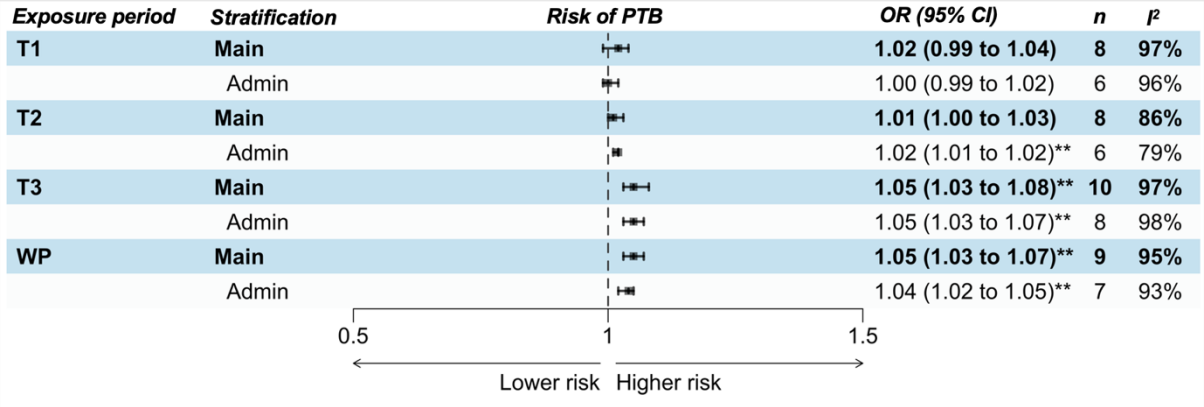


Figure 63. Forest plot for the summary estimates of the meta-analyses conducted for the association between $PM_{2.5}$ exposure at trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP), and the risk of PTB from studies using HRs, stratified by the source of data of the study: Administrative (Admin) or Research Cohort (Cohort). The summary effect sizes (per 10 µg/m³ increase in exposure) and 95%CIs, I^2 statistics, and number of estimates (n) are presented. The forest plots which show the individual studies included in the sensitivity analyses are available in Appendix X.

5.3.7.5.3. Stratified analysis by data source for NO_2 exposure (ORs)

A summary of the results is shown in **Figure 64**, with p-values obtained from the Q-test shown in **Table 44**. No significant subgroup differences were found. The results suggested that data source could be a partial driver of the high heterogeneity found. For studies that used research cohorts, a moderate to strong reduction in heterogeneity was found in trimester 1, 3, and whole pregnancy exposure, suggesting that data source explained some of the heterogeneity in trimester 1, trimester 3, and whole pregnancy exposure. The increase in risk found for trimester 3 exposure dissipated in both strata, whereas an increase in risk was found for whole pregnancy among studies that used research cohorts (OR: 1.08 [1.01 - 1.06]).

Table 44. P-values obtained from Q-test of subgroup differences for stratification by data source for meta-analyses of studies of NO_2 exposure reporting ORs.

Stratification	T1	T2	T3	WP
Data source	0.64	0.06	0.42	0.89

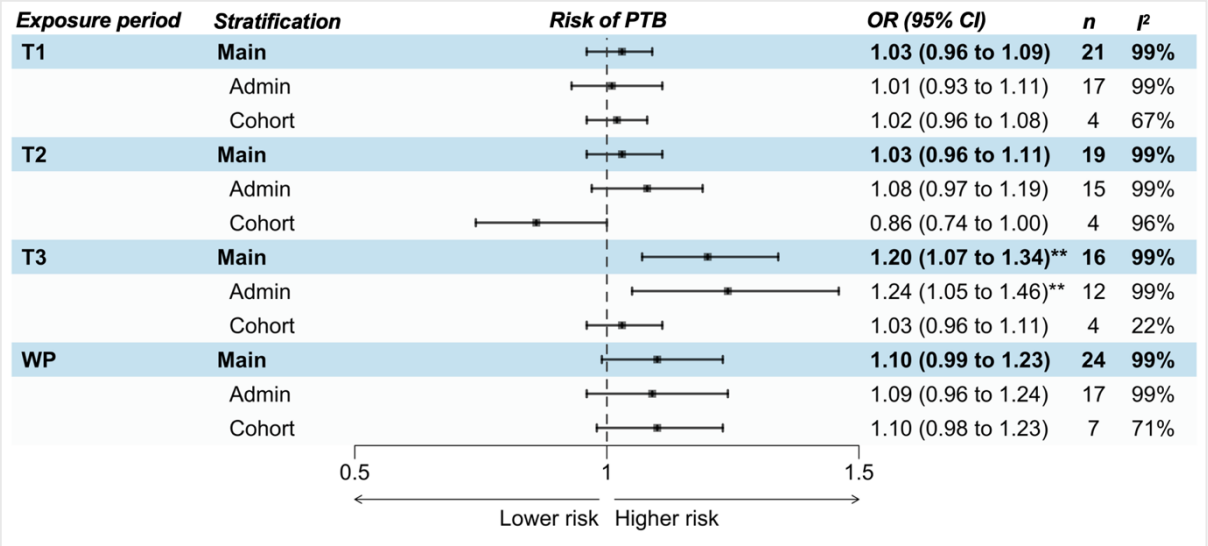


Figure 64. Forest plot for the summary estimates of the meta-analyses conducted for the association between NO_2 exposure at trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP), and the risk of PTB from studies using ORs, stratified by the source of data of the study: Administrative (Admin) or Research Cohort (Cohort). The summary effect sizes (per 10 µg/m^3 increase in exposure) and 95%CIs, I^2 statistics, and number of estimates (n) are presented. The forest plots which show the individual studies included in the sensitivity analyses are available in Appendix 22.

5.3.7.5.4. Stratified analysis by data source for NO_2 exposure (HRs)

A summary of the results is shown in **Figure 65**, with p-values obtained from the Q-test shown in **Table 45**. No significant subgroup differences were found. The results suggested that source of data could be a partial driver of the heterogeneity found, but large uncertainty was present due to the few studies available. A moderate to strong decrease in heterogeneity was found for studies that used data from administrative sources in trimesters 1, 2, and whole pregnancy.

Table 45. P-values obtained from Q-test of subgroup differences for stratification by data source for meta-analyses of studies of NO_2 exposure reporting HRs. *P-value: <0.01

Stratification	T1	T2	T3	WP
Data source	/	/	/	0.83

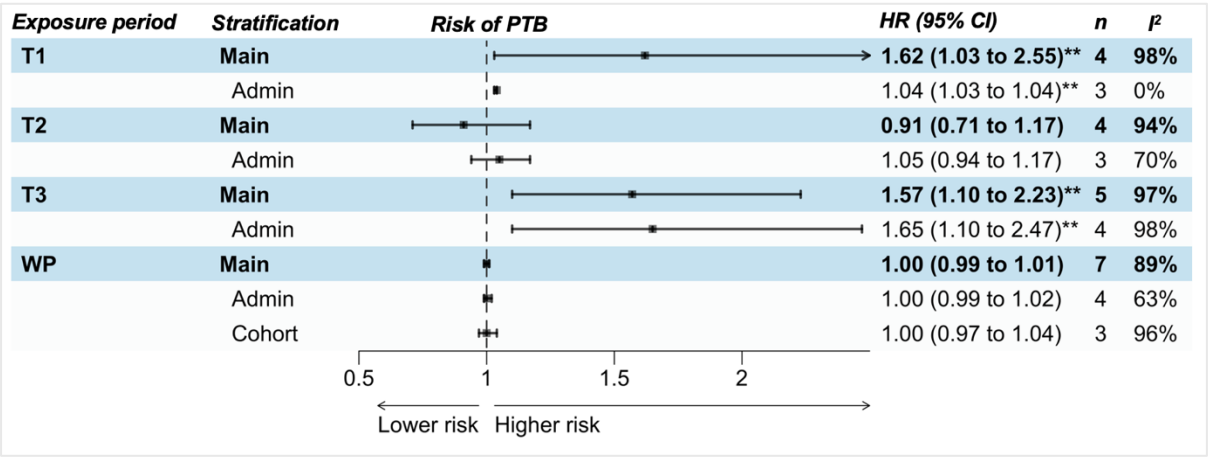


Figure 65. Forest plot for the summary estimates of the meta-analyses conducted for the association between NO_2 exposure at trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP), and the risk of PTB from studies using HRs, stratified by the source of data of the study: Administrative (Admin) or Research Cohort (Cohort). The summary effect sizes (per 10 µg/m³ increase in exposure) and 95%CIs, I^2 statistics, and number of estimates (n) are presented. The forest plots which show the individual studies included in the sensitivity analyses are available in Appendix 22.

5.3.7.5.5. Stratified analysis by data source for O_3 exposure (ORs)

A summary of the results is shown in **Figure 66**. There was no evidence that source of data was a driver of the heterogeneity found here.

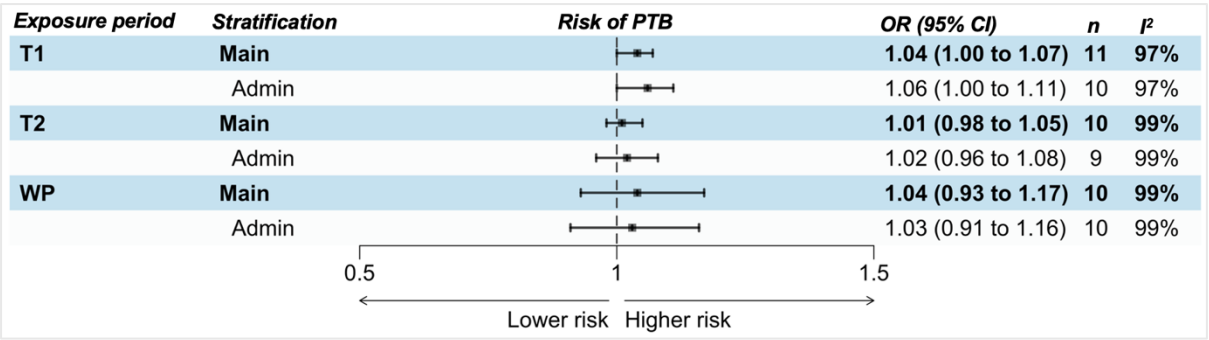


Figure 66. Forest plot for the summary estimates of the meta-analyses conducted for the association between O_3 exposure at trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP), and the risk of PTB from studies using ORs, stratified by the source of data of the study: Administrative (Admin) or Research Cohort (Cohort). The summary effect sizes (per 10 µg/m³ increase in exposure) and 95%CIs, I^2 statistics, and number of estimates (n) are presented. The forest plots which show the individual studies included in the sensitivity analyses are available in Appendix 22. All studies in T3 used administrative data sources.

5.3.7.5.6. *Stratified analysis by data source for O_3 exposure (HRs)*

Only the exposure window for whole pregnancy had sufficient studies to conduct stratification (**Figure 67**). The summary estimate did not differ after being restricted to studies that have used research cohorts as the data source.

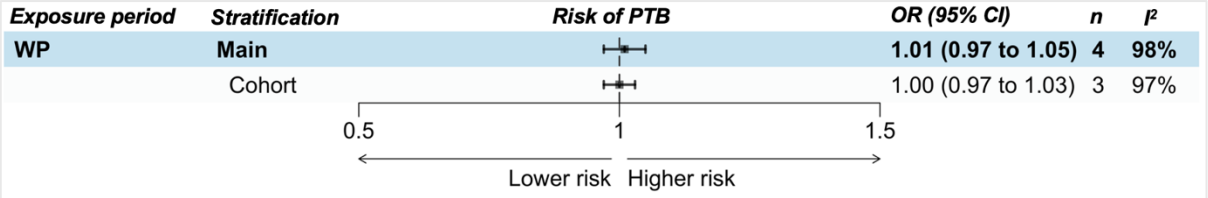


Figure 67. Forest plot for the summary estimates of the meta-analyses conducted for the association between O_3 exposure at trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP), and the risk of PTB from studies using HRs, stratified by the source of data of the study: Administrative (Admin) or Research Cohort (Cohort). The summary effect sizes (per 10 µg/m^3 increase in exposure) and 95%CIs, I^2 statistics, and number of estimates (n) are presented. The forest plots which show the individual studies included in the sensitivity analyses are available in Appendix 22.

5.3.8. Publication bias and funnel plots

Funnel plots and Egger's test for assessing funnel plot asymmetry were used for the exposure-outcome pairs with more than 10 studies available, and the results are available in **Appendix 19** and **20** respectively. Visual inspection of the funnel plots suggested that there were small degrees of asymmetry present for most of the plots. However, Egger's tests showed no support for the presence of asymmetry or small study effect, suggesting that the asymmetry could be a result of the high heterogeneity of studies included in the meta-analysis.

5.4. Discussion

In this section, a summary of the findings from the main meta-analyses and the stratified analyses are first provided. Then, the implications of the findings are explored alongside the results of preceding works. Afterwards, causality is discussed guided by the Bradford Hill Criteria, incorporating findings from Chapter 4 and the wider literature. An explanation of the decision to conduct a health impact assessment in Chapter 6, and other related decisions, for instance, the choice of concentration-response function (CRF), are included in the discussion on causality. The strengths and limitations of the work conducted in this chapter are described before an interim conclusion is made.

5.4.1. Summary of findings of the main meta-analyses

Whole pregnancy $PM_{2.5}$ exposure was consistently shown to be associated with an increased risk of PTB, mlPTB, and vPTB (both ORs and HRs) as well as its subgroups (**Figure 39** and **Figure 40**). The increase in risk from studies using HRs was smaller for PTB and mlPTB but was greater for vPTB when compared to studies using ORs. Some inconsistencies were present for trimester-specific exposures for PTB and its subgroups. Trimester 1 exposure was associated with an increased risk of vPTB only among studies reported HRs. Trimester 2 exposure was found to be associated with an increased risk of mlPTB from studies using ORs and HRs, and an increased risk of vPTB but only among studies that used ORs. Trimester 3 exposure was associated with an increased risk of PTB (in HRs and ORs), and an increased risk of mlPTB (in ORs).

Trimester 3 NO_2 exposure was associated with an increased risk of PTB (in ORs and HRs) (**Figure 41**). However, the risk was not replicated in vPTB (in HRs). No association was found for trimester 2 exposure. An increase in risk of vPTB in HRs was found for trimester 1 and whole pregnancy exposure, but it was not replicated using ORs.

No increase in risk of PTB or its subgroups was found to be associated with O_3 exposure. However, the findings were sensitive to the exclusion of Zhou et al. (2022), with an increase in risk of trimester 2 and whole pregnancy exposure being found after its exclusion.

Generally, there was higher certainty regarding the findings for PTB from studies reporting ORs due to greater number of studies available, compared to the findings for subgroups of PTB and those that used studies reporting HRs. There was also lower certainty in the evidence for O_3 exposure due to fewer studies available.

5.4.2. Summary of findings of the stratified analyses

A summary of the results of the stratified analyses is shown in **Table 46**. Most factors were suggested to be partial drivers of heterogeneity as reduction in heterogeneity shown by I^2 was observed in some but not all strata, or some but not all exposure windows. They may explain some of the heterogeneity observed in the main result, but its effect was not consistent across other exposure windows. Climate region explained much of the heterogeneity for studies conducted in China. However, it is highly uncertain if it is truly a driver of heterogeneity as no other strata were available for comparison. Some inconsistencies were found between pollutants. For instance, risk of bias was not suggested to be a driver of heterogeneity for $PM_{2.5}$ and O_3, but was a partial driver for NO_2 exposure. There was also large uncertainty in the findings for some factors with only 1 stratum that contained sufficient studies to be meta-analysed for most exposure windows. None of the factors was suggested to be a partial or major driver of the heterogeneity observed across all pollutants.

Table 46. Summary of the results of the stratified analyses, according to the potential strength of a factor being a driver of the heterogeneity observed. EA: Exposure assessment. (each): comparison of the respective exposure assessment methods.

Factor	Levels of the factor	$PM_{2.5}$	NO_2	O_3
Risk of bias	Low/Moderate/High	✗	✓ (Partial)	✗
	Confounder adjustment (All / Not all)			
Location	Regional	✓ (Partial)	✓ (Partial)	✗
	Country	✗		
	Climate region (US)	✓ (Partial)	✗	✓ (Partial)
	Climate region (China)	✓*		✗
Exposure assessment method	Monitor/Model-based	✗	✗	✓ (Partial)
	Individual methods	✓ (Partial)		✗
Concentration level	AQG values/Quartile	✓ (Partial)		
Data source	Administrative data / Research cohorts	✓ (Partial)		✗

✗: Not a driver of heterogeneity
✓ (Partial): Partial driver of heterogeneity
✓*: Driver of heterogeneity but high uncertainty was present

5.4.3. Implication of findings and comparison with preceding work

5.4.3.1. $PM_{2.5}$

A forest plot showing the increase in risk of PTB found in preceding meta-analyses as well as this work associated with $PM_{2.5}$ exposure is shown in **Figure 68.** The increase in risk of whole pregnancy exposure (from ORs) was consistent with preceding MAs, especially those published in the past 3 years, except for Wang et al. (2023) which restricted their analysis to studies conducted in China. The increase in risk derived from HR estimates were smaller than that of preceding MAs, but its range heavily overlapped with the 95% CIs of preceding works. Moreover, it showed a relative precise range of effect (HR: 1.05 [1.03 – 1.07]) despite being based on fewer studies than the preceding work. The increase in risk of PTB associated with trimester 3 exposure was consistent and in line with two recent publications (Ju et al., 2023; Yu et al., 2022), and in line with the findings of the narrative synthesis of the studies on the risk of PTB by gestational months or weeks, which showed greater risk of PTB at later stages of pregnancy (Chapter 4 section 4.3.3.5.4). There was no evidence of an association with trimester 2 exposure. However, the summary estimate (OR: 1.03 [1.00 - 1.06]) was very similar in of Ju et al. (2023) (OR: 1.03 [1.01-1.06]). The null finding in trimester 1 is consistent with all published MAs except Liu et al. (2017) which was based on fewer studies.

The association between $PM_{2.5}$ exposure and the risk of PTB subgroups was investigated by two recent published MAs (Ju et al., 2023; Yu et al., 2022). A forest plot comparing the results of the two publications and the meta-analyses conducted in this thesis is shown in **Appendix 23**. None of the works compared found an increase in risk of mlPTB for trimester 1 and 3 exposure, or risk of vPTB for trimester 3 exposure. On the other hand, an increase in risk of mlPTB for trimester 2 exposure, and the risk of vPTB for whole pregnancy exposure was found in all works compared. An increase in risk of mlPTB for whole pregnancy exposure was also found in the current meta-analysis and Yu et al. (2022). The increase in risk of mlPTB associated with trimester 2 exposure obtained from all works were very similar across all meta-analyses, whereas the increase in risk of vPTB for whole pregnancy exposure found in this work, regardless of the type of estimate (OR/HR), were much higher than those reported in Ju et al. (2023) or Yu et al. (2022). Mixed findings were present for the other exposure-outcome pairs. This suggests that there is some consistency in the findings of the current meta-analysis and past publications. However, Ju et al. (2023) and Yu et al. (2022) were based on cohort studies only and earlier cut-offs. Moreover, they suffered from the methodological issues mentioned in Chapter 3. For instance, Yu et al. (2022) claimed to have tackled study dependencies, but the exact approach was unclear. Hyder et al. (2014) and Kloog et al. (2012)

which were both geographically and temporally dependent were both present in Yu et al. (2022) but was removed in this work as the dependencies was deemed “intolerable”. Therefore, the current meta-analysis was able to provide more rigorous summary estimates that was based on studies of different study designs for the risks of PTB and its subgroups.

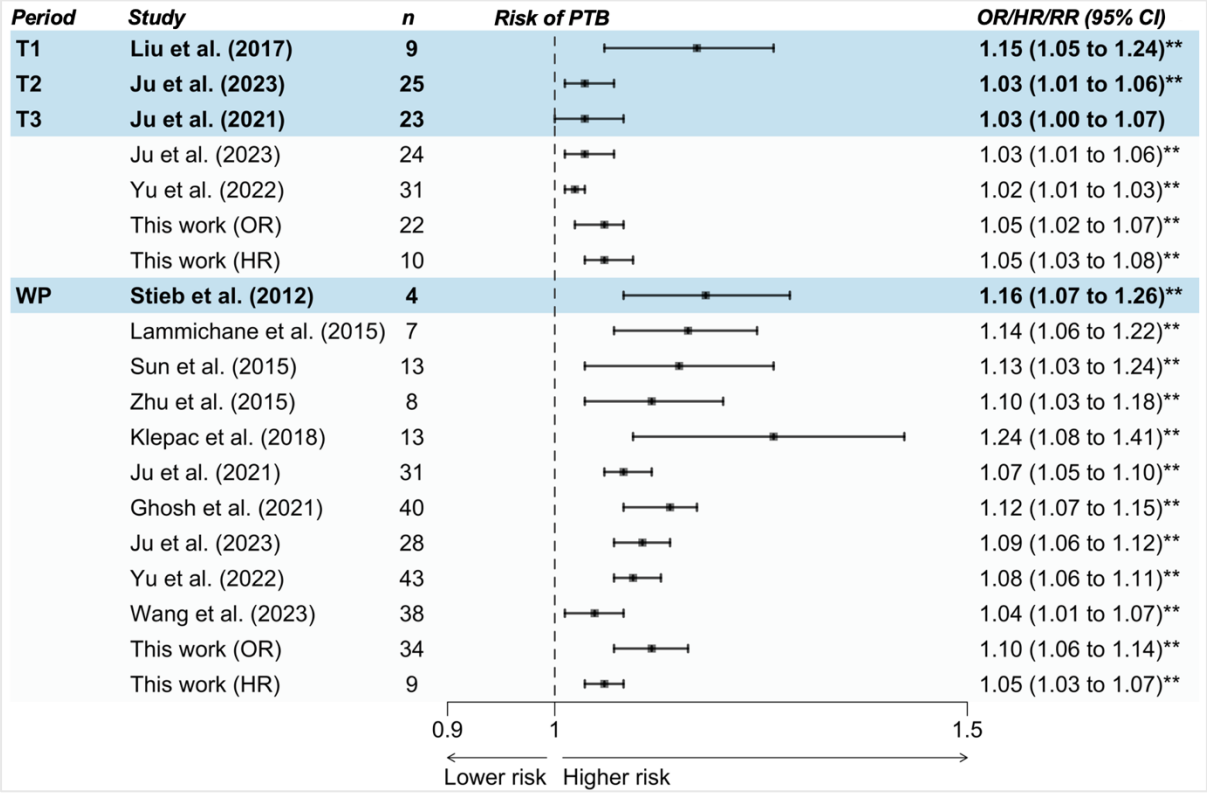


Figure 68. Forest plot of the results from preceding meta-analyses and this work for exposure periods that found an increase in risk of PTB associated with per 10 μg/m^3 increase in$PM_{2.5}$ exposure at trimester 1 (T1), 2 (T2), 3 (T3), or whole pregnancy (WP) exposure. n: number of studies. **: Summary estimate showing an increase in risk.

Two previously published meta-analyses have provided a summary estimate based on two studies for the effect of preconception $PM_{2.5}$ exposure (Blanc et al., 2023; Yu et al., 2022). The current meta-analysis identified essentially the same groups of studies with the addition of Mekonnen et al. (2021) (Chen et al., 2021b; Mendola et al., 2016; William et al., 2021). Nevertheless, meta-analysis was not carried out as there were <3 studies available after study selection and sorting of dependencies, but the conclusion drawn from reviewing the results of the studies were broadly consistent with other publications, showing an inconclusive result. Therefore, more studies are needed to ascertain the effect of preconception $PM_{2.5}$ exposure.

5.4.3.2. NO_2 and O_3

The effects of $PM_{2.5}$ exposure have received more attention in the literature compared to gaseous pollutants such as NO_2 and O_3. The effect of NO_2 exposure on the risk of PTB was explored in 5 meta-analyses (Ju et al., 2021; Klepac et al. 2018; Simoncic et al. 2020; Stieb et al. 2012; Wang et al., 2023). Similarly, there were 5 meta-analyses on the effect of O_3 exposure (Ju et al., 2021; Klepac et al. 2018; Stieb et al. 2012; Rappazzo et al., 2021; Wang et al., 2023). As Simoncic et al. (2020) used only studies from Europe, Rappazzo et al. (2021) focused on trimester 1 and 2 exposure, Wang et al. (2023) used only studies from China, Ju et al. (2021) only used cohort studies, and methodological shortcomings were present in Klepac et al. (2018), this makes the current work a major update of the overall literature landscape since Stieb et al. (2012) that was conducted more than 10 years ago. Forest plots showing the results from preceding work and the current meta-analyses are shown in **Figure 69** and **Figure 70** for NO_2 and O_3 exposure respectively.

Results for NO_2 exposure were broadly consistent with previous MAs, showing no association for trimester 2 and whole pregnancy exposure. The increase in risk of trimester 3 exposure shown in the current meta-analysis was much stronger than that of Ju et al. (2021), especially for the risk derived from studies reporting HRs. However, the latter was based on far fewer studies using HRs thus there was uncertainty present regarding its magnitude. The increase in risk from trimester 1 NO_2 exposure, based on only 4 studies using HRs and with a wide CI, has similar uncertainty. Moreover, it was not consistent with preceding publications, nor the result from studies reporting ORs in the current meta-analysis that found no association. The risk of PTB subgroups associated with NO_2 exposure were explored in Ju et al. (2021), but the authors found more than 3 studies for whole pregnancy exposure only. Thus, this is the first meta-analysis that has provided summary estimates for trimester-specific NO_2 exposures, and found an increase in risk of vPTB from trimester 1 exposure.

A US cohort study in Philadelphia found no association between whole pregnancy NO_2 exposure and the risk of spontaneous or medically indicated PTB (Escoto et al., 2022). Similarly, a study using the North Carolina Birth Cohort found generally null association with PTB for NO_2 exposure across all gestational weeks (Krajewski et al., 2023), supporting the lack of evidence for an association found for whole pregnancy NO_2 exposure and the risk of PTB found in the current meta-analysis. When viewed in conjunction with the findings for NO_2 exposure of the present study (Chapter 5 Figure 31), it is possible that the risk of PTB is associated with acute instead of cumulative or chronic gestational NO_2 exposure with its effect more pronounced in late pregnancy. Therefore, when whole pregnancy was used as

an exposure window and PTB was used as an outcome, the risk was diluted and hence not detected.

Jones et al. (2023) used a cohort in the southern US and found that NO_2 exposure was associated with increased risk of all subgroups of spontaneous PTB, with the risk becoming greater as gestational age at delivery became earlier. There is considerable uncertainty in the findings for the effect of NO_2 exposure and PTB subgroups in the present review. Whole pregnancy NO_2 exposure may increase the risk of PTB, but it is only evident among subgroups of PTB that were spontaneously delivered, as there were limited studies that compared the risk between those that were spontaneously delivered and medically indicated (Chapter 4 Figure 11), and their findings were inconclusive.

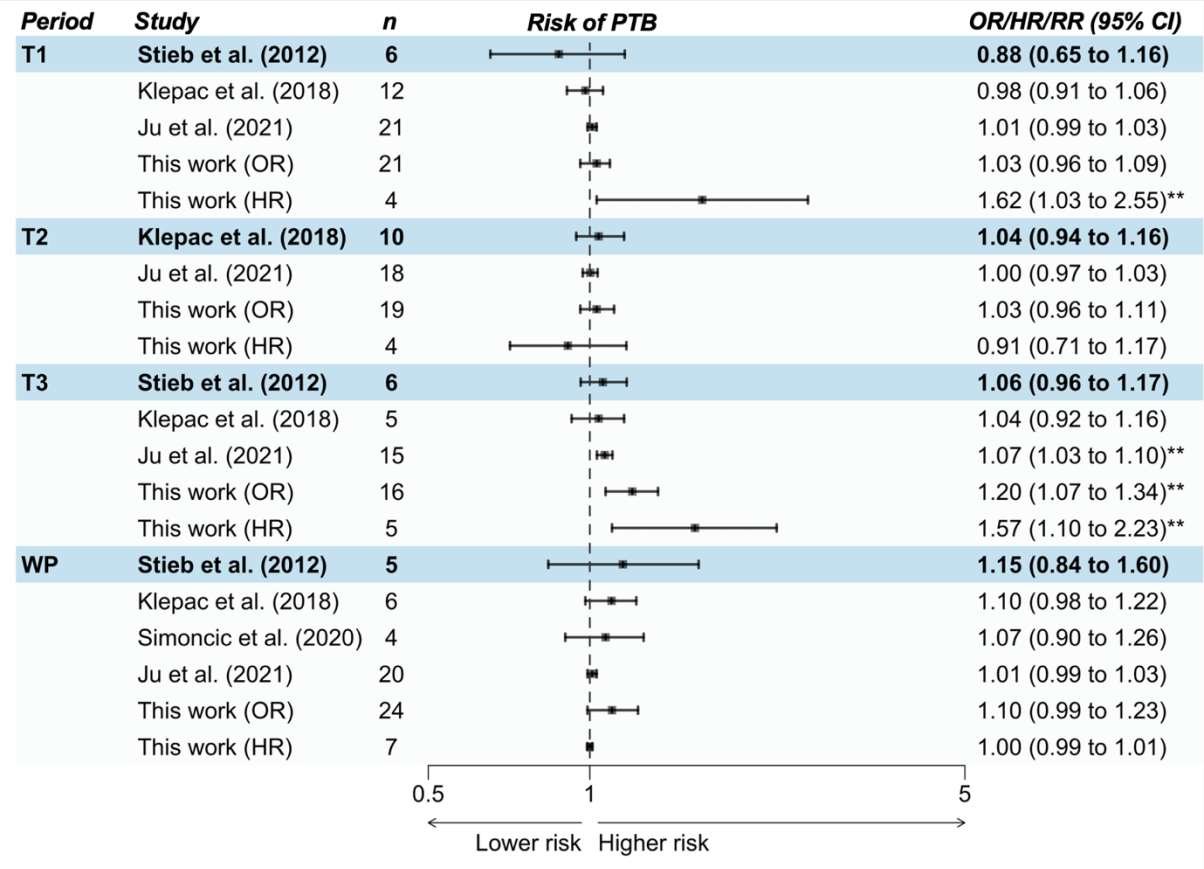


Figure 69. Forest plot of the results from preceding meta-analyses and this work for exposure periods that found an increase in risk of PTB associated with per 10 μg/m^3 increase in NO_2 exposure at trimester 1 (T1), 2 (T2), 3 (T3), or whole pregnancy (WP) exposure. n: number of studies

The null findings for the risk of PTB from O_3 exposure in the current meta-analysis were inconsistent with recent publications that find increase in risk across all exposure periods. None of the preceding publications except Stieb et al. (2012) considered the comparability of different O_3 measurement methods. Primary studies that measured exposures using mean or median of 24-hour averages typically report lower O_3 levels than those that used daily 8-

hour mean (10 am to 6 pm) or maximum. The associations reported from these different types of measurement are incomparable (Lange et al., 2018). Including them in the same meta-analysis would increase the uncertainty of the estimate, as the lower levels of exposure assigned may artificially elevate the risk estimate if the risk of PTB is associated with peak exposures. This may explain the differences in the findings between the current meta-analysis and recent publications such as Ju et al. (2021). Therefore, to ensure comparability of the estimates, it was decided to conduct separate meta-analyses for studies using different exposure metrics. This is the first meta-analysis that has provided summary estimates based on daily 8-hour mean (10am to 6pm) or maximum, which was different from previous meta-analyses that did not differentiate according to measurement metric, or used 24-hour averages (Stieb et al., 2012, with results for trimester 1 and 3 only). Future studies can consider reporting the findings in both types of metrics. More work is required to study the variability the conversion factor under different conditions (climate, location, variation in precursors over time). Therefore, authors of primary studies that chose to convert their

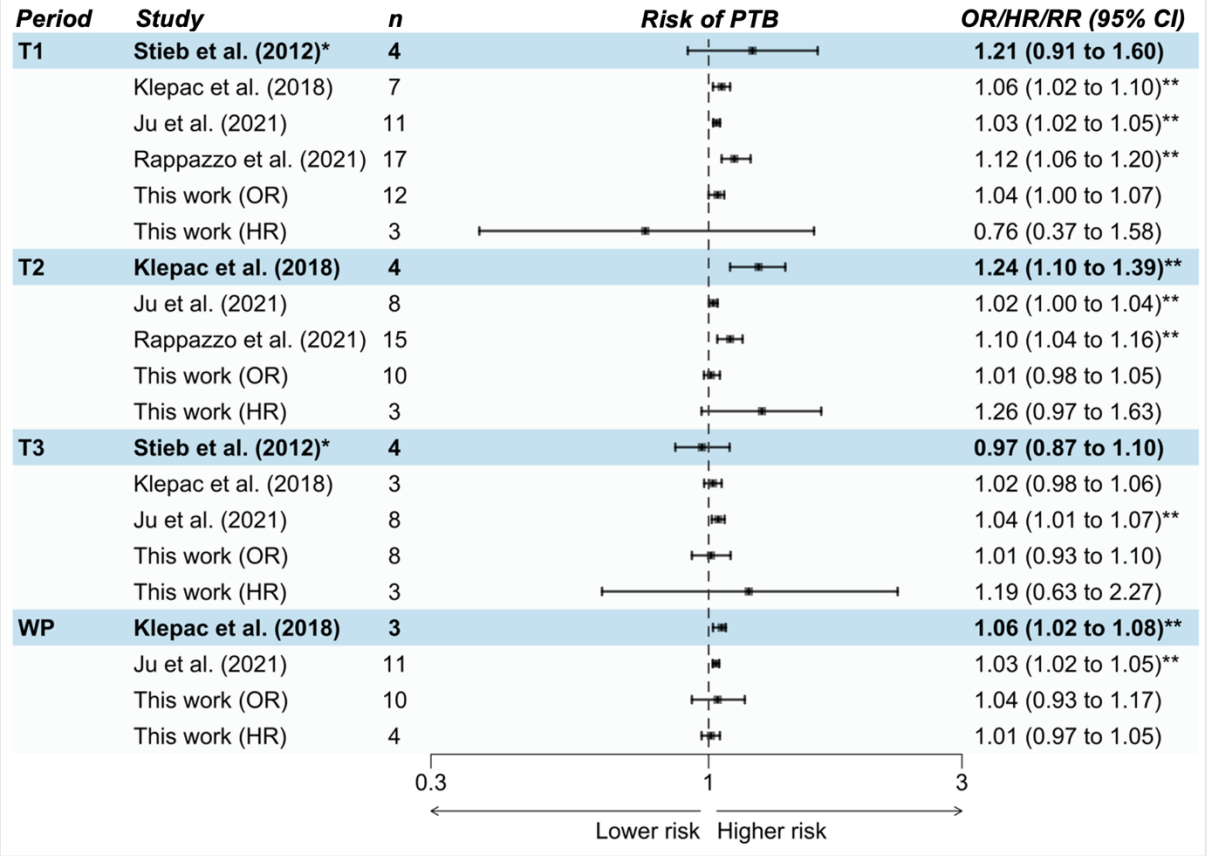


Figure 70. Forest plot of the results from preceding meta-analyses and this work for exposure periods that found an increase in risk of PTB associated with per 10 µg/m³ increase in O_3 exposure at trimester 1 (T1), 2 (T2), 3 (T3), or whole pregnancy (WP) exposure. n: number of studies. *: Stieb et al. (2012) only used studies that used 24-hour average O_3. Wang et al. (2023) was not included as it only used studies from China and studies that provided estimates in a per 10 µg/m³ increment.

findings or wish to aid others in converting their findings. can discuss the reliability of the conversion factor used in their work.

5.4.3.3. Stratification by risk of bias

This is the first meta-analysis that has used a quality assessment framework incorporating specific consideration of PTB, with preceding publications using relatively simple frameworks such as the Newcastle-Ottawa Scale (NOS) (Li et al., 2017), more comprehensive frameworks for observational studies (Yu et al., 2022), or self-set criteria (Stieb et al., 2012) (more details on their shortcomings in Chapter 3 section 3.3). The results of their quality assessment suggested that the primary studies used were mainly of high quality. Exclusion of studies of low quality or high risk of bias (RoB) thus had little effect on the summary estimate and its heterogeneity found in their work. Findings for $PM_{2.5}$ and O_3 exposure in the current meta-analysis were broadly consistent, showing no evidence that RoB or confounder adjustment contributed to the heterogeneity observed. However, they were suggested to be partial drivers for NO_2 findings (**section 5.3.7.1**).

Only one other meta-analysis has explored the impact of confounder adjustment (Ghosh et al., 2021), suggesting that accounting for confounder adjustment increased the summary estimate and is a potential source of bias and heterogeneity. The difference in findings of the current meta-analysis could be due to differences in methodological choices. Both meta-analyses used a dichotomised approach: Ghosh et al. (2021) categorised confounder adjustment as “sufficient” or “insufficient”, and this meta-analysis categorised levels of confounder adjustment as “adjusted for all critical confounders” or “did not adjust for all critical confounders. In the current meta-analysis, confounder adjustment is a sub-stratification of the RoB level, which likely reduced the number of studies available to examine its impact.

The differences in the findings for NO_2 exposure, compared to those of $PM_{2.5}$ and O_3, are difficult to interpret. Studies assigned a low RoB found a decrease in risk of PTB associated with trimester 1 and 2 NO_2 exposure, which their 95% CIs barely overlapped with that of the null findings of the main results. In addition, the increase in risk found for trimester 3 exposure dissipated among studies assigned a low RoB, but remained among studies of moderate RoB.

As studies assigned a low RoB were from varied locations, used different sources of data, and of different concentration levels, these are unlikely the reason behind the decrease in risk observed. Studies were predominantly moved from low to moderate RoB due to

shortcomings in confounder adjustment or exposure assessment. One way of interpreting the results is that studies of moderate RoB may have used exposure assessment methods that have systematically overestimated NO_2 exposure, leading to elevated risk estimates. However, in the stratification by individual exposure assessment methods, those that used more comprehensive or model-based exposure assessment methods did not necessarily show a decrease in risk, albeit there were few studies available in the strata. This makes exposure assessment unlikely the factor driving the decrease in risk observed. Primary studies that have reported a decrease in risk offered different explanations, such as the impact of spatial variations (hospitals with higher air pollution concentrations have lower rates of spontaneous PTBs) (Johnson et al., 2016), inverse correlation with O_3 (Lin et al., 2015), or the impact of co-pollutant exposures with $PM_{2.5}$ (Smith et al., 2020). The last explanation echoes the findings of section 4.4.3.2.1, which the effect of NO_2 exposure could be partly driven by $PM_{2.5}$. If $PM_{2.5}$ exposure has a stronger effect on PTB in early pregnancy than NO_2, then an apparent protective effect can be observed. This suggests the need of adopting co/multi-pollutant models for NO_2 to disentangle its effect from other pollutants

5.4.3.4. Stratified analysis by location

Location at a broad level (WHO region) was a partial driver of heterogeneity for the association between $PM_{2.5}$ and NO_2 exposure and the risk of PTB (**section 5.4.3.4**). The reduction in heterogeneity observed for WHO regions was in the stratum for European studies. The studies included were mainly those in Western Europe, which were relatively homogeneous compared to studies that were included in the strata for Western Pacific or the Americas, which could differ drastically in terms of population characteristics or healthcare policies. Countries in Europe were also of cold temperate climate with similar seasonal patterns of pollutants (Aas et al., 2024; Metzger et al., 2013), while the US and China are large countries that themselves cover four or more climate zones (Beck et al., 2018). For O_3 exposure, there were insufficient studies to be meta-analysed in the stratum for European studies, so no decrease in heterogeneity was observed at all at WHO region level. The present meta-analysis found no evidence of an association between whole pregnancy NO_2 exposure and the risk of PTB (OR: 1.07 [0.98-1.18], I^2 = 51%, n = 8). The summary estimate was identical to that of Simoncic et al. (2020) (OR: 1.07 [0.98 - 1.28], I^2 = 72%, n = 4), but there was less heterogeneity shown by I^2.

Location at country level was not a driver of heterogeneity observed across all pollutants. However, changes in summary estimates occurred when restricted to the US or China. For instance, the increase in risk of trimester 3 $PM_{2.5}$ exposure was driven by Chinese studies,

where it dissipated for studies in the US. An increase in risk was also found for trimester 2 $PM_{2.5}$ exposure for Chinese studies, which was different from the null finding in the main analysis.

Climate region in the US was a partial driver of the heterogeneity observed for $PM_{2.5}$ and O_3 but not NO_2 exposure. The strongest reduction in heterogeneity was mainly in the western region of the US, which could be due to all of them being from the same state (California). However, the findings were less consistent for other climate regions in the US. Climate region in China was a partial driver of the heterogeneity observed for $PM_{2.5}$ exposure among studies reported HRs, with studies conducted in the “hot summer cold winter” (HSCW) region reported drastically lower heterogeneity in most exposure windows. California and the HSCW region share similar range of temperature in summer and winter, suggesting that having a hot summer may contribute to the association found, as there tend to be higher levels of $PM_{2.5}$ or O_3 being formed. There could be modifying effects of high temperature or heatwaves, which was suggested by the results of the narrative synthesis in Chapter 4 section 4.3.3.2.4. However, these strata did not report a greater risk than others, and for some exposure-outcome pairs only one stratum had enough studies to be meta-analysed. Hence, more studies are needed to ascertain the findings.

5.4.3.5. Stratified analysis by exposure assessment method

Exposure assessment method was found to be a partial driver of heterogeneity observed in the association between all pollutants and the risk of PTB (**section 5.3.7.3)**. The heterogeneity was mostly driven by the differences in respective exposure assessment methods, rather than if the method was model-based or monitor-based.

The stratum for LUR models had generally reported much lower heterogeneity shown by I^2 compared to the other exposure assessment methods for $PM_{2.5}$ and NO_2 exposure (there were insufficient studies for O_3 exposure). LUR models may produce more accurate exposure estimates of the effect of $PM_{2.5}$ and NO_2 compared to other methods. They are more robust than monitor-based approaches as they can accurately capture spatial variabilities and incorporate multiple data sources and variables (Ryan & LeMasters, 2007; Zhang et al., 2021b). Similar advantages can be present for other models such as dispersion models, and the current meta-analysis cannot confirm if such observation would be present if there were sufficient (>3) studies available for the strata for dispersion models. On the other hand, studies that used LUR models may share other underlying similarities that were not measured.

Two preceding MAs have explored the impact of exposure assessment methods. Ghosh et al. (2021) suggested that adjustment for exposure assessment methods increased the summary estimate, whereas Yu et al. (2022) showed no substantial changes. Results from this meta-analysis were inconclusive, showing that large fluctuations in the summary estimates were present for the respective exposure assessment methods, partly due to the relatively few studies available in each stratum. The finding was consistent with the finding in Chapter 4 section 4.3.3.1.1 which suggested that the method of exposure estimation may contribute to differences in estimates derived.

It is difficult to detect the reduction in heterogeneity in the stratum for studies that used hybrid models, as the nature of hybrid models mean that the estimates may inherently be more heterogeneous due to the different combinations of models available, which may explain the high heterogeneity observed in the strata of hybrid models. Nonetheless, hybrid models overcome the limitations of respective exposure assessment methods, such as the low ability for LUR models to capture temporal variations by incorporating atmospheric dispersion models that can offer better temporal resolution, and thus should be preferred in future studies.

5.4.3.6. Stratified analysis by concentration levels

Mean/median pollutant concentration over the course of pregnancy was found to be partial drivers of the heterogeneity observed. Reduction of heterogeneity mainly occurred in strata of lower ranges of concentrations for all pollutants (details in **section 5.3.7.4**).

For $PM_{2.5}$, in the range of mean exposure available, the risk of PTB followed a U-shaped distribution, with a higher risk in <10 μgm^{-3}/Q1 and >50 μgm^{-3}/Q4 (**Figure 57** and **Figure 58**). The higher risk found above 50 $\mu g/m^3$ was consistent with the preceding publication which plotted a risk-response curve suggesting that the relationship increases curvilinearly then either levels off or continues to increase above 50 μgm^{-3} (Chen et al., 2021a; Li et al., 2021; Sun et al., 2019). The greater risk found at lower levels (<10 $\mu g/m^3$) were counterintuitive, but recent findings regarding air pollution's general health effects showed that the relationship can be supralinear and persist below 10 $\mu g/m^3$ (Chen & Hoek, 2020; Lee et al., 2022b; Papadogeorgou et al., 2019). Other causes could be reduced confounding from other pollutants or environmental factors, population characteristics, or statistical technique used which could contribute to the increased risk observed at lower concentrations (COMEAP, 2022). Though it is difficult to capture exposure contrast at low concentrations, the finding was unlikely to be caused by measurement error, as it has been shown to have an attenuating instead of exponentiating effect on estimates (Samoli et al., 2020). This

finding potentially reaffirms the need for a lower threshold for $PM_{2.5}$ levels at 5 μg/m^3 annual average set by the 2021 WHO Air Quality Guidelines (Hoffmann et al., 2021).

Compared to $PM_{2.5}$, the pattern found for different levels of NO_2 exposure was less clear, with the risk estimate following an inverted U- or S-shape relationship in the range of mean exposure available. The finding was broadly consistent with previous publications suggesting a nonlinear or J-shaped relationship (Huang et al., 2020; Llop et al., 2010; Sun et al., 2019). The increase in risk of trimester 3 from exposure to NO_2 persisted in the range of 10 to 30 μg/m^3 (Q2) and >50 μg/m^3 (Q4) but not >30 to 50 μg/m^3 (Q3), with those at the highest levels of exposure showing the greatest risk (OR: 1.34 [1.01 - 1.77]). The increases in risk echo the necessity to reduce the WHO guideline levels from 40 to 10 μg/m^3 annual average However, only one exposure window (whole pregnancy) had enough studies at the <10 μg/m^3 range. Therefore, it is uncertain if the risk of PTB found for trimester 3 exposure would persist in this range. Findings from **section 4.3.3.2.1** as well as the wider literature suggested that the effect observed for NO_2 exposure could be partly driven by $PM_{2.5}$ (COMEAP, 2018a). Therefore, it is possible that its dose-response relationship with PTB could be better elucidated using estimates derived from co/multi-pollutant models.

5.4.3.7. Stratified analysis by source of data

Source of data was a partial driver of the heterogeneity observed for $PM_{2.5}$ and NO_2 exposure, whereas it was inconclusive for O_3 exposure as the studies were predominantly retrospective cohorts that used administrative data. For data sources, moderate to strong reduction in heterogeneity occurred in the stratum for research cohorts but not administrative data sources, suggesting that source of data explained some of the heterogeneity in the main result in that exposure window. However, uncertainties remain for the findings as there were far fewer studies in the strata with low heterogeneity. Nonetheless, it is consistent with the findings of two preceding MAs which explored the impact of study designs or data sources (Ghosh et al., 2021; Yu et al., 2022).

High heterogeneity may have remained for the stratum of administrative data sources because of the underlying variation of administrative data sources available, which could be in terms of available information, data quality, or record keeping practices. For instance, some may have collected information on smoking while some did not. Most countries included in the present meta-analysis are middle to high-income countries with birth certificates and registries with high completion rates and coverage, making them a reliable source of information with low levels of missing data (Laurent et al., 2016; Nealy et al., 2018). Still, future meta-analyses incorporating studies from low-to-middle-income countries

with varying degrees of comprehensiveness in these data sources might show differences (Bhatia et al., 2019). Another reason could be poorer adjustment of confounders such as smoking, which could be present for the stratum of administrative data source, compared to those that used research cohorts.

The stratified analyses suggested that data obtained from research cohorts likely more consistent with less variability than that of administrative data sources or retrospective cohorts. Thus, it highlights the need of high-quality prospective cohorts to investigate the impact of air pollution exposure on the risk of PTB, and their data should be made available for future studies to be used retrospectively.

5.4.4. Findings from ORs and HRs: advantages and shortfalls

The findings from HRs and ORs were mainly consistent in the current meta-analysis, although there were differences in magnitude of summary estimate depending on the pollutant. The results derived from ORs compared to HRs differed in two instances: for whole pregnancy $PM_{2.5}$ exposure, the summary estimate for OR is greater than that of HR, whereas the opposite pattern was present for trimester 3 NO_2 exposure). As discussed in **section 5.2.1**, there are advantages of using HRs over ORs. Moreover, ORs overestimates RR when RR is greater than 1 (Mittinty & Lynch, 2023). Among the three exposure-outcome relationships that found an increase in risk of PTB from studies reporting HRs as well as ORs (trimester 3 $PM_{2.5}$, whole pregnancy $PM_{2.5}$, and trimester 3 NO_2 exposure), no pattern of overestimation was found.

The interpretation of the two types of effect estimates is different. For whole pregnancy $PM_{2.5}$ exposure, the finding based on ORs was expressed as 10% greater odds of PTB, whereas the finding based on HRs showed 5% greater risk/hazard/rate of PTB per 10 $\mu g/m^3$ increase in exposure, over the course of pregnancy. In terms of usage in heath impact assessments (HIA), it can be argued that concentration-response functions based on HRs are preferred over ORs, as HIAs aimed to understand future health impacts over a certain time periods, and HRs are time-based estimates as well. Hence, the increase in risk of whole pregnancy $PM_{2.5}$ exposure expressed in HR can potentially be used in HIAs. However, it is based on far fewer studies compared to the summary estimate based on ORs. Hence, more future studies are needed to ascertain if the group of studies have showed distinctly different results.

5.4.5. Causality discussion based on the Bradford Hill criteria

Two exposure-outcome pairs were analysed using the Bradford Hill criteria (Chapter 1 Table 4) to determine if there is sufficient evidence supporting the presence of causality. The first is the increase in risk of PTB associated with per 10 μg/m^3 increase in whole pregnancy $PM_{2.5}$ exposure (OR: 1.10 [1.06 - 1.14]), as it was based on the largest number of studies. The second was increase in risk of PTB associated with per 10 μg/m^3 increase in trimester 3 NO_2 exposure (OR: 1.20 [1.07 - 1.34]), as it was the only association showing an increase in risk of PTB out of all the exposure windows examined for NO_2. The increase in risk for whole pregnancy $PM_{2.5}$ exposure based on HR was not used as it was based on relatively few studies. A summary of the results is shown in **Table 47**. Each criterion is discussed in the following sections. Phrasing from the USEPA Integrated Science Assessments (ISA) were used for the conclusion of the causality determination to facilitate comparison (Owens et al., 2017). The definition of the conclusions is made available in **Appendix 24**.

Table 47. Summary of the results of the causality discussion based on Bradford Hill criteria.

<table>
<tr><th rowspan="2">Criteria</th><th colspan="2">Pollutant</th></tr>
<tr><th>$PM_{2.5}$</th><th>NO_2</th></tr>
<tr><td>Temporality</td><td colspan="2">✓</td></tr>
<tr><td>Specificity</td><td colspan="2">N/A</td></tr>
<tr><td>Experiment</td><td colspan="2">Inconclusive</td></tr>
<tr><td>Plausibility</td><td colspan="2">✓#</td></tr>
<tr><td>Strength</td><td colspan="2">✓#</td></tr>
<tr><td>Coherence</td><td rowspan="2">✓</td><td rowspan="3">Suggestive</td></tr>
<tr><td>Analogy</td></tr>
<tr><td>Consistency</td><td>✓#</td></tr>
<tr><td>Biological gradient</td><td>✓</td><td>✓#</td></tr>
<tr><td>Conclusion</td><td>“Likely to be a causal relationship”</td><td>“Suggestive of, but not sufficient to infer, a causal relationship”</td></tr>
</table>

✓#: criteria fulfilled but with caveats (Details in the following sections)

5.4.5.1. Temporality

The criterion of temporality was fulfilled for all pollutants based on the findings of the work conducted in this thesis. Most studies used a retrospective cohort design (Chapter 4 Table

1) that did not follow the study population as the outcome developed. However, a temporal sequence was established by the use of monitoring data collected before delivery, or exposures modelled before delivery. Data from birth certificates or registries regarding date of birth is relatively accurate, hence the temporal sequences were usually deemed reasonably accurate.

5.4.5.2. Specificity

The criterion of specificity may not be applicable to the association between air pollution exposure and the risk of PTB. It is because the health effect of air pollution exposure is extensive and complex. Moreover, the risk of PTB is known to be influenced by various factors (more details in Chapter 1 section 1.3.2.1 and 1.3.2.2), thus it is challenging to establish a specific pathway that connects air pollution exposure and the risk of PTB, but not other health outcomes. While specificity to particular pollutants can be investigated using multipollutant models (MPMs) with some caveats to their interpretations, it is in any case possible for more than one pollutant to have a causal relationship with PTB, making specificity in terms of exposure hard to establish. There are chamber studies that looked into the effect of specific air pollutants or mixtures of air pollutants (Madden et al., 2014; Orach et al., 2021), but to the best of my knowledge, there are none on pregnant women due to ethical concerns (Rom et al., 2013). Therefore, other criteria would be more applicable in the causality discussion.

5.4.5.3. Experiment

In the context of air pollution exposure and the risk of PTB, the criterion of experiment refers to the presence of evidence of the effect of the implementation of air quality controls or regulations that subsequently result in decreased risk of PTB. The systematic review conducted in this thesis (Chapter 4 section 4.3.3.5.3) retrieved two natural experiment studies (Huang et al., 2015; Liu et al., 2022b). Neither found an association between $PM_{2.5}$ exposure and the risk of PTB, but one of them found an association between NO_2 exposure and the risk of PTB.

As the evidence available to date is scarce, it is insufficient to support the presence or absence of a causal relationship between $PM_{2.5}$ and NO_2 exposure and the risk of PTB. More future studies are needed to ascertain the presence of a causal relationship. However, there are methodological challenges in using quasi-experimental approaches to investigate the exposure-outcome relationship of interest. First, the changes in air pollution exposure

may not be exogenous, therefore correlates with confounding factors such as socioeconomic status, leading to a biased estimate of the causal effect. Second, changes in air pollution avoidance behaviour could be triggered by the intervention, thus the effect of the air quality intervention could be overestimated if changes in such behaviour were not adjusted for. Lastly, defining an appropriate control group may be challenging, as air quality interventions may span across neighbouring regions that share similar population characteristics, making it difficult to obtain a well-matched population group that did not experience the intervention to be used in the analysis. Therefore, caution must be exercised when interpreting findings from studies that used quasi-experimental approaches.

5.4.5.4. Plausibility

The criterion plausibility was fulfilled (with some caveats) for the increase in risk of PTB associated with $PM_{2.5}$ and NO_2 exposure. A plethora of mechanistic evidence has shown that air pollution exposure was associated with inflammatory responses, leads to higher oxidative stress, induces epigenetic changes, and results in endocrine disruption in pregnant women. These adverse physiological responses were in turn associated with greater risk of PTB (more details in Chapter 1 section 1.3.2.1). Furthermore, animal and *in vitro* studies found adverse effects such as changes in placental morphology, increased oxidative stress and inflammatory responses associated with $PM_{2.5}$ or NO_2 exposure, although the doses used were much higher than the usual levels of human exposure (Fussell et al., 2024; Veras et al., 2008). Fussell et al. (2024) concluded that the mechanistic evidence available suggested air pollution exposure could affect a variety of processes during pregnancy and leads to adverse birth outcomes. However, no clear pathway is currently shown to link all events that could occur in the pathway between air pollution exposure and the risk of PTB, thus the greater risk of PTB was often inferred by investigating the potential mechanisms. Nonetheless, this criterion does not require a fully defined mechanistic pathway between the exposure and the outcome, but rather a reasonable biological explanation, which was fulfilled here.

The evidence could be strengthened in two ways. First, more studies or cohorts should be designed to integrate mechanistic and epidemiological investigations. Currently, studies investigating effects of air pollution exposure on biomarkers or physiological processes do not have PTB as an outcome, whereas epidemiological studies would only focus on the exposure and the outcome without considering the processes in between. Through this, candidate mechanisms aforementioned can potentially be linked to PTB and outline a pathway of action. Second, metabolomic/proteomics research has several advantages over

traditional mechanistic studies and could offer valuable insights. It allows global profiling of metabolites which leads to unbiased discovery of candidate biomarkers, has the ability to capture complex biological interactions which are present in the relationship between air pollution exposure and the risk of PTB, and is less prone to confounding (Liang et al., 2023). For instance, significantly higher glycoprotein N-acetyl resonance, a marker of inflammation, was found in women who delivered pre-term (Fussell et al., 2024; Maitre et al., 2014; You et al., 2024). Nevertheless, the field of metabolomic research is still rapidly evolving, with challenges in the standardisation of methods, generalisability, specificity to air pollution, and ways to integrate the findings with mechanistic and epidemiological evidence.

5.4.5.5. Strength

The strength of the increase in risk of PTB found per 10 $\mu g/m^3$ increase in whole pregnancy $PM_{2.5}$ exposure (OR: 1.10 [1.06 - 1.14], n = 35) or trimester 3 NO_2 exposure (OR: 1.20 [1.07 - 1.34], n = 16) were statistically significant and broadly consistent in size with all-cause mortality which has established links with air pollution exposure (RR: 1.08 [1.06 - 1.08], per 10 $\mu g/m^3$ increase in annual average $PM_{2.5}$) (Chen & Hoek, 2020), or risk factors that are known to be associated with increased risk of PTB, such as higher maternal age or parity (Fuchs et al., 2018; Koullali et al., 2020). It is much smaller than the effect for other stronger risk factors for PTB, such as smoking or amphetamine use (Ladhani et al., 2011; Liu et al., 2020), but the population exposed for air pollution is much larger. The effect of bias and confounding, when not being adequately accounted for, could lead an over- or underestimation of the association, and could change how a summary estimate was interpreted when the increase in risk is close to null. For instance, there could be underreporting/misreporting of prevalence of smoking among pregnant women. Hence, the observed association could be due to, or an attenuated estimate of smoking, if smoking is associated with exposure to air pollutants. Nonetheless, the risk was relatively consistent across various stratifications and supported by studies of other designs (more details in the coherence section), and showed a potential concentration-response relationship.

5.4.5.6. Coherence

The increase in risk found to be associated with $PM_{2.5}$ or NO_2 exposure was in line with the existing knowledge available on the impact of their exposure on a variety of other pregnancy outcomes, such as low birth weight (LBW), foetal growth restriction (FGR), stillbirth, miscarriage, or premature rupture of membranes (PROM).

Similar to PTB, considerable epidemiological evidence has suggested a link between air pollution exposure and increased risk of birth outcomes, such as LBW ($PM_{2.5}$ and NO_2) (Simoncic et al., 2020), and stillbirth ($PM_{2.5}$ only) (Zhang et al., 2021a). Outcomes such as LBW and PTB could occur concurrently in an infant (Pusdekar et al., 2020; Younes et al., 2021). Thus, it is possible that they are caused by similar mechanistic pathways, some of those mentioned in section 5.4.5.4. Preterm PROM occurs in 3% of pregnancies and results in approximately 40% of PTB (Mercer, 2003). It is thus a strong precursor of PTB, and its risk has been shown to be linked with increased exposure to $PM_{2.5}$ and NO_2 (Li et al., 2017; Liang et al., 2024; Wallace et al., 2016). As the evidence regarding $PM_{2.5}$ exposure and these linked health outcomes was more established than that for NO_2 exposure, the criterion of coherence was fulfilled for $PM_{2.5}$ exposure, but was only suggestive for NO_2 exposure.

5.4.5.7. Analogy

While the argument for analogy was supported by associations found in studies of short-term exposure, or indoor or household air pollution for both pollutants, the evidence for $PM_{2.5}$ exposure was strengthened further by the findings from occupational and industrial settings, wildfire exposures, smoking, as well as constituents of $PM_{2.5}$. Hence, the criterion of analogy was fulfilled for $PM_{2.5}$ exposure, but suggestive for NO_2 exposure.

There were time-series studies that investigated the short-term effects of $PM_{2.5}$ and NO_2 exposure and found an increased risk of PTB (Zhang et al., 2020; Yang et al., 2023). In addition, $PM_{2.5}$ and NO_2 can be generated in indoor or household environments through the use of household cooking appliances, heating, or cleaning (Tran et al., 2020). Though no studies identified have directly linked indoor NO_2 exposure to the risk of PTB, some have found linkage between use of cooking stoves that generate NO_2 and elevated risk of PTB (Kephart et al., 2021; Younger et al., 2022).

Smoking, as well as second-hand smoke exposure, has a well-established causal link with greater risk of PTB (Liu et al., 2020; Selvarathnam et al., 2023). It provides a useful analogy as some of the harmful substances inhaled during smoking are found in the ambient air, such as $PM_{2.5}$, NO_2, polycyclic aromatic hydrocarbons (PAHs), and CO (Braun et al., 2019; Vu et al., 2015; Wald et al., 1981). Similarly, the emerging evidence for an effect of wildfire-generated $PM_{2.5}$ on the risk of PTB strengthened the analogy argument, as it shares similar constituents with ambient $PM_{2.5}$, although higher toxicity has been observed (Chapter 4 **section 4.3.3.2.2**) (Jiang et al., 2024; Picciotto et al., 2024). Moreover, constituents of ambient $PM_{2.5}$, such as, PAH, heavy metals, and ammonium are shown to be associated with an increased risk of PTB (Chapter 4 **section 4.3.3.2.3**). These constituents, in

occupational or industrial settings, have been demonstrated to increase the risk of PTB as well (Adane et al., 2023; Veber et al., 2022) but at higher concentrations.

5.4.5.8. Consistency

The criterion of consistency was fulfilled for $PM_{2.5}$ but remained suggestive for NO_2 exposure due to the following reasons.

First, there was greater consistency for the increase in risk found for whole pregnancy $PM_{2.5}$ exposure found across trimesters, statistical method used (OR/HR), and subgroups of PTB (**Figure 39** and **Figure 40**). The finding was replicated in trimester 3 for both statistical methods, whereas the finding for trimester 3 NO_2 exposure was not replicated in any other exposure periods except in trimester 1 using studies reporting HRs. Moreover, the increase in risk remained in all subgroups of PTB for whole pregnancy $PM_{2.5}$ exposure, whereas it was not replicated in very PTB for trimester 3 NO_2 exposure, with insufficient studies available for the other PTB subgroups.

Second, the increase in risk for whole pregnancy $PM_{2.5}$ exposure persisted among studies that were assigned a low risk of bias, or used data from research cohorts. However, the increase in risk found for trimester 3 NO_2 exposure dissipated after the same stratified analyses, and no increase in risk was found for whole pregnancy NO_2 exposure. It implied that the findings for $PM_{2.5}$ exposure were more consistent and remained robust among studies that were of better quality or design. Nonetheless, it is worth noting that between-study heterogeneity remained high in most instances, with stratifications only able to reduce heterogeneity in certain strata. It is possible that some of the heterogeneity was from “true” heterogeneity, i.e. differences in time period, pollution mixtures, follow-up times, or age structure of the studies (Boogaard et al., 2023), but it was likely that some unexplained heterogeneity was present due to reasons such as unmeasured confounding. High heterogeneity does not necessarily reduce the confidence in the finding, but it could reduce the confidence in applying the finding in HIA, as the summary estimate (and 95% CI) may not be generalisable to the population of interest.

Third, the studies that have used multipollutant models (MPM) to disentangle the effect of individual pollutants have found that the findings for $PM_{2.5}$ exposure were robust to adjustment for NO_2 in most cases (Chapter 4 **section 4.3.3.2.1**, Figure 5). Yet, the increase in risk found for NO_2 exposure dissipated after adjustment for $PM_{2.5}$ in some cases, showing that the findings derived for $PM_{2.5}$ were more consistent after accounting for other pollutants. Similarly, studies that compared successive pregnancies found increased risk of PTB for

pregnancies exposed to higher $PM_{2.5}$ levels among the same mothers, but there were no studies available for NO_2 exposure (**section 4.3.3.5.2**) However, the findings were inconclusive due to the few studies available.

5.4.5.9. Biological gradient

The criterion of biological gradient was fulfilled as the increase in risk of PTB found in the current meta-analysis for whole pregnancy $PM_{2.5}$ and trimester 3 NO_2 exposure were inherently based on continuous exposure (per 10 μg/m^3 increase). However, it is possible that the strength, magnitude or even the direction of the dose-response relationship can change under different ranges of concentration. In addition, the results of the stratified analysis by pollutant concentration (**section 5.3.7.4**) and studies of categorical exposures (**section 4.3.3.4.1**) both found more consistent evidence for a greater risk of PTB associated with $PM_{2.5}$ exposure at lower (<10 μg/m^3) and higher concentrations (>50 μg/m^3), compared to NO_2 exposure where the pattern was less clear.

5.4.5.10. Result of the causality discussion

When all criteria were viewed in conjunction, the strength of the risk of PTB associated with whole pregnancy $PM_{2.5}$ exposure was in line with other adverse effects of air pollution, and the finding was mostly consistent across studies of different study designs, co-pollutant adjustment, and stratification such as risk of bias, although limited reduction in heterogeneity was observed. There was a plethora of mechanistic evidence available, but this did not demonstrate a clear link between air pollution exposure and the risk of PTB. Lastly, the findings from the small number of natural experiments were inconclusive. The increase in risk of PTB found for whole pregnancy $PM_{2.5}$ exposure was therefore determined to be "*likely to be causal*".

For trimester 3 NO_2 exposure, it was deemed to be "*suggestive of, but not sufficient to infer, a causal relationship*" (**Table 47**). It suffered from similar shortcomings as $PM_{2.5}$ exposure, with the additional issue of having less support for a causal relationship in several Bradford Hill criteria, as its analysis was based on fewer studies available and of greater uncertainty. Therefore, the conclusion could shift with emergence of new epidemiological or mechanistic evidence. The conclusion for NO_2 exposure may explain the reason why health impact assessments have been conducted for the impact of particulate matter exposure and reduction on preterm birth, but to the best of my knowledge, have not yet been conducted for gaseous pollutants (Kim et al., 2019; Pereira et al., 2021; Trasande et al., 2016). Therefore,

in the following chapter, a health impact assessment was conducted for the effect of whole pregnancy $PM_{2.5}$ exposure only. However, as other health outcomes included in health impact assessments often have much more established mechanistic pathways, e.g. cardiovascular and respiratory outcomes (USEPA, 2019), it is recommended that preterm birth be incorporated in future health impact assessment as a sensitivity analysis/add-on.

5.4.6. Strength and limitations

To the best of my knowledge, this is the first meta-analysis of the area that has included in-depth consideration into issues of study dependencies, performed meta-analysis separately for studies using ORs and HRs, and conducted quality assessment using a tailored framework for studies of the area. The meta-analyses contained fewer studies than some preceding publications for these important reasons. The findings of this meta-analysis reaffirmed the results of previous meta-analyses that found increased in risk of trimester 3 and whole pregnancy $PM_{2.5}$ exposure, as well as trimester 3 NO_2 exposure. moreover, it is novel as it is the first meta-analysis which found an increase in risk of trimester 3 and whole pregnancy $PM_{2.5}$ exposure among studies that used HRs. Comprehensive stratified analyses were conducted which yielded valuable insights on potential contributors of heterogeneity and improvements future studies could make. A discussion on the presence of causality was also made through integrating evidence from previous chapters, which allow findings outside the meta-analyses conducted to be used. Despite these advantages, some limitations remain.

First, high heterogeneity was present for most meta-analyses conducted in the primary analyses, and it remained high even after most sensitivity analyses were conducted to account for a portion of the observed heterogeneity. This suggests that sources of heterogeneity either originate from aspects not investigated in this study, such as differences in $PM_{2.5}$ constituents, the spatial resolution of models used, or other individual factors, or a combination of the factors aforementioned. In some situations, these factors may not be available as they were not recorded. High heterogeneity is also expected given the diversities in populations, method of exposure assessment (and thus variations in spatiotemporal resolution), and levels of covariate adjustment. On the other hand, it might be possible that there is no true association, and the heterogeneity is driven by variation in control for confounding. Meta-regression could be conducted to understand which factor and how much a factor explains the heterogeneity observed. However, it was not conducted due to time constraints, as it would require additional preparation and exploration of the dataset to minimise the risk of overfitting models. It is also likely that some of the variables to be

introduced to the model, such as location and concentration level, will have some extent of multicollinearity, which will also threaten the validity of the model (Geissbühler et al., 2021). Whereas preceding publications have reported using meta-regression to explain heterogeneity, few have shown if the model assumptions are met and if the model is fitted based on factors selected a priori (Yu et al., 2022; Ghosh et al., 2021). Nevertheless, future studies should use a combination of subgroup analyses and meta-regression to investigate the drivers of heterogeneity.

Second, the health effects of different air pollutants are often measured and reported independently. However, air pollutants are inhaled as a mixture, often correlating highly with each other (Hurley, 2018). There were far fewer studies that used multipollutant models compared to single pollutant models. Hence, there may not be enough studies available to conduct a meta-analysis, and it was decided to focus on estimates from single pollutant models, with the assumption that the pollutant measured is representative of the effect of the whole mixture. However, the results reported by single pollutant analysis can be confounded by other pollutants. It is thus becoming increasingly common for researchers to adopt a co/multi-pollutant approach when investigating the effect of air pollutants on various health outcomes (Cassidy-Bushrow et al., 2020; Liang et al., 2019b; Siddika et al., 2019). Future studies could consider conducting meta-analyses on effect estimates derived from multipollutant analyses, as results from Chapter 4 section 4.3.3.2.1 suggested a potential increase in risk of PTB. This approach was adopted for outcomes such as hospital admissions and mortality and was found to produce smaller but more robust summary estimates (Mills et al., 2015; Stieb et al., 2021).

Some other limitations include some additional restrictions being imposed on the meta-analysis in an ad-hoc manner (Chapter 4 PRISMA). For instance, this has resulted in exclusion of time-series studies which could offer useful insights into the short-term effects of air pollution exposure on the risk of PTB, and studies in languages rather than Chinese that were not translated. In addition, the climate region analysis used two distinct climate classification systems for the US and China, but global level classification system, such as the Köppen climate classification system could have been used (Beck et al., 2018). This may allow the inclusion of studies outside the US or China in the meta-analysis, thus increase the number of studies available in each stratum.

5.4.7. Second interim conclusion

This comprehensive and rigorously conducted meta-analysis with a discussion on potential of causality found an increase in risk of PTB that was “likely to be causal” for whole pregnancy $PM_{2.5}$ exposure, and “suggestive to be causal” for trimester 3 NO_2 exposure. The findings were mostly consistent with earlier literature. Stratified analyses revealed that various factors such as pollution concentrations or data sources were likely partial drivers of the high heterogeneity observed. Limitations such as the use of single pollutant models remain, and the high heterogeneity observed was not well accounted for. Based on the findings of the causality discussion, the use of the finding of an increased risk of PTB associated with whole pregnancy $PM_{2.5}$ exposure, derived using studies that reported ORs, is recommended to be used for quantification in health impact assessments as a sensitivity analysis.

Chapter 6. Health impact assessment

6.1. Introduction

Using the increase in risk of PTB found for whole pregnancy $PM_{2.5}$ exposure in **Chapter 5.** (OR = 1.10, 95%CI: 1.06 - 1.14, per 10 μg/m^3 increase), a health impact assessment (HIA) was conducted to quantify the number of PTBs that would be averted by air pollution reduction measures under different emission scenarios in England. The emissions scenarios include a counterfactual scenario, business-as-usual (BAU), and balanced net-zero pathway (BNZP). The BAU scenario is based on existing UK Government's policies, and the BNZP scenario is based on the Balanced Pathway recommended in the Sixth Carbon Budget by the Climate Change Committee (CCC) for the UK to achieve net-zero by 2050 (CCC, 2020).

Modelled and projected $PM_{2.5}$ concentrations were obtained from a linked project investigating how climate change mitigation and adaptation policies impact mortality and morbidities outcomes through reduction in air pollution in the UK (Beevers et al., forthcoming; Walton et al., forthcoming). In the following section, the data sources retrieved from the linked project and additional sources for conducting the HIA are first described. Afterwards, the methods of quantification and assigning economic costs were explained.

Table 48 gives a brief comparison of the analysis done in this chapter and the linked project (Beevers et al., forthcoming; Walton et al., forthcoming). The base year of this HIA is set to be 2019, and the results are presented as number of PTBs averted up to 2050 and 2154. The result to 2050 was presented due to relevance to the UK Government's goal to achieve Net-Zero by 2050, while the result to 2154 allows comparison with calculations for other health outcomes in the separate project (Walton et al., forthcoming), which emphasised results to 2154 to allow proper projection of the potential benefits of the policies over time and population changes, as needed when incorporating lifetable analyses to examine impacts on life expectancy.

Table 48. Comparison of the analysis presented in this Chapter and the linked project (Walton et al., forthcoming).

Similarities / Differences	PTB analysis	Walton et al. (forthcoming)
Modelling and Projection of $PM_{2.5}$ levels	Obtained from Walton et al. (forthcoming)	
Emission scenarios	Details of the scenarios used in this analysis are obtained from Walton et al. Additional scenarios are included in Walton et al.	
Spatial and temporal resolution of $PM_{2.5}$ levels	9-month preceding concentrations at LA level	Annual exposures at ward level
Location	England	UK
Lifetable analysis	No	Yes

6.2. Methods

6.2.1. Health impact assessment (HIA)

6.2.1.1. Base year $PM_{2.5}$ levels and its projections

Hourly $PM_{2.5}$ concentrations in England in 2019 were modelled from several sources. A detailed description of the modelling can be found in Walton et al. (forthcoming). In brief, the authors modelled vehicle emissions using the established USEPA Community Multiscale Air Quality Modelling System coupled with a road source dispersion (CMAQ-Urban) model (Beevers et al., 2012; Dajnak et al., 2023). Anthropogenic emissions were taken from the UK National Atmospheric Emissions Inventory (NAEI, 2019). European emissions that could affect UK $PM_{2.5}$ concentrations were sourced from the European Monitoring and Evaluation Programme (EMEP) Centre on Emission Inventories and Projections (EMEP, 2023). The authors categorised non-vehicle emissions into 11 selected nomenclatures for air pollution (SNAP) sectors, and the UK anthropogenic emissions were produced at 2 km grid resolution and subsequently shaped to local authority level (Solazzo et al., 2017). For London, $PM_{2.5}$ emissions in 2019 were obtained from the London Atmospheric Emission Inventory (LAEI) (Greater London Authority, 2019).

Three scenarios were used in this HIA: counterfactual, Business-as-usual (BAU), and Balanced Net-Zero Pathway (BNZP). The counterfactual scenario represents an alternative situation in which no government policies or interventions were implemented, thus $PM_{2.5}$ concentrations in 2019 was assumed to have remain unchanged until 2154. The BAU and BNZP scenarios featured different lower $PM_{2.5}$ concentrations in 2030 and 2040 compared to the counterfactual. $PM_{2.5}$ levels for the BAU scenario were modelled primarily based on existing UK government emission reduction policies, energy forecasts, and additional policies outlined in the London Environment Strategy, which was projected to cause changes in trends in $PM_{2.5}$ originating from road transport use and non-road transport sources (BEIS, 2019; Mayor of London, 2018; Walton et al., forthcoming). The BNZP scenario is anticipated to cause a larger reduction of $PM_{2.5}$ levels than the BAU scenario. Adopting this scenario requires changes in various areas, including but not limited to uptakes of electric vehicles, building policies, heat networks and behavioural changes (CCC, 2020). The reduction of $PM_{2.5}$ levels modelled in the linked project was based on changes in transport behaviours and building policies, as they are anticipated to be the areas of changes affecting air pollution levels the most (Beevers et al., forthcoming; Walton et al., forthcoming).

To assess the impact of ambient $PM_{2.5}$ exposure on the risks of morbidities that have long latencies, annual levels of $PM_{2.5}$ exposures are usually used. However, the pregnancy exposure period (of interest for PTBs in the current project) is shorter than a year, so it would be inappropriate to use annual exposures. The hourly $PM_{2.5}$ concentrations were thus averaged to monthly concentrations using the “openair” package in R (Carslaw & Ropkins, 2012). Afterwards, average $PM_{2.5}$ levels were produced for the 9-months preceding birth. Rolling 9-month averages of $PM_{2.5}$ concentrations were derived and assigned to births being born from October 2019 onwards until December 2154.

Concentrations of $PM_{2.5}$ in the scenarios used were modelled for the years 2019, 2030, and 2040, excluding the interim years. Therefore, in this HIA, the concentrations between the modelled years were calculated assuming that there would be a constant proportional decrease in $PM_{2.5}$ levels between the same calendar months of the periods between 2019 and 2030, as well as 2030 and 2040, on the basis that monthly or seasonal patterns of $PM_{2.5}$ remain unchanged. For the years between 2040 and 2154, the $PM_{2.5}$ concentration in the same calendar month was assumed to be the same as in 2040.

6.2.1.2. Birth data

Number of births and PTBs per month in England were obtained from the NHS Maternity Services Monthly Statistics, sourced from the NHS Maternity Services Data Set (MSDS) (NHS Digital, 2024). Prior to September 2019, information on number of births and PTBs per month were only available at Clinical Commissioning Group[3] level. Afterwards, it is available at local authority (LA) level (NHS Digital, 2024). Therefore, birth data for future years was used to back calculate the number of births and PTB for the missing months. The year 2023 was selected over 2020, 2021, or 2022 due to the potential influences of the COVID-19 pandemic (ONS, 2022; Pomar et al., 2022). Thus, birth data at LA level from January to August 2023 was used to estimate birth data at LA level in the same months of 2019. This is achieved by deriving the proportions of births and PTBs at each LA in each month from January to August 2023, then applying the proportions to the number of births and PTBs at national level at each month from January to August 2019, assuming that the proportions remain unchanged between 2019 and 2023.

MSDS has a wide range of information from the initiation of care of pregnant women to their discharge. It is classified as experimental statistics, which are new official statistics undergoing evaluation (NHS Digital, 2024). The data is open for the public to download without any access restrictions. Suppression was applied to the datasets due to statistical disclosure risk by NHS Digital using the following rules: counts of zero were shown as zero, counts of 1 to 7 were converted to 5, all other counts converted to the nearest 5, with exceptions for rare events (NHS Digital, 2024). Further correspondence with the NHS Digital Team confirmed that if a LA is missing from the dataset i.e. with both births and PTBs information missing, it was implied that no reporting of births or PTBs was done, rather than that the number of PTBs were necessarily zero (**Appendix 27**). In the months of the base year constructed, this occurred in 6 LAs only and fewer than half of the months (**Appendix 25**).

6.2.1.2.1. Missing data

Data exploration revealed missing data were present in the MSDS datasets used to construct the births in base year 2019. Data were missing due to three reasons: (1) Data for the LA not being reported at all in the dataset, (2) The number of term births in an LA of that particular month was missing, and (3) the number of PTBs in an LA of that particular month

[3] Clinical Commissioning Groups (CCGs) (now called Integrated care systems (ICSs)) are at an organisation level above local authorities that are responsible for planning healthcare services (NHS England, 2024)

was missing. The percentage of LAs with missing data due to the reasons mentioned are shown in **Table 49**.

The total number of LAs in England in 2019 and 2023 differs due to structural changes in the local government during the years, which led to some LAs being abolished or united. For instance, the Buckinghamshire local authority (E10000002) was formed after 31st March 2020 and hence was not present in August to December 2019 base data (UK Government, 2019). When the months (January to July) adjusted using 2023 birth data were compared with those without adjustment (August to December), some patterns were found. There were no missing data for term births and low levels of missingness was found for PTBs in the adjusted months, whereas unadjusted months have much higher percentage of LAs with missing PTB information, which could reflect the improvement in data collection practice over time.

In the base year, around 64% of LAs have no missing data at all, whereas around 36% of LAs have some level of missingness. **Figure 71** shows the number of months with missing data and its spatial distribution. For 19 LAs, more than 50% of the data are missing. The respective reasons for missing data in these 19 LAs are shown in **Appendix 26**. 15 out of the 19 LAs having high levels of 'missing' data were due to the structural changes in LA between 2019 and 2023 meaning that it was unclear how to allocate data for a 2023 LA to those for 2019 when back calculating.

Table 49. Percentage of LA with missing data present in each month of the base year (2019) and the type of information that was missing.

Year the data is from	Month	Percentage of LA with missing data		
		LA completely missing	Missing term births	Missing PTBs
2023	January	3.47%	0.00%	1.58%
	February	3.79%	0.00%	0.63%
	March	3.79%	0.00%	0.95%
	April	4.10%	0.00%	0.63%
	May	8.83%	0.00%	0.63%
	June	8.83%	0.00%	1.58%
	July	8.83%	0.00%	2.21%
2019	August	5.36%	0.95%	16.09%
	September	4.73%	1.89%	12.93%
	October	5.68%	0.63%	13.56%

	November	5.68%	0.32%	11.36%
	December	4.73%	0.95%	8.83%

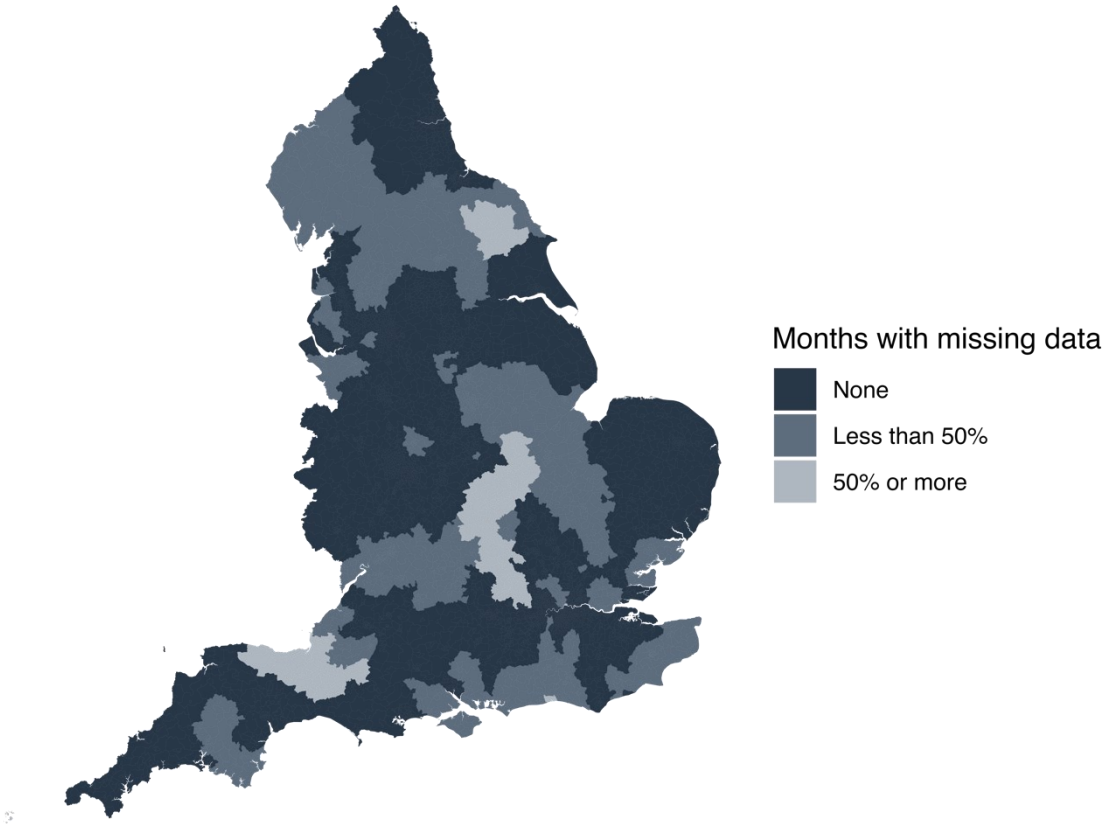


Figure 71. The level of missingness for each local authority in England by month in the base year (2019).

Information for number of term births and PTBs were obtained from the "GestationLengthBirthGroup37" dimension of the MSDS dataset. Another relevant variable "TotalBabies" which shows the total number of babies were used to back calculate either the number of term births or PTBs, if either one was missing. However, after replacement, particular months of some LAs showed very high percentage of PTBs, with the prevalence of PTB in a particular month range from 0% to 100%. Months of LA with 50% or 100% of births being born preterm was accepted due to the impact of suppression on small numbers, where 50% is derived when both term births and PTBs were coded as 5, and 100% when both number of PTBs and total number of babies were coded as 5. As this was very rare (0.29% of months of LAs used across the whole base year), they were not excluded from the

analysis. On the contrary, months of LAs showing a high prevalence of PTB but did not suffer from the issue (0.75% of months of LA used across the whole base year) were excluded from the analysis.

6.2.1.3. Birth projection

The projected number of births per month at LA level from 2020 to 2043 was derived from the data on population projections by the Office for National Statistics (ONS), which has such data available but at an annual level. Hence, the proportions of births (and PTBs) at each month of the base year (2019) were applied to the annual number of births from 2020 to 2043, to derive the monthly number of births at LA level from 2020 to 2043, assuming the proportions of births in each month of the year remain unchanged over time. Data on actual births from January 2020 to July 2023 from the MSDS dataset was available (**section 6.2.1.2**). However, this data was not used in this instance to maintain consistency with the linked project, which utilized birth projections from 2019 and the same population projection data and base year. This allowed effective comparison between numbers of PTB avoided and cases averted for other health outcomes in the companion project.

6.2.2. Economic evaluation

6.2.2.1. Overall methodology

After deriving the estimated number of PTBs averted per year from year 2020 to 2154, the cost for each PTB averted per year was calculated from three aspects: medical cost, lost productivity, and loss of utility. Healthcare cost is the cost incurred from use of medical services, such as hospital admissions, intensive care, or outpatient services. Lost productivity refers to the cost incurred from productivity loss of either the parents of the child that had a preterm birth, or the child directly, in forms such as absenteeism, disabilities, or loss of earnings (Mitchell & Bates, 2011). Lost utility is the hypothetical cost allocated to represent loss of quality of life, which can be derived from surveys of people's willingness-to-pay (WTP) to avoid the risk of occurrence of an unfavourable outcome (in this case PTB), or the consequences associated with the outcome (Ajzen, 1991; Stern et al., 1993). It is expressed as the sum of two components:

a) The WTP to avoid death due to being born preterm, expressed as the value of a statistical life (VSL) (Andersson, 2020)

b) The WTP to avoid non-fatal consequences of PTB e.g., illnesses in later life, expressed as the value of a statistical case (VSC) (Robinson et al., 2022)

The derivation of these costs from primary sources is outside the scope of this thesis. Therefore, a literature search was conducted in March 2024 using Ovid Medline, assisted with further internet searching, to find relevant, pre-existing information from published reports or articles. Afterwards, the article or report to be used was chosen based on the following criteria:

1. Date of publication, with preference given to more recent publications
2. Setting of the publication, with preference given to UK publications

6.2.2.2. Healthcare costs

Two studies in the UK were deemed most relevant after screening of articles retrieved from the literature search (Hua et al., 2023; Mangham et al., 2009). The study by Mangham et al. (2009) has information on community health and social care cost, as well as education and public sector costs in the UK. On the other hand, Hua et al. (2023) only has hospital admission costs, but it is much more recent and conducted in England. Hence, it was chosen to be the source of medical cost used in this analysis. **Table 50** shows the costs used in this analysis from Hua et al., 2023. The analysis utilised the incremental cost, which compares the expenses associated with preterm births to those of term births.

Table 50. Hospital admission costs for respective PTB types and term births over different time horizons (adapted from Hua et al., 2023, Table 5).

	Cost per birth (in £2018)				
Hospital admission cost	**Late PTB (34-36 weeks)**	**Moderate PTB (32-33 weeks)**	**Very PTB (28-31 weeks)**	**Extremely PTB (<28 weeks)**	**Term birth (≥37 weeks)**
1-year total	£17,608	£32,190	£68,147	£71,997	£1,153
5-year total	£21,870	£36,293	£73,837	£78,432	£1,936
8-year total	£23,920	£37,916	£76,625	£80,559	£2,347

6.2.2.3. Lost productivity

Three studies were deemed to be relevant after article screening (Institute of Medicine Committee on Understanding Premature Birth Assuring Health Outcomes, 2007; Patel et al., 2024; Waitzman et al., 2021). They were all conducted in the US, and all looked at the lost productivity of the parents of the PTBs, except one that looked at lifetime productivity loss of those that were born preterm (Waitzman et al., 2021). Although Patel et al. (2024) was

published more recently, Waitzman et al. (2021) was used in this analysis because Patel et al. (2024) has only provided monetary evaluation regarding medical-related absenteeism and disability of the parents. The cost of lost productivity per PTB used in the analysis was $12,236 (in $2016).

6.2.2.4. Lost utility

No studies or reports retrieved by the literature search have derived the WTP for reducing fatal or non-fatal consequences of PTB. Reports previously published by national or international bodies on consequences of premature mortality and related outcomes were hence used.

Economic evaluation of the fatality resulted from PTB (VSL) used in this analysis was from the UK Department of Transport WTP value for avoiding a death incurred by road traffic accidents (UK Department of Transport, 2018). The value was applied to preterm infants that would have died within the first year of life, using case-fatality risk derived from a global epidemiological study (Blencowe et al., 2013). The Organisation for Economic Co-operation and Development (OECD) report for values of later-life consequences of very low birth weight was used as the VSC incurred from non-fatal consequences of PTB (Ščasný et al., 2023). It was deemed appropriate in this case due to scarce information available, and infants born with a very low birth weight were likely preterm as well. Two other monetary values were considered and used in the sensitivity analyses. First was the WTP value for reducing the overall consequences for very low birth weight (including death). Second was the WTP value for premature mortality from another OECD report that was published earlier (OECD, 2012). **Table 51** shows the monetary values provided by the reports aforementioned.

Table 51. The Willingness-to-pay (WTP) values used in this analysis and their respective sources.

	Cost of lost utility per case of PTB	
	VSL	**VSC**
UK Department of Transport	£1,537,896 (£2010*)	
OECD, 2012	$3,600,000 ($2005*)	
Ščasný et al. (2023)		$456,000 ($2022*)
	$602,000 ($2022*)	

*The value of the currency (and thus the purchasing power) of that particular year, to allow comparison of real costs rather than nominal costs that can fluctuate due to inflation.

6.2.3. Method of quantification

6.2.3.1. Health impact assessment (HIA)

Using the concentration-response function (CRF) of 1.10 (95%CI: 1.06 - 1.14) per 10 μg/m^3 increase in $PM_{2.5}$ exposure, the steps of quantifying the number of PTBs averted due to the air pollution reduction are explained below. The steps are adapted from the COMEAP report for air pollution exposure and chronic bronchitis, as the CRF used in the report and that of this analysis are both CRFs as ORs (COMEAP, 2016). The following calculation is a hypothetical example where the decrease in $PM_{2.5}$ level between the counterfactual and BAU scenario is 5 μg/m^3, with 100 births in the LA, and the probability of reporting PTB was 15% in a given month for that LA (**Table 52**).

Table 52. Steps of quantification used to derive the number of PTBs averted

Steps	Description	Calculations
1	Baseline odds of PTB	$Odds_{base} = \frac{0.15}{(1-0.15)} = 0.17647059$
2	Baseline log odds of PTB	$\ln(Odds_{base}) = -1.73460106$
3	CRF of choice	$OR_{per\ 10} = 1.10$
4	Log of the CRF per 10 μg/m^3 increase	$\ln(OR_{per\ 10}) = 0.09531$
5	Log of the CRF per 1 μg/m^3 increase	$\ln(OR_{per\ 1}) = \frac{\ln(OR_{per\ 10})}{10} = 0.009531$
6	Change in log odds per concentration change (cc)	$\ln(OR_{per\ cc}) = 0.009531 * -5 = -0.04765509$
7	Log odds at BAU	

		$\ln(Odds_{base}) + \ln(OR_{per\ cc}) = -1.78225615$
8	Odds at BAU	$Odds_{BAU} = exp(-1.78225615) = 0.168258104$
9	Probability of having a PTB at BAU	$p(PTB)_{BAU} = \frac{Odds_{BAU}}{(1 + Odds_{BAU})} = 0.144024769$
10	Change in probability of having a PTB in BAU compared to baseline	$0.15 - p(PTB)_{BAU} = 0.005975231$
11	Number of PTBs averted	$PTB_{avoided} = 100 \times 0.005975231 = 0.597523064$

This process was repeated for each month from October 2019 up to 2154 for each LA in England, to obtain the total number of PTBs averted from 2019 to 2154 due to the implementation of the BAU scenario, compared to the counterfactual. The process was then repeated by comparing the BNZP scenario and the counterfactual, and the BAU scenario and the BNZP scenario. As the number of singleton PTBs was not available in the MSDS and the CRF applies to singleton births, the number of PTBs averted was adjusted using the percentage of singleton births out of the total number of births from the ONS birth characterisation dataset (ONS, 2022).

6.2.3.1.1. Incorporating a 5 μg/m^3 cut-off

The WHO Air Quality Guideline for 24-hour $PM_{2.5}$ exposure is set at 5 μg/m^3, as most studies informing the WHO AQG used concentration levels above 5 μg/m^3. There are hence larger uncertainties of the health effects of $PM_{2.5}$ exposure below that. A 5 μg/m^3 cut-off was thus incorporated to the calculations shown in the previous section by replacing concentration below 5 μg/m^3 to 5 μg/m^3, hence the impact of concentrations below the cut-off was not considered. This is also consistent with what was done for the other health outcomes considered in the linked project (Walton et al., forthcoming). The analyses were also repeated without a 5 μg/m^3 cut-off being used.

6.2.3.2. Economic evaluation

After deriving the number of PTBs averted per year from 2019 to 2154, costs incurred from healthcare, lost productivity, and lost utility were assigned to the PTBs averted. The monetary values were first inflation-adjusted to £2022 prices using the ONS Consumer Price Inflation Table (Table 57) (ONS, 2024) to facilitate comparison with the monetary valuation of the other health outcomes in the accompanying project (ref). As the monetary value from some sources used involved those from countries not using Great British Pounds (GBP), it was first converted to the corresponding value in GBP of that particular year using the OECD purchasing power parity converter (OECD, 2024), before being inflation-adjusted. The total cost per year was then derived.

For economic evaluation of policies, a discount is usually applied to the monetised future benefits of the policies of interest. It is because people generally prefer to receive goods and services now compared to in the future. The discount rates, or more precisely the Social Time Preference Rate (STPR) used in this work, were obtained from the Treasury Green Book (HM Treasury, 2022). For healthcare costs and lost productivity, a discount rate of 3.5 percent was used. An uplift of two percent, based on assumed GDP per capita growth per year in the UK, was also applied (HM Treasury, 2022). For monetised benefits for lost utility averted per year, a discount rate of 1.5 instead of 3.5 percent was chosen, which incorporates the assumed increase in willingness-to-pay (WTP) for health interventions with capital growth over time.

6.2.3.2.1. Sensitivity analyses

Sensitivity analyses were conducted using different sources of medical costs and lost utility (**Table 53**). Two other combinations of lost utility were used, (1) VSL of $3.6 million per death within first year of life following a PTB (OECD, 2012), and the VSC of $456,000 per PTBs that survive beyond first year of life (Ščasný et al., 2023), (2) VSC of $602,000 per individual (including fatal consequences) (Ščasný et al., 2023). For medical costs, the cumulative hospital admission costs up to 1-year or 5-year of life were used.

Table 53. The sources of medical cost, lost productivity, and lost utility in main results and sensitivity analyses.

<table>
<tr><th colspan="2" rowspan="2"></th><th rowspan="2">Medical Cost</th><th rowspan="2">Lost productivity</th><th colspan="2">Lost utility</th></tr>
<tr><th>VSL</th><th>VSC</th></tr>
<tr><td colspan="2">Main result</td><td rowspan="3">Hua et al. (2023) (up to 8th year of life)</td><td rowspan="5">Waitzman et al. (2021)</td><td>UK DoT</td><td rowspan="2">Ščasný et al. (2023) (non-fatal consequences)</td></tr>
<tr><td rowspan="4">Sensitivity analysis</td><td>1</td><td>OECD (mortality)</td></tr>
<tr><td>2</td><td colspan="2">Ščasný et al. (2023) (fatal and non-fatal consequences)</td></tr>
<tr><td>3</td><td>Hua et al. (2023) (up to 1st year of life)</td><td rowspan="2">UK DoT</td><td rowspan="2">Ščasný et al. (2023) (non-fatal consequences)</td></tr>
<tr><td>4</td><td>Hua et al. (2023) (up to 5th year of life)</td></tr>
</table>

6.3. Results

6.3.1. Characteristics of HIA inputs

6.3.1.1. Characteristics of births at base year (2019)

Across the 317 LAs in 2019, the number of births varied greatly. Birmingham had the highest number of births, followed by Leeds and Bradford. This was expected as they are LAs in major cities in England that have high population sizes. LAs with the smallest number of births in 2019 were mostly due to unavailable 2023 data in months that required back-calculation to 2019 using January to August 2023 data (details in next section). The geographical distribution of births and prevalence of PTBs at LA level in the base year are shown in **Figure 72**.

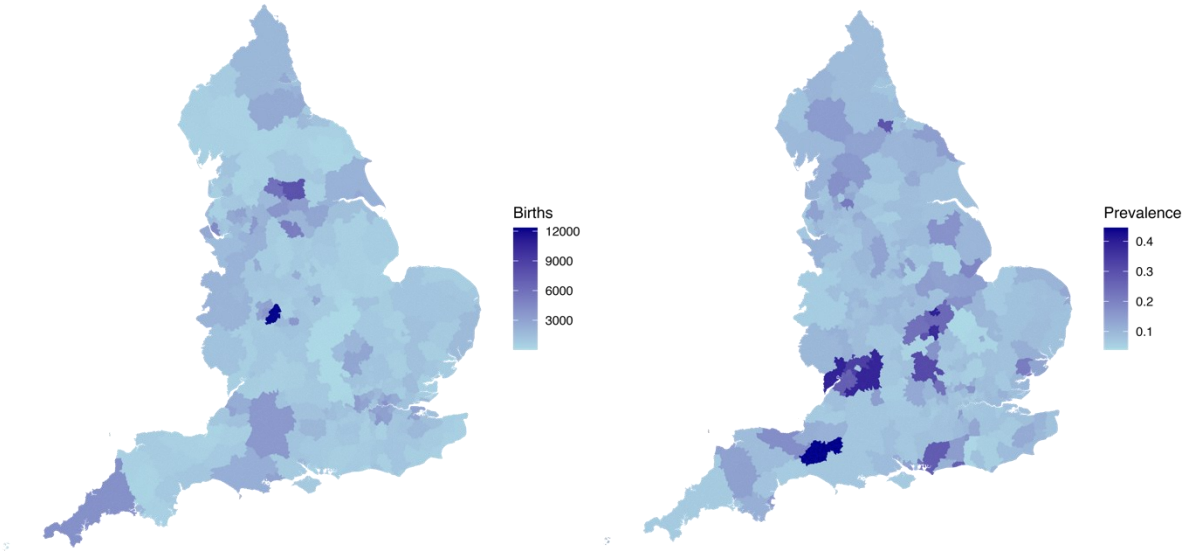


Figure 72. Averaged number of births (left) and prevalence of preterm birth (right) in the 317 LAs in England at 2019 (averaged across all months).

6.2.3.2.2. Impact of missing data

Removal of LAs with more than half of the months with missing data due to the structural changes in LAs between 2019 and 2023 (**Appendix 26**) was not influential to the results of the HIA and economic evaluation presented in the following sections.

6.3.1.2. Projected trends in $PM_{2.5}$ concentration

The annual average $PM_{2.5}$ concentration at base year 2019 was 8.81 μg/m^3. In the baseline scenario, it remained unchanged up to the end of the study period (2154). In the BAU scenario, annual $PM_{2.5}$ concentrations decreased to 6.49 μg/m^3 in 2030, then increased slightly to 6.58 μg/m^3 in 2040. The decreases in $PM_{2.5}$ concentrations were larger in the BNZP scenario. Annual $PM_{2.5}$ concentration decreased to 6.34 μg/m^3 in 2030, then further decreased to 6.03 μg/m^3 in 2040. The overall trend is shown in **Figure 73**.

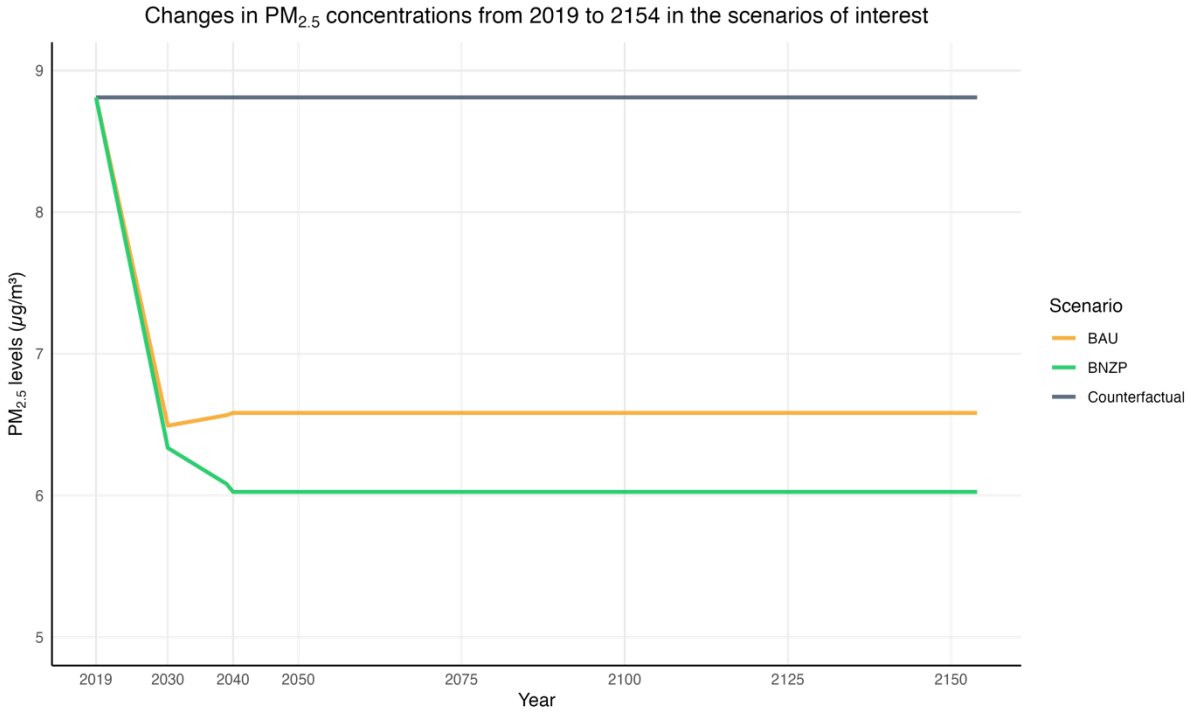


Figure 73. Projected $PM_{2.5}$ concentrations for the counterfactual, BAU, and BNZP scenarios from 2019 to 2154.

The annual $PM_{2.5}$ level in each LA in 2019 (baseline), 2050 (BAU scenario), and 2050 (BNZP scenario) are graphically shown in **Appendix 28**, whereas the differences in $PM_{2.5}$ levels between the baseline and the BAU and BNZP scenarios are shown in **Appendix 29** and **Figure 74** respectively. In the base year, most regions of England except parts of the Northern region and Southwest have levels of PM2.5 exceeding 10 μg/m^3 (data not shown). Implementation of the BNZP or BAU scenario produce a noticeable reduction in $PM_{2.5}$ levels, with the greatest reduction in $PM_{2.5}$ levels occurring in the East of England and South East regions.

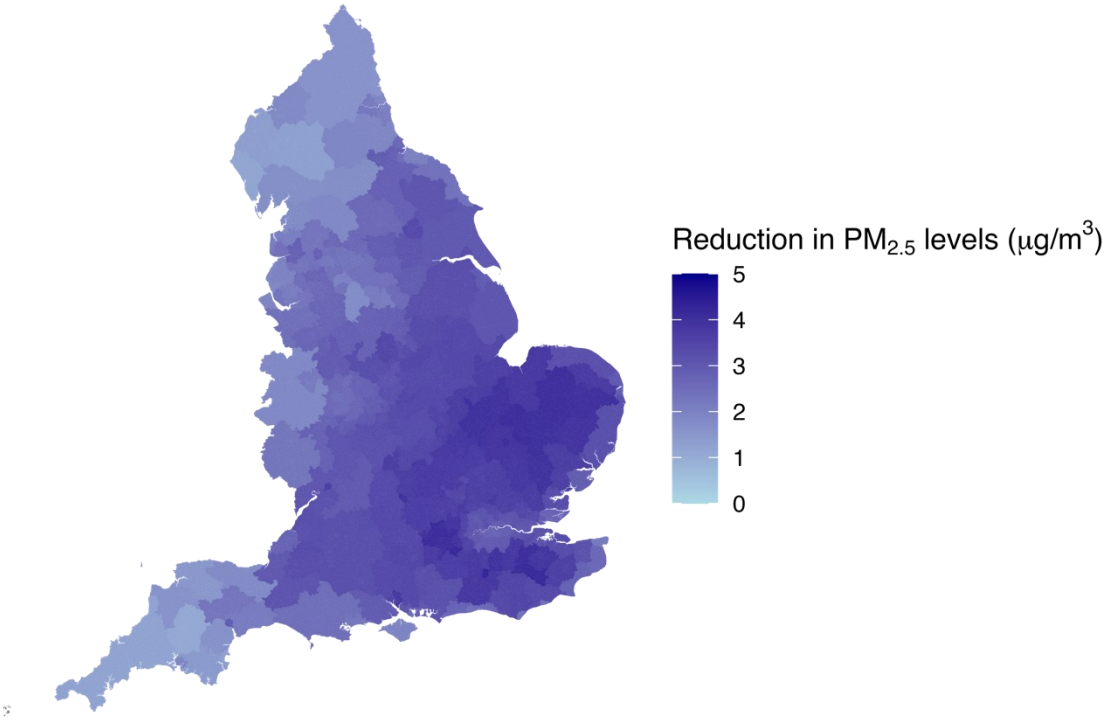


Figure 74. The reduction in $PM_{2.5}$ concentrations at 2050 if the BNZP scenario was implemented (compared to the base year 2019).

6.3.2. Number of PTBs averted

The number of singleton PTBs averted up to year 2050 and 2154 due to the air pollution concentration differences between the scenarios are shown in **Table 54** and **Table 55**. When the counterfactual scenario was compared to the BAU or BNZP scenarios, the number of PTBs averted were much higher than a comparison between the BAU and BNZP scenarios. This is due to the larger concentration differences between the counterfactual and the BAU or BNZP scenarios. The presence of a 5 µg/m^3 cut-off slightly reduced the number of PTBs averted. However, the results did not substantially change as mean $PM_{2.5}$ levels were above 5 µg/m^3 even for the BNZP scenario which would cause further $PM_{2.5}$ reduction than the BAU scenario.

With the presence of a 5 µg/m^3 cut-off, the number of PTBs averted consist of around 0.24% or 0.39% of the total number of PTBs being born up to 2050 or 2154 respectively in the BAU vs. BNZP scenario. The average number of PTBs averted per year from 2019 to 2154 would be approximately 200 per year. The majority of the PTBs avoided would occur between the periods where $PM_{2.5}$ levels were assumed to remain unchanged i.e. between 2040 to 2154.

Table 54. Number of PTBs averted based on the concentration differences between the scenarios of choice from 2019 up to 2050.

Comparison	Cut-off	Up to Year 2050	
		Number of PTBs averted	% of total PTBs
Counterfactual vs. BAU	/	18,294 (16,675 – 19,994)	1.22 (1.11 – 1.33)
	5 μg/m^3	18,103 (16,493 – 19,792)	1.21 (1.10 – 1.32)
Counterfactual vs. BNZP	/	21,876 (19,923 – 23,926)	1.45 (1.33 – 1.60)
	5 μg/m^3	21,579 (19,642 – 23,613)	1.44 (1.31 – 1.57)
BAU vs. BNZP	/	3,632 (3,284 – 3,996)	0.24 (0.22 – 0.27)
	5 μg/m^3	3,526 (3,184 – 3,884)	0.24 (0.21 – 0.26)

Table 55. Number of PTBs averted based on the concentration differences between the scenarios of choice from 2019 up to 2154.

Scenarios	Cut-off	Up to Year 2154	
		Number of PTBs averted	% of total PTBs
Counterfactual vs. BAU	/	108,272 (100,088 – 116,865)	1.60 (1.48 – 1.72)
	5 μg/m^3	106,275 (98,611 – 115,244)	1.58 (1.46 – 1.70)
Counterfactual vs. BNZP	/	135,402 (124,955 – 146,374)	1.99 (1.84 – 2.16)
	5 μg/m^3	132,910 (122,578 – 143,759)	1.96 (1.81 – 2.12)
BAU vs. BNZP	/	27,572 (25,199 – 30,057)	0.41 (0.37 – 0.44)
	5 μg/m^3	26,612 (24,288 – 29,045)	0.39 (0.36 – 0.43)

The PTBs averted are mainly in the regions of South East and East Midlands (**Figure 75**). The Eastbourne, Epsom and Ewell, and Oadby and Wigston Borough Councils are the 3 LAs with the highest number of PTBs averted. This is similar to the regions showing the greatest reduction in $PM_{2.5}$ levels, but there were LAs with noticeably high number of PTBs averted, suggesting that they benefit the most from the emission reduction policies introduced. LAs with close to zero number of PTBs averted are those with missing data.

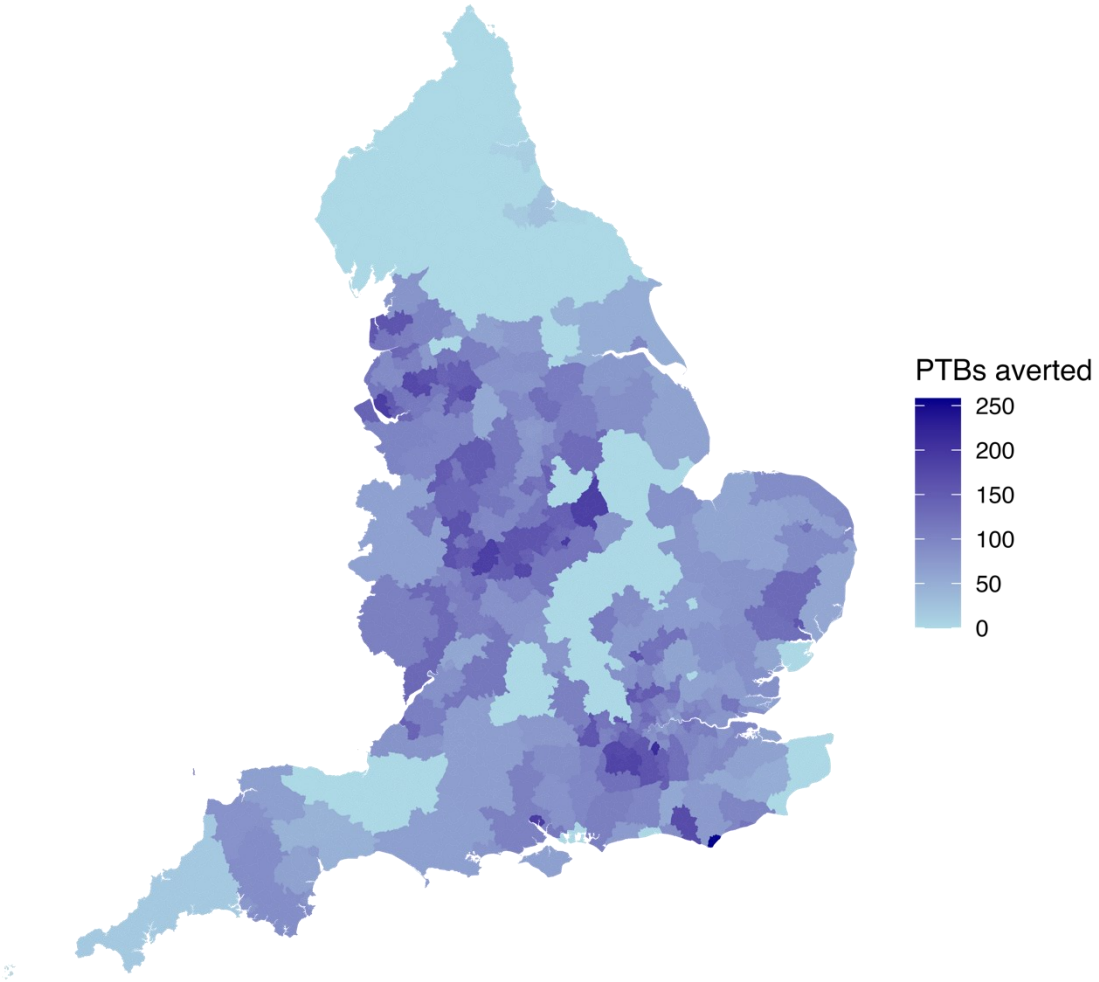


Figure 75. Spatial distribution of the PTBs averted across the LAs of England over the whole study period (135 years).

6.3.3. Economic evaluation

6.3.3.1. Medical costs

Based on the chosen CRF and information from Hua et al., the cumulative hospital admission cost up to first, fifth, and eighth year of life for the PTBs compared to term births were derived and shown in **Table 56**. Most of the cost was incurred in the first year of life. The presence of a 5 μg/m^3 cut-off had little influence on the findings. Hence, the results without a 5 μg/m^3 cut-off are presented separately in **Appendix 30**. When the BNZP instead of BAU scenario was implemented, the medical cost (up to eighth year of life) per PTB (total cost / total number of PTBs averted) was approximately £18,000 (£16,000 - £19,000), and

the medical cost per year (total cost / policy window i.e. 135 years) was £3,500,000 (£3,200,000 - £3,800,000).

Table 56. Cumulative hospital costs (in £2022, millions) at different time horizons of the life of a PTB based on comparison between the scenarios of interest (with 5 µg/m^3 cut-off)

Scenarios	Cumulative hospital admission costs (in £2022, millions)		
	1st year of life	5th year of life	8th year of life
Counterfactual vs. BAU	1,637 (1,511 – 1,768)	1,864 (1,721 – 2,014)	1,965 (1,814 – 2,123)
Counterfactual vs. BNZP	2,024 (1,865 – 2,190)	2,304 (2,124 – 2,493)	2,429 (2,239 – 2,628)
BAU vs. BNZP	393 (359 – 429)	448 (409 – 488)	472 (431 – 515)

6.3.3.2. Lost productivity

Based on the chosen CRF and estimate for the cost of productivity loss for an individual being born preterm from Waitzman et al. (2021), monetised lost productivity from PTBs averted were derived and shown in **Table 57**. The difference in lost productivity between BAU and BNZP scenarios with cut-off was estimated to be around £129 million, which was smaller than the hospital admission cost in the 1st year of life (shown above). Similar to medical costs, the presence of a cut-off was not influential to the results. Monetised lost productivity expressed per PTB averted and per year when the BNZP instead of BAU scenario was implemented were approximately £4,800 (£4,400 - £5,300), and £955,000 (£874,000 - £1,040,000) respectively.

Table 57. Monetised value of the lost productivity (in £2022, millions) due to delivery of PTBs based on comparison between the scenarios of interest from 2019 to 2154.

Scenarios	Monetised value of lost productivity (in £2022, millions)		
	5 µg/m^3 cut-off	Central estimate	Range
Counterfactual vs. BAU	No	544	502 – 587
	Yes	536	495 – 580
Counterfactual vs. BNZP	No	675	623 – 730
	Yes	663	611 – 718
BAU vs. BNZP	No	133	122 – 145
	Yes	129	118 – 141

6.3.3.3. Lost utility

Lost utility was calculated from the sum of VSL and VSC from different sources. The monetised values of the different VSLs and VSCs are shown in **Appendix 31**. The cost of lost utility of PTBs avoided due to the concentration changes between the scenarios of interest, using the VSL from Department of Transport and the VSC from OECD is presented in **Table 58**. The presence of a cut-off was not influential to the results. As the majority of PTBs can survive beyond first year of life, the value of VSC quantified was much greater than VSL, which was only assigned to those that died within the first year of life. Monetised lost utility per PTB averted and per year when the BNZP instead of the BAU scenario was implemented were approximately £210,000 (£191,000 - £228,000) and £41,100,000 (£37,600,000 - £44,900,000).

Table 58. Monetised value of lost utility (in £2022, millions) derived from PTBs averted from 2019 to 2154.

	5 µg/m^3 cut-off	VSL[4] (DoT)	VSC[5] (OECD)	Lost utility (VSL + VSC)
Counterfactual vs. BAU	No	3,076 (2,841-33,23)	20,666 (19,087-22,324)	23,742 (21,929–25,647)
	Yes	3,035 (2,802-3,280)	20,389 (18,823-22,032)	23,423 (21,624–25,312)
Counterfactual vs. BNZP	No	3,809 (3,513-4,120)	25,592 (23,602-27,681)	29,401 (27,116–31,801)
	Yes	3,743 (3,450-4,051)	25,147 (23,178-27,214)	28,889 (26,628–31,265)
BAU vs. BNZP	No	745 (681-812)	5,005 (4,574-5,456)	5,750 (5,255–6,268)
	Yes	720 (657-785)	4,835 (4,413-5,277)	5,555 (5,070–6,062)

6.3.3.4. Total monetised benefits and sensitivity analyses

Results for the main and sensitivity analyses that show the monetised benefit (in £2022, millions) of reduction in PTB numbers due to the difference in $PM_{2.5}$ levels between the BAU and the BNZP scenario (with a 5 µg/m^3 cut-off) were shown in **Table 59**. The monetised values of the respective components in different comparison of scenarios are shown in **Appendix 32**.

Lost utility, in particular VSC that concerns the long-term consequences of PTB, constituted the largest component of the monetised benefits for PTB reduction. It was followed by VSL, medical cost, and lost productivity. How VSC was quantified was thus very significant in the value of monetised benefit in the sensitivity analyses. The VSC used in the second sensitivity analysis (SA2) sourced from Ščasný et al. (2023), which included WTP to reduce fatal as well as non-fatal consequences, resulted in the highest monetised benefits. The

[4] VSL: Value of a statistical life (see section 6.2.2.4 for more details)
[5] VSC: Value of a statistical case (see section 6.2.2.4 for more details)

second highest monetised benefit was shown in the first sensitivity analysis (SA1), where VSL was quantified using the much higher estimate from OECD, instead of the more conservative estimate from the UK Department of Transport that was used in the main result.

Table 59. The sum of monetised benefits (in £2022, millions) from its respective components in the main and sensitivity analysis (comparison between the BAU and BNZP scenarios) from 2019 to 2154.

<table>
<tr><th colspan="2" rowspan="2"></th><th rowspan="2">Medical Cost</th><th rowspan="2">Lost productivity</th><th colspan="2">Lost utility</th><th rowspan="2">Total</th></tr>
<tr><th>VSL</th><th>VSC</th></tr>
<tr><td colspan="2">Main result</td><td rowspan="5">472 (431-515)</td><td rowspan="5">129 (118-141)</td><td>720 (657-785)</td><td rowspan="2">4,835 (4,413-5,277)</td><td>6,156 (5,618-6,718)</td></tr>
<tr><td rowspan="4">SA*</td><td>1</td><td>1,366 (1,246-1,490)</td><td>6,801 (6,208-7,423)</td></tr>
<tr><td>2</td><td colspan="2">6,383 (5,824-6,966)</td><td>6,984 (6,375-7,622)</td></tr>
<tr><td>3</td><td rowspan="2">720 (657-785)</td><td rowspan="2">4,835 (4,413-5,277)</td><td>6,077 (5,547-6,632)</td></tr>
<tr><td>4</td><td>448 (409-488)</td><td>6,131 (5,596-6,691)</td></tr>
</table>

**SA: Sensitivity analysis*

6.4. Discussion

6.4.1. Summary of findings

Based on the decrease in $PM_{2.5}$ concentrations from 2019-2154, around 27,000 PTBs would be avoided in England if the UK Government adopts the BNZP instead of BAU pathway to achieve Net-Zero by 2050. This corresponds to a saving of approximately £6,000,000,000 (in £2022) based on averted medical costs, monetised lost productivity, and monetised lost utility. The presence of a cut-off at 5 $\mu g/m^3$ did not substantially influence the overall findings. Lost utility was the largest constituent of the monetised benefits, and its monetised value was most affected by how VSC was quantified, as the majority of PTBs would survive beyond first year of life.

6.4.2. Implications of findings

6.4.2.1. Comparison with other health outcomes

An informative comparison can be made with Walton et al. (forthcoming), as the present work has used one of the same policy scenarios used in Walton et al. (forthcoming), and has used the same source of $PM_{2.5}$ levels projected onward until 2154. However, Walton et al. covered the whole of the UK and used ward-level concentrations that were subsequently population-weighted to local authority or country, whereas the local-authority level concentrations in England were directly used in the current work. Moreover, for the health endpoints included in Walton et al., differences between the annual concentrations of the scenarios compared were used. In the present work, 9-month preceding concentrations were deemed most appropriate. Walton et al. have included a variety of morbidities but not PTB. A comparison can be made with the other morbidities included for the number of averted cases, as well as the corresponding monetised benefits.

The number of PTBs averted (n = 26,612) was the smallest when compared to the impact of $PM_{2.5}$ reduction on other health outcomes analysed in Walton et al, potentially due to the smaller baseline risk for PTB compared to the other health outcomes, as well as the smaller population-at-risk (pregnant mothers resulting in singleton live births) used. For instance, it was approximately half the number of lung cancer incidences (n = 51,000), which has a greater population-at-risk (adults aged 30 or above). The number of PTBs averted is likely an underestimation for the whole of the UK, as the analysis was for England only, and information from the whole UK was used for other morbidities. According to birth

characteristics dataset from ONS, births in England constitutes approximately 86% of all births in the UK in 2021 (ONS, 2023). Assuming that the proportion of PTBs and concentration reduction is the same, it would increase the number of PTBs averted to around 30,000, which would still make it the smallest unless the impact of other pollutants (O_3) on other health outcomes were considered.

The monetised benefits corresponding to the PTBs averted was not the smallest compared to the other health outcomes investigated in Walton et al. (forthcoming). The 8-year cumulative medical cost derived from Hua et al. was £23,920 per PTB, which was higher than all morbidities except dementia, lung cancer, and stroke. The monetised benefit of the lost productivity averted was close to diabetes, ischaemic heart disease (IHD), and chronic obstructive pulmonary disease (COPD). For lost utility, the cost per PTB was £209,000, which was similar to diabetes (£210,000) and lower than IHD and stroke. The total monetised benefits of 6.16 billion is higher than lung cancer similar to that of COPD, albeit lower than child health outcomes such as asthma incidences. The result of this analysis indicates that the monetised benefits for PTBs averted, when viewed alongside other health outcomes, would be relatively small.

6.4.2.2. Comparison with literature on PTB

How interventions improving future air quality would affect projected number of PTBs and its associated economic benefits has received scant attention in the literature. In addition, comparison of existing work in terms of the number of PTBs averted, or the monetised benefits are difficult as they were often conducted in different countries, on different policies, and used different baseline rates. The corresponding monetised benefits/averted costs are thus difficult to compare as well. Available literature mainly focused on either deriving the burden of PTB attributable to air pollution, or the economic costs of PTB, without linking them to each other using policies that reduce levels of air pollution (Achana et al., 2022; Deguen et al., 2020; Liu et al., 2023; Marc-Hodek et al., 2011).

One study investigated the impact of removing all combustion-related emissions on children's health outcomes, namely asthma incidence and preterm birth, in 16 cities in different countries (including London and Glasgow) (Milner et al., 2023a). Using country-level birth data from the Global Burden of Disease (GBD) 2019 study that was downscaled to city-level, they found that the policies would lead to 483 cases of PTB being averted in London and Glasgow per year, which was higher than what was found in the present study (approximately 200 per year). However, the present work's estimate was more realistic as it

did not assume removal of all combustion-related emission. Moreover, it covered the whole of England, was able to leverage monthly local authority level birth data, and therefore calculate averted PTBs at a higher granularity across England.

Another study in the UK has explored the impact of implementing a low emission zone within the outer ring roads of Leeds and Bradford on the health-related quality of life and healthcare costs, via its impact on air quality (Lomas et al., 2016). Using a CRF from Sapkota et al. (2012), they found that each PTB avoided would save approximately £50,000 per year (£2013/2014) (around £62,000 in £2022). The monetary evaluation was much lower than that found in this study, likely because Lomas et al. did not consider lost productivity and lost utility, and the source for medical costs (Mangham et al., 2009) has lower values than that used here (Hua et al., 2023). One study in Sweden estimated the socioeconomic cost per late PTB as approximately €34,000 (€2016) (Kriit et al., 2024), using WTP estimates that were subsequently adjusted by QALY weight derived from US national data (Craig et al., 2016)

There are several differences in the methodological approach of how number of PTBs averted were calculated in the present analysis and the other two studies (Lomas et al., 2016; Milner et al., 2023).

First, different CRFs were used. The CRF used in the present analysis (OR: 1.10 [1.06 - 1.14]) is lower than that of Milner et al. (2023) (RR: 1.12 [1.06 - 1.19]) and Lomas et al. (2016) (OR: 1.15 [1.14 - 1.16]). The meta-analysis used by Lomas et al. (2016) is most dated (Sapkota et al., 2012), and hence was based on the fewest number of studies. The CRF used in the present analysis was derived from a meta-analysis which included similar number of studies as that of Milner et al. (2023a), which was derived from Ghosh et al. (2021), and thus was likely to be a more accurate representation of the actual risk than Lomas et al.

Second, different methodological choices were made regarding the presence and level of a cut-off. A 5 $\mu g/m^3$ cut-off was used in this work to aid comparison with other morbidities. Lomas et al. did not use a cut-off while Milner et al. (2023a) used a lower theoretical minimum risk exposure level (TMREL) of 4.2 $\mu g/m^3$ derived from Ghosh et al. (2021). It was likely used as Milner et al. (2023a) included only two child health outcomes and used different cut-offs for them. On the contrary, there is lack of consensus regarding the minimum risk exposure level for some of the other comorbidities included in the linked project e.g. dementia, hence using a 5 $\mu g/m^3$ cut-off based on the WHO AQG would be more methodologically sound.

Third, different methodological choices were made for exposure and study population. Annual exposures were used by both Lomas et al. (2016) and Milner et al. (2023a), which introduces the risk of exposure misclassification as previously discussed in section 6.2.1.1. 9-month preceding concentration used in this work reduced the risk of exposure misclassification and thus assigned more accurate exposure to the study population. For study populations, it was unclear what method was used by Lomas et al. (2016), whereas Milner et al. (2023a) used the Global Burden of Disease (GBD) 2019 (IHME, 2019) counts of births and PTB at country-level. The study population in this work was derived at local authority level, which was in much higher granularity than Milner et al. (2023a).

Lastly, the method of quantification was different in this work and Milner et al. Cost was the focus of Lomas et al. (2016), so how the number of PTBs averted were derived was not shown which could be due to journalistic constraints. For Milner et al. (2023a), the standard population-attributable fraction (PAF) approach was used: $PAF = (RR - 1)/RR$, where the PAF derived would be applied to the baseline rate to obtain the number of attributable cases. The method used in this work is slightly different as ORs were used. RRs are easier to interpret and thus more intuitive in public health messaging. However, a CRF based on OR was used in this instance due to larger uncertainty around the increase in risk found using HR.

There are also differences in how the monetised benefits were calculated. Lomas et al. estimated the monetised benefits based on QALYs lost per each PTB, which was an approach more often adopted by healthcare decision-making. On the other hand, the present work has used the WTP and VSL approaches that are more commonly used in environmental economics (Lomas et al., 2016). Their estimates were derived based on preferences or wage-risk trade-offs of individuals (Adamowicz et al., 1998; Hammit, 2002), which is more relevant when the intervention studied has effects beyond health than QALYs which rely on expert evaluation of the differences in health states (Hammit, 2002). Recent studies that have derived PTB-related QALYs in the UK used data from clinical trials or cohorts, which were not representative of the national population of births and focused on specific healthcare interventions (Grosso et al., 2020; Hunter et al., 2022; Stock et al., 2021).

The approach of adjusting country-specific WTP estimates with QALYs weights has been explored and adopted by Kriit et al., which is useful as it showed how future studies can potentially derive the WTP for a QALY, which allows a condition's impact on mortality and morbidity outcomes to be comprehensively captured. Thus, valuations can be more accurately linked to interventions that have both health and environmental impacts (Kriit et al., 2024). However, limitations are present for this methodology. First, QALYs derived from

the US used by Kriit et al. may not be transferrable to the English population used in this analysis. Second, how the QALYs were derived depends heavily on how well the immediate and lifelong consequences of PTB were characterised, as well as methodological choices made by the authors. While immediate consequences of PTB such as hearing loss or vision impairment are relatively well-studied and thus often included in time trade-off studies, lifelong consequences are not currently well-characterised (Prior and Modi, 2020). Thus, conditions such as learning disability or dementia are potential lifetime consequences of PTB, but learning disability was not selected by the authors to adjust the QALY weights, and information for dementia is not available from time trade-off studies (Craig et al., 2016).

6.4.3. Limitations

6.4.3.1. Health impact assessment (HIA)

6.4.3.1.1. Data availability

As previously mentioned, birth data from January to July 2023 was used to back-calculate the January to July 2019 birth rates due to data unavailability at local authority level before August 2019. The structural changes that have occurred in the English local government between 2019 and 2023 mean that some LAs no longer exist. Therefore, information for births remained unavailable after the back-calculation for 15 LAs. This likely underestimated the number of births and PTBs averted, especially when they are projected forward for a long period of time.

The number of PTBs and term births available from the MSDS dataset are suppressed due to the risk of statistical disclosure of personal identity. The method of suppression introduces systematic error that disproportionately affect smaller numbers, and reduces precision of summary statistics such as mean and variance. This is partly why summary statistics for birth data was not reported, and prevalence of PTB tended to be higher for LAs with smaller numbers. For instance, if term births and PTBs were both reported as 5, then the prevalence of PTB is reported as 50%. Random error was also introduced due to the method used for rounding, which can be cancelled out to a certain degree with the high number of observations. The number of PTBs averted derived thus should be treated as an approximation rather than a precise estimate.

6.4.3.1.2. Population-at-risk

The population at risk in the HIA conducted was assumed to be the unborn babies, as it was assumed that the foetus experiences the same exposure as the mother during the pregnancy. It was also justified as the HIA (as well as the CRF used) was for singleton PTBs, hence the number of births would equal the number of mothers. The ideal population-at-risk would be women of child-bearing age who are pregnant, as it accounts for multiple pregnancies, and foetuses could be lost due to miscarriage or death (stillbirth). However, information regarding the proportion of women of child-bearing age who are pregnant, or the proportion of women of child-bearing age with pregnancies that would result in a singleton live birth were not available from the data obtained. The smaller population-at-risk may lead to an underestimation of the true number of PTBs (or related outcomes such as stillbirths) being averted. The adjustment by percentage to obtain number of singleton live births did not consider the spatial differences in prevalence of singleton live births across England that could be present, which could lead to either an over-estimation or underestimations of the number of PTBs averted. No statistics on it is currently available, but it could be deduced that if more populated LAs have a lower prevalence of singleton live births, then it would lead to an overestimation of the number of PTBs averted.

6.4.3.1.3. Exposure misclassification

$PM_{2.5}$ concentrations were modelled for 2019, 2030, and 2040 only. Between these years, it was assumed that there would be a scaled constant decrease of $PM_{2.5}$ levels. However, that may not be true as fluctuations of $PM_{2.5}$ levels between the years would be possible due to short-to-medium term variations in weather and in policy implementation/uptake. The same applies to the assumed constant $PM_{2.5}$ levels after 2040 until 2154.

In addition, $PM_{2.5}$ concentrations were averaged to the 9-month preceding concentration before the births occurred. While this is better than using annual exposures and takes seasonal variations account, two issues persist: (1) all births are assumed to have occurred on the first day of the month, and (2) the births would have different gestational lengths but were assigned the same exposure. Differential exposure misclassification is present as the extent of misclassification is stronger in births with shorter gestational lengths or occurred at later days of the months.

6.4.3.2. Economic evaluation

The source of healthcare cost used was purely based on hospital admission costs (Hua et al., 2023). Hence, it is probably an underestimation of the healthcare cost incurred from PTB, as other sources of healthcare costs, such as outpatient services, community care, or prescription costs were not considered. Likewise, cost of lost productivity used is likely an underestimation as it did not consider lost productivity of the parents of the PTBs, which could incur from areas such as time spent on informal care. However, these are currently not quantified and not available in the UK.

For loss utility, the WTP estimate used for long term consequences of PTB was not developed specifically for PTB but rather very low birth weight. Although very low birth weight infants are highly likely to be born preterm and share the same long-term consequences, there could be subtle differences in the WTP estimate used. The WTP estimates used are also not QALY-weight-adjusted and thus could be an overestimation of the long-term consequences, as it did not consider the duration of the long-term consequences.

6.4.4. Future work and future perspectives

As the method of suppression introduces both systematic and random errors, future work conducted in England could consider accessing unsuppressed births data, or birth data that do not round large numbers to reduce random error. Requesting access for the use of birth data at greater spatial resolution, for instance at ward level from NHS England or ONS to avoid use of suppressed numbers could improve the precision of the number of PTBs averted. The final results published would be based on aggregated numbers across the years (policy window of interest) and thus the risk of statistical disclosure of personal information would be minimal, although information published alongside would still be subjected to approval by data holders. The work can be extended to the wider UK including Wales, Scotland, and Northern Ireland so the results can be better compared to the linked project. In addition, it could also be extended to other adverse birth outcomes, such as term low birth weight (TLBW), or other pollutants e.g. NO_2, which has been shown to have links with gestational air pollution exposure in this work or others (Gong et al., 2022).

The discussion on causality of the association in Chapter 5 suggested that PTB should be incorporated into future health impact assessments as a sensitivity analysis. Findings in this chapter echoes the recommendation, showing that PTB may not be the most substantial contributor to the total monetised health benefits for air pollution reduction. However, the

finding is specific to England and would not necessarily be applicable to other countries, especially those that have very different burden of diseases and $PM_{2.5}$ levels (Southerland et al., 2022; Vos et al., 2020; Yu et al., 2023b). If similar climate-change policies were introduced and analysed for countries that have a high burden of PTB attributable to air pollution, it is possible that the addition of PTB results to the analysis would be influential to the results. This would strengthen the support for emission reduction policies that combat climate change, but at the same time acknowledge that some health endpoints, in certain situations, would be more influential than others. The use of a CRF based on the association between whole pregnancy $PM_{2.5}$ exposure and the risk of PTB derived using HR can also be considered to derive the number of PTBs averted.

The Paris Agreement marks the beginning of transitioning to Net-Zero by 2050 for many countries (UN, 2016). Since then, there is an emerging body of literature both globally and in the UK that has analysed the health co-benefits in terms of mortality and morbidities improvements that would result from the transition (Hamilton et al., 2021; Milner et al., 2023; William et al., 2018; Walton et al., forthcoming). Many of the morbidities included in the analyses are those that appear later in life, for instance ischaemic heart disease (IHD) and dementia. Children's health outcomes have received less attention. Hence, this work has demonstrated to future work that outcomes at the start of life could be considered and included as well. Furthermore, adverse birth outcomes are associated with deprivation, and emission reduction policies with a wide range of impacts could produce greater benefits in more deprived populations (Genin et al., 2022; Ponce et al., 2005).

Chapter 7. Wider discussion and concluding remarks

7.1. Outline

This Chapter serves as an overarching discussion of the results from Chapter 3 to 6. A summary of the findings from these chapters is first provided, then the implications of the findings, when viewed in conjunction with each other as well as the wider literature, are discussed. Subsequently, how the findings could be applied in wider settings, in the context of policies and practices are explored. Then, some concluding remarks highlighting the contribution of this thesis with potential research directions are made.

7.2. Summary of key findings and consistency with recent literature

The key findings outlined in the present study are as follows: Methodological shortcomings were found in the analysis of preceding SR/MAs. Improvements to address previous methodological shortcomings were made, and the meta-analyses conducted showed that whole pregnancy $PM_{2.5}$ exposure was associated with increased risk of PTB, mlPTB, and vPTB based on studies reporting ORs and HRs. Trimester 3 NO_2 exposure was associated with increased risk of PTB based on studies reporting ORs and HRs. Stratified analyses showed that most factors were partial drivers of the heterogeneity observed. Using findings from the narrative synthesis and the wider literature, the association between whole pregnancy $PM_{2.5}$ exposure and the risk of PTB was determined to be “likely to be causal”, while that of trimester 3 NO_2 exposure and the risk of PTB was “suggestive to be causal”. The increase in risk of PTB associated with $PM_{2.5}$ exposure was used as a CRF in a health impact assessment quantifying the impact of air pollution reduction caused by Net-Zero policies in the England. Approximately 27,000 PTBs would be averted if the UK Government adopts the BNZP instead of the BAU scenario, and it corresponds to a monetised benefit of £6,000,000,000 from 2019 to 2154.

The above conclusion was based on a systematic literature search conducted in April 2022. It is possible that recent evidence could show different findings. Therefore, an updated literature search was conducted using the Ovid Medline search string of the present study in October 2024. Close to 400 additional articles were retrieved after the original cut-off date (11th April 2022). A brief screen of the article titles found around 90 studies could be relevant, demonstrating the importance of regular updates of high quality of the study area

(more of this discussed in section 7.4.1.). Results of the retrieved studies were broadly consistent with the findings in this thesis.

7.3. Wider discussion

7.3.1. Comparison of findings and the approach of causality determination

The findings of this thesis can be broadly compared to that of 2 preceding works, namely the USEPA Integrated Science Assessment (ISA) for $PM_{2.5}$ and NO_2, as they both assessed the likelihood of a causal relationship between $PM_{2.5}$/NO_2 and the risk of adverse birth outcomes based on the Bradford Hill Criteria (Owens et al., 2017).

The conclusion of "*likely to be a causal relationship*" for the effect of $PM_{2.5}$ exposure in this thesis offered more certainty than the "*suggestive of, but not sufficient to infer, a causal relationship"* in the USEPA ISA for $PM_{2.5}$ exposure for birth outcomes (USEPA, 2019). On the other hand, this work in this thesis and the USEPA ISA for NO_2 exposure both suggested that the evidence was "*suggestive of, but not sufficient to infer, a causal relationship"* (USEPA, 2016).

There are some differences between the USEPA ISAs and the work in this thesis. The USEPA ISAs were on pregnancy and birth outcomes in general, whereas this thesis focused on PTB only. The cut-off dates, which were approximately January 2018 for the $PM_{2.5}$ ISA (Sacks et al., 2022), and unclear for the NO_2 ISA, were earlier than this thesis. Moreover, the USEPA ISA for $PM_{2.5}$ and NO_2 used only studies from North America and Europe, and animal toxicological studies. On the other hand, this thesis included more recently published epidemiological evidence from all locations, as well as studies on possible mechanistic pathways on both humans and animal studies from which a higher risk of PTB was inferred using changes in biomarkers and occurrence of related adverse events. These could lead to the differences in conclusions made.

Some of the differences in the conclusions made for the effect of $PM_{2.5}$ exposure was from the greater confidence assigned to evidence from multipollutant models (MPMs), as the conclusion in this thesis was drawn from a greater pool of studies (Chapter 4 Figure 5) compared to only 2 MPM studies only in the USEPA ISA (Ha et al., 2014; Stieb et al., 2016a). In addition, in this thesis there was greater confidence on the biological plausibility when human instead of animal mechanistic evidence was used, as they are of greater relevance to human biology and avoid species-specific responses in animals that may not be

comparable to humans, and the biological doses used in the animal studies were not in the range of usual human exposure (Fussell et al., 2024). Lastly, several Bradford Hill criteria were strengthened due to evidence from quasi-experimental methods, constituents and sources of $PM_{2.5}$, which were not covered in the USEPA ISA. All in all, the two conclusions do not conflict with each other as the conclusion from the USEPA ISA was for all pregnancy outcomes, and the conclusion in this thesis was for PTB only.

The findings of this thesis resulted in the same conclusion as the USEPA NO_2 ISA published in 2016, despite the present work having utilised 6 additional years of evidence available. Both works concluded that there were inconsistencies in the epidemiological evidence. However, there were differences in how the two works arrived at the same conclusion. The conclusion of the ISA was based on 25 studies and did not conduct a meta-analysis, whereas this thesis' conclusion is based on a meta-analysis of close to 30 studies available, and 89 studies in the narrative synthesis. There are uncertainties in the increase in risk found in trimester 3, which was based on fewer studies than that of whole pregnancy exposure, which found a null association. In addition, in the current work the relatively low robustness of findings for NO_2 exposure in MPMs was considered and decreased confidence in the consistency of the evidence.

7.3.2. Insights from the meta-analyses of $PM_{2.5}$ exposure and different types of PTB

The findings from the increase in risk of PTB and its subtypes associated with $PM_{2.5}$ exposures can be summarised below:

1. Early exposure (at trimester 1) was associated with vPTB but not mlPTB and PTB as an overall category
2. Mid-gestation exposure (at trimester 2) was associated with vPTB, mlPTB, but not PTB as an overall category
3. Late exposure (at trimester 3) was associated with mlPTB and PTB
4. Whole gestation exposure was associated with all categories of PTB

At first glance, consistent with the literature discussed in **section 4.3.3.4.4**, the critical exposure window appears to be in late pregnancy. Consequently, more robust findings were observed in the third trimester and for whole pregnancy compared to the earlier trimesters. However, the apparent critical exposure window may actually be the period when an association between $PM_{2.5}$ exposure and the risk of PTB is most likely to be detected, rather than representing the actual critical period during which ambient $PM_{2.5}$ presents the greatest

risk to preterm delivery (when a PTB occurs). The stronger and more consistent findings for the third trimester and overall pregnancy exposure may be attributed to three factors:

- Rarity of the outcome: The effect on vPTB is more challenging to detect than on moderate to late preterm birth mlPTB due to its rarity, necessitating greater statistical power.
- Cumulative effect of exposure over time: Authors often do not account for the correlation of pollutant levels across trimesters or adjust for exposure levels in other trimesters.
- Timing between exposure and outcome: In the third trimester, there is inherently a higher likelihood of PTB, as the probability of delivery increases as pregnancy progresses.

Thus, one interpretation of the results is that there may not be a single critical window of exposure where the associated risk of PTB is greatest. Instead, the timing of exposure could influence the type of PTB that occurs. Early exposure may disrupt critical developmental processes in the placenta and foetus during the first trimester, making it difficult to sustain the pregnancy and leading to vPTB. Conversely, exposure during middle to late pregnancy may adversely affect later developmental processes. When they were considered together for the entire duration of pregnancy, an increased risk of PTB (from 24 to 37 weeks) can be observed. However, the issue of high inter-trimester correlation in exposure remains, and the risk associated with mlPTB in the later stages of pregnancy could still be due to cumulative effects across trimesters. This remains a hypothesis and would require support from future studies that would overcome some general limitations of the findings of the thesis shown below.

7.3.3. General limitations of this thesis

Some limitations are discussed below, and the strengths and contribution of this thesis are shown in the next section.

Most of the primary studies used in this thesis focused on the impact of air pollution exposure on the risk of PTB or its subgroups *during* pregnancy. However, one could argue that there is a high probability of a correlation between maternal exposure *before* pregnancy (preconception) and *during* pregnancy. Hence, it is possible that the observed risk reflects the effect of long-term, chronic higher exposure to certain air pollutants. In this thesis, there were insufficient studies available to provide a quantitative relationship between

preconception exposure and the risk of PTB, with available studies showing inconclusive findings. Nonetheless, studies of preconception exposure could offer a useful comparison of the risk before and during pregnancy.

Most of the adverse physiological changes brought by long-term air pollution exposure are not sex specific. Studies have suggested that a negative association between air pollution exposure and the quality of sperm as well as risk of infertility diagnosis among men (Kumar & Singh, 2022; Sørensen et al., 2024). It is therefore possible that parents with greater air pollution exposure are of different baseline health status compared to those with lower exposures even before pregnancy, and it is a combination of their exposures that resulted in a greater risk of PTB. One study that looked into paternal exposure to air pollution before pregnancy, in combination with maternal exposure, found an increased risk of PTB (Guo et al., 2023), with a mechanistic study suggesting paternal sperm DNA methylation as a mediator between paternal exposure and the risk of PTB (Cheng et al., 2024).

A persisting issue in the studies of air pollution exposure and the risk of PTB is the presence of live birth bias. Miscarriage and stillbirth are competing risks of PTB as they precede its occurrence. Most studies have only included pregnancies resulting in live births, which may underestimate the true effect of air pollution exposure on PTB, even when air pollution may not influence the risk of miscarriage and/or stillbirth, or when it increases the risk of miscarriage and/or stillbirth using a different mechanistic pathway to PTB. Only studies that have included information on miscarriage and/or stillbirth in addition to PTB in their analysis would be able to overcome the issue (Smith et al., 2020).

As miscarriage is often underreported (Department of Health and Social Care, 2023), it could be difficult to study the impact on miscarriage. However, there could be better reporting of data for stillbirth, and future primary studies can consider the impact of air pollution exposure on both stillbirth and PTB (Singh et al., 2024).

Globally, the likelihood of PTBs to survive into adulthood without major complications is drastically different in high versus low-income settings. The percentage of vPTBs that survived without impairments could be as high as 90% in the former, but less than 30% in the latter (March of Dimes et al., 2012). The difference in case-fatality risk reduced the generalisability of the findings to HIA in low-income settings, as the majority of primary studies used were based on high income settings. Moreover, the long-term consequences of PTB are not well-characterised. Emerging evidence suggests that being born preterm can be associated with a higher risk of all-cause mortality in later life in high income settings (Crump et al., 2020; Risnes et al., 2021), but evidence from other high income countries or other settings are required to corroborate the findings.

7.3.4. Implications of the findings for policymakers and pregnant women

In recent years, the scientific community have published various statements on the adverse health effects of air pollution which mentioned PTB (RCOG, 2022; RCPCH, 2024; Rajagopalan et al., 2020; Thurston et al., 2017). The conclusions from the meta-analyses in this thesis align with current knowledge. When viewed in conjunction with mechanistic evidence and studies of other study designs, it strengthened the evidence for a causal relationship between gestational $PM_{2.5}$ exposure and the risk of PTB, supporting its recognition as a significant health outcome in future statements. Pregnant women are recognised by the UK government as a group vulnerable to air pollution. However, more national level action is needed as limited individual actions can be taken by pregnant women to reduce their air pollution exposure (RCOG, 2022). The climate mitigation and adaptation policies used in the HIA of this thesis are policies that could reduce pregnant women's exposure to air pollution. Thus, the health and economic benefits of such policies could be underestimated if the impact on PTB is ignored.

The findings from the HIA and corresponding economic evaluation suggested that PTB should be considered in future HIAs of air pollution reduction, as the monetised benefits could be comparable to other health outcomes such as COPD. However, the evidence available may not be sufficient to advise pregnant women to undertake active measures to avoid air pollution exposure, as there are gaps shown in the causality criteria where the evidence needs to be strengthened. There is clinical advice for pregnant women to not smoke during pregnancy, alongside other advice such as stopping alcohol consumption and maintaining a healthy diet (NHS, 2024). The evidence of such advice or behavioural intervention is based on randomised controlled trials of interventions against standard care/advice (Poston et al., 2015). This is not available for air pollution and PTB.

There are various ways to reduce air pollution exposure, such as avoiding outdoor activities at high traffic areas or when ambient air pollution levels are high, using portable air cleaners, installing high-efficiency home air filtration systems, or using air-purifying respirators (Laumbach et al., 2015; Rajagopalan et al., 2020). Most of them target indoor rather than ambient air pollution exposure. It is possible to develop behavioural interventions focusing on avoidance of ambient air pollution levels, with evidence from control of asthma symptoms using the air quality index (AQI) (Rosser et al., 2023), and a quasi-experimental study investigating the effect of air quality alerts (Chen et al., 2018b). However, to the best of my knowledge, to date there is no intervention being tested on pregnant women to reduce exposure to air pollution. It is likely because most personal actions for pregnant women to

avoid ambient air pollution would require substantial life changes, such as moving residence to somewhere with low air pollution levels, which may not be feasible for most mothers. Giving out such advice is impractical and could lead to undue stress which could be harmful to the pregnancy. The use of face mask could be a relatively inexpensive method for reducing air pollution exposure. However, the efficacy of face mask wearing is heavily influenced by the type of mask used, fitting of the mask, and mask wearing practices (Cherrie et al., 2018). The filtration rate of face masks can also decrease when the filter has gotten wet, and a face mask may induce discomfort with long-use, and even give pregnant women a false sense of security and indirectly increase their exposure by staying in areas of higher pollution longer than they have intended (Huang and Morwska, 2019). Furthermore, there is an interplay between different risk factors for PTB that must be considered. Such advice may cause pregnant women to miss out on the benefits of physical activity and outdoor activities, while low physical activity and obesity are also risk factors for PTB.

Therefore, at this moment, policymakers may need to rely on quasi-experimental studies to develop the evidence on the use of tools such as air quality alerts, which can be coupled with enhanced messaging to encourage adherence to the accompanying advice on behavioural change (D'Antoni et al., 2019). Nonetheless, the effectiveness of these interventions could be limited, as they may reduce PTB attributable to short-term peaks in air pollution but not the cumulative exposure over the course of pregnancy.

In addition, the evidence developed in this thesis targets ambient air pollution exposure. Yet, pregnant women are exposed to pollution from households and other microenvironments. In these settings, $PM_{2.5}$, NO_2, and O_3, as well as other pollutants can be produced through processes such as cooking, heating, printing appliances or burning of candles or incense (Rivas et al., 2019; Wong et al., 2020). Numerous studies have investigated the impact of household air pollution exposure and the risk of PTB, but they mostly focused on certain sources such as unclean cooking fuel in developing countries, or certain events such as household renovation, with few investigating a joint effect (Liu et al., 2022c; Pan et al., 2022; Younger et al., 2022). There are far fewer studies available compared to the effect of ambient air pollution exposure. However, indoor environments could be where potential interventions such as the use of portable air cleaners could be tested. Hence, if clinical advice specifically targeting pregnant women were needed, evidence might be built from the impact of interventions targeting indoor air pollution. Before such evidence is deemed sufficient and causal, it is likely that policies at a top-down level that aim to regulate and reduce emissions would be the main measure available to reduce ambient air pollution exposure of pregnant women (RCOG, 2022).

7.4. Implications of this thesis for future research

7.4.1. Methodological decisions are critical in systematic reviews and should be well-informed

The analysis on the findings and practices of the preceding systematic reviews and meta-analyses provided a solid foundation for the methodological choices in terms of study retrieval, study selection, and strategies for reducing study dependencies. Such analysis, albeit not advanced, provided a useful comparison and visualisation of the issues. The recurrent methodological issues found suggested that preceding work may not have paid sufficient attention to these issues, which therefore undermined the robustness of their findings.

The modified WHO Risk of Bias assessment instrument was developed to provide a detailed and comprehensive quality assessment of the primary studies eligible to be included in the narrative synthesis. It showed that most studies have some level of methodological shortcomings in exposure assessment and adjustment of confounding, contrasting what was suggested by preceding work. Risk of bias was determined to not be a driver of heterogeneity for the effect of $PM_{2.5}$ and O_3 exposure, but was a partial driver for NO_2 exposure, as the studies assigned a low risk of bias showed different findings compared to those that were not. Hence, there is likely a need for more high-quality studies of NO_2 exposure, as most studies retrieved as well as the recent studies still mainly focused on $PM_{2.5}$ or its constituents. Therefore, at the current state of the evidence, it is difficult to differentiate if the effect of NO_2 exposure is indeed smaller or less specific than $PM_{2.5}$, or if more evidence would make the likelihood of a causal effect similar to that of $PM_{2.5}$.

The strategy for reducing study dependencies was unique as it incorporated risk of bias into consideration, which none of the preceding work has except the review by HEI that took confounder adjustment into account. Therefore, the present study was able to keep studies of lower risk of bias in the meta-analysis and maintain robustness of the findings. However, one of the issues this thesis did not explore was the effect of using different criteria in the process of reducing study dependencies, or the extent of “tolerable” study dependencies, on the final summary estimate. Therefore, future systematic reviews can consider investigating the impact, as different strategies, such as emphasising sample size over risk of bias, may produce different results compared to the one adopted in this thesis.

One of the methodological shortcomings of previous reviews identified in Chapter 3 was the lack of transparent reporting, or actual contact of, authors of primary studies, with only two reviews stating that attempts were made (HEI, 2022; Stieb et al., 2012). It was unclear what

the impact of these shortcomings would be. In the present study, the authors contacted responded at different times over the course of the study, the results of the meta-analyses were continuously updated, and a comparison of the results with or without authors' replies was not possible for all pollutant-exposure period-outcome relationships except for the effect of whole pregnancy $PM_{2.5}$ exposure and the risk of PTB, inclusive of studies assigned a high RoB. Contacting authors led to seven additional estimates being available in the meta-analysis, although only a slight shift in the summary estimate from 1.09 (Chapter 5 Figure 32) to 1.08 (data not shown) was observed. The limited change observed was likely due to the relatively high number of studies already available for the specific exposure-outcome pair. However, the effect on exposure-outcome pairs with fewer studies available may become larger as the new primary studies included would have a relatively heavier weighting within a small group of studies than within a large group. Thus, efforts should be made to contact authors for additional information in systematic reviews and meta-analyses. In addition, reviewers or editors of journals can improve reporting of articles by requiring information such as exposure increment to be necessary when reviewing studies of air pollution exposure and birth outcomes, to minimise additional efforts required by future systematic reviews.

All in all, future researchers should critically appraise the preceding work before conducting their own evidence syntheses, as it can help advance the research area by providing more robust work and maximise completeness by having the whole set of studies available.

7.4.2. Comprehensive coverage of literature in the narrative synthesis improves causality determination

The systematic review in this thesis is more thorough than preceding ones as it provided a narrative synthesis of the studies that may not be covered in detail in meta-analyses, through the use of a list of questions that these studies had aimed to answer (**section 4.3.3.1 to 4.3.3.4**). For instance, if the association between $PM_{2.5}$ exposure and the risk of PTB was robust to the adjustment for other pollutants. Most of the questions were not given a conclusive answer due to the relatively few studies available. Moreover, a quantitative relationship was not derived when there were more than 3 studies available due to time-constraints. Hence, future systematic reviews and meta-analyses can consider adopting a similar approach, and use the findings from the narrative synthesis to decide if additional meta-analyses can be conducted to derive a quantitative relationship for specific groups of estimates. For instance, estimates from co-exposures or mothers with certain conditions can be used to understand if there are specific population groups with greater susceptibility to

the effect of air pollution exposure. In the next syntheses of evidence of the area, there could be other environmental exposures of interest, such as greenness measured by Normalised Difference Vegetation Index (NDVI) that shows the density of vegetation of an area. In the updated literature search, there were a few studies that studied how exposure to greenness has negatively modified the association between air pollution exposure and the risk of PTB (Mi et al., 2024; Sinsamala et al., 2024; Yu et al., 2024).

The causality discussion using the Bradford Hill criteria was partly based on the information derived from these questions. Thus, the evidence for some of the criteria used would likely be strengthened in the future. For the criterion of analogy, the evidence from constituents of $PM_{2.5}$ would likely be strengthened with newer studies (Meng et al., 2023; Qiao et al., 2022), warranting a meta-analysis that could supplement the findings of total $PM_{2.5,}$ and provide a useful comparison with the findings from the HEI traffic review, the only meta-analysis conducted for the effects of $PM_{2.5}$ constituents currently available. This found no association between EC exposure and the risk of PTB (HEI, 2022). For the experiment criterion, the updated search retrieved two natural experiment studies (Jiang et al., 2022; Ni et al., 2024), with one on the impact of COVID-19 lockdown (Ni et al., 2024). Since lockdowns had occurred in various countries during the COVID-19 pandemic, it presents an opportunity for researchers to study its impact using a natural experiment approach. However, the lockdown introduced drastic changes in the daily life of most individuals and thus introduce unusual confounders, and COVID-19 infection could increase the risk of PTB to a different extent depending on its severity (Smith et al., 2022).

There is a rise in the use of other causal inference methods, including instrumental variable (IV) analysis and g-computation. While they do not contribute to the criterion of experiment, they can contribute to various other criteria, such as consistency, strength, coherence, or plausibility. IV analysis estimates the causal effect of an exposure on an outcome, by emulating randomisation through the use of a variable that is related to the exposure but not the outcome (Iwashyna & Kennedy, 2013). Two studies investigated the effect of O_3 exposure using IV analysis and found an increase in risk of adverse birth outcomes (Bai et al., 2022; Noghanibehambari et al., 2023). They suggested the possibility that traditional observational studies used in the present review may have failed to detect or underestimated the effect of O_3. On the other hand, g-computation estimates the causal effect through the use of a parametric regression model to estimate the potential counterfactual outcomes (e.g. exposed to higher or lower levels of air pollution) (Robins, 1986; Snowden et al., 2011). Two studies that used g-computation found an increase in risk of PTB due to $PM_{2.5}$ exposure at either mid-gestation or late pregnancy (Leung et al., 2024; Shi et al., 2024). Nonetheless, very few of these studies are available for PTB to date, and

the causal effect estimated by g-computation assumes no unmeasured confounding, and thus, is heavily reliant on the robustness of the model.

In the updated literature search, several studies using populations at locations not found in the present study were found, including South Africa (Mitku et al., 2023), Colombia (Márquez-Lázaro et al., 2024), Northern Ireland (Jahanshahi et al., 2024), and rural regions of China (Li et al., 2024). Incorporation of these studies into future systematic reviews and meta-analyses would enhance generalisability of findings, and the ability of stratified analyses to detect regional heterogeneity.

7.4.3. High quality evidence synthesis is necessary

The present study reaffirms the necessity of the presence of high-quality systematic reviews and meta-analyses. Through the analysis presented in Chapter 3, methodological considerations shown in Chapter 4, and the meta-analytical results of Chapter 5, it demonstrated that the presence of numerous systematic reviews and meta-analyses, which showed some agreement with each other, is not necessarily a good indication of consensus (Ioannidis et al., 2016; Ioannidis et al., 2018). A rigorously conducted, high-quality systematic review and meta-analysis could arrive at different conclusions. The narrative synthesis in Chapter 4 also generated insights which were not covered by any preceding systematic reviews and meta-analyses. It is evident that conducting an evidence synthesis using a similar process to that conducted from Chapter 3 to 6 requires monumental effort and is subject to having sufficient time available. Yet, such effort is of paramount importance with a rapid rise in number of systematic reviews and meta-analyses that may be less thorough and thus give misleading results, and the increasing complexity in study methodologies and findings. The identification of errors in previous works by the forthcoming COMEAP report has resulted in the retraction of 2 published meta-analyses, and the publication of a correction of a third meta-analysis paper (COMEAP, forthcoming).

Development of methods that can shorten the time required to produce comprehensive evidence syntheses without sacrificing their quality is important, as more than 70% of systematic reviews were outdated by the time they were published (Shojania et al., 2007). These methods should be applicable to both rapid or traditional systematic reviews, as regulatory agencies or policymakers require recommendations to be made within a relatively short timeframe, and rapid methods are prone to error and may produce less reliable results (Marshall et al., 2019; Moons et al., 2021). Some of the methods are discussed below.

First, a living systematic review, which allows continuous or regular updates using new and relevant evidence can be conducted. Using the data extracted from the present study, a database can be developed to hold information of the retrieved studies, then new evidence retrieved from regular searches of databases can be incorporated (Cochrane, 2019). A living systematic review could drastically enhance research efficiency by utilising previous efforts on screening and data extraction, thus enabling greater focus on evidence synthesis. Moreover, it could shorten the delay between evidence availability, integration into review, and publication of the review, enabling policymakers to use up-to-date evidence in a timely manner (Millard et al., 2019; Winter et al., 2014). In addition, evidence on air pollution exposure and the risk of PTB broadly fits Cochrane's recommendation for adopting a "living" review approach, as there are uncertainties remaining in the existing evidence (as shown in Chapter 5, particularly for gaseous pollutants), high research interest (as shown in Chapter 3), as well as potential influences in decision-making (as shown in Chapter 6) (Cochrane, 2019). However, such an approach is resource-intensive and requires consistent funding, maintenance and staffing for the project to avoid inconsistencies.

Second, the use of data repositories for systematic reviews of environmental health needs to be advocated. There are repositories available for health risk assessments (Health Assessment Work Collaborative, 2024) or systematic reviews in general (AHRQ, 2024; Saldanha et al., 2019). However, none of the SR/MAs retrieved in Chapter 3, as well as the present study, had used online data repositories. The possibility of doing so should be explored to further improve reproducibility of findings.

Third, the possibility of using artificial intelligence (AI) tools should be carefully explored. Large-language models (LLMs) are transformers which were trained on large amounts of text data to learn language structures and patterns. When a prompt is used, it generates responses by predicting the next token (word) which is most coherently and relevant to the context of the prompt (Vaswani et al., 2017). Since the release of OpenAI's LLM Chat-GPT (OpenAI, 2022), a variety of tools for assisting various stages of systematic review became available to researchers (Fabiano et al., 2024). While the use of these tools can drastically reduce the time needed for researchers to conduct systematic reviews, it has raised concerns in reporting of use, production of erroneous information, and risk of non-expert use (Cacciamani et al., 2023; Qureshi et al., 2023). As the technology is still at its infancy, their accuracy in performing tasks such as data extraction should be validated and tested against manual performances, and incorporating of AI tools would likely require additional steps to safeguard methodological rigour and quality (Gartlehner et al., 2024; van Dijk et al., 2023).

7.4.4. PTB should be one of the health outcomes considered in HIA for air pollution

The present study put findings for PTB into the context of a wide range of other health outcomes in a health impact assessment of net zero policies. This was an advance over Milner et al. (2023a) who only compared PTB impacts with other adverse birth or child health outcomes. As described in Chapter 6, the number of PTBs averted was small compared to other health outcomes included in a similar project. However, the monetised benefits were similar to outcomes such as COPD and ischaemic heart disease. Uncertainty emphatically exists for the results of Chapter 6, due to the choice of the concentration-response function (CRF), for instance a CRF based on HR could have been used as an alternative, and inputs for the economic evaluation, as demonstrated in the sensitivity analyses. This would reduce the number of PTBs averted and change the monetised benefits quantified. The longer-term consequences of PTB were not considered in this work, which potentially leads to an underestimation of the health benefits quantified. Nonetheless, the results of the HIA suggested a need to broaden the range of health outcomes that should be considered by policymakers when a health impact assessment related to air pollution is conducted.

References

Aas, W., Fagerli, H., Alastuey, A., Cavalli, F., Degorska, A., Feigenspan, S., Brenna, H., Gliß, J., Heinesen, D., Hueglin, C., Holubová, A., Jaffrezo, J.-L., Mortier, A., Murovec, M., Putaud, J.-P., Rüdiger, J., Simpson, D., Solberg, S., Tsyro, S., Tørseth, K. & Yttri, K. E. 2024. Trends in Air Pollution in Europe, 2000–2019. Aerosol and Air Quality Research, 24, 4, 230237. http://dx.doi.org/10.4209/aaqr.230237

Abdo, M., Ward, I., O'dell, K., Ford, B., Pierce, J. R., Fischer, E. V. & Crooks, J. L. 2019. Impact of Wildfire Smoke on Adverse Pregnancy Outcomes in Colorado, 2007-2015. Int J Environ Res Public Health, 16, 19. https://www.mdpi.com/1660-4601/16/19/3720

Achana, F., Johnson, S., Ni, Y., Marlow, N., Wolke, D., Khan, K. & Petrou, S. 2022. Economic costs and health utility values associated with extremely preterm birth: Evidence from the EPICure2 cohort study. Paediatr Perinat Epidemiol, 36, 5, 696-705.

Achilleos, S., Kioumourtzoglou, M. A., Wu, C. D., Schwartz, J. D., Koutrakis, P. & Papatheodorou, S. I. 2017. Acute effects of fine particulate matter constituents on mortality: A systematic review and meta-regression analysis. Environ Int, 109, 89-100.

Adamowicz, W., Boxall, P., Williams, M. & Louviere, J. 1998. Stated Preference Approaches for Measuring Passive Use Values: Choice Experiments and Contingent Valuation. American Journal of Agricultural Economics, 80, 1, 64-75. http://www.jstor.org/stable/3180269

Adane, H. A., Iles, R., Boyle, J. A., Gelaw, A. & Collie, A. 2023. Maternal Occupational Risk Factors and Preterm Birth: A Systematic Review and Meta-Analysis. Public Health Reviews, 44. https://www.ssph-journal.org/journals/public-health-reviews/articles/10.3389/phrs.2023.1606085

Agency for Healthcare Research and Quality (Ahrq). 2024. *About the Systematic Review Data Repository Plus (SRDR+)* [Online]. Available: https://srdrplus.ahrq.gov/about

Ajzen, I. 1991. The theory of planned behavior. Organizational Behavior and Human Decision Processes, 50, 2, 179-211. https://www.sciencedirect.com/science/article/pii/074959789190020T

Alman, B. L., Stingone, J. A., Yazdy, M., Botto, L. D., Desrosiers, T. A., Pruitt, S., Herring, A. H., Langlois, P. H., Nembhard, W. N., Shaw, G. M., Olshan, A. F. & Luben, T. J. 2019. Associations between PM(2.5) and risk of preterm birth among liveborn infants. Ann Epidemiol, 39, 46-53.e2.

Anderson, G. B. & Bell, M. L. 2010. Does one size fit all? The suitability of standard ozone exposure metric conversion ratios and implications for epidemiology. J Expo Sci Environ Epidemiol, 20, 1, 2-11.

Andersson, H. 2020. Chapter Three - The value of a statistical life. *In:* Mouter, N. (ed.) *Advances in Transport Policy and Planning.* Academic Press. Available: https://www.sciencedirect.com/science/article/pii/S2543000920300172

Andrawus, M., Sharvit, L. & Atzmon, G. 2022. Epigenetics and Pregnancy: Conditional Snapshot or Rolling Event. Int J Mol Sci, 23, 20.

Anenberg, S. C., Mohegh, A., Goldberg, D. L., Kerr, G. H., Brauer, M., Burkart, K., Hystad, P., Larkin, A., Wozniak, S. & Lamsal, L. 2022. Long-term trends in urban NO2 concentrations and associated paediatric asthma incidence: estimates from global datasets. The Lancet Planetary Health, 6, 1, e49-e58. https://doi.org/10.1016/S2542-5196(21)00255-2

Appel, K. W., Napelenok, S. L., Foley, K. M., Pye, H. O. T., Hogrefe, C., Luecken, D. J., Bash, J. O., Roselle, S. J., Pleim, J. E., Foroutan, H., Hutzell, W. T., Pouliot, G. A., Sarwar, G., Fahey, K. M., Gantt, B., Gilliam, R. C., Heath, N. K., Kang, D., Mathur, R., Schwede, D. B., Spero, T. L., Wong, D. C. & Young, J. O. 2017. Description and evaluation of the Community Multiscale Air Quality (CMAQ) modeling system version 5.1. Geosci Model Dev, 10, 4, 1703-1732.

Apsimon, H., Oxley, T., Woodward, H., Mehlig, D., Holland, M. & Reeves, S. 2023. Integrated Assessment Modelling of Future Air Quality in the UK to 2050 and Synergies with Net-Zero Strategies. Atmosphere, 14, 3, 525.

Arroyave, W. D., Mehta, S. S., Guha, N., Schwingl, P., Taylor, K. W., Glenn, B., Radke, E. G., Vilahur, N., Carreón, T., Nachman, R. M. & Lunn, R. M. 2021. Challenges and recommendations on the conduct of systematic reviews of observational epidemiologic studies in environmental and occupational health. J Expo Sci Environ Epidemiol, 31, 1, 21-30.

Ashmore, M. 2005. Assessing the future global impacts of ozone on vegetation. Plant, Cell & Environment, 28, 8, 949-964. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-3040.2005.01341.x

Atkinson, R. W., Kang, S., Anderson, H. R., Mills, I. C. & Walton, H. A. 2014. Epidemiological time series studies of PM2.5 and daily mortality and hospital admissions: a systematic review and meta-analysis. Thorax, 69, 7, 660-5.
Bachwenkizi, J., Liu, C., Meng, X., Zhang, L., Wang, W., Van Donkelaar, A., Martin, R. V., Hammer, M. S., Chen, R. & Kan, H. 2022. Maternal exposure to fine particulate matter and preterm birth and low birth weight in Africa. Environment International, 160, 107053. Available: https://www.sciencedirect.com/science/article/pii/S0160412021006784
Bai, S., Du, S., Liu, H., Lin, S., Zhao, X., Wang, Z. & Wang, Z. 2022. The causal and independent effect of ozone exposure during pregnancy on the risk of preterm birth: Evidence from northern China. Environmental Research, 214, 113879. Available: https://www.sciencedirect.com/science/article/pii/S0013935122012063
Baklanov, A. & Zhang, Y. 2020. Advances in air quality modeling and forecasting. Global Transitions, 2, 261-270. https://www.sciencedirect.com/science/article/pii/S2589791820300232
Baldo, C., Zhong, J., Hall, J., Muller, C. L., Bartington, S. E. & Bloss, W. J. 2023. New Air Quality Targets & Interim Goals for Fine Particulate Matter–PM2. 5: Implications for the West Midlands. NA.
Balduzzi, S., Rücker, G. & Schwarzer, G. 2019. How to perform a meta-analysis with {R}: a practical tutorial. Evidence-Based Mental Health, 22, 153-160.
Barba-Vasseur, M., Bernard, N., Pujol, S., Sagot, P., Riethmuller, D., Thiriez, G., Houot, H., Defrance, J., Mariet, A. S., Luu, V. P., Barbier, A., Benzenine, E., Quantin, C. & Mauny, F. 2017. Does low to moderate environmental exposure to noise and air pollution influence preterm delivery in medium-sized cities? Int J Epidemiol, 46, 6, 2017-2027.
Basu, R., Pearson, D., Ebisu, K. & Malig, B. 2017. Association between PM(2.5) and PM(2.5) Constituents and Preterm Delivery in California, 2000-2006. Paediatr Perinat Epidemiol, 31, 5, 424-434. Available: https://doi.org/10.1111/ppe.12380
Baxter, L. K., Dionisio, K. L., Burke, J., Ebelt Sarnat, S., Sarnat, J. A., Hodas, N., Rich, D. Q., Turpin, B. J., Jones, R. R., Mannshardt, E., Kumar, N., Beevers, S. D. & Özkaynak, H. 2013. Exposure prediction approaches used in air pollution epidemiology studies: key findings and future recommendations. J Expo Sci Environ Epidemiol, 23, 6, 654-9.
Beck, H. E., Zimmermann, N. E., Mcvicar, T. R., Vergopolan, N., Berg, A. & Wood, E. F. 2018. Present and future Köppen-Geiger climate classification maps at 1-km resolution. Sci Data, 5, 180214.
Beelen, R., Hoek, G., Vienneau, D., Eeftens, M., Dimakopoulou, K., Pedeli, X., Tsai, M.-Y., Künzli, N., Schikowski, T., Marcon, A., Eriksen, K. T., Raaschou-Nielsen, O., Stephanou, E., Patelarou, E., Lanki, T., Yli-Tuomi, T., Declercq, C., Falq, G., Stempfelet, M., Birk, M., Cyrys, J., Von Klot, S., Nádor, G., Varró, M. J., Dėdelė, A., Gražulevičienė, R., Mölter, A., Lindley, S., Madsen, C., Cesaroni, G., Ranzi, A., Badaloni, C., Hoffmann, B., Nonnemacher, M., Krämer, U., Kuhlbusch, T., Cirach, M., De Nazelle, A., Nieuwenhuijsen, M., Bellander, T., Korek, M., Olsson, D., Strömgren, M., Dons, E., Jerrett, M., Fischer, P., Wang, M., Brunekreef, B. & De Hoogh, K. 2013. Development of NO2 and NOx land use regression models for estimating air pollution exposure in 36 study areas in Europe – The ESCAPE project. Atmospheric Environment, 72, 10-23. Available: https://www.sciencedirect.com/science/article/pii/S1352231013001386
Beevers, S., Assareh, N., Beddows, A., Stewart, G., Holland, M., Fecht, D., Liu, Y., Goodman, A., Walton, H., Brand, C., Evangelopoulos, D., Wood, D., Vu, T. & Dajnak, D. forthcoming. Climate change policies can reduce air pollution, improve health and inequalities, while providing economic benefits.
Bekkar, B., Pacheco, S., Basu, R. & Denicola, N. 2020. Association of Air Pollution and Heat Exposure With Preterm Birth, Low Birth Weight, and Stillbirth in the US: A Systematic Review. JAMA Network Open, 3, 6, e208243-e208243. Available: https://doi.org/10.1001/jamanetworkopen.2020.8243
Belur, J., Tompson, L., Thornton, A. & Simon, M. 2018. Interrater Reliability in Systematic Review Methodology: Exploring Variation in Coder Decision-Making. Sociological Methods & Research, 50, 2, 837-865. Available: https://doi.org/10.1177/0049124118799372
Benmarhnia, T., Huang, J., Basu, R., Wu, J. & Bruckner, T. A. 2017. Decomposition Analysis of Black–White Disparities in Birth Outcomes: The Relative Contribution of Air Pollution and Social Factors in California. Environmental Health Perspectives, 125, 10, 107003. Available: https://ehp.niehs.nih.gov/doi/abs/10.1289/EHP490
Berman, N. G. & Parker, R. A. 2002. Meta-analysis: Neither quick nor easy. BMC Medical Research Methodology, 2, 1, 10. Available: https://doi.org/10.1186/1471-2288-2-10

Bertin, M., Chevrier, C., Serrano, T., Monfort, C., Rouget, F., Cordier, S. & Viel, J. F. 2015. 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? Environ Res, 142, 17-24.
Bhatia, A., Krieger, N., Beckfield, J., Barros, A. J. D. & Victora, C. 2019. Are inequities decreasing? Birth registration for children under five in low-income and middle-income countries, 1999–2016. BMJ Global Health, 4, 6, e001926. Available: http://gh.bmj.com/content/4/6/e001926.abstract
Blair, A., Stewart, P., Lubin, J. H. & Forastiere, F. 2007. Methodological issues regarding confounding and exposure misclassification in epidemiological studies of occupational exposures. Am J Ind Med, 50, 3, 199-207.
Blanc, N., Liao, J., Gilliland, F., Zhang, J., Berhane, K., Huang, G., Yan, W. & Chen, Z. 2023. A systematic review of evidence for maternal preconception exposure to outdoor air pollution on Children's health. Environmental Pollution, 318, 120850. Available: https://www.sciencedirect.com/science/article/pii/S0269749122020656
Blencowe, H., Cousens, S., Chou, D., Oestergaard, M., Say, L., Moller, A. B., Kinney, M. & Lawn, J. 2013. Born too soon: the global epidemiology of 15 million preterm births. Reprod Health, 10 Suppl 1, Suppl 1, S2.
Bobak, M. & Leon, D. A. 1999. Pregnancy outcomes and outdoor air pollution: an ecological study in districts of the Czech Republic 1986-8. Occup Environ Med, 56, 8, 539-43.
Boing, A. F., Desouza, P., Boing, A. C., Kim, R. & Subramanian, S. V. 2022. Air Pollution, Socioeconomic Status, and Age-Specific Mortality Risk in the United States. JAMA Network Open, 5, 5, e2213540-e2213540. Available: https://doi.org/10.1001/jamanetworkopen.2022.13540
Bongaerts, E., Lecante, L. L., Bové, H., Roeffaers, M. B. J., Ameloot, M., Fowler, P. A. & Nawrot, T. S. 2022. 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, 6, 10, e804-e811. Available: https://doi.org/10.1016/S2542-5196(22)00200-5
Bonzini, M., Carugno, M., Grillo, P., Mensi, C., Bertazzi, P. & Pesatori, A. 2010. Impact of ambient air pollution on birth outcomes: Systematic review of the current evidences. La Medicina del lavoro, 101, 341-63.
Boogaard, H., Atkinson, R. W., Brook, J. R., Chang, H. H., Hoek, G., Hoffmann, B., Sagiv, S. K., Samoli, E., Smargiassi, A., Szpiro, A. A., Vienneau, D., Weuve, J., Lurmann, F. W. & Forastiere, F. 2023. Evidence Synthesis of Observational Studies in Environmental Health: Lessons Learned from a Systematic Review on Traffic-Related Air Pollution. Environmental Health Perspectives, 131, 11, 115002. Available: https://ehp.niehs.nih.gov/doi/abs/10.1289/EHP11532
Bové, H., Bongaerts, E., Slenders, E., Bijnens, E. M., Saenen, N. D., Gyselaers, W., Van Eyken, P., Plusquin, M., Roeffaers, M. B. J., Ameloot, M. & Nawrot, T. S. 2019. Ambient black carbon particles reach the fetal side of human placenta. Nature Communications, 10, 1, 3866. Available: https://doi.org/10.1038/s41467-019-11654-3
Brauer, M., Lencar, C., Tamburic, L., Koehoorn, M., Demers, P. & Karr, C. 2008. A cohort study of traffic-related air pollution impacts on birth outcomes. Environmental Health Perspectives, 116, 5, 680-686. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-45749117202&doi=10.1289%2fehp.10952&partnerID=40&md5=c9addfae2370842f9f89fb6fd63b6b9e
Brauer, M., Roth, G. A., Aravkin, A. Y., Zheng, P., Abate, K. H., Abate, Y. H., Abbafati, C., Abbasgholizadeh, R., Abbasi, M. A., Abbasian, M., Abbasifard, M., Abbasi-Kangevari, M., Abd Elhafeez, S., Abd-Elsalam, S., Abdi, P., Abdollahi, M., Abdoun, M., Abdulah, D. M., Abdullahi, A., Abebe, M., Abedi, A., Abedi, A., Abegaz, T. M., Abeldaño Zuñiga, R. A., Abiodun, O., Abiso, T. L., Aboagye, R. G., Abolhassani, H., Abouzid, M., Aboye, G. B., Abreu, L. G., Abualruz, H., Abubakar, B., Abu-Gharbieh, E., Abukhadijah, H. J. J., Aburuz, S., Abu-Zaid, A., Adane, M. M., Addo, I. Y., Addolorato, G., Adedoyin, R. A., Adekanmbi, V., Aden, B., Adetunji, J. B., Adeyeoluwa, T. E., Adha, R., Adibi, A., Adnani, Q. E. S., Adzigbli, L. A., Afolabi, A. A., Afolabi, R. F., Afshin, A., Afyouni, S., Afzal, M. S., Afzal, S., Agampodi, S. B., Agbozo, F., Aghamiri, S., Agodi, A., Agrawal, A., Agyemang-Duah, W., Ahinkorah, B. O., Ahmad, A., Ahmad, D., Ahmad, F., Ahmad, N., Ahmad, S., Ahmad, T., Ahmed, A., Ahmed, A., Ahmed, A., Ahmed, L. A., Ahmed, M. B., Ahmed, S., Ahmed, S. A., Ajami, M., Akalu, G. T., Akara, E. M., Akbarialiabad, H., Akhlaghi, S., Akinosoglou, K., Akinyemiju, T., Akkaif, M. A.,

Akkala, S., Akombi-Inyang, B., Al Awaidy, S., Al Hasan, S. M., Alahdab, F., Al-Ahdal, T. M. A., Alalalmeh, S. O., Alalwan, T. A., Al-Aly, Z., Alam, K., Alam, N., Alanezi, F. M., Alanzi, T. M., Albakri, A., Albataineh, M. T., Aldhaleei, W. A., Aldridge, R. W., et al. 2024. Global burden and strength of evidence for 88 risk factors in 204 countries and 811 subnational locations, 1990-2021: a systematic analysis for the Global Burden of Disease Study 2021. The Lancet, 403, 10440, 2162-2203. Available: https://doi.org/10.1016/S0140-6736(24)00933-4

Braun, M., Koger, F., Klingelhöfer, D., Müller, R. & Groneberg, D. A. 2019. Particulate Matter Emissions of Four Different Cigarette Types of One Popular Brand: Influence of Tobacco Strength and Additives. Int J Environ Res Public Health, 16, 2.

Brook, R. D., Rajagopalan, S., Pope, C. A., Brook, J. R., Bhatnagar, A., Diez-Roux, A. V., Holguin, F., Hong, Y., Luepker, R. V., Mittleman, M. A., Peters, A., Siscovick, D., Smith, S. C., Whitsel, L. & Kaufman, J. D. 2010. Particulate Matter Air Pollution and Cardiovascular Disease. Circulation, 121, 21, 2331-2378. Available: https://doi.org/10.1161/CIR.0b013e3181dbece1

Brown, J. M., Harris, G., Pantea, C., Hwang, S.-A. & Talbot, T. O. 2015. Linking Air Pollution Data and Adverse Birth Outcomes: Environmental Public Health Tracking in New York State. Journal of Public Health Management and Practice, 21. Available: https://journals.lww.com/jphmp/fulltext/2015/03001/linking_air_pollution_data_and_adverse_birth.10.aspx

Cacciamani, G. E., Chu, T. N., Sanford, D. I., Abreu, A., Duddalwar, V., Oberai, A., Kuo, C. C. J., Liu, X., Denniston, A. K., Vasey, B., Mcculloch, P., Wolff, R. F., Mallett, S., Mongan, J., Kahn, C. E., Sounderajah, V., Darzi, A., Dahm, P., Moons, K. G. M., Topol, E., Collins, G. S., Moher, D., Gill, I. S. & Hung, A. J. 2023. PRISMA AI reporting guidelines for systematic reviews and meta-analyses on AI in healthcare. Nature Medicine, 29, 1, 14-15. Available: https://doi.org/10.1038/s41591-022-02139-w

Cai, J., Zhao, Y., Kan, J., Chen, R., Martin, R., Van Donkelaar, A., Ao, J., Zhang, J., Kan, H. & Hua, J. 2020. Prenatal Exposure to Specific PM2.5 Chemical Constituents and Preterm Birth in China: A Nationwide Cohort Study. Environmental Science & Technology, 54, 22, 14494-14501. Available: https://doi.org/10.1021/acs.est.0c02373

Cao, K., Jin, H., Li, H., Tang, M., Ge, J., Li, Z., Wang, X. & Wei, X. 2022. Associations of maternal exposure to fine particulate matter with preterm and early-term birth in high-risk pregnant women. Genes and Environment, 44, 1, 9. https://doi.org/10.1186/s41021-022-00239-0

Capobussi, M., Tettamanti, R., Marcolin, L., Piovesan, L., Bronzin, S., Gattoni, M. E., Polloni, I., Sabatino, G., Tersalvi, C. A., Auxilia, F. & Castaldi, S. 2016. Air Pollution Impact on Pregnancy Outcomes in Como, Italy. Journal of Occupational and Environmental Medicine, 58, 1, 47-52. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84958793833&doi=10.1097%2fJOM.0000000000000630&partnerID=40&md5=e3d431a3ea0dee88d026e5450762fe8a

Carslaw, D. C. & Ropkins, K. 2012. openair — An R package for air quality data analysis. Environmental Modelling & Software, 27-28, 52-61. Available: https://www.sciencedirect.com/science/article/pii/S1364815211002064

Cassidy-Bushrow, A. E., Burmeister, C., Lamerato, L., Lemke, L. D., Mathieu, M., O'leary, B. F., Sperone, F. G., Straughen, J. K. & Reiners, J. J., Jr. 2020. Prenatal airshed pollutants and preterm birth in an observational birth cohort study in Detroit, Michigan, USA. Environ Res, 189, 109845.

Chang, H. H., Reich, B. J. & Miranda, M. L. 2012. Time-to-Event Analysis of Fine Particle Air Pollution and Preterm Birth: Results From North Carolina, 2001–2005. American Journal of Epidemiology, 175, 2, 91-98. Available: https://doi.org/10.1093/aje/kwr403

Chang, H. H., Warren, J. L., Darrow, L. A., Reich, B. J. & Waller, L. A. 2015. Assessment of critical exposure and outcome windows in time-to-event analysis with application to air pollution and preterm birth study. Biostatistics, 16, 3, 509-521. Available: https://doi.org/10.1093/biostatistics/kxu060

Chawanpaiboon, S., Vogel, J. P., Moller, A. B., Lumbiganon, P., Petzold, M., Hogan, D., Landoulsi, S., Jampathong, N., Kongwattanakul, K., Laopaiboon, M., Lewis, C., Rattanakanokchai, S., Teng, D. N., Thinkhamrop, J., Watananirun, K., Zhang, J., Zhou, W. & Gulmezoglu, A. M. 2019. Global, regional, and national estimates of levels of preterm birth in 2014: a systematic review and modelling analysis. Lancet Glob Health, 7, 1, e37-e46. Available: https://www.ncbi.nlm.nih.gov/pubmed/30389451

Chen, G., Guo, Y., Abramson, M. J., Williams, G. & Li, S. 2018a. Exposure to low concentrations of air pollutants and adverse birth outcomes in Brisbane, Australia, 2003–2013. Science of The

Total Environment, 622-623, 721-726. Available: https://www.sciencedirect.com/science/article/pii/S0048969717334691
Chen, H., Li, Q., Kaufman, J. S., Wang, J., Copes, R., Su, Y. & Benmarhnia, T. 2018b. Effect of air quality alerts on human health: a regression discontinuity analysis in Toronto, Canada. The Lancet Planetary Health, 2, 1, e19-e26. Available: https://doi.org/10.1016/S2542-5196(17)30185-7
Chen, J., Fang, J., Zhang, Y., Xu, Z., Byun, H. M., Li, P. H., Deng, F., Guo, X., Guo, L. & Wu, S. 2021a. Associations of adverse pregnancy outcomes with high ambient air pollution exposure: Results from the Project ELEFANT. Science of the Total Environment, 761. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85096148597&doi=10.1016%2fj.scitotenv.2020.143218&partnerID=40&md5=81e37a14eaaa4bff79691f53238ca98f
Chen, J. & Hoek, G. 2020. Long-term exposure to PM and all-cause and cause-specific mortality: A systematic review and meta-analysis. Environ Int, 143, 105974. Available: https://www.ncbi.nlm.nih.gov/pubmed/32703584
Chen, Q., Ren, Z., Liu, Y., Qiu, Y., Yang, H., Zhou, Y., Wang, X., Jiao, K., Liao, J. & Ma, L. 2021b. The Association between Preterm Birth and Ambient Air Pollution Exposure in Shiyan, China, 2015-2017. Int J Environ Res Public Health, 18, 8.
Cheng, Y., Feng, J., Wang, J., Zhou, Y., Bai, S., Tang, Q., Li, J., Pan, F., Xu, Q., Lu, C., Wu, W. & Xia, Y. 2024. Alterations in sperm DNA methylation may as a mediator of paternal air pollution exposure and offspring birth outcomes: Insight from a birth cohort study. Environmental Research, 244, 117941. Available: https://www.sciencedirect.com/science/article/pii/S0013935123027457
Cherrie, J. W., Apsley, A., Cowie, H., Steinle, S., Mueller, W., Lin, C., Horwell, C. J., Sleeuwenhoek, A. & Loh, M. 2018. Effectiveness of face masks used to protect Beijing residents against particulate air pollution. Occupational and Environmental Medicine, 75, 6, 446-452. Available: https://oem.bmj.com/content/oemed/75/6/446.full.pdf
Chersich, M. F., Pham, M. D., Areal, A., Haghighi, M. M., Manyuchi, A., Swift, C. P., Wernecke, B., Robinson, M., Hetem, R., Boeckmann, M. & Hajat, S. 2020. Associations between high temperatures in pregnancy and risk of preterm birth, low birth weight, and stillbirths: systematic review and meta-analysis. BMJ, 371, m3811.
Chief Medical Officer (CMO) 2022. Chief Medical Officer's Annual Report 2022: Air Pollution. (DHSC), Department of Health and Social Care. Available: https://assets.publishing.service.gov.uk/media/639aeb81e90e0721889bbf2f/chief-medical-officers-annual-report-air-pollution-dec-2022.pdf
Chu, C., Zhu, Y., Liu, C., Chen, R., Yan, Y., Ren, Y., Li, X., Wang, J., Ge, W., Kan, H. & Gui, Y. 2021. Ambient fine particulate matter air pollution and the risk of preterm birth: A multicenter birth cohort study in China. Environmental Pollution, 287. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85109549369&doi=10.1016%2fj.envpol.2021.117629&partnerID=40&md5=f30074755d760f96e4f231b448c77163
Cleary, E. G., Cifuentes, M., Grinstein, G., Brugge, D. & Shea, T. B. 2018. Association of Low-Level Ozone with Cognitive Decline in Older Adults. J Alzheimers Dis, 61, 1, 67-78.
Climate Change Committee (CCC) 2020. The Sixth Carbon Budget - The UK's Path to Net Zero. Available: https://www.theccc.org.uk/publication/sixth-carbon-budget/
Cochrane uidance for the production and publication of Cochrane living systematic reviews: Cochrane Reviews in living mode (Version December 2019). Available: https://community.cochrane.org/sites/default/files/uploads/inline-files/Transform/201912_LSR_Revised_Guidance.pdf
Committee on the Medical Effects of Air Pollutants (COMEAP) 2016. Long-term Exposureto Air Pollution and Chronic Bronchitis. Available: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/541745/COMEAP_chronic_bronchitis_report_2016__rev_07-16_.pdf
Committee on the Medical Effects of Air Pollutants (COMEAP) 2018a. Associations of long-term average concentrations of nitrogen dioxide with mortality. Available: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/734799/COMEAP_NO2_Report.pdf
Committee on the Medical Effects of Air Pollutants (COMEAP). 2018b. *Committee on the Medical Effects of Air Pollutants: Minutes [Online]*. Available: https://app.box.com/s/gv2xjsp6g6ffp1zhm72igzjlwtrt7uf7/file/552229534077

Committee on the Medical Effects of Air Pollutants (COMEAP) 2022. Cognitive decline, dementia and air pollution. Available: https://assets.publishing.service.gov.uk/media/62ceccdc8fa8f50c012d1406/COMEAP-dementia-report-2022.pdf
Committee on the Medical Effects of Air Pollutants (COMEAP) 2023. Summary of COMEAP recommendations for the quantification of health effects associated with air pollutants. Available: https://assets.publishing.service.gov.uk/media/64fadfdea78c5f0014265847/COMEAP_Quantification_recommendations.pdf
Committee on the Medical Effects of Air Pollutants (COMEAP) forthcoming. Effects of maternal exposure to ambient air pollution on adverse birth outcomes.
Cook, C. N., Nichols, S. J., Webb, J. A., Fuller, R. A. & Richards, R. M. 2017. Simplifying the selection of evidence synthesis methods to inform environmental decisions: A guide for decision makers and scientists. Biological Conservation, 213, 135-145. https://www.sciencedirect.com/science/article/pii/S000632071630982X
Cossi, M., Zuta, S., Padula, A. M., Gould, J. B., Stevenson, D. K. & Shaw, G. M. 2015. Role of infant sex in the association between air pollution and preterm birth. Annals of Epidemiology, 25, 11, 874-876. https://www.scopus.com/inward/record.uri?eid=2-s2.0-84944886072&doi=10.1016%2fj.annepidem.2015.08.005&partnerID=40&md5=0838b770d39b8de8592c957469b56f2b
Costello, J. M., Steurer, M. A., Baer, R. J., Witte, J. S. & Jelliffe-Pawlowski, L. L. 2022. Residential particulate matter, proximity to major roads, traffic density and traffic volume as risk factors for preterm birth in California. Paediatr Perinat Epidemiol, 36, 1, 70-79.
Covidence. 2024a. *About Us* [Online]. Available: https://www.covidence.org/about-us-covidence/
Covidence. 2024b. *Best practice guidelines for abstract screening – systematic reviews* [Online]. Available: https://www.covidence.org/blog/best-practice-guidelines-for-abstract-screening/
Cowell, W., Colicino, E., Zhang, X., Ledyard, R., Burris, H. H., Hacker, M. R., Kloog, I., Just, A., Wright, R. O. & Wright, R. J. 2021. Spatially and Temporally Resolved Ambient PM(2.5) in Relation to Preterm Birth. Toxics, 9, 12.
Craig, B. M., Hartman, J. D., Owens, M. A. & Brown, D. S. 2016. Prevalence and Losses in Quality-Adjusted Life Years of Child Health Conditions: A Burden of Disease Analysis. Maternal and Child Health Journal, 20, 4, 862-869. https://doi.org/10.1007/s10995-015-1874-z
Crump, C. 2020. Preterm birth and mortality in adulthood: a systematic review. J Perinatol, 40, 6, 833-843. https://www.ncbi.nlm.nih.gov/pubmed/31767981
D'antoni, D., Auyeung, V., Walton, H., Fuller, G. W., Grieve, A. & Weinman, J. 2019. The effect of evidence and theory-based health advice accompanying smartphone air quality alerts on adherence to preventative recommendations during poor air quality days: A randomised controlled trial. Environment International, 124, 216-235. https://www.sciencedirect.com/science/article/pii/S0160412018321871
Dajnak, D., Kitwiroon, N., Assareh, N., Stewart, G., Hicks, W., Evangelopoulos, D., Wood, D., Walton, H. & Beevers, S. 2022. Pathway to WHO: Achieving clean air in the UK - Modelling air quality costs and benefits. London. Available: https://www.imperial.ac.uk/media/imperial-college/medicine/sph/environmental-research-group/CAF-Technical-Report-Final.pdf
Dėdelė, A., Gražulevičienė, R. & Bendokienė, I. 2011. Individual Exposure to Nitrogen Dioxide and Preterm Birth Risk in Kaunas. Individuali ekspozicija azoto dioksidu ir priešlaikinio gimdymo rizika Kaune., 56, 2, 49-56. https://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,shib,url,uid&db=8gh&AN=63893942&site=ehost-live
Dedele, A., Grazuleviciene, R. & Miskinyte, A. 2017. Individual exposure to nitrogen dioxide and adverse pregnancy outcomes in Kaunas study. International Journal of Environmental Health Research, 27, 3, 230-240. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85019654143&doi=10.1080%2f09603123.2017.1332348&partnerID=40&md5=36f3d9e183ca9568b267675fd388c9c6
Deeks, J., Higgins, J. & Altman, D. 2022. Chapter 10: Analysing data and undertaking meta-analyses. *In:* Higgins, J., Thomas, J., Crumpston, M., Li, T., Page, M. & Welch, V. (eds.) *Cochrane Handbook for Systematic Reviews of Interventions version 6.4.* Cochrane. Available: https://training.cochrane.org/handbook/current/chapter-10
Deeks, J. J., Dinnes, J., D'amico, R., Sowden, A. J., Sakarovitch, C., Song, F., Petticrew, M. & Altman, D. G. 2003. Evaluating non-randomised intervention studies. Health Technol Assess, 7, 27, iii-x, 1-173.

DeFranco, E., Moravec, W., Xu, F., Hall, E., Hossain, M., Haynes, E. N., Muglia, L. & Chen, A. 2016. Exposure to airborne particulate matter during pregnancy is associated with preterm birth: a population-based cohort study. Environ Health, 15, 6.
Deguen, S., Marchetta, G. P. & Kihal-Talantikite, W. 2020. Measuring burden of disease attributable to air pollution due to preterm birth complications and infant death in paris using disability-adjusted life years (DALYs). International Journal of Environmental Research and Public Health, 17, 21, 1-14. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85094193611&doi=10.3390%2fijerph17217841&partnerID=40&md5=8d72fbe4bd3f80aac2f27e107a6c619e
Demendi, C., Börzsönyi, B., Pajor, A., Rigó, J., Nagy, Z. B., Szentpéteri, I. & Joó, J. G. 2012. Abnormal fetomaternal glucocorticoid metabolism in the background of premature delivery: placental expression patterns of the 11β-hydroxysteroid dehydrogenase 2 gene. European Journal of Obstetrics & Gynecology and Reproductive Biology, 165, 2, 210-214. Available: https://www.sciencedirect.com/science/article/pii/S0301211512003673
Department for Business Energy & Industrial Strategy (Beis) 2022. United Kingdom of Great Britain and Northern Ireland's Nationally Determined Contribution. Available: https://assets.publishing.service.gov.uk/media/633d937d8fa8f52a5803e63f/uk-nationally-determined-contribution.pdf
Department for Business Energy & Industry Strategy (Beis). 2019. *Updated Energy and Emission Projections 2018* [Online]. Available: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/794590/updated-energy-and-emissions-projections-2018.pdf
Department for Environment Food & Rural Affairs (Defra). 2012. *Conversion Factors Between ppb and µg m-3 and ppm and mgm-3* [Online]. Available: https://uk-air.defra.gov.uk/assets/documents/reports/cat06/0502160851_Conversion_Factors_Between_ppb_and.pdf
Department of Health and Social Care (Dhsc). 2023. *Pregnancy Loss Review* [Online]. Available: https://www.gov.uk/government/publications/pregnancy-loss-review
Dersimonian, R. & Laird, N. 1986. Meta-analysis in clinical trials. Control Clin Trials, 7, 3, 177-88.
Dettori, J. R., Norvell, D. C. & Chapman, J. R. 2021. Risks, Rates and Odds: What's the Difference and Why Does It Matter? Global Spine J, 11, 7, 1156-1158.
Dibben, C. & Clemens, T. 2015. Place of work and residential exposure to ambient air pollution and birth outcomes in Scotland, using geographically fine pollution climate mapping estimates. Environmental Research, 140, 535-541. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84930678000&doi=10.1016%2fj.envres.2015.05.010&partnerID=40&md5=22f506aaedc85740753bab6833711d1d
doNascimento, F. P., De Almeida, M. F. & Gouveia, N. 2022. Individual and contextual socioeconomic status as effect modifier in the air pollution-birth outcome association. Science of the Total Environment, 803. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85113981292&doi=10.1016%2fj.scitotenv.2021.149790&partnerID=40&md5=2c577cb23ed397bf6b0e4f9226a41710
Doi, S. A. & Thalib, L. 2008. A quality-effects model for meta-analysis. Epidemiology, 19, 1, 94-100.
Doiron, D., De Hoogh, K., Probst-Hensch, N., Fortier, I., Cai, Y., De Matteis, S. & Hansell, A. L. 2019. Air pollution, lung function and COPD: results from the population-based UK Biobank study. Eur Respir J, 54, 1.
Downs, S. H. & Black, N. 1998. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. J Epidemiol Community Health, 52, 6, 377-84.
Dzekem, B. S., Aschebrook-Kilfoy, B. & Olopade, C. O. 2024. Air Pollution and Racial Disparities in Pregnancy Outcomes in the United States: A Systematic Review. J Racial Ethn Health Disparities, 11, 1, 535-544.
Dzhambov, A. M., Markevych, I. & Lercher, P. 2019. Associations of residential greenness, traffic noise, and air pollution with birth outcomes across Alpine areas. Science of the Total Environment, 678, 399-408. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85065205135&doi=10.1016%2fj.scitotenv.2019.05.019&partnerID=40&md5=d19b742b5651975febdc075e2bea7bd6
EBSCO. 2024a. *CINAHL Database* [Online]. Available: https://www.ebsco.com/products/research-databases/cinahl-database
EBSCO. 2024b. GREENFile. https://www.ebsco.com/products/research-databases/greenfile

Eeftens, M., Tsai, M.-Y., Ampe, C., Anwander, B., Beelen, R., Bellander, T., Cesaroni, G., Cirach, M., Cyrys, J., De Hoogh, K., De Nazelle, A., De Vocht, F., Declercq, C., Dėdelė, A., Eriksen, K., Galassi, C., Gražulevičienė, R., Grivas, G., Heinrich, J., Hoffmann, B., Iakovides, M., Ineichen, A., Katsouyanni, K., Korek, M., Krämer, U., Kuhlbusch, T., Lanki, T., Madsen, C., Meliefste, K., Mölter, A., Mosler, G., Nieuwenhuijsen, M., Oldenwening, M., Pennanen, A., Probst-Hensch, N., Quass, U., Raaschou-Nielsen, O., Ranzi, A., Stephanou, E., Sugiri, D., Udvardy, O., Vaskövi, É., Weinmayr, G., Brunekreef, B. & Hoek, G. 2012. Spatial variation of PM2.5, PM10, PM2.5 absorbance and PMcoarse concentrations between and within 20 European study areas and the relationship with NO2 – Results of the ESCAPE project. Atmospheric Environment, 62, 303-317. Available: https://www.sciencedirect.com/science/article/pii/S1352231012008102

Egger, M., Davey Smith, G., Schneider, M. & Minder, C. 1997. Bias in meta-analysis detected by a simple, graphical test. Bmj, 315, 7109, 629-34.

El-Sikaily, A., Helal, M., Nsonwu-Anyanwu, A. C., Azab, H., Abd Elmoneim, N., Farahat, E. O. S. & Saad, A. 2023. Impacts of PAH accumulation on reproductive hormones, indices of oxidative stress and BPDE-albumin adduct in women with recurrent pregnancy loss. Toxicological Research, 39, 3, 517-531. Available: https://doi.org/10.1007/s43188-023-00181-5

Elseiver. 2024. *Embase content coverage* [Online]. Available: https://www.elsevier.com/solutions/embase-biomedical-research/coverage-and-content?dgcid=RN_CM_Sourced_400001094

Escoto, K. M., Mullin, A. M., Ledyard, R., Rovit, E., Yang, N., Tripathy, S., Burris, H. H. & Clougherty, J. E. 2022. Benzene and NO2 Exposure during Pregnancy and Preterm Birth in Two Philadelphia Hospitals, 2013–2017. *International Journal of Environmental Research and Public Health* [Online], 19.

Estarlich, M., Ballester, F., Davdand, P., Llop, S., Esplugues, A., Fernández-Somoano, A., Lertxundi, A., Guxens, M., Basterrechea, M., Tardón, A., Sunyer, J. & Iñiguez, C. 2016. Exposure to ambient air pollution during pregnancy and preterm birth: A Spanish multicenter birth cohort study. Environmental Research, 147, 50-58. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84956668084&doi=10.1016%2fj.envres.2016.01.037&partnerID=40&md5=4d62314eeaf1d9fd4dc1d5fe09feb697

Etzel, R. A. 2020. Is the Environment Associated With Preterm Birth? JAMA Network Open, 3, 4, e202239-e202239. Available: https://doi.org/10.1001/jamanetworkopen.2020.2239

European Monitoring and Evaluation Programme (EMEP) Centre on Emission Inventories and Projections. 2024. *EMEP Centre on Emission Inventories and Projections* [Online]. Available: https://www.ceip.at/

European Union (EU) 2008. Directive 2008/50/EC of the European Parliament and of the Council of 21st May 2008 on ambient air quality and cleaner air for Europe. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:02008L0050-20150918&from=DE

Fabiano, N., Gupta, A., Bhambra, N., Luu, B., Wong, S., Maaz, M., Fiedorowicz, J. G., Smith, A. L. & Solmi, M. 2024. How to optimize the systematic review process using AI tools. JCPP Advances, 4, 2, e12234. https://acamh.onlinelibrary.wiley.com/doi/abs/10.1002/jcv2.12234

Fang, J., Kang, C.-M., Osorio-Yáñez, C., Barrow, T. M., Zhang, R., Zhang, Y., Li, C., Liu, H., Li, P.-H., Guo, L. & Byun, H.-M. 2020. Prenatal PM2.5 exposure and the risk of adverse births outcomes: Results from Project ELEFANT. Environmental Research, 191, 110232. https://www.sciencedirect.com/science/article/pii/S0013935120311294

Fang, J., Yang, Y., Zou, X., Xu, H., Wang, S., Wu, R., Jia, J., Xie, Y., Yang, H., Yuan, N., Hu, M., Deng, Y., Zhao, Y., Wang, T., Zhu, Y., Ma, X., Fan, M., Wu, J., Song, X. & Huang, W. 2022. Maternal exposures to fine and ultrafine particles and the risk of preterm birth from a retrospective study in Beijing, China. Science of The Total Environment, 812, 151488. https://www.sciencedirect.com/science/article/pii/S0048969721065669

Fedak, K. M., Bernal, A., Capshaw, Z. A. & Gross, S. 2015. Applying the Bradford Hill criteria in the 21st century: how data integration has changed causal inference in molecular epidemiology. Emerg Themes Epidemiol, 12, 14.

Fenech, S., Doherty, R. M., O'connor, F. M., Heaviside, C., Macintyre, H. L., Vardoulakis, S., Agnew, P. & Neal, L. S. 2021. Future air pollution related health burdens associated with RCP emission changes in the UK. Sci Total Environ, 773, 145635.

Fisher, S. 2022. *Create online surveys with our free online survey maker tool* [Online]. Available: https://www.qualtrics.com/blog/create-online-survey/

Fleischer, N. L., Merialdi, M., Van Donkelaar, A., Vadillo-Ortega, F., Martin Randall, V., Betran Ana, P. & Souza João, P. 2014. 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, 122, 4, 425-430. https://doi.org/10.1289/ehp.1306837
Forastiere, F., Spadaro, J. V., Ancona, C., Jovanovic Andersen, Z., Cozzi, I., Gumy, S., Loncar, D., Mudu, P., Medina, S., Perez Velasco, R., Walton, H., Zhang, J. & Krzyzanowski, M. 2024. Choices of morbidity outcomes and concentration–response functions for health risk assessment of long-term exposure to air pollution. Environmental Epidemiology, 8, 4. https://journals.lww.com/environepidem/fulltext/2024/08000/choices_of_morbidity_outcomes_and.4.aspx
Freije, S. L., Enquobahrie, D. A., Day, D. B., Loftus, C., Szpiro, A. A., Karr, C. J., Trasande, L., Kahn, L. G., Barrett, E., Kannan, K., Bush, N. R., Lewinn, K. Z., Swan, S., Alex Mason, W., Robinson, M. & Sathyanarayana, S. 2022. Prenatal exposure to polycyclic aromatic hydrocarbons and gestational age at birth. Environment International, 164, 107246. https://www.sciencedirect.com/science/article/pii/S0160412022001726
Fuchs, F., Monet, B., Ducruet, T., Chaillet, N. & Audibert, F. 2018. Effect of maternal age on the risk of preterm birth: A large cohort study. PLoS One, 13, 1, e0191002.
Fung, I. C. 2008. Chinese journals: a guide for epidemiologists. Emerg Themes Epidemiol, 5, 20.
Fussell, J. C., Jauniaux, E., Smith, R. B. & Burton, G. J. 2024. Ambient air pollution and adverse birth outcomes: A review of underlying mechanisms. Bjog, 131, 5, 538-550.
Gan, Q., Ye, W., Zhao, X., Teng, Y., Mei, S., Long, Y., Ma, J., Rehemutula, R., Zhang, X., Zeng, F., Jin, H., Liu, F., Huang, Y., Gao, X. & Zhu, C. 2022. Mediating effects of gut microbiota in the associations of air pollutants exposure with adverse pregnancy outcomes. Ecotoxicology and Environmental Safety, 234, 113371. https://www.sciencedirect.com/science/article/pii/S0147651322002111
Garshol, B. F., Aamodt, G., Madsen, C., Vatn, M. H. & Bengtson, M.-B. 2020. The effect of nitrogen dioxide on low birth weight in women with inflammatory bowel disease: a Norwegian pregnancy cohort study (MoBa). Scandinavian journal of gastroenterology, 55, 3, 272-278. http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=med17&NEWS=N&AN=32064969
Gartlehner, G., Kahwati, L., Hilscher, R., Thomas, I., Kugley, S., Crotty, K., Viswanathan, M., Nussbaumer-Streit, B., Booth, G., Erskine, N., Konet, A. & Chew, R. 2024. Data extraction for evidence synthesis using a large language model: A proof-of-concept study. Res Synth Methods, 15, 4, 576-589.
Gehring, U., Tamburic, L., Sbihi, H., Davies, H. W. & Brauer, M. 2014. Impact of noise and air pollution on pregnancy outcomes. Epidemiology, 25, 3, 351-8.
Gehring, U., Wijga, A. H., Fischer, P., De Jongste, J. C., Kerkhof, M., Koppelman, G. H., Smit, H. A. & Brunekreef, B. 2011a. Traffic-related air pollution, preterm birth and term birth weight in the PIAMA birth cohort study. Environmental Research, 111, 1, 125-135. https://www.sciencedirect.com/science/article/pii/S001393511000174X
Gehring, U., Van Eijsden, M., Dijkema, M. B. A., Van Der Wal, M. F., Fischer, P. & Brunekreef, B. 2011b. Traffic-related air pollution and pregnancy outcomes in the Dutch ABCD birth cohort study. Occupational and Environmental Medicine, 68, 1, 36-43. https://www.scopus.com/inward/record.uri?eid=2-s2.0-78650963766&doi=10.1136%2foem.2009.053132&partnerID=40&md5=9b4cd2d1f2c5db25fb8d5fd6a5e3ec5d
Geissbühler, M., Hincapié, C. A., Aghlmandi, S., Zwahlen, M., Jüni, P. & Da Costa, B. R. 2021. Most published meta-regression analyses based on aggregate data suffer from methodological pitfalls: a meta-epidemiological study. BMC Medical Research Methodology, 21, 1, 123. https://doi.org/10.1186/s12874-021-01310-0
Genin, M., Lecoeuvre, A., Cuny, D., Subtil, D., Chevalier, G., Ficheur, G., Occelli, F. & Garabedian, C. 2022. The association between the incidence of preterm birth and overall air pollution: A nationwide, fine-scale, spatial study in France from 2012 to 2018. Environmental Pollution, 311, 120013. https://www.sciencedirect.com/science/article/pii/S0269749122012271
Ghosh, R., Causey, K., Burkart, K., Wozniak, S., Cohen, A. & Brauer, M. 2021. Ambient and household PM2.5 pollution and adverse perinatal outcomes: A meta-regression and analysis of attributable global burden for 204 countries and territories. PLoS Med, 18, 9, e1003718. https://www.ncbi.nlm.nih.gov/pubmed/34582444
Ghosh, R. E., Close, R., Mccann, L. J., Crabbe, H., Garwood, K., Hansell, A. L. & Leonardi, G. 2016. Analysis of hospital admissions due to accidental non-fire-related carbon monoxide poisoning

in England, between 2001 and 2010. Journal of Public Health, 38, 1, 76-83. https://doi.org/10.1093/pubmed/fdv026
Gierisch, J., Beadles, C., Shapiro, A. & Al., E. 2014. APPENDIX BNEWCASTLE-OTTAWA SCALE CODING MANUAL FOR COHORT STUDIES. *Health Disparities in Quality Indicators of Healthcare Among Adults with Mental Illness.* Washington (DC). Available: https://www.ncbi.nlm.nih.gov/books/NBK299087/
Giorgis-Allemand, L., Pedersen, M., Bernard, C., Aguilera, I., Beelen, R. M., Chatzi, L., Cirach, M., Danileviciute, A., Dedele, A., Van Eijsden, M., Estarlich, M., Fernández-Somoano, A., Fernández, M. F., Forastiere, F., Gehring, U., Grazuleviciene, R., Gruzieva, O., Heude, B., Hoek, G., De Hoogh, K., Van Den Hooven, E. H., Håberg, S. E., Iñiguez, C., Jaddoe, V. W., Korek, M., Lertxundi, A., Lepeule, J., Nafstad, P., Nystad, W., Patelarou, E., Porta, D., Postma, D., Raaschou-Nielsen, O., Rudnai, P., Siroux, V., Sunyer, J., Stephanou, E., Sørensen, M., Eriksen, K. T., Tuffnell, D., Varró, M. J., Vrijkotte, T. G., Wijga, A., Wright, J., Nieuwenhuijsen, M. J., Pershagen, G., Brunekreef, B., Kogevinas, M. & Slama, R. 2017. The Influence of Meteorological Factors and Atmospheric Pollutants on the Risk of Preterm Birth. Am J Epidemiol, 185, 4, 247-258.
Glinianaia, S. V., Rankin, J., Bell, R., Pless-Mulloli, T. & Howel, D. 2004. Particulate Air Pollution and Fetal Health: A Systematic Review of the Epidemiologic Evidence. Epidemiology, 15, 1. https://journals.lww.com/epidem/fulltext/2004/01000/particulate_air_pollution_and_fetal_health__a.8.aspx
Goin, D. E., Casey, J. A., Kioumourtzoglou, M. A., Cushing, L. J. & Morello-Frosch, R. 2021. Environmental hazards, social inequality, and fetal loss: Implications of live-birth bias for estimation of disparities in birth outcomes. Environ Epidemiol, 5, 2, e131.
Grant, R. L. 2014. Converting an odds ratio to a range of plausible relative risks for better communication of research findings. Bmj, 348, f7450.
Gray, S. C., Edwards, S. E., Schultz, B. D. & Miranda, M. L. 2014. Assessing the impact of race, social factors and air pollution on birth outcomes: a population-based study. Environ Health, 13, 1, 4.
Greater London Authority (GLA). 2019. *London Atmospheric Emissions Inventory (LAEI) 2019* [Online]. Available: https://data.london.gov.uk/dataset/london-atmospheric-emissions-inventory--laei--2019.
Grosso, A., Neves De Faria, R. I., Bojke, L., Donohue, C., Fraser, C. I., Harron, K. L., Oddie, S. J. & Gilbert, R. 2020. Cost-effectiveness of strategies preventing late-onset infection in preterm infants. Archives of Disease in Childhood, 105, 5, 452-457. https://adc.bmj.com/content/archdischild/105/5/452.full.pdf
Guo, L., Lin, H., Li, H., Jin, X., Zhao, L., Li, P., Xu, N., Xu, S., Fang, J., Wu, S. & Liu, Q. 2023. Exposure of ambient PM2.5 during gametogenesis period affects the birth outcome: Results from the project ELEFANT. Environmental Research, 220, 115204. https://www.sciencedirect.com/science/article/pii/S0013935122025312
Guo, L. Q., Chen, Y., Mi, B. B., Dang, S. N., Zhao, D. D., Liu, R., Wang, H. L. & Yan, H. 2019. Ambient air pollution and adverse birth outcomes: a systematic review and meta-analysis. J Zhejiang Univ Sci B, 20, 3, 238-252.
Guo, T., Wang, Y., Zhang, H., Zhang, Y., Zhao, J., Wang, Q., Shen, H., Wang, Y., Xie, X., Wang, L., Xu, Z., Zhang, Y., Yan, D., He, Y., Yang, Y., Xu, J., Peng, Z. & Ma, X. 2018. The association between ambient PM2.5 exposure and the risk of preterm birth in China: A retrospective cohort study. Science of The Total Environment, 633, 1453-1459. https://www.sciencedirect.com/science/article/pii/S004896971831088X
Ha, S., Hu, H., Roussos-Ross, D., Haidong, K., Roth, J. & Xu, X. 2014. The effects of air pollution on adverse birth outcomes. Environmental Research, 134, 198-204. https://www.sciencedirect.com/science/article/pii/S001393511400259X
Hamilton, I., Kennard, H., Mcgushin, A., Höglund-Isaksson, L., Kiesewetter, G., Lott, M., Milner, J., Purohit, P., Rafaj, P., Sharma, R., Springmann, M., Woodcock, J. & Watts, N. 2021. The public health implications of the Paris Agreement: a modelling study. Lancet Planet Health, 5, 2, e74-e83.
Hammer, M. S., Swinburn, T. K. & Neitzel, R. L. 2014. Environmental noise pollution in the United States: developing an effective public health response. Environ Health Perspect, 122, 2, 115-9.
Hammitt, J. K. 2002. QALYs Versus WTP. Risk Analysis, 22, 5, 985-1001. https://onlinelibrary.wiley.com/doi/abs/10.1111/1539-6924.00265

Han, Y., Jiang, P., Dong, T., Ding, X., Chen, T., Villanger, G. D., Aase, H., Huang, L. & Xia, Y. 2018. Maternal air pollution exposure and preterm birth in Wuxi, China: Effect modification by maternal age. Ecotoxicol Environ Saf, 157, 457-462.
Hannam, K., Mcnamee, R., Baker, P., Sibley, C. & Agius, R. 2014. Air pollution exposure and adverse pregnancy outcomes in a large UK birth cohort: use of a novel spatio-temporal modelling technique. Scand J Work Environ Health, 40, 5, 518-30.
Hao, H., Chang, H. H., Holmes, H. A., Mulholland, J. A., Klein, M., Darrow, L. A. & Strickland, M. J. 2016. 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, 124, 6, 875-880. https://ehp.niehs.nih.gov/doi/abs/10.1289/ehp.1409651
Hartling, L., Milne, A., Hamm, M. P., Vandermeer, B., Ansari, M., Tsertsvadze, A. & Dryden, D. M. 2013. Testing the Newcastle Ottawa Scale showed low reliability between individual reviewers. J Clin Epidemiol, 66, 9, 982-93.
Hassan Bhat, T., Jiawen, G. & Farzaneh, H. 2021. Air Pollution Health Risk Assessment (AP-HRA), Principles and Applications. Int J Environ Res Public Health, 18, 4.
He, Y., Jiang, Y., Yang, Y., Xu, J., Wang, Q., Shen, H., Zhang, Y., Yan, D., Peng, Z., Liu, C., Wang, W., Schikowski, T., Li, H., Yan, B., Ji, J. S., Chen, A., Van Donkelaar, A., Martin, R., Chen, R., Kan, H., Cai, J. & Ma, X. 2022. Composition of fine particulate matter and risk of preterm birth: A nationwide birth cohort study in 336 Chinese cities. Journal of Hazardous Materials, 425, 127645. https://www.elsevier.com/locate/jhazmat
Health Assessment Workspace Collaborative (HAWC). 2024. *Health Assessment Workspace Collaborative (HAWC)* [Online]. USEPA. Available: https://www.epa.gov/risk/health-assessment-workspace-collaborative-hawc .
Health Effects Institute (HEI). 2022. Systematic Review and Meta-analysis of Selected Health Effects of Long-Term Exposure to Traffic-Related Air Pollution. Boston MA. Available: https://www.healtheffects.org/system/files/hei-special-report-23_6.pdf
Health Protection Agency (Currently Known as UKHSA) 2011. Report of a Workshop to Identify Needs for Research on the Health Effects of Nitrogen Dioxide - London, 2-3 March 2011. Available: https://assets.publishing.service.gov.uk/media/5a75b381ed915d6faf2b5077/HPA-CRCE-026_for_website.pdf
Heft-Neal, S., Driscoll, A., Yang, W., Shaw, G. & Burke, M. 2022. Associations between wildfire smoke exposure during pregnancy and risk of preterm birth in California. Environmental Research, 203. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85113478562&doi=10.1016%2fj.envres.2021.111872&partnerID=40&md5=e1285808d0fe0b01d4c3f236464df0a9
Helfer, B., Prosser, A., Samara, M. T., Geddes, J. R., Cipriani, A., Davis, J. M., Mavridis, D., Salanti, G. & Leucht, S. 2015. Recent meta-analyses neglect previous systematic reviews and meta-analyses about the same topic: a systematic examination. BMC Medicine, 13, 1, 82. https://doi.org/10.1186/s12916-015-0317-4
Héroux, M. E., Anderson, H. R., Atkinson, R., Brunekreef, B., Cohen, A., Forastiere, F., Hurley, F., Katsouyanni, K., Krewski, D., Krzyzanowski, M., Künzli, N., Mills, I., Querol, X., Ostro, B. & Walton, H. 2015. Quantifying the health impacts of ambient air pollutants: recommendations of a WHO/Europe project. Int J Public Health, 60, 5, 619-27.
Higgins, J., Savović, J., Page, M., Elbers, R. & Sterne, J. 2024. Chapter 8: Assessing risk of bias in a randomized trial [last updated October 2019]. Available: https://training.cochrane.org/handbook/current/chapter-08
Higgins, J., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. & Welch, V. 2022. Cochrane Handbook for Systematic Reviews of Interventions version 6.3 (updated February 2022). Cochrane. https://training.cochrane.org/handbook/current/chapter-10
Higgins, J. P., Thompson, S. G. & Spiegelhalter, D. J. 2009. A re-evaluation of random-effects meta-analysis. J R Stat Soc Ser A Stat Soc, 172, 1, 137-159.
Hill, A. B. 1965. THE ENVIRONMENT AND DISEASE: ASSOCIATION OR CAUSATION? Proc R Soc Med, 58, 5, 295-300.
HM Government 2021. Net Zero Strategy: Build Back Greener. Available: https://assets.publishing.service.gov.uk/media/6194dfa4d3bf7f0555071b1b/net-zero-strategy-beis.pdf
HM Treasury 2022. The Green Book: appraisal and evaluation in central government.

Hodek, J. M., Von Der Schulenburg, J. M. & Mittendorf, T. 2011. Measuring economic consequences of preterm birth - Methodological recommendations for the evaluation of personal burden on children and their caregivers. Health Econ Rev, 1, 1, 6.
Hodgson, S., Lurz, P. W., Shirley, M. D., Bythell, M. & Rankin, J. 2015. Exposure misclassification due to residential mobility during pregnancy. Int J Hyg Environ Health, 218, 4, 414-21.
Hodgson, S., Shirley, M., Bythell, M. & Rankin, J. 2009. Residential mobility during pregnancy in the north of England. BMC Pregnancy and Childbirth, 9, 1, 52. https://doi.org/10.1186/1471-2393-9-52
Hoffmann, B., Boogaard, H., De Nazelle, A., Andersen, Z. J., Abramson, M., Brauer, M., Brunekreef, B., Forastiere, F., Huang, W., Kan, H., Kaufman, J. D., Katsouyanni, K., Krzyzanowski, M., Kuenzli, N., Laden, F., Nieuwenhuijsen, M., Mustapha, A., Powell, P., Rice, M., Roca-Barceló, A., Roscoe, C. J., Soares, A., Straif, K. & Thurston, G. 2021. WHO Air Quality Guidelines 2021-Aiming for Healthier Air for all: A Joint Statement by Medical, Public Health, Scientific Societies and Patient Representative Organisations. Int J Public Health, 66, 1604465.
Howards, P. P., Hertz-Picciotto, I. & Poole, C. 2007. Conditions for bias from differential left truncation. Am J Epidemiol, 165, 4, 444-52.
Howe, C. G., Eckel, S. P., Habre, R., Girguis, M. S., Gao, L., Lurmann, F. W., Gilliland, F. D. & Breton, C. V. 2018. Association of Prenatal Exposure to Ambient and Traffic-Related Air Pollution With Newborn Thyroid Function: Findings From the Children's Health Study. JAMA Netw Open, 1, 5, e182172.
Hua, X. Y., S, P., V., C., Carson, C., J.J., K., Quigley, M. A., E., B., Johnson, S., Macfarlane, A. & Rivero-Arias, O. 2023. Gestational age and hospital admission costs from birth to childhood: a population-based record linkage study in England. Archives of Disease in Childhood - Fetal and Neonatal Edition, 108, 5, 485-491. Available: https://fn.bmj.com/content/fetalneonatal/108/5/485.full.pdf
Huang, C., Nichols, C., Liu, Y., Zhang, Y., Liu, X., Gao, S., Li, Z. & Ren, A. 2015. Ambient air pollution and adverse birth outcomes: a natural experiment study. Population Health Metrics, 13, 1, 17. Available: https://doi.org/10.1186/s12963-015-0050-4
Huang, H., Woodruff, T. J., Baer, R. J., Bangia, K., August, L. M., Jellife-Palowski, L. L., Padula, A. M. & Sirota, M. 2018. Investigation of association between environmental and socioeconomic factors and preterm birth in California. Environment International, 121, 1066-1078. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85059338918&doi=10.1016%2fj.envint.2018.07.027&partnerID=40&md5=3cc8ae146b2479c0e23fca713f4d99ba
Huang, J. V., Leung, G. M. & Schooling, C. M. 2017. The association of air pollution with birthweight and gestational age: Evidence from Hong Kong's 'Children of 1997' birth cohort. Journal of Public Health (United Kingdom), 39, 3, 476-484. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85032273103&doi=10.1093%2fpubmed%2ffdw068&partnerID=40&md5=8d426770efdb399ac1a2032a8936c7de
Huang, W. & Morwska, L. 2019. Face masks could raise pollution risks Nature Comments, 574, 29-30. https://www-nature-com.iclibezp1.cc.ic.ac.uk/articles/d41586-019-02938-1#ref-CR6
Hung, T. H., Chen, P. H., Tung, T. H., Hsu, J., Hsu, T. Y. & Wan, G. H. 2023. Risks of preterm birth and low birth weight and maternal exposure to NO(2)/PM(2.5) acquired by dichotomous evaluation: a systematic review and meta-analysis. Environ Sci Pollut Res Int, 30, 4, 9331-9349.
Hunter, R., Beardmore-Gray, A., Greenland, M., Linsell, L., Juszczak, E., Hardy, P., Placzek, A., Shennan, A., Marlow, N. & Chappell, L. C. 2022. Cost-Utility Analysis of Planned Early Delivery or Expectant Management for Late Preterm Pre-eclampsia (PHOENIX). Pharmacoecon Open, 6, 5, 723-733.
Hurley, J. 2018. Working Paper 2: A viewpoint on using adjusted coefficients for NO 2 and PM2.5. Working Paper for COMEAP Report “Associations of long-term average concentrations of nitrogen dioxide with mortality” published July 2018. Available: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/734805/COMEAP_NO2_Working_Paper_2.pdf
Huynh, M., Woodruff, T. J., Parker, J. D. & Schoendorf, K. C. 2006. Relationships between air pollution and preterm birth in California. Paediatric and Perinatal Epidemiology, 20, 6, 454-461. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-3016.2006.00759.x
Hyder, A., Lee, H. J., Ebisu, K., Koutrakis, P., Belanger, K. & Bell, M. L. 2014. PM2.5 exposure and birth outcomes: use of satellite- and monitor-based data. Epidemiology, 25, 1, 58-67.

Hystad, P., Davies, H. W., Frank, L., Loon, J. V., Gehring, U., Tamburic, L. & Brauer, M. 2014. Residential Greenness and Birth Outcomes: Evaluating the Influence of Spatially Correlated Built-Environment Factors. Environmental Health Perspectives, 122, 10, 1095-1102.
Institute of Health Metrics and Evaluation (Ihme). 2019. *GBD Results Tool* [Online]. Available: https://www.healthdata.org/data-tools-practices/interactive-visuals/gbd-results
Institute of Medicine Committee on Understanding Premature Birth Assuring Healthy Outcomes 2007. The National Academies Collection: Reports funded by National Institutes of Health. *In:* Behrman, R. E. & Butler, A. S. (eds.) *Preterm Birth: Causes, Consequences, and Prevention.* Washington (DC): National Academies Press (US).
IntHout, J., Ioannidis, J. P. & Borm, G. F. 2014. The Hartung-Knapp-Sidik-Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard DerSimonian-Laird method. BMC Med Res Methodol, 14, 25.
IntHout, J., Ioannidis, J. P., Rovers, M. M. & Goeman, J. J. 2016. Plea for routinely presenting prediction intervals in meta-analysis. BMJ Open, 6, 7, e010247.
Ioannidis, J. P. 2016. The Mass Production of Redundant, Misleading, and Conflicted Systematic Reviews and Meta-analyses. Milbank Q, 94, 3, 485-514.
Ioannidis, J. P. A. 2018. Meta-analyses in environmental and occupational health. Occup Environ Med, 75, 6, 443-445.
Iwashyna, T. J. & Kennedy, E. H. 2013. Instrumental variable analyses. Exploiting natural randomness to understand causal mechanisms. Ann Am Thorac Soc, 10, 3, 255-60.
Jack, D. W., Asante, K. P., Wylie, B. J., Chillrud, S. N., Whyatt, R. M., Ae-Ngibise, K. A., Quinn, A. K., Yawson, A. K., Boamah, E. A., Agyei, O., Mujtaba, M., Kaali, S., Kinney, P. & Owusu-Agyei, S. 2015. Ghana randomized air pollution and health study (GRAPHS): study protocol for a randomized controlled trial. Trials, 16, 1, 420. https://doi.org/10.1186/s13063-015-0930-8
Jackson, C. H., Best, N. G. & Richardson, S. 2009. Bayesian graphical models for regression on multiple data sets with different variables. Biostatistics, 10, 2, 335-51.
Jacobs, M., Zhang, G., Chen, S., Mullins, B., Bell, M., Jin, L., Guo, Y., Huxley, R. & Pereira, G. 2017. The association between ambient air pollution and selected adverse pregnancy outcomes in China: A systematic review. Science of The Total Environment, 579, 1179-1192. https://www.sciencedirect.com/science/article/pii/S0048969716325396
Jahanshahi, B., Johnston, B., Mcgovern, M. E., Mcvicar, D., O'reilly, D., Rowland, N. & Vlachos, S. 2024. Prenatal exposure to particulate matter and infant birth outcomes: Evidence from a population-wide database. Health Economics, 33, 9, 2182-2200. https://onlinelibrary.wiley.com/doi/abs/10.1002/hec.4862
Jalaludin, B., Mannes, T., Morgan, G., Lincoln, D., Sheppeard, V. & Corbett, S. 2007. Impact of ambient air pollution on gestational age is modified by season in Sydney, Australia. Environmental Health, 6, 1, 16. https://doi.org/10.1186/1476-069X-6-16
Janssen, B. G., Saenen, N. D., Roels, H. A., Madhloum, N., Gyselaers, W., Lefebvre, W., Penders, J., Vanpoucke, C., Vrijens, K. & Nawrot, T. S. 2017. Fetal Thyroid Function, Birth Weight, and in Utero Exposure to Fine Particle Air Pollution: A Birth Cohort Study. Environ Health Perspect, 125, 4, 699-705.
Jedrychowski, W. A., Perera, F. P., Maugeri, U., Mroz, E., Klimaszewska-Rembiasz, M., Flak, E., Edwards, S. & Spengler, J. D. 2010. Effect of prenatal exposure to fine particulate matter on ventilatory lung function of preschool children of non-smoking mothers. Paediatr Perinat Epidemiol, 24, 5, 492-501.
Ji, X., Meng, X., Liu, C., Chen, R., Ge, Y., Kan, L., Fu, Q., Li, W., Tse, L. A. & Kan, H. 2019. Nitrogen dioxide air pollution and preterm birth in Shanghai, China. Environmental Research, 169, 79-85. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85056214559&doi=10.1016%2fj.envres.2018.11.007&partnerID=40&md5=c369c093a0aaf1a29111972d3be8bcb6
Jiang, L., Xiao, Q., Zhang, J., Zhao, Y., Chen, L. & Lu, S. 2022. Association between fetal exposure to polycyclic aromatic hydrocarbons and low birth weight: a case–control study in Shenzhen, China. Environmental Science and Pollution Research, 29, 59, 88779-88787. https://doi.org/10.1007/s11356-022-21965-6
Jiang, P., Li, Y., Tong, M. K., Ha, S., Gaw, E., Nie, J., Mendola, P. & Wang, M. 2024. Wildfire particulate exposure and risks of preterm birth and low birth weight in the Southwestern United States. Public Health, 230, 81-88.
Johnson, J. B. 2022. An Introduction to Atmospheric Pollutant Dispersion Modelling. Environmental Sciences Proceedings, 19, 1, 18. https://www.mdpi.com/2673-4931/19/1/18

Johnson, S., Bobb, J. F., Ito, K., Savitz, D. A., Elston, B., Shmool, J. L., Dominici, F., Ross, Z., Clougherty, J. E. & Matte, T. 2016. Ambient Fine Particulate Matter, Nitrogen Dioxide, and Preterm Birth in New York City. Environ Health Perspect, 124, 8, 1283-90.
Jones, S. I., Pruszynski, J. E., Spong, C. Y. & Nelson, D. B. 2023. Traffic-related air pollution is associated with spontaneous extremely preterm birth and other adverse perinatal outcomes. American Journal of Obstetrics and Gynecology, 229, 4, 455.e1-455.e7. https://www.sciencedirect.com/science/article/pii/S0002937823005124
Ju, L., Hua, L., Xu, H., Li, C., Sun, S., Zhang, Q., Cao, J. & Ding, R. 2023. Maternal atmospheric particulate matter exposure and risk of adverse pregnancy outcomes: A meta-analysis of cohort studies. Environmental Pollution, 317, 120704. https://www.sciencedirect.com/science/article/pii/S0269749122019182
Ju, L., Li, C., Yang, M., Sun, S., Zhang, Q., Cao, J. & Ding, R. 2021. Maternal air pollution exposure increases the risk of preterm birth: Evidence from the meta-analysis of cohort studies. Environ Res, 202, 111654.
Karlsson, L. & Takahashi, R. 2017. Evidence syntheses in support of policy-making. *A resource for developing an evidence synthesis report for policy-making.* Copenhagen: WHO Regional Office for Europe. Available: https://www.ncbi.nlm.nih.gov/books/NBK453546/
Kelly, F. J. & Fussell, J. C. 2012. Size, source and chemical composition as determinants of toxicity attributable to ambient particulate matter. Atmospheric Environment, 60, 504-526. https://www.sciencedirect.com/science/article/pii/S1352231012006024
Kelly, M. M. & Griffith, P. B. 2020. Umbrella Review of School Age Health Outcomes of Preterm Birth Survivors. Journal of Pediatric Health Care, 34, 5, e59-e76. https://www.sciencedirect.com/science/article/pii/S0891524520301516
Kephart, J. L., Fandiño-Del-Rio, M., Williams, K. N., Malpartida, G., Lee, A., Steenland, K., Naeher, L. P., Gonzales, G. F., Chiang, M., Checkley, W. & Koehler, K. 2021. Nitrogen dioxide exposures from LPG stoves in a cleaner-cooking intervention trial. Environment International, 146, 106196. https://www.sciencedirect.com/science/article/pii/S0160412020321516
Khader, Y., Abdelrahman, M., Abdo, N., Awad, S., Al-Sharif, M., Elbetieha, A. & Malkawi, M. 2016. Exposure to Air Pollution and Pregnancy Outcomes in the East Mediterranean Region: a Systematic Review. Journal of Pediatric Perspectives, 4, 1, 1255-1271. https://ijp.mums.ac.ir/article_6285_575595763ec13b740faa3f234b59eb35.pdf
Khreis, H., Kelly, C., Tate, J., Parslow, R., Lucas, K. & Nieuwenhuijsen, M. 2017. Exposure to traffic-related air pollution and risk of development of childhood asthma: A systematic review and meta-analysis. Environ Int, 100, 1-31.
Kim, H. J., Kim, J. H., Chay, D. B., Park, J. H. & Kim, M. A. 2017. Association of isolated single umbilical artery with perinatal outcomes: Systemic review and meta-analysis. Obstet Gynecol Sci, 60, 3, 266-273.
Kim, J. H., Choi, Y. Y., Yoo, S. I. & Kang, D. R. 2021. Association between ambient air pollution and high-risk pregnancy: A 2015–2018 national population-based cohort study in Korea. Environmental Research, 197. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85103328547&doi=10.1016%2fj.envres.2021.110965&partnerID=40&md5=5f147f047d406de02f396923b96da532
Kim, J. J., Axelrad, D. A. & Dockins, C. 2019. Preterm birth and economic benefits of reduced maternal exposure to fine particulate matter. Environmental Research, 170, 178-186. https://www.sciencedirect.com/science/article/pii/S001393511830642X
Kingsley, S. L., Eliot, M. N., Glazer, K., Awad, Y. A., Schwartz, J. D., Savitz, D. A., Kelsey, K. T., Marsit, C. J. & Wellenius, G. A. 2017. 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, 71, 12, 1131-1136. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85037060242&doi=10.1136%2fjech-2017-208963&partnerID=40&md5=e78b6f40bb30d50eafc97f20b2ad364a
Kirwa, K., Feric, Z., Manjourides, J., Alshawabekeh, A., Vega, C. M. V., Cordero, J. F., Meeker, J. D. & Suh, H. H. 2021. Preterm birth and PM2.5 in Puerto Rico: evidence from the PROTECT birth cohort. Environmental Health: A Global Access Science Source, 20, 1. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85107790214&doi=10.1186%2fs12940-021-00748-5&partnerID=40&md5=2600654942c0979243d09fdc242b1fef
Klepac, P., Locatelli, I., Korošec, S., Künzli, N. & Kukec, A. 2018. Ambient air pollution and pregnancy outcomes: A comprehensive review and identification of environmental public health challenges. Environ Res, 167, 144-159.

Kloog, I., Melly, S. J., Ridgway, W. L., Coull, B. A. & Schwartz, J. 2012. Using new satellite based exposure methods to study the association between pregnancy $PM_{2.5}$ exposure, premature birth and birth weight in Massachusetts. Environ Health, 11, 40.
Kolstad, C. D. 2011. *Environmental Economics (Chapter 7 and 10),* New York, Oxford University Press
Kornartit, C., Sokhi, R. S., Burton, M. A. & Ravindra, K. 2010. Activity pattern and personal exposure to nitrogen dioxide in indoor and outdoor microenvironments. Environment International, 36, 1, 36-45. https://www.sciencedirect.com/science/article/pii/S0160412009002037
Koullali, B., Van Zijl, M. D., Kazemier, B. M., Oudijk, M. A., Mol, B. W. J., Pajkrt, E. & Ravelli, A. C. J. 2020. The association between parity and spontaneous preterm birth: a population based study. BMC Pregnancy and Childbirth, 20, 1, 233. https://doi.org/10.1186/s12884-020-02940-w
Krajewski, A. K., Luben, T. J., Warren, J. L. & Rappazzo, K. M. 2023. Associations between weekly gestational exposure of fine particulate matter, ozone, and nitrogen dioxide and preterm birth in a North Carolina Birth Cohort, 2003–2015. Environmental Epidemiology, 7, 6. https://journals.lww.com/environepidem/fulltext/2023/12000/associations_between_weekly_gestational_exposure.7.aspx
Kriit, H. K., Sommar, J. N. & Åström, S. 2024. Socioeconomic per-case costs of stroke, myocardial infarction, and preterm birth attributable to air pollution in Sweden. PLOS ONE, 19, 1, e0290766. https://doi.org/10.1371/journal.pone.0290766
Kumar, N. & Singh, A. K. 2022. Impact of environmental factors on human semen quality and male fertility: a narrative review. Environmental Sciences Europe, 34, 1, 6. https://doi.org/10.1186/s12302-021-00585-w
Kumar, S., Sharma, S. & Thaker, R. 2017. Occupational, Environmental, and Lifestyle Factors and their Contribution to Preterm Birth - An Overview. Indian J Occup Environ Med, 21, 1, 9-17.
Kwag, Y., Kim, M. H., Ye, S., Oh, J., Yim, G., Kim, Y. J., Kim, E., Lee, S., Koh, T. K. & Ha, E. 2021a. The Combined Effects of Fine Particulate Matter and Temperature on Preterm Birth in Seoul, 2010-2016. Int J Environ Res Public Health, 18, 4.
Kwag, Y., Kim, M. H., Oh, J., Shah, S., Ye, S. & Ha, E. H. 2021b. Effect of heat waves and fine particulate matter on preterm births in Korea from 2010 to 2016. Environment International, 147. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85098079016&doi=10.1016%2fj.envint.2020.106239&partnerID=40&md5=adb51877b725d510f8e07cfa52e06f2e
Ladhani, N. N., Shah, P. S. & Murphy, K. E. 2011. Prenatal amphetamine exposure and birth outcomes: a systematic review and metaanalysis. Am J Obstet Gynecol, 205, 3, 219.e1-7.
Lamichhane, D. K., Leem, J. H., Lee, J. Y. & Kim, H. C. 2015. A meta-analysis of exposure to particulate matter and adverse birth outcomes. Environ Health Toxicol, 30, e2015011.
Lange, S. S. 2018. Comparing apples to oranges: Interpreting ozone concentrations from observational studies in the context of the United States ozone regulatory standard. Science of The Total Environment, 644, 1547-1556. https://www.sciencedirect.com/science/article/pii/S0048969718324483
Larkin, A. & Hystad, P. 2017. Towards Personal Exposures: How Technology Is Changing Air Pollution and Health Research. Curr Environ Health Rep, 4, 4, 463-471.
Lasserson Tj, Thomas J & Jpt, H. 2024. Chapter 1: Starting a review [last updated August 2021]. Available: https://training.cochrane.org/handbook/current/chapter-01
Laumbach, R., Meng, Q. & Kipen, H. 2015. What can individuals do to reduce personal health risks from air pollution? J Thorac Dis, 7, 1, 96-107.
Laurent, O., Hu, J., Li, L. & Et Al. 2016. A Statewide Nested Case-Control Study of Preterm Birth and Air Pollution by Source and Composition: California, 2001-2008. Environmental Health Perspectives, 124, 9. http://ehp.niehs.nih.gov/15-10133/
Lavigne, É., R.T., B., Stieb, D. M., G.J., E., Godri Pollitt, K. J., H., C., Van Rijswijk, D. & Weichenthal, S. 2018. Fine Particulate Air Pollution and Adverse Birth Outcomes: Effect Modification by Regional Nonvolatile Oxidative Potential. Environmental Health Perspectives, 126, 7, 077012. https://ehp.niehs.nih.gov/doi/abs/10.1289/EHP2535
Lavigne, E., Yasseen, A. S., Stieb, D. M., Hystad, P., Van Donkelaar, A., Martin, R. V., Brook, J. R., Crouse, D. L., Burnett, R. T., Chen, H., Weichenthal, S., Johnson, M., Villeneuve, P. J. & Walker, M. 2016. Ambient air pollution and adverse birth outcomes: Differences by maternal comorbidities. Environmental Research, 148, 457-466. https://www.sciencedirect.com/science/article/pii/S0013935116301517

Le, H. Q., Batterman, S. A., Wirth, J. J., Wahl, R. L., Hoggatt, K. J., Sadeghnejad, A., Hultin, M. L. & Depa, M. 2012. Air pollutant exposure and preterm and term small-for-gestational-age births in Detroit, Michigan: Long-term trends and associations. Environment International, 44, 1, 7-17. https://www.scopus.com/inward/record.uri?eid=2-s2.0-84860589094&doi=10.1016%2fj.envint.2012.01.003&partnerID=40&md5=d62bafad4eff91b36e944315ffc2cb44

Lee, P. C., Roberts, J. M., Catov, J. M., Talbott, E. O. & Ritz, B. 2013. First trimester exposure to ambient air pollution, pregnancy complications and adverse birth outcomes in Allegheny County, PA. Matern Child Health J, 17, 3, 545-55.

Lee, S., Walton, H., Smith, R., Exley, K. & Mcenery, K. 2022a. A systematic review and meta-analysis exploring the impact of ambient air pollution exposure during pregnancy on the risk of preterm birth. PROSPERO CRD42022301743. https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42022301743

Lee, W., Wu, X., Heo, S., Fong, K., Son, J., Sabath, M., Braun, D., Park, Y., Kim, Y., Lee, J., Schwartz, J., Kim, H., Dominici, F. & Bell, M. 2022b. Associations between long term air pollution exposure and first hospital admission for kidney and total urinary system diseases in the US Medicare population: nationwide longitudinal cohort study. BMJ Medicine, 1, 1, e000009. https://bmjmedicinesite-bmj.vercel.app/content/1/1/e000009

Leem, J., Kaplan, B. M., Shim, Y. K., Pohl, H. R., Gotway, C. A., Bullard, S. M., Rogers, J. F., Smith, M. M. & Tylenda, C. A. 2006. Exposures to air pollutants during pregnancy and preterm delivery. Environmental Health Perspectives, 114, 6, 905-910. https://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,shib,url,uid&db=cin20&AN=106147751&site=ehost-live

Leung, D. Y. C. 2015. Outdoor-indoor air pollution in urban environment: challenges and opportunity. Frontiers in Environmental Science, 2. https://www.frontiersin.org/journals/environmental-science/articles/10.3389/fenvs.2014.00069

Leung, M., Kioumourtzoglou, M. A., Raz, R. & Weisskopf, M. G. 2021. Bias due to Selection on Live Births in Studies of Environmental Exposures during Pregnancy: A Simulation Study. Environ Health Perspect, 129, 4, 47001.

Leung, M., Weisskopf, M., Modest, A., Hacker, M., Iyer, H., Hart, J., Wei, Y., Schwartz, J., Coull, B., Laden, F. & Papatheodorou, S. 2024. Using Parametric g-Computation for Time-to-Event Data and Distributed Lag Models to Identify Critical Exposure Windows for Preterm Birth: An Illustrative Example Using PM2.5 in a Retrospective Birth Cohort Based in Eastern Massachusetts (2011-2016). Environmental Health Perspectives, 132, 7, 077002. https://ehp.niehs.nih.gov/doi/abs/10.1289/EHP13891

Li, C., Xu, J. J., He, Y. C., Chen, L., Dennis, C. L., Huang, H. F. & Wu, Y. T. 2021. Effects of acute ambient pollution exposure on preterm prelabor rupture of membranes: A time-series analysis in Shanghai, China. Environmental Pollution, 276. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85101785895&doi=10.1016%2fj.envpol.2021.116756&partnerID=40&md5=65a1e8dab27c35c70638a44a84ce497a

Li, J., Gu, J., Liu, L., Cao, M., Wang, Z., Tian, X. & He, J. 2024. The relationship between air pollutants and preterm birth and blood routine changes in typical river valley city. BMC Public Health, 24, 1, 1677. https://doi.org/10.1186/s12889-024-19140-2

Li, Q., Wang, Y. Y., Guo, Y., Zhou, H., Wang, X., Wang, Q., Shen, H., Zhang, Y., Yan, D., Zhang, Y., Zhang, H., Li, S., Chen, G., Lin, L., Zhao, J., He, Y., Yang, Y., Xu, J., Wang, Y., Peng, Z., Wang, H. J. & Ma, X. 2018. Effect of airborne particulate matter of 2.5 μm or less on preterm birth: A national birth cohort study in China. Environ Int, 121, Pt 2, 1128-1136.

Li, Q., Wang, Y. Y., Guo, Y., Zhou, H., Wang, X., Wang, Q. M., Shen, H. P., Zhang, Y. P., Yan, D. H., Li, S., Chen, G., Lin, L., He, Y., Yang, Y., Peng, Z. Q., Wang, H. J. & Ma, X. 2020. Folic acid supplementation and the association between maternal airborne particulate matter exposure and preterm delivery: A national birth cohort study in China. Environmental Health Perspectives, 128, 12, 127010-1-127010-9. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85098630037&doi=10.1289%2fEHP6386&partnerID=40&md5=376bf5c47b6696e10e18d36b6d122dab

Li, S., Zhang, M., Tian, H., Liu, Z., Yin, X. & Xi, B. 2014. Preterm birth and risk of type 1 and type 2 diabetes: systematic review and meta-analysis. Obes Rev, 15, 10, 804-11.

Li, X., Huang, S., Jiao, A., Yang, X., Yun, J., Wang, Y., Xue, X., Chu, Y., Liu, F., Liu, Y., Ren, M., Chen, X., Li, N., Lu, Y., Mao, Z., Tian, L. & Xiang, H. 2017. Association between ambient fine

particulate matter and preterm birth or term low birth weight: An updated systematic review and meta-analysis. Environ Pollut, 227, 596-605.
Liang, D., Li, Z., Vlaanderen, J., Tang, Z., Jones, D. P., Vermeulen, R. & Sarnat, J. A. 2023. A State-of-the-Science Review on High-Resolution Metabolomics Application in Air Pollution Health Research: Current Progress, Analytical Challenges, and Recommendations for Future Direction. Environ Health Perspect, 131, 5, 56002.
Liang, Y., Li, M., Lyu, Q., Li, P., Lyu, Y., Yu, Y. & Peng, W. 2024. The relationship between maternal exposure to ambient air pollutants and premature rupture of membranes: A systematic review and meta-analysis. Environmental Pollution, 347, 123611. https://www.sciencedirect.com/science/article/pii/S0269749124003257
Liang, Z., Yang, Y., Qian, Z., Ruan, Z., Chang, J., Vaughn, M. G., Zhao, Q. & Lin, H. 2019a. Ambient PM2.5 and birth outcomes: Estimating the association and attributable risk using a birth cohort study in nine Chinese cities. Environment International, 126, 329-335. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85062236180&doi=10.1016%2fj.envint.2019.02.017&partnerID=40&md5=2384a31a816622b98044e3ad8a4e04f7
Liang, Z., Yang, Y., Li, J., Zhu, X., Ruan, Z., Chen, S., Huang, G., Lin, H., Zhou, J. Y. & Zhao, Q. 2019b. 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, 222, 7, 1047-1053. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85068816467&doi=10.1016%2fj.ijheh.2019.07.004&partnerID=40&md5=448dd7d6d88ebbf5f5d53ae6208415c8
Liang, Z., Zhao, L., Qiu, J., Zhu, X., Jiang, M., Liu, G. & Zhao, Q. 2022. PM2.5 exposure increases the risk of preterm birth in pre-pregnancy impaired fasting glucose women: A cohort study in a Southern province of China. Environmental Research, 204. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85119529233&doi=10.1016%2fj.envres.2021.112403&partnerID=40&md5=f18c8a096791f622009c9aa0e4306b3c
Liew, Z., Olsen, J., Cui, X., Ritz, B. & Arah, O. A. 2015. Bias from conditioning on live birth in pregnancy cohorts: an illustration based on neurodevelopment in children after prenatal exposure to organic pollutants. International Journal of Epidemiology, 44, 1, 345-354. https://doi.org/10.1093/ije/dyu249
Lin, Y. T., Jung, C. R., Lee, Y. L. & Hwang, B. F. 2015. Associations between ozone and preterm birth in women who develop gestational diabetes. American Journal of Epidemiology, 181, 4, 280-287. https://www.scopus.com/inward/record.uri?eid=2-s2.0-84927978464&doi=10.1093%2faje%2fkwu264&partnerID=40&md5=ef7fa60c7983df07c26b34738e7a5dba
Lisonkova, S. & Joseph, K. S. 2015. Left truncation bias as a potential explanation for the protective effect of smoking on preeclampsia. Epidemiology, 26, 3, 436-40.
Liu, B., Xu, G., Sun, Y., Qiu, X., Ryckman, K. K., Yu, Y., Snetselaar, L. G. & Bao, W. 2020. Maternal cigarette smoking before and during pregnancy and the risk of preterm birth: A dose-response analysis of 25 million mother-infant pairs. PLoS Med, 17, 8, e1003158.
Liu, C., Sun, J., Liu, Y., Liang, H., Wang, M., Wang, C. & Shi, T. 2017. Different exposure levels of fine particulate matter and preterm birth: a meta-analysis based on cohort studies. Environ Sci Pollut Res Int, 24, 22, 17976-17984.
Liu, S., Krewski, D., Shi, Y. & Et Al. 2003. Association between gaseous ambient air pollutants and adverse pregnancy outcomes in Vancouver, Canada. Environmental Health Perspectives, 111, 14, 1773-1778. http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=mwic&NEWS=N&AN=20031218 42
Liu, S., Zhang, Y., Ma, R., Liu, X., Liang, J., Lin, H., Shen, P., Zhang, J., Lu, P., Tang, X., Li, T. & Gao, P. 2022a. Long-term exposure to ozone and cardiovascular mortality in a large Chinese cohort. Environ Int, 165, 107280.
Liu, X., Miao, H., Behrman, J. R., Hannum, E., Liang, Z. & Zhao, Q. 2022b. The Asian Games, air pollution and birth outcomes in South China: An instrumental variable approach. Economics & Human Biology, 44, 101078. https://www.sciencedirect.com/science/article/pii/S1570677X21001039
Liu, X., Ye, Y., Chen, Y., Li, X., Feng, B., Cao, G., Xiao, J., Zeng, W., Li, X., Sun, J., Ning, D., Yang, Y., Yao, Z., Guo, Y., Wang, Q., Zhang, Y., Ma, W., Du, Q., Zhang, B. & Liu, T. 2019a. 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, 133. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85073292115&doi=10.1016%2fj.envint.2019.105177&partnerID=40&md5=8b114f45eee337b41280e55f9f27d2f0

Liu, X.C., Strodl, E., Huang, L.H., Hu, B.J. & Chen, W.Q. 2022c. Effect of Prenatal Exposure to Household Air Pollution from Multiple Sources on Risk of Preterm Birth. Atmosphere, 13, 12, 2022. https://www.mdpi.com/2073-4433/13/12/2022

Liu, X.-X., Fan, S.-J., Luo, Y.-N., Hu, L.-X., Li, C.-C., Zhang, Y.-D., Li, J.-X., Qiu, H.-L., Dong, G.-H. & Yang, B.-Y. 2023. Global, regional, and national burden of preterm birth attributable to ambient and household PM2.5 from 1990 to 2019: Worsening or improving? Science of The Total Environment, 871, 161975. https://www.sciencedirect.com/science/article/pii/S0048969723005909

Liu, Y., Xu, J., Chen, D. & Et Al. 2019b. The association between air pollution and preterm birth and low birth weight in Guangdong, China. BMC Public Health, 19, 3. https://doi.org/10.1186/s12889-018-6307-7

Llop, S., Ballester, F., Estarlich, M., Esplugues, A., Rebagliato, M. & Iñiguez, C. 2010. Preterm birth and exposure to air pollutants during pregnancy. Environmental Research, 110, 8, 778-785. https://www.scopus.com/inward/record.uri?eid=2-s2.0-78049291386&doi=10.1016%2fj.envres.2010.09.009&partnerID=40&md5=c59cbaaeaa6d209cfd32f43d522069ed

Lo, C. K.-L., Mertz, D. & Loeb, M. 2014. Newcastle-Ottawa Scale: comparing reviewers' to authors' assessments. BMC Medical Research Methodology, 14, 1, 45. https://doi.org/10.1186/1471-2288-14-45

Lock, K. 2000. Health impact assessment. Bmj, 320, 7246, 1395-8.

Lomas, J., Schmitt, L., Jones, S., Mcgeorge, M., Bates, E., Holland, M., Cooper, D., Crowther, R., Ashmore, M., Rojas-Rueda, D., Weatherly, H., Richardson, G. & Bojke, L. 2016. A pharmacoeconomic approach to assessing the costs and benefits of air quality interventions that improve health: a case study. BMJ Open, 6, 6, e010686. https://bmjopen.bmj.com/content/bmjopen/6/6/e010686.full.pdf

Lu, C., Cao, L., Norbäck, D., Li, Y., Chen, J. & Deng, Q. 2019. Combined effects of traffic air pollution and home environmental factors on preterm birth in China. Ecotoxicology and Environmental Safety, 184. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85072024848&doi=10.1016%2fj.ecoenv.2019.109639&partnerID=40&md5=baf652d1031ef149b2c35ba343e427ad

Luchini, C., Veronese, N., Nottegar, A., Shin, J. I., Gentile, G., Granziol, U., Soysal, P., Alexinschi, O., Smith, L. & Solmi, M. 2021. Assessing the quality of studies in meta-research: Review/guidelines on the most important quality assessment tools. Pharm Stat, 20, 1, 185-195.

Luo, D., Kuang, T., Chen, Y. X., Huang, Y. H., Zhang, H. & Xia, Y. Y. 2021. Air pollution and pregnancy outcomes based on exposure evaluation using a land use regression model: A systematic review. Taiwan J Obstet Gynecol, 60, 2, 193-215.

Lupo, P. J., Symanski, E., Chan, W., Mitchell, L. E., Waller, D. K., Canfield, M. A. & Langlois, P. H. 2010. Differences in exposure assignment between conception and delivery: the impact of maternal mobility. Paediatr Perinat Epidemiol, 24, 2, 200-8.

Macintyre, E. A., Gehring, U., Mölter, A., Fuertes, E., Klümper, C., Krämer, U., Quass, U., Hoffmann, B., Gascon, M., Brunekreef, B., Koppelman, G. H., Beelen, R., Hoek, G., Birk, M., De Jongste, J. C., Smit, H. A., Cyrys, J., Gruzieva, O., Korek, M., Bergström, A., Agius, R. M., De Vocht, F., Simpson, A., Porta, D., Forastiere, F., Badaloni, C., Cesaroni, G., Esplugues, A., Fernández-Somoano, A., Lerxundi, A., Sunyer, J., Cirach, M., Nieuwenhuijsen, M. J., Pershagen, G. & Heinrich, J. 2014. Air pollution and respiratory infections during early childhood: an analysis of 10 European birth cohorts within the ESCAPE Project. Environ Health Perspect, 122, 1, 107-13.

Madden, M. C., Stevens, T., Case, M., Schmitt, M., Diaz-Sanchez, D., Bassett, M., Montilla, T. S., Berntsen, J. & Devlin, R. B. 2014. Diesel exhaust modulates ozone-induced lung function decrements in healthy human volunteers. Part Fibre Toxicol, 11, 37.

Maitre, L., Fthenou, E., Athersuch, T., Coen, M., Toledano, M. B., Holmes, E., Kogevinas, M., Chatzi, L. & Keun, H. C. 2014. Urinary metabolic profiles in early pregnancy are associated with preterm birth and fetal growth restriction in the Rhea mother-child cohort study. BMC Med, 12, 110.

Mangham, L. J., Petrou, S., Doyle, L. W., Draper, E. S. & Marlow, N. 2009. The Cost of Preterm Birth Throughout Childhood in England and Wales. Pediatrics, 123, 2, e312-e327. https://doi.org/10.1542/peds.2008-1827
Manisalidis, I., Stavropoulou, E., Stavropoulos, A. & Bezirtzoglou, E. 2020. Environmental and Health Impacts of Air Pollution: A Review. Front Public Health, 8, 14. https://pmc.ncbi.nlm.nih.gov/articles/PMC7044178/
Manuck, T. A. 2017. Racial and ethnic differences in preterm birth: A complex, multifactorial problem. Semin Perinatol, 41, 8, 511-518.
March of Dimes, Pmmch, Save the Children & Who 2012. Born Too Soon: The Global action report on preterm Birth. Geneva. Available: https://iris.who.int/bitstream/handle/10665/44864/9789241503433_eng.pdf?sequence=1
Markandya, A., Sampedro, J., Smith, S. J., Van Dingenen, R., Pizarro-Irizar, C., Arto, I. & González-Eguino, M. 2018. Health co-benefits from air pollution and mitigation costs of the Paris Agreement: a modelling study. The Lancet Planetary Health, 2, 3, e126-e133. https://doi.org/10.1016/S2542-5196(18)30029-9
Maroziene, L. & Grazuleviciene, R. 2002. Maternal exposure to low-level air pollution and pregnancy outcomes: Apopulation-based study. Environmental Health: A Global Access Science Source, 1, 1-7. https://www.scopus.com/inward/record.uri?eid=2-s2.0-34248377786&doi=10.1186%2f1476-069X-1-1&partnerID=40&md5=30d011cb90eb4703c2781a97110c8d16
Márquez-Lázaro, J., Madera, M. & Bernabe, E. 2024. Particulate matter 2.5 exposure during pregnancy and birth outcomes: Evidence from Colombia. Science of The Total Environment, 927, 172369. https://www.sciencedirect.com/science/article/pii/S0048969724025154
Marshall, I. J., Marshall, R., Wallace, B. C., Brassey, J. & Thomas, J. 2019. Rapid reviews may produce different results to systematic reviews: a meta-epidemiological study. Journal of Clinical Epidemiology, 109, 30-41. https://www.sciencedirect.com/science/article/pii/S089543561830893X
Martin, A., Faes, C., Debevec, T., Rytz, C., Millet, G. & Pialoux, V. 2018. Preterm birth and oxidative stress: Effects of acute physical exercise and hypoxia physiological responses. Redox Biol, 17, 315-322.
Mayor of London. 2018. *London Environmental Strategy* [Online]. London. Available: https://www.london.gov.uk/sites/default/files/london_environment_strategy_0.pdf
Mcelroy, S., Ilango, S., Dimitrova, A., Gershunov, A. & Benmarhnia, T. 2022. Extreme heat, preterm birth, and stillbirth: A global analysis across 14 lower-middle income countries. Environment International, 158, 106902. https://www.sciencedirect.com/science/article/pii/S0160412021005274
Mcgartland, A., Revesz, R., Axelrad, D. A., Dockins, C., Sutton, P. & Woodruff, T. J. 2017. Estimating the health benefits of environmental regulations. Science, 357, 6350, 457-458.
Mchugh, M. L. 2012. Interrater reliability: the kappa statistic. Biochem Med (Zagreb), 22, 3, 276-82. https://www.biochemia-medica.com/assets/images/upload/xml_tif/McHugh_ML_Interrater_reliability.pdf
Mehlig, D., Woodward, H., Oxley, T., Holland, M. & Apsimon, H. 2021. Electrification of Road Transport and the Impacts on Air Quality and Health in the UK. Atmosphere, 12, 11, 1491.
Mekonnen, Z. K., Oehlert, J. W., Eskenazi, B., Shaw, G. M., Balmes, J. R. & Padula, A. M. 2021. The relationship between air pollutants and maternal socioeconomic factors on preterm birth in California urban counties. Journal of Exposure Science & Environmental Epidemiology, 31, 3, 503-513. https://doi.org/10.1038/s41370-021-00323-7
Melody, S., Wills, K., Knibbs, L. D., Ford, J., Venn, A. & Johnston, F. 2020. Adverse birth outcomes in Victoria, Australia in association with maternal exposure to low levels of ambient air pollution. Environmental Research, 188. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85086563633&doi=10.1016%2fj.envres.2020.109784&partnerID=40&md5=09b3af1d941a9be568a47790f285a79b
Melody, S. M., Ford, J., Wills, K., Venn, A. & Johnston, F. H. 2019a. Maternal exposure to fine particulate matter from a coal mine fire and birth outcomes in Victoria, Australia. Environment International, 127, 233-242. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85063455879&doi=10.1016%2fj.envint.2019.03.028&partnerID=40&md5=f2ffd153b2ec1478b4cd509f66dee5fc
Melody, S. M., Ford, J., Wills, K., Venn, A. & Johnston, F. H. 2019b. Maternal exposure to short-to medium-term outdoor air pollution and obstetric and neonatal outcomes: A systematic review. Environ Pollut, 244, 915-925.

Mendola, P., Nobles, C., Williams, A., Sherman, S., Kanner, J., Seeni, I. & Grantz, K. 2019. 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, 16, 18. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85072554748&doi=10.3390%2fijerph16183365&partnerID=40&md5=b971e649eac1ad4688d20181d82a949e

Mendola, P., Wallace, M., Hwang, B. S., Liu, D., Robledo, C., Männistö, T., Sundaram, R., Sherman, S., Ying, Q. & Grantz, K. L. 2016. Preterm birth and air pollution: Critical windows of exposure for women with asthma. J Allergy Clin Immunol, 138, 2, 432-440.e5.

Meng, Q., Liu, J., Shen, J., Rosario, I., Lakey, P., Shiraiwa, M., Su, J., Weichenthal, S., Zhu, Y., Oroumiyeh, F., Paulson, S., Jerrett, M. & Ritz, B. 2023. Fine Particulate Matter Metal Composition, Oxidative Potential, and Adverse Birth Outcomes in Los Angeles. Environmental Health Perspectives, 131, 10, 107012. https://ehp.niehs.nih.gov/doi/abs/10.1289/EHP12196

Mercer, B. M. 2003. Preterm premature rupture of the membranes. Obstet Gynecol, 101, 1, 178-93.

Mercer, L. D., Szpiro, A. A., Sheppard, L., Lindström, J., Adar, S. D., Allen, R. W., Avol, E. L., Oron, A. P., Larson, T., Liu, L. J. & Kaufman, J. D. 2011. Comparing universal kriging and land-use regression for predicting concentrations of gaseous oxides of nitrogen (NOx) for the Multi-Ethnic Study of Atherosclerosis and Air Pollution (MESA Air). Atmos Environ (1994), 45, 26, 4412-4420.

Metzger, M. J., Bunce, R. G. H., Jongman, R. H. G., Sayre, R., Trabucco, A. & Zomer, R. 2013. A high-resolution bioclimate map of the world: a unifying framework for global biodiversity research and monitoring. Global Ecology and Biogeography, 22, 5, 630-638. https://doi.org/10.1111/geb.12022

Mi, T., Qiu, Z., Li, C., Li, W., Gao, Y., Chen, Z., Xu, W., Liu, Z., Li, Q., Jiang, M., Liu, H., Dai, L. & Zhan, Y. 2024. Joint effects of green space and air pollutant exposure on preterm birth: evidence from a nationwide study in China. Environmental Science and Pollution Research, 31, 24, 35149-35160. https://doi.org/10.1007/s11356-024-33561-x

Midirs. 2024. *About MIDIRS* [Online]. Available: https://www.midirs.org/about-us/

Millard, T., Synnot, A., Elliott, J., Green, S., Mcdonald, S. & Turner, T. 2019. Feasibility and acceptability of living systematic reviews: results from a mixed-methods evaluation. Systematic Reviews, 8, 1, 325. https://doi.org/10.1186/s13643-019-1248-5

Mills, I. C., Atkinson, R. W., Kang, S., Walton, H. & Anderson, H. R. 2015. Quantitative systematic review of the associations between short-term exposure to nitrogen dioxide and mortality and hospital admissions. BMJ Open, 5, 5, e006946.

Milner, J., Hughes, R., Chowdhury, S., Picetti, R., Ghosh, R., Yeung, S., Lelieveld, J., Dangour, A. D. & Wilkinson, P. 2023a. Air pollution and child health impacts of decarbonization in 16 global cities: Modelling study. Environ Int, 175, 107972.

Milner, J., Turner, G., Ibbetson, A., Colombo, P. E., Green, R., Dangour, A. D., Haines, A. & Wilkinson, P. 2023b. Impact on mortality of pathways to net zero greenhouse gas emissions in England and Wales: a multisectoral modelling study. The Lancet Planetary Health, 7, 2, e128-e136.

Mitchell, R. J. & Bates, P. 2011. Measuring health-related productivity loss. Popul Health Manag, 14, 2, 93-8.

Mitku, A. A., Zewotir, T., North, D., Jeena, P., Asharam, K., Muttoo, S., Tularam, H. & Naidoo, R. N. 2023. Impact of ambient air pollution exposure during pregnancy on adverse birth outcomes: generalized structural equation modeling approach. BMC Public Health, 23, 1, 45. https://doi.org/10.1186/s12889-022-14971-3

Mittinty, M. N. & Lynch, J. 2022. Reflection on modern methods: risk ratio regression—simple concept yet complex computation. International Journal of Epidemiology, 52, 1, 309-314. https://doi.org/10.1093/ije/dyac220

Moghaddamhosseini, V., Ebrahimi Aval, H., Lari Najafi, M., Lotfi, H., Heydari, H., Miri, M. & Dadvand, P. 2023. The association between exposure to polycyclic aromatic hydrocarbons and birth outcomes: A systematic review and meta-analysis of observational studies. Science of The Total Environment, 905, 166922. https://www.sciencedirect.com/science/article/pii/S004896972305547X

Moons, P., Goossens, E. & Thompson, D. R. 2021. Rapid reviews: the pros and cons of an accelerated review process. European Journal of Cardiovascular Nursing, 20, 5, 515-519. https://doi.org/10.1093/eurjcn/zvab041

Morgen, C. S., Bjørk, C., Andersen, P. K., Mortensen, L. H. & Nybo Andersen, A.-M. 2008. Socioeconomic position and the risk of preterm birth—a study within the Danish National Birth Cohort. International Journal of Epidemiology, 37, 5, 1109-1120. https://doi.org/10.1093/ije/dyn112
Morniroli, D., Tiraferri, V., Maiocco, G., De Rose, D. U., Cresi, F., Coscia, A., Mosca, F. & Giannì, M. L. 2023. Beyond survival: the lasting effects of premature birth. Frontiers in Pediatrics, 11. https://www.frontiersin.org/journals/pediatrics/articles/10.3389/fped.2023.1213243
Morshed, S. R., Fattah, M. A., Kafy, A. A., Alsulamy, S., Almulhim, A. I., Shohan, A. a. A. & Khedher, K. M. 2024. Decoding seasonal variability of air pollutants with climate factors: A geostatistical approach using multimodal regression models for informed climate change mitigation. Environ Pollut, 345, 123463.
Moster, D., Lie, R. T. & Markestad, T. 2008. Long-term medical and social consequences of preterm birth. N Engl J Med, 359, 3, 262-73.
Mumby, S., Chung, K. F. & Adcock, I. M. 2019. Transcriptional Effects of Ozone and Impact on Airway Inflammation. Front Immunol, 10, 1610.
National Environmental Technology Centre (NETCEN) 2003. Air Pollution in the UK: 2003 - A report prepared by the National Environmental Technology Centre for the Department for Environment, Food and Rural Affairs and the Devolved Administrations (Part 3 - Appendices). Available: https://naei.energysecurity.gov.uk/sites/default/files/cat05/0502251136_Report_Air_Pollution_in_the_UK_Part_3_appendices.pdf
National Health Services (NHS). 2021. *Stillbirth* [Online]. Available: https://www.nhs.uk/conditions/stillbirth/
National Health Services (NHS). 2024. *Keeping well in pregnancy* [Online]. Available: https://www.nhs.uk/pregnancy/keeping-well/
National Institue of Health (NIH). 2024. *Study Quality Assessment Tools* [Online]. Available: https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools
National Institute for Health and Care Research (NIHR). 2020. *The air quality health and economic costs and benefits of a zero carbon UK* [Online]. Available: https://www.fundingawards.nihr.ac.uk/award/NIHR129406
National Institute for Health and Clinical Excellence (NICE). 2020. *Miscarriage: What is it?* [Online]. Available: https://cks.nice.org.uk/topics/miscarriage/background-information/definition/
National Library of Medicine (NLM). 2023. *MEDLINE, PubMed, and PMC (PubMed Central): How are they different?* [Online]. Available: https://www.nlm.nih.gov/bsd/difference.html
National Oceanic and Atmsopheric Administration (NOAA). 2024. *U.S. Climate Regions* [Online]. Available: https://www.ncei.noaa.gov/access/monitoring/reference-maps/us-climate-regions
Nealy, B. E., Warren, J. L., Strickland, M. J., Darrow, L. A. & Chang, H. H. 2018. Impacts of gestational age uncertainty in estimating associations between preterm birth and ambient air pollution. Environ Epidemiol, 2, 4, e031.
Neophytou, A. M., Kioumourtzoglou, M.-A., Goin, D. E., Darwin, K. C. & Casey, J. A. 2020. Educational note: addressing special cases of bias that frequently occur in perinatal epidemiology. International Journal of Epidemiology, 50, 1, 337-345. https://doi.org/10.1093/ije/dyaa252
Nhs Digital. 2024. *Maternity Services Monthly Statistics* [Online]. Available: https://digital.nhs.uk/data-and-information/publications/statistical/maternity-services-monthly-statistics
Ni, W., Xing, Y., Li, G., Du, Z., Yang, P., Wang, Q., Yang, X., Lyu, B., Fa, H., Shi, Q. & Xing, Q. 2024. Windows of sensitivity for risk of adverse birth outcomes related to gestational PM2.5 exposure: Evidence from a natural experiment. Environmental Pollution, 347, 123759. https://www.sciencedirect.com/science/article/pii/S0269749124004731
Noghanibehambari, H., Bagheri, H., Salari, M., Tavassoli, N., Javid, R. & Toranji, M. 2023. Breathing in the future: prenatal exposure to air pollution and infants' health outcomes in the USA. Public Health, 225, 198-205. https://www.sciencedirect.com/science/article/pii/S0033350623003803
Norman, J. E., Morris, C. & Chalmers, J. 2009. The effect of changing patterns of obstetric care in Scotland (1980-2004) on rates of preterm birth and its neonatal consequences: perinatal database study. PLoS Med, 6, 9, e1000153.
Nyadanu, S. D., Dunne, J., Tessema, G. A., Mullins, B., Kumi-Boateng, B., Lee Bell, M., Duko, B. & Pereira, G. 2022. Prenatal exposure to ambient air pollution and adverse birth outcomes: An umbrella review of 36 systematic reviews and meta-analyses. Environmental Pollution, 306, 119465. https://www.sciencedirect.com/science/article/pii/S0269749122006790

O'connor, S. R., Tully, M. A., Ryan, B., Bradley, J. M., Baxter, G. D. & Mcdonough, S. M. 2015. Failure of a numerical quality assessment scale to identify potential risk of bias in a systematic review: a comparison study. BMC Res Notes, 8, 224.
Office for National Statistics (ONS). 2020. *Population projections – local authorities: SNPP Z1 - 2018 based edition* [Online]. Available: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationprojections/datasets/localauthoritiesinenglandz1
Office for National Statistics (ONS). 2022. *Births QMI* [Online]. Available: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/livebirths/methodologies/birthsqmi
Office for National Statistics (Ons). 2023. *Birth Characteristics (2021 Edition)* [Online]. Available: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/livebirths/datasets/birthcharacteristicsinenglandandwales .
Office for National Statistics (ONS). 2024. *Consumer price inflation tables* [Online]. Available: https://www.ons.gov.uk/economy/inflationandpriceindices/datasets/consumerpriceinflation
Ohuma, E. O., Moller, A.-B., Bradley, E., Chakwera, S., Hussain-Alkhateeb, L., Lewin, A., Okwaraji, Y. B., Mahanani, W. R., Johansson, E. W., Lavin, T., Fernandez, D. E., Domínguez, G. G., De Costa, A., Cresswell, J. A., Krasevec, J., Lawn, J. E., Blencowe, H., Requejo, J. & Moran, A. C. 2023. National, regional, and global estimates of preterm birth in 2020, with trends from 2010: a systematic analysis. The Lancet, 402, 10409, 1261-1271. https://doi.org/10.1016/S0140-6736(23)00878-4
Olsson, D., Ekström, M. & Forsberg, B. 2012. Temporal variation in air pollution concentrations and preterm birth-a population based epidemiological study. International Journal of Environmental Research and Public Health, 9, 1, 272-285. https://www.scopus.com/inward/record.uri?eid=2-s2.0-84856314863&doi=10.3390%2fijerph9010272&partnerID=40&md5=988b336d680fd532ad78c2c994b5c45f
Olsson, D., Mogren, I. & Forsberg, B. 2013. Air pollution exposure in early pregnancy and adverse pregnancy outcomes: A register-based cohort study. BMJ Open, 3, 2. https://www.scopus.com/inward/record.uri?eid=2-s2.0-84874644368&doi=10.1136%2fbmjopen-2012-001955&partnerID=40&md5=02b93dbc8b65ef2f301464e46b42b651
OpenAI. 2022. *Introducing ChatGPT* [Online]. Available: https://openai.com/index/chatgpt/
Orach, J., Rider, C. F. & Carlsten, C. 2021. Concentration-dependent health effects of air pollution in controlled human exposures. Environment International, 150, 106424. https://www.sciencedirect.com/science/article/pii/S0160412021000489
Organisation for Economic Co-Operation and Development (OECD) 2012. *Mortality Risk Valuation in Environment, Health and Transport Policies,* Available: https://www.oecd-ilibrary.org/content/publication/9789264130807-en
Organisation for Economic Co-Operation and Development (OECD). 2024. *Purchasing power parities (PPP) (indicator)* [Online]. Available: https://data.oecd.org/conversion/purchasing-power-parities-ppp.htm
Ottone, M., Broccoli, S., Parmagnani, F., Giannini, S., Scotto, F., Bonvicini, L., Luberto, F., Bacco, D., Trentini, A., Poluzzi, V., Angelini, P., Colacci, A., Giorgi Rossi, P. & Ranzi, A. 2020. Source-related components of fine particulate matter and risk of adverse birth outcomes in Northern Italy. Environmental Research, 186. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85088155986&doi=10.1016%2fj.envres.2020.109564&partnerID=40&md5=d585bce200bb1ba01f465af09a5e8a4a
Owens, E. O., Patel, M. M., Kirrane, E., Long, T. C., Brown, J., Cote, I., Ross, M. A. & Dutton, S. J. 2017. Framework for assessing causality of air pollution-related health effects for reviews of the National Ambient Air Quality Standards. Regul Toxicol Pharmacol, 88, 332-337.
Padula, A. M., Huang, H., Baer, R. J., August, L. M., Jankowska, M. M., Jellife-Pawlowski, L. L., Sirota, M. & Woodruff, T. J. 2018. Environmental pollution and social factors as contributors to preterm birth in Fresno County. Environmental Health: A Global Access Science Source, 17, 1. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85052725545&doi=10.1186%2fs12940-018-0414-x&partnerID=40&md5=e9cf14692819ac7f50653ffe12389721
Padula, A. M., Mortimer, K. M., Tager, I. B., Hammond, S. K., Lurmann, F. W., Yang, W., Stevenson, D. K. & Shaw, G. M. 2014. Traffic-related air pollution and risk of preterm birth in the San Joaquin Valley of California. Annals of Epidemiology, 24, 12, 888-895.e4.

https://www.scopus.com/inward/record.uri?eid=2-s2.0-84927175492&doi=10.1016%2fj.annepidem.2014.10.004&partnerID=40&md5=c59b96156f9c9accf31661afbf22ee31
Padula, A. M., Yang, W., Lurmann, F. W., Balmes, J., Hammond, S. K. & Shaw, G. M. 2019. Prenatal exposure to air pollution, maternal diabetes and preterm birth. Environmental Research, 170, 160-167. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85058684858&doi=10.1016%2fj.envres.2018.12.031&partnerID=40&md5=b322593679b1951f2c2ac8bf7cdbf3bd
Pan, D., Liu, S., Huang, D., Zeng, X., Zhang, Y., Pang, Q., Wu, H., Tan, H. J. J., Liang, J., Sheng, Y. & Qiu, X. 2022. Effects of household environmental exposure and ventilation in association with adverse birth outcomes: A prospective cohort study in rural China. Science of The Total Environment, 822, 153519. https://www.sciencedirect.com/science/article/pii/S0048969722006118
Panasevich, S., Håberg, S. E., Aamodt, G., London, S. J., Stigum, H., Nystad, W. & Nafstad, P. 2016. Association between pregnancy exposure to air pollution and birth weight in selected areas of Norway. Archives of Public Health, 74, 1. https://www.scopus.com/inward/record.uri?eid=2-s2.0-84976510612&doi=10.1186%2fs13690-016-0138-8&partnerID=40&md5=9f0c2b9d5e9a82a84a067d5dac9b59e6
Papadogeorgou, G., Kioumourtzoglou, M. A., Braun, D. & Zanobetti, A. 2019. Low Levels of Air Pollution and Health: Effect Estimates, Methodological Challenges, and Future Directions. Curr Environ Health Rep, 6, 3, 105-115.
Parkinson, J. R., Hyde, M. J., Gale, C., Santhakumaran, S. & Modi, N. 2013. Preterm birth and the metabolic syndrome in adult life: a systematic review and meta-analysis. Pediatrics, 131, 4, e1240-63.
Parry, J. & Stevens, A. 2001. Prospective health impact assessment: pitfalls, problems, and possible ways forward. Bmj, 323, 7322, 1177-82.
Patel, V. P., Davis, M., Li, J., Hwang, S., Johnson, S., Kondejewski, J., Croft, D., Rood, K. & Simhan, H. N. 2024. Workplace Productivity Loss and Indirect Costs Associated With Preterm Birth in the United States. Obstet Gynecol, 143, 1, 23-34.
Pereira, G. 2022. A Simple Method to Establish Sufficiency and Stability in Meta-Analyses: With Application to Fine Particulate Matter Air Pollution and Preterm Birth. International Journal of Environmental Research and Public Health, 19, 4. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85124323065&doi=10.3390%2fijerph19042036&partnerID=40&md5=2bc2e5804f4780f4ef224fe1eec508b4
Pereira, G., Belanger, K., Ebisu, K. & Bell, M. L. 2014a. Fine Particulate Matter and Risk of Preterm Birth in Connecticut in 2000–2006: A Longitudinal Study. American Journal of Epidemiology, 179, 1, 67-74. https://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,shib,url,uid&db=cin20&AN=104125976&site=ehost-live
Pereira, G., Bell, M. L., Belanger, K. & De Klerk, N. 2014b. 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, 73, 143-149. https://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,shib,url,uid&db=8gh&AN=98809523&site=ehost-live
Pereira, G., Bell, M. L., Hyung Joo, L., Koutrakis, P. & Belanger, K. 2014c. Sources of Fine Particulate Matter and Risk of Preterm Birth in Connecticut, 2000–2006: A Longitudinal Study. Environmental Health Perspectives, 122, 10, 1117-1122. https://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,shib,url,uid&db=cin20&AN=103902708&site=ehost-live
Pereira, G., Bracken, M. B. & Bell, M. L. 2016. Particulate air pollution, fetal growth and gestational length: The influence of residential mobility in pregnancy. Environmental Research, 147, 269-274. https://www.sciencedirect.com/science/article/pii/S0013935116300445
Perera, F., Berberian, A., Cooley, D., Shenaut, E., Olmstead, H., Ross, Z. & Matte, T. 2021. Potential health benefits of sustained air quality improvements in New York City: A simulation based on air pollution levels during the COVID-19 shutdown. Environmental Research, 193, 110555. https://www.sciencedirect.com/science/article/pii/S0013935120314523
Perera, F., Cooley, D., Berberian, A., Mills, D. & Kinney, P. 2020. Co-Benefits to Children's Health of the U.S. Regional Greenhouse Gas Initiative. Environmental Health Perspectives, 128, 7, 077006. https://ehp.niehs.nih.gov/doi/abs/10.1289/EHP6706

Picciotto, S., Huang, S., Lurmann, F., Pavlovic, N., Ying Chang, S., Mukherjee, A., Goin, D. E., Sklar, R., Noth, E., Morello-Frosch, R. & Padula, A. M. 2024. Pregnancy exposure to PM2.5 from wildland fire smoke and preterm birth in California. Environment International, 186, 108583. https://www.sciencedirect.com/science/article/pii/S0160412024001697

Poirier, A., Dodds, L., Dummer, T., Rainham, D., Maguire, B. & Johnson, M. 2015. Maternal Exposure to Air Pollution and Adverse Birth Outcomes in Halifax, Nova Scotia. Journal of Occupational and Environmental Medicine, 57, 12. https://journals.lww.com/joem/fulltext/2015/12000/maternal_exposure_to_air_pollution_and_adverse.6.aspx

Pomar, L., Favre, G., De Labrusse, C., Contier, A., Boulvain, M. & Baud, D. 2022. Impact of the first wave of the COVID-19 pandemic on birth rates in Europe: a time series analysis in 24 countries. Hum Reprod, 37, 12, 2921-2931.

Ponce, N. A., Hoggatt, K. J., Wilhelm, M. & Et Al. 2005. Preterm birth: the interaction of traffic-related air pollution with economic hardship in Los Angeles neighborhoods. American Journal of Epidemiology, 162, 2, 140-148. http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=mwic&NEWS=N&AN=20050080171

Poston, L., Bell, R., Croker, H., Flynn, A. C., Godfrey, K. M., Goff, L., Hayes, L., Khazaezadeh, N., Nelson, S. M., Oteng-Ntim, E., Pasupathy, D., Patel, N., Robson, S. C., Sandall, J., Sanders, T. a. B., Sattar, N., Seed, P. T., Wardle, J., Whitworth, M. K. & Briley, A. L. 2015. Effect of a behavioural intervention in obese pregnant women (the UPBEAT study): a multicentre, randomised controlled trial. The Lancet Diabetes & Endocrinology, 3, 10, 767-777. https://doi.org/10.1016/S2213-8587(15)00227-2

Power, M. C., Adar, S. D., Yanosky, J. D. & Weuve, J. 2016. Exposure to air pollution as a potential contributor to cognitive function, cognitive decline, brain imaging, and dementia: A systematic review of epidemiologic research. Neurotoxicology, 56, 235-253.

Pražnikar, J. 2017. Particulate matter time-series and Köppen-Geiger climate classes in North America and Europe. Atmospheric Environment, 150, 136-145. https://www.sciencedirect.com/science/article/pii/S1352231016309463

Prior, E. & Modi, N. 2020. Adult outcomes after preterm birth. Postgraduate Medical Journal, 96, 1140, 619-622. https://doi.org/10.1136/postgradmedj-2020-137707

Pryor, J. T., Cowley, L. O. & Simonds, S. E. 2022. The Physiological Effects of Air Pollution: Particulate Matter, Physiology and Disease. Front Public Health, 10, 882569.

Pusdekar, Y. V., Patel, A. B., Kurhe, K. G., Bhargav, S. R., Thorsten, V., Garces, A., Goldenberg, R. L., Goudar, S. S., Saleem, S., Esamai, F., Chomba, E., Bauserman, M., Bose, C. L., Liechty, E. A., Krebs, N. F., Derman, R. J., Carlo, W. A., Koso-Thomas, M., Nolen, T. L., Mcclure, E. M. & Hibberd, P. L. 2020. Rates and risk factors for preterm birth and low birthweight in the global network sites in six low- and low middle-income countries. Reprod Health, 17, Suppl 3, 187.

Qian, Z., Liang, S., Yang, S., Trevathan, E., Huang, Z., Yang, R., Wang, J., Hu, K., Zhang, Y., Vaughn, M., Shen, L., Liu, W., Li, P., Ward, P., Yang, L., Zhang, W., Chen, W., Dong, G., Zheng, T. & Xu, S. 2016. Ambient air pollution and preterm birth: A prospective birth cohort study in Wuhan, China. International Journal of Hygiene & Environmental Health, 219, 2, 195-203. https://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,shib,url,uid&db=cin20&AN=112630605&site=ehost-live

Qiao, P., Fan, K., Bao, Y., Yuan, L., Kan, H., Zhao, Y., Cai, J. & Ying, H. 2022. Prenatal exposure to fine particulate matter and the risk of spontaneous preterm birth: A population-based cohort study of twins. Frontiers in Public Health, 10. https://www.frontiersin.org/journals/public-health/articles/10.3389/fpubh.2022.1002824

Qiu, X., Fong, K. C., Shi, L., Papatheodorou, S., Di, Q., Just, A., Kosheleva, A., Messerlian, C. & Schwartz, J. D. 2020. Prenatal exposure to particulate air pollution and gestational age at delivery in Massachusetts neonates 2001–2015: A perspective of causal modeling and health disparities. Environmental Epidemiology, 4, 5. https://journals.lww.com/environepidem/fulltext/2020/10000/prenatal_exposure_to_particulate_air_pollution_and.3.aspx

Qureshi, R., Shaughnessy, D., Gill, K. a. R., Robinson, K. A., Li, T. & Agai, E. 2023. Are ChatGPT and large language models “the answer” to bringing us closer to systematic review automation? Systematic Reviews, 12, 1, 72. https://doi.org/10.1186/s13643-023-02243-z

R Core Team 2023. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria. https://www.R-project.org/
Rajagopalan, S., Brauer, M., Bhatnagar, A., Bhatt, D. L., Brook, J. R., Huang, W., Münzel, T., Newby, D., Siegel, J., Brook, R. D., On Behalf of the American Heart Association Council On, L., Cardiometabolic, H., Council on Arteriosclerosis, T., Vascular, B., Council on Clinical, C., Council On, C., Stroke, N. & Stroke, C. 2020. Personal-Level Protective Actions Against Particulate Matter Air Pollution Exposure: A Scientific Statement From the American Heart Association. Circulation, 142, 23, e411-e431. https://doi.org/10.1161/CIR.0000000000000931
Rappazzo, K. M., Daniels, J. L., Messer, L. C., Poole, C. & Lobdell, D. T. 2014. 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, 122, 9, 992-997. https://www.scopus.com/inward/record.uri?eid=2-s2.0-84906875843&doi=10.1289%2fehp.1307456&partnerID=40&md5=4ddb032b5121a8337bd86ebc0d25e41d
Rappazzo, K. M., Daniels, J. L., Messer, L. C., Poole, C. & Lobdell, D. T. 2015. Exposure to elemental carbon, organic carbon, nitrate, and sulfate fractions of fine particulate matter and risk of preterm birth in New Jersey, Ohio, and Pennsylvania (2000–2005). Environmental Health Perspectives, 123, 10, 1059-1065. https://www.scopus.com/inward/record.uri?eid=2-s2.0-84943279712&doi=10.1289%2fehp.1408953&partnerID=40&md5=73ad558bd4e353446ad06b6feea92277
Rappazzo, K. M., Lobdell, D. T., Messer, L. C., Poole, C. & Daniels, J. L. 2016. Comparison of gestational dating methods and implications for exposure-outcome associations: An example with PM2.5 and preterm birth. Occupational and Environmental Medicine. 74, 138-143. Available: https://doi.org/10.1136/oemed-2016-103833
Rappazzo, K. M., Nichols, J. L., Rice, R. B. & Luben, T. J. 2021. Ozone exposure during early pregnancy and preterm birth: A systematic review and meta-analysis. Environ Res, 198, 111317.
Raz, R., Kioumourtzoglou, M. A. & Weisskopf, M. G. 2018. Live-Birth Bias and Observed Associations Between Air Pollution and Autism. Am J Epidemiol, 187, 11, 2292-2296.
Reis, L. A., Drouet, L. & Tavoni, M. 2022. Internalising health-economic impacts of air pollution into climate policy: a global modelling study. The Lancet Planetary Health, 6, 1, e40-e48. https://doi.org/10.1016/S2542-5196(21)00259-X
Richardson, K., Steffen, W., Lucht, W., Bendtsen, J., Cornell, S. E., Donges, J. F., Drüke, M., Fetzer, I., Bala, G., Von Bloh, W., Feulner, G., Fiedler, S., Gerten, D., Gleeson, T., Hofmann, M., Huiskamp, W., Kummu, M., Mohan, C., Nogués-Bravo, D., Petri, S., Porkka, M., Rahmstorf, S., Schaphoff, S., Thonicke, K., Tobian, A., Virkki, V., Wang-Erlandsson, L., Weber, L. & Rockström, J. 2023. Earth beyond six of nine planetary boundaries. Science Advances, 9, 37, eadh2458. https://www.science.org/doi/abs/10.1126/sciadv.adh2458
Riddell, C. A., Goin, D. E., Morello-Frosch, R., Apte, J. S., Glymour, M. M., Torres, J. M. & Casey, J. A. 2021. Hyper-localized measures of air pollution and risk of preterm birth in Oakland and San Jose, California. International Journal of Epidemiology, 50, 6, 1875-1885. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85123461935&doi=10.1093%2fije%2fdyab097&partnerID=40&md5=043d295cc43910dd17f505dead245c70
Risnes, K., Bilsteen, J. F., Brown, P., Pulakka, A., Andersen, A.-M. N., Opdahl, S., Kajantie, E. & Sandin, S. 2021. Mortality Among Young Adults Born Preterm and Early Term in 4 Nordic Nations. JAMA Network Open, 4, 1, e2032779-e2032779. https://doi.org/10.1001/jamanetworkopen.2020.32779
Ritz, B., Wilhelm, M., Hoggatt, K. J. & Ghosh, J. K. C. 2007. Ambient Air Pollution and Preterm Birth in the Environment and Pregnancy Outcomes Study at the University of California, Los Angeles. American Journal of Epidemiology, 166, 9, 1045-1052. https://doi.org/10.1093/aje/kwm181
Ritz, B., Yu, F., Chapa, G. & Fruin, S. 2000. Effect of air pollution on preterm birth among children born in Southern California between 1989 and 1993. Epidemiology, 11, 5, 502-11.
Rivas, I., Fussell, J. C., Kelly, F. J. & Querol, X. 2019. Indoor Sources of Air Pollutants. *In:* Harrison, R. M. & Hester, R. E. (eds.) *Indoor Air Pollution.* The Royal Society of Chemistry. Available: https://doi.org/10.1039/9781788016179-00001
Robertson, S. A., Chin, P. Y., Femia, J. G. & Brown, H. M. 2018. Embryotoxic cytokines-Potential roles in embryo loss and fetal programming. J Reprod Immunol, 125, 80-88.

Robins, J. 1986. A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the healthy worker survivor effect. Mathematical Modelling, 7, 9, 1393-1512. https://www.sciencedirect.com/science/article/pii/0270025586900886
Robinson, L. A., Eber, M. R. & Hammitt, J. K. 2022. Valuing COVID-19 Morbidity Risk Reductions. J Benefit Cost Anal, 13, 2, 247-268.
Rohatgi, A. 2024. *WebPlotDigitizer (automeris.lo v5)* [Online]. Available: https://automeris.io/docs/
Rom, W. N., Boushey, H. & Caplan, A. 2013. Experimental human exposure to air pollutants is essential to understand adverse health effects. Am J Respir Cell Mol Biol, 49, 5, 691-6.
Romanello, M., Di Napoli, C., Drummond, P., Green, C., Kennard, H., Lampard, P., Scamman, D., Arnell, N., Ayeb-Karlsson, S., Ford, L. B., Belesova, K., Bowen, K., Cai, W., Callaghan, M., Campbell-Lendrum, D., Chambers, J., Van Daalen, K. R., Dalin, C., Dasandi, N., Dasgupta, S., Davies, M., Dominguez-Salas, P., Dubrow, R., Ebi, K. L., Eckelman, M., Ekins, P., Escobar, L. E., Georgeson, L., Graham, H., Gunther, S. H., Hamilton, I., Hang, Y., Hänninen, R., Hartinger, S., He, K., Hess, J. J., Hsu, S.-C., Jankin, S., Jamart, L., Jay, O., Kelman, I., Kiesewetter, G., Kinney, P., Kjellstrom, T., Kniveton, D., Lee, J. K. W., Lemke, B., Liu, Y., Liu, Z., Lott, M., Batista, M. L., Lowe, R., Macguire, F., Sewe, M. O., Martinez-Urtaza, J., Maslin, M., Mcallister, L., Mcgushin, A., Mcmichael, C., Mi, Z., Milner, J., Minor, K., Minx, J. C., Mohajeri, N., Moradi-Lakeh, M., Morrissey, K., Munzert, S., Murray, K. A., Neville, T., Nilsson, M., Obradovich, N., O'hare, M. B., Oreszczyn, T., Otto, M., Owfi, F., Pearman, O., Rabbaniha, M., Robinson, E. J. Z., Rocklöv, J., Salas, R. N., Semenza, J. C., Sherman, J. D., Shi, L., Shumake-Guillemot, J., Silbert, G., Sofiev, M., Springmann, M., Stowell, J., Tabatabaei, M., Taylor, J., Triñanes, J., Wagner, F., Wilkinson, P., Winning, M., Yglesias-González, M., Zhang, S., Gong, P., Montgomery, H. & Costello, A. 2022. The 2022 report of the Lancet Countdown on health and climate change: health at the mercy of fossil fuels. The Lancet, 400, 10363, 1619-1654. https://doi.org/10.1016/S0140-6736(22)01540-9
Rombout, P. J. A., Lioy, P. J. & Goldstein, B. D. 1986. Rationale for an Eight-Hour Ozone Standard. Journal of the Air Pollution Control Association, 36, 8, 913-917. https://doi.org/10.1080/00022470.1986.10466130
Rooney, A. A., Boyles, A. L., Wolfe, M. S., Bucher, J. R. & Thayer, K. A. 2014. Systematic review and evidence integration for literature-based environmental health science assessments. Environ Health Perspect, 122, 7, 711-8.
Rosser, F. J., Rothenberger, S. D., Han, Y. Y., Forno, E. & Celedón, J. C. 2023. Air Quality Index and Childhood Asthma: A Pilot Randomized Clinical Trial Intervention. Am J Prev Med, 64, 6, 893-897.
Royal College of Obstetricans & Gynaecologists (Rcog) 2013. Small-for-Gestational-Age Fetus, Investigation and Management (Green-top Guideline No. 31) Available: https://www.rcog.org.uk/guidance/browse-all-guidance/green-top-guidelines/small-for-gestational-age-fetus-investigation-and-management-green-top-guideline-no-31/
Royal College of Obstetricans & Gynaecologists (Rcog). 2022. *UK Government consultation on environmental targets: target proposals for air quality - RCOG response* [Online]. Available: https://www.rcog.org.uk/media/ujeb143x/rcog-response-dhsc-consultation-on-environmental-targets.pdf
Royal College of Physicians (Rcp) 2016. Every breath we take: the lifelong impact of air pollution. London: party, R. o. a. w. Available: https://rcp.ac.uk/media/jzul5jgn/every-breath-we-take-the-lifelong-impact-of-air-pollution-full-report.pdf
Ruan, T., Yue, Y., Lu, W., Zhou, R., Xiong, T., Jiang, Y., Ying, J., Tang, J., Shi, J., Wang, H., Xiao, G., Li, J., Qu, Y. & Mu, D. 2023. Association between low ambient temperature during pregnancy and adverse birth outcomes: A systematic review and meta-analysis. Chinese Medical Journal, 136, 19, 2307-2315.
Rudra, C. B., Williams, M. A., Sheppard, L., Koenig, J. Q. & Schiff, M. A. 2011. Ambient carbon monoxide and fine particulate matter in relation to preeclampsia and preterm delivery in western Washington State. Environ Health Perspect, 119, 6, 886-92.
Ruiz, M., Goldblatt, P., Morrison, J., Kukla, L., Švancara, J., Riitta-Järvelin, M., Taanila, A., Saurel-Cubizolles, M.-J., Lioret, S., Bakoula, C., Veltsista, A., Porta, D., Forastiere, F., Van Eijsden, M., Vrijkotte, T. G. M., Eggesbø, M., White, R. A., Barros, H., Correia, S., Vrijheid, M., Torrent, M., Rebagliato, M., Larrañaga, I., Ludvigsson, J., Olsen Faresjö, Å., Hryhorczuk, D., Antipkin, Y., Marmot, M. & Pikhart, H. 2015. Mother's education and the risk of preterm and small for gestational age birth: a DRIVERS meta-analysis of 12 European cohorts. Journal of Epidemiology and Community Health, 69, 9, 826-833. https://jech.bmj.com/content/jech/69/9/826.full.pdf

Ryan, P. H. & Lemasters, G. K. 2007. A review of land-use regression models for characterizing intraurban air pollution exposure. Inhal Toxicol, 19 Suppl 1, Suppl 1, 127-33.
Sacks J, Buckley B, Deflorio-Barker S & Al, E. 2022. Supplement to the 2019 Integrated Science Assessment for Particulate Matter. (1). Introduction and Scope. Research Triangle Park (NC). Available: https://www.ncbi.nlm.nih.gov/books/NBK588509/
Saigal, S. & Doyle, L. W. 2008. An overview of mortality and sequelae of preterm birth from infancy to adulthood. Lancet, 371, 9608, 261-9.
Saldanha, I. J., Smith, B. T., Ntzani, E., Jap, J., Balk, E. M. & Lau, J. 2019. The Systematic Review Data Repository (SRDR): descriptive characteristics of publicly available data and opportunities for research. Systematic Reviews, 8, 1, 334. https://doi.org/10.1186/s13643-019-1250-y
Saldiva, S. R. D. M., Barrozo, L. V., Leone, C. R., Failla, M. A., Bonilha, E. A., Bernal, R. T. I., De Oliveira, R. C. & Saldiva, P. H. N. 2018. Small-scale variations in urban air pollution levels are significantly associated with premature births: A case study in Sao Paulo, Brazil. International Journal of Environmental Research and Public Health, 15, 10, 2236. https://www.mdpi.com/1660-4601/15/10/2236/pdf
Salihu, H., Ghaji, N., Mbah, A., Alio, A., August, E. & Boubakari, I. 2012. Particulate Pollutants and Racial/Ethnic Disparity in Feto-Infant Morbidity Outcomes. Maternal & Child Health Journal, 16, 8, 1679-1687. https://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,shib,url,uid&db=cin20&AN=104442313&site=ehost-live
Samoli, E., Atkinson, R. W., Analitis, A., Fuller, G. W., Green, D. C., Mudway, I., Anderson, H. R. & Kelly, F. J. 2016. Associations of short-term exposure to traffic-related air pollution with cardiovascular and respiratory hospital admissions in London, UK. Occupational and Environmental Medicine, 73, 5, 300. http://oem.bmj.com/content/73/5/300.abstract
Samoli, E., Butland, B. K., Rodopoulou, S., Atkinson, R. W., Barratt, B., Beevers, S. D., Beddows, A., Dimakopoulou, K., Schwartz, J. D., Yazdi, M. D. & Katsouyanni, K. 2020. The impact of measurement error in modeled ambient particles exposures on health effect estimates in multilevel analysis: A simulation study. Environ Epidemiol, 4, 3, e094.
Sapkota, A., Chelikowsky, A. P., Nachman, K. E., Cohen, A. J. & Ritz, B. 2012. Exposure to particulate matter and adverse birth outcomes: a comprehensive review and meta-analysis. Air Quality, Atmosphere & Health, 5, 4, 369-381. https://doi.org/10.1007/s11869-010-0106-3
Savitz, D. A., Eliot, M. N., Ito, K., Johnson, S., Manjourides, J., Danilack, V. A. & Wellenius, G. A. 2019. Should we adjust for delivery hospital in studies of air pollution and pregnancy outcomes? Environ Epidemiol, 3, 5, e064.
Ščasný, M., Zvěřinová, I. & Dussaux, D. 2023. Valuing a reduction in the risk of very low birth weight. https://www.oecd-ilibrary.org/content/paper/dfd159a1-en
Schifano, P., Asta, F., Dadvand, P., Davoli, M., Basagana, X. & Michelozzi, P. 2016. Heat and air pollution exposure as triggers of delivery: A survival analysis of population-based pregnancy cohorts in Rome and Barcelona. Environment International, 88, 153-159. https://www.scopus.com/inward/record.uri?eid=2-s2.0-84952884161&doi=10.1016%2fj.envint.2015.12.013&partnerID=40&md5=bc24f4d272e4ad7ba02dd8beb705faa7
Schraufnagel, D. E. 2020. The health effects of ultrafine particles. Exp Mol Med, 52, 3, 311-317.
Scopus. 2024. *About Scopus* [Online]. Available: https://blog.scopus.com/about
Seabrook, J. A., Smith, A., Clark, A. F. & Gilliland, J. A. 2019. Geospatial analyses of adverse birth outcomes in Southwestern Ontario: Examining the impact of environmental factors. Environmental Research, 172, 18-26. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85061340758&doi=10.1016%2fj.envres.2018.12.068&partnerID=40&md5=76f58308cb09b39a8e73af9e77cb04c0
Selvaratnam, R. J., Sovio, U., Cook, E., Gaccioli, F., Charnock-Jones, D. S. & Smith, G. C. S. 2023. Objective measures of smoking and caffeine intake and the risk of adverse pregnancy outcomes. International Journal of Epidemiology, 52, 6, 1756-1765. Available: https://doi.org/10.1093/ije/dyad123
Sexton, K. & Ryan, P. 1988. *Assessment of Human Exposure to Air Pollution: Methods, Measurements, and Models,* Washington (DC), National Academies Press (US). Available: https://www.ncbi.nlm.nih.gov/books/NBK218147/

Shah, P. S., Balkhair, T. & Knowledge Synthesis Group on Determinants of Preterm/Lbw Births 2011. Air pollution and birth outcomes: a systematic review. Environ Int, 37, 2, 498-516.
Shao, J., Zosky, G. R., Hall, G. L., Wheeler, A. J., Dharmage, S., Melody, S., Dalton, M., Foong, R. E., O'sullivan, T., Williamson, G. J., Chappell, K., Abramson, M. J. & Johnston, F. H. 2020. Early life exposure to coal mine fire smoke emissions and altered lung function in young children. Respirology, 25, 2, 198-205.
Shepard, D. 1968. A two-dimensional interpolation function for irregularly-spaced data. *Proceedings of the 1968 23rd ACM national conference.* Association for Computing Machinery.
Sheridan, P., Ilango, S., Bruckner, T. A., Wang, Q., Basu, R. & Benmarhnia, T. 2019. Ambient Fine Particulate Matter and Preterm Birth in California: Identification of Critical Exposure Windows. American Journal of Epidemiology, 188, 9, 1608-1615. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85072058725&doi=10.1093%2faje%2fkwz120&partnerID=40&md5=bdb1eb013eaeb7d551e58655bb5d1f5c
Shezi, B., Jafta, N., Asharam, K., Tularam, H., Jeena, P. & Naidoo, R. N. 2022. Maternal exposure to indoor PM(2.5) and associated adverse birth outcomes in low socio-economic households, Durban, South Africa. Indoor Air, 32, 1, e12934.
Shi, T. S., Ma, H. P., Li, D. H., Pan, L., Wang, T. R., Li, R. & Ren, X. W. 2024. Prenatal exposure to PM2.5 components and the risk of different types of preterm birth and the mediating effect of pregnancy complications: a cohort study. Public Health, 227, 202-209. Available: https://www.sciencedirect.com/science/article/pii/S0033350623004869
Shi, W., Jiang, M., Kan, L., Zhang, T., Yu, Q., Wu, Z., Xue, S., Fei, X. & Jin, C. 2021. Association Between Ambient Air Pollutants Exposure and Preterm Birth in Women Who Underwent in vitro Fertilization: A Retrospective Cohort Study From Hangzhou, China. Frontiers in Medicine, 8. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85121809030&doi=10.3389%2ffmed.2021.785600&partnerID=40&md5=c4d3fa31d11ce76eb40dfee29141af9d
Shimonovich, M., Pearce, A., Thomson, H., Keyes, K. & Katikireddi, S. V. 2021. Assessing causality in epidemiology: revisiting Bradford Hill to incorporate developments in causal thinking. Eur J Epidemiol, 36, 9, 873-887.
Shojania, K. G., Sampson, M., Ansari, M. T., Ji, J., Doucette, S. & Moher, D. 2007. How quickly do systematic reviews go out of date? A survival analysis. Ann Intern Med, 147, 4, 224-33.
Siddika, N., Rantala, A. K., Antikainen, H., Balogun, H., Amegah, A. K., Ryti, N. R. I., Kukkonen, J., Sofiev, M., Jaakkola, M. S. & Jaakkola, J. J. K. 2019. 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. Environ Res, 176, 108549.
Siddika, N., Rantala, A. K., Antikainen, H., Balogun, H., Amegah, A. K., Ryti, N. R. I., Kukkonen, J., Sofiev, M., Jaakkola, M. S. & Jaakkola, J. J. K. 2020. Short-term prenatal exposure to ambient air pollution and risk of preterm birth - A population-based cohort study in Finland. Environmental Research, 184. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85080065657&doi=10.1016%2fj.envres.2020.109290&partnerID=40&md5=9d5377ded2fb8c74d5f5aca76db3df0b
Simoncic, V., Enaux, C., Deguen, S. & Kihal-Talantikite, W. 2020. Adverse Birth Outcomes Related to NO(2) and PM Exposure: European Systematic Review and Meta-Analysis. Int J Environ Res Public Health, 17, 21.
Singh, T., Green Donna, L., Meissner, K. & Jegasothy, E. 2024. Association of stillbirths and preterm birth with wildfires in Sydney, Australia between 2010 and 2020. ISEE Conference Abstracts, 2024, 2024, 1. Available: https://doi.org/10.1289/isee.2024.0938
Sinsamala, R. M., Johannessen, A., Bertelsen, R. J., Accordini, S., Brandt, J., Frohn, L. M., Geels, C., Gislason, T., Holm, M., Janson, C., Markevych, I., Orru, H., Real, F. G., Sigsgaard, T., Skulstad, S. M., Svanes, C. & Marcon, A. 2024. Pregnancy outcomes as related to in utero exposure to air pollution and greenness: The Life-GAP Project. Environmental Epidemiology, 8, 4. Available: https://journals.lww.com/environepidem/fulltext/2024/08000/pregnancy_outcomes_as_related_to_in_utero_exposure.3.aspx
Smith, L. H., Dollinger, C. Y., Vanderweele, T. J., Wyszynski, D. F. & Hernández-Díaz, S. 2022. Timing and severity of COVID-19 during pregnancy and risk of preterm birth in the International Registry of Coronavirus Exposure in Pregnancy. BMC Pregnancy Childbirth, 22, 1, 775.

Smith, R. B., Beevers, S. D., Gulliver, J., Dajnak, D., Fecht, D., Blangiardo, M., Douglass, M., Hansell, A. L., Anderson, H. R., Kelly, F. J. & Toledano, M. B. 2020. Impacts of air pollution and noise on risk of preterm birth and stillbirth in London. Environment International, 134, 105290. Available: https://www.sciencedirect.com/science/article/pii/S0160412019314734
Snowden, J. M., Rose, S. & Mortimer, K. M. 2011. Implementation of G-Computation on a Simulated Data Set: Demonstration of a Causal Inference Technique. American Journal of Epidemiology, 173, 7, 731-738. Available: https://doi.org/10.1093/aje/kwq472
Solazzo, E., Bianconi, R., Hogrefe, C., Curci, G., Tuccella, P., Alyuz, U., Balzarini, A., Barô, R., Bellasio, R., Bieser, J., Brandt, J., Christensen, J. H., Colette, A., Francis, X., Fraser, A., Vivanco, M. G., Jiménez-Guerrero, P., Im, U., Manders, A., Nopmongcol, U., Kitwiroon, N., Pirovano, G., Pozzoli, L., Prank, M., Sokhi, R. S., Unal, A., Yarwood, G. & Galmarini, S. 2017. Evaluation and error apportionment of an ensemble of atmospheric chemistry transport modeling systems: multivariable temporal and spatial breakdown. Atmos Chem Phys, 17, 4, 3001-3054.
Sørensen, M., Poulsen, A. H., Nøhr, B., Khan, J., Ketzel, M., Brandt, J., Raaschou-Nielsen, O. & Jensen, A. 2024. Long term exposure to road traffic noise and air pollution and risk of infertility in men and women: nationwide Danish cohort study. BMJ, 386, e080664. Available: https://www.bmj.com/content/bmj/386/bmj-2024-080664.full.pdf
Southerland, V. A., Brauer, M., Mohegh, A., Hammer, M. S., Van Donkelaar, A., Martin, R. V., Apte, J. S. & Anenberg, S. C. 2022a. Global urban temporal trends in fine particulate matter (PM2.5) and attributable health burdens: estimates from global datasets. The Lancet Planetary Health, 6, 2, e139-e146. Available: https://doi.org/10.1016/S2542-5196(21)00350-8
Southerland, V. A., Brauer, M., Mohegh, A., Hammer, M. S., Van Donkelaar, A., Martin, R. V., Apte, J. S. & Anenberg, S. C. 2022b. Global urban temporal trends in fine particulate matter (PM2.5) and attributable health burdens: estimates from global datasets. The Lancet Planetary Health, 6, 2, e139-e146. Available: https://doi.org/10.1016/S2542-5196(21)00350-8
Spruance, S. L., Reid, J. E., Grace, M. & Samore, M. 2004. Hazard ratio in clinical trials. Antimicrob Agents Chemother, 48, 8, 2787-92.
Stang, A. 2010. Critical evaluation of the Newcastle-Ottawa scale for the assessment of the quality of nonrandomized studies in meta-analyses. European Journal of Epidemiology, 25, 9, 603-605. Available: https://doi.org/10.1007/s10654-010-9491-z
Stern, P. C., Dietz, T. & Kalof, L. 1993. Value orientations, gender, and environmental concern. Environment and Behavior, 25, 3, 322-348.
Sterne, J. a. C., Sutton, A. J., Ioannidis, J. P. A., Terrin, N., Jones, D. R., Lau, J., Carpenter, J., Rücker, G., Harbord, R. M., Schmid, C. H., Tetzlaff, J., Deeks, J. J., Peters, J., Macaskill, P., Schwarzer, G., Duval, S., Altman, D. G., Moher, D. & Higgins, J. P. T. 2011. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. BMJ, 343, d4002. Available: https://www.bmj.com/content/bmj/343/bmj.d4002.full.pdf
Stewart, L., Moher, D. & Shekelle, P. 2012. Why prospective registration of systematic reviews makes sense. Syst Rev, 1, 7.
Stieb, D. M., Berjawi, R., Emode, M., Zheng, C., Salama, D., Hocking, R., Lyrette, N., Matz, C., Lavigne, E. & Shin, H. H. 2021. Systematic review and meta-analysis of cohort studies of long term outdoor nitrogen dioxide exposure and mortality. PLOS ONE, 16, 2, e0246451. Available: https://doi.org/10.1371/journal.pone.0246451
Stieb, D. M., Chen, L., Beckerman, B. S., Jerrett, M., Crouse, D. L., Omariba, D. W. R., P.A., P., Van Donkelaar, A., Martin, R. V., Burnett, R. T., Gilbert, N. L., Tjepkema, M., Liu, S. & Dugandzic, R. M. 2016a. Associations of Pregnancy Outcomes and PM2.5 in a National Canadian Study. Environmental Health Perspectives, 124, 2, 243-249. Available: https://ehp.niehs.nih.gov/doi/abs/10.1289/ehp.1408995
Stieb, D. M., Chen, L., Eshoul, M. & Judek, S. 2012. Ambient air pollution, birth weight and preterm birth: a systematic review and meta-analysis. Environ Res, 117, 100-11. Available: https://www.ncbi.nlm.nih.gov/pubmed/22726801
Stieb, D. M., Chen, L., Hystad, P., Beckerman, B. S., Jerrett, M., Tjepkema, M., Crouse, D. L., Omariba, D. W., Peters, P. A., Van Donkelaar, A., Martin, R. V., Burnett, R. T., Liu, S., Smith-Doiron, M. & Dugandzic, R. M. 2016b. A national study of the association between traffic-related air pollution and adverse pregnancy outcomes in Canada, 1999-2008. Environ Res, 148, 513-526.
Stock, S. J., Horne, M., Bruijn, M., White, H., Boyd, K. A., Heggie, R., Wotherspoon, L., Aucott, L., Morris, R. K., Dorling, J., Jackson, L., Chandiramani, M., David, A. L., Khalil, A., Shennan, A.,

Van Baaren, G.-J., Hodgetts-Morton, V., Lavender, T., Schuit, E., Harper-Clarke, S., Mol, B. W., Riley, R. D., Norman, J. E. & Norrie, J. 2021. Development and validation of a risk prediction model of preterm birth for women with preterm labour symptoms (the QUIDS study): A prospective cohort study and individual participant data meta-analysis. PLOS Medicine, 18, 7, e1003686. Available: https://doi.org/10.1371/journal.pmed.1003686

Strand, L. B., Barnett, A. G. & Tong, S. 2011. Maternal Exposure to Ambient Temperature and the Risks of Preterm Birth and Stillbirth in Brisbane, Australia. American Journal of Epidemiology, 175, 2, 99-107. Available: https://doi.org/10.1093/aje/kwr404

Suhaimi, N. F. & Jalaludin, J. 2015. Biomarker as a research tool in linking exposure to air particles and respiratory health. Biomed Res Int, 2015, 962853.

Sun, X., Luo, X., Zhao, C., Chung Ng, R. W., Lim, C. E. D., Zhang, B. & Liu, T. 2015. The association between fine particulate matter exposure during pregnancy and preterm birth: a meta-analysis. BMC Pregnancy and Childbirth, 15, 1, 300. Available: https://doi.org/10.1186/s12884-015-0738-2

Sun, Y., Ilango, S. D., Schwarz, L., Wang, Q., Chen, J. C., Lawrence, J. M., Wu, J. & Benmarhnia, T. 2020. Examining the joint effects of heatwaves, air pollution, and green space on the risk of preterm birth in California. Environ Res Lett, 15, 10.

Sun, Z., Yang, L., Bai, X., Du, W., Shen, G., Fei, J., Wang, Y., Chen, A., Chen, Y. & Zhao, M. 2019. Maternal ambient air pollution exposure with spatial-temporal variations and preterm birth risk assessment during 2013–2017 in Zhejiang Province, China. Environment International, 133, 105242. Available: https://www.sciencedirect.com/science/article/pii/S0160412019327783

Sutradhar, R. & Austin, P. C. 2018. Relative rates not relative risks: addressing a widespread misinterpretation of hazard ratios. Annals of Epidemiology, 28, 1, 54-57. Available: https://www.sciencedirect.com/science/article/pii/S1047279717305057

Swamy, G. K., Østbye, T. & Skjærven, R. 2008. Association of Preterm Birth With Long-term Survival, Reproduction, and Next-Generation Preterm Birth. JAMA, 299, 12, 1429-1436. Available: https://doi.org/10.1001/jama.299.12.1429

Symanski, E., Davila, M., Mchugh, M. K., Waller, D. K., Zhang, X. & Lai, D. 2014. Maternal exposure to fine particulate pollution during narrow gestational periods and newborn health in harris county, Texas. Maternal and Child Health Journal, 18, 8, 2003-2012. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84907166735&doi=10.1007%2fs10995-014-1446-7&partnerID=40&md5=ea9c7e8fe30acb344af5ac8d95d5a028

Symanski, E., Mchugh, M. K., Zhang, X., Craft, E. S. & Lai, D. 2016. Evaluating narrow windows of maternal exposure to ozone and preterm birth in a large urban area in Southeast Texas. Journal of Exposure Science and Environmental Epidemiology, 26, 2, 167-172. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84958661462&doi=10.1038%2fjes.2015.32&partnerID=40&md5=4cff5ae9d4801b7392d1c67416d9b5e0

Tanriver-Ayder, E., Faes, C., Van De Casteele, T., Mccann, S. K. & Macleod, M. R. 2021. Comparison of commonly used methods in random effects meta-analysis: application to preclinical data in drug discovery research. BMJ Open Sci, 5, 1, e100074.

Tapia, V. L., Vasquez, B. V., Vu, B., Liu, Y., Steenland, K. & Gonzales, G. F. 2020. Association between maternal exposure to particulate matter (PM2.5) and adverse pregnancy outcomes in Lima, Peru. Journal of Exposure Science & Environmental Epidemiology, 30, 4, 689-697. Available: https://doi.org/10.1038/s41370-020-0223-5

Thangavel, P., Park, D. & Lee, Y. C. 2022. Recent Insights into Particulate Matter (PM(2.5))-Mediated Toxicity in Humans: An Overview. Int J Environ Res Public Health, 19, 12.

Thayamballi, N., Habiba, S., Laribi, O. & Ebisu, K. 2021. Impact of Maternal Demographic and Socioeconomic Factors on the Association Between Particulate Matter and Adverse Birth Outcomes: a Systematic Review and Meta-analysis. J Racial Ethn Health Disparities, 8, 3, 743-755.

Thompson, R., Stewart, G., Vu, T., Jephcote, C., Lim, S., Barratt, B., Smith, R. B., Karim, Y. B., Mussa, A., Mudway, I., Fisher, H. L., Dumontheil, I., Thomas, M. S. C., Gulliver, J., Beevers, S., Kelly, F. J. & Toledano, M. B. 2024. Air pollution, traffic noise, mental health, and cognitive development: A multi-exposure longitudinal study of London adolescents in the SCAMP cohort. Environment International, 191, 108963. Available: https://www.sciencedirect.com/science/article/pii/S016041202400549X

Thorlund, K., Imberger, G., Johnston, B. C., Walsh, M., Awad, T., Thabane, L., Gluud, C., Devereaux, P. J. & Wetterslev, J. 2012. Evolution of heterogeneity (I2) estimates and their 95% confidence intervals in large meta-analyses. PLoS One, 7, 7, e39471.
Thurston, G. D., Kipen, H., Annesi-Maesano, I., Balmes, J., Brook, R. D., Cromar, K., De Matteis, S., Forastiere, F., Forsberg, B., Frampton, M. W., Grigg, J., Heederik, D., Kelly, F. J., Kuenzli, N., Laumbach, R., Peters, A., Rajagopalan, S. T., Rich, D., Ritz, B., Samet, J. M., Sandstrom, T., Sigsgaard, T., Sunyer, J. & Brunekreef, B. 2017. A joint ERS/ATS policy statement: what constitutes an adverse health effect of air pollution? An analytical framework. European Respiratory Journal, 49, 1, 1600419. Available: https://publications.ersnet.org//content/erj/49/1/1600419.full.pdf
Tran, V. V., Park, D. & Lee, Y. C. 2020. Indoor Air Pollution, Related Human Diseases, and Recent Trends in the Control and Improvement of Indoor Air Quality. Int J Environ Res Public Health, 17, 8.
Trasande, L., Malecha, P. & Attina, T. M. 2016. Particulate matter exposure and preterm birth: Estimates of U.S. attributable burden and economic costs. Environmental Health Perspectives, 124, 12, 1913-1918. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85002584202&doi=10.1289%2fehp.1510810&partnerID=40&md5=a049b325d4b41ae5c3a9f083e3fc69b2
Trasande, L., Wong, K., Roy, A., Savitz, D. A. & Thurston, G. 2013. Exploring prenatal outdoor air pollution, birth outcomes and neonatal health care utilization in a nationally representative sample. Journal of Exposure Science and Environmental Epidemiology, 23, 3, 315-321. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84876778058&doi=10.1038%2fjes.2012.124&partnerID=40&md5=5cb79acf4298b92228ed7ecc746c2af5
Tripathy, S., Marsland, A. L., Kinnee, E. J., Tunno, B. J., Manuck, S. B., Gianaros, P. J. & Clougherty, J. E. 2021. Long-Term Ambient Air Pollution Exposures and Circulating and Stimulated Inflammatory Mediators in a Cohort of Midlife Adults. Environ Health Perspect, 129, 5, 57007.
Tu, J. & Tu, W. 2018. How the relationships between preterm birth and ambient air pollution vary over space: A case study in Georgia, USA using geographically weighted logistic regression. Applied Geography, 92, 31-40. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85041460409&doi=10.1016%2fj.apgeog.2018.01.007&partnerID=40&md5=4ff3b21bf6e08437363b973e0345b939
Turnock, S. T., Reddington, C. L., West, J. J. & O'connor, F. M. 2023. The Air Pollution Human Health Burden in Different Future Scenarios That Involve the Mitigation of Near-Term Climate Forcers, Climate and Land-Use. Geohealth, 7, 8, e2023GH000812.
UK Department of Transport (UK DoT). 2018. Transport Analysis Guidance (TAG) data book. Available: https://www.gov.uk/government/publications/tag-data-book
UK Government. 2019. *The Buckinghamshire (Structural Changes) Order 2019* [Online]. Available: https://www.legislation.gov.uk/uksi/2019/957/contents/made
UK National Atmospheric Inventory (UK NAEI). 2024. *2019 Concentrations* [Online]. Available: https://naei.beis.gov.uk/
Ulziikhuu, B., Gombojav, E., Banzrai, C., Batsukh, S., Enkhtuya, E., Boldbaatar, B., Bellinger, D. C., Lanphear, B. P., Mccandless, L. C., Tamana, S. K. & Allen, R. W. 2022. Portable HEPA Filter Air Cleaner Use during Pregnancy and Children's Cognitive Performance at Four Years of Age: The UGAAR Randomized Controlled Trial. Environmental Health Perspectives, 130, 6, 067006. Available: https://ehp.niehs.nih.gov/doi/abs/10.1289/EHP10302
United Nations Framework Convention on Climate Change (UNFCC) 2015. Paris Agreement.
United States Environmental Protection Agency (USEPA). 2019. Integrated Science Assessment for Particulate Matter. Research Triangle Park, North Carolina. Available: http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=216546
United States Environmental Protection Agency (USEPA). 2013. Integrated Science Assessment of Ozone and Related Photochemical Oxidants (Final Report). Research Triangle Park, North Carolina. Available: http://cfpub.epa.gov/ncea/isa/recordisplay.cfm?deid=247492
United States Environmental Protection Agency (USEPA). 2016. Integrated Science Assessment (ISA) for Oxides of Nitrogen - Health Criteria. Washington, DC. Available: https://cfpub.epa.gov/ncea/isa/recordisplay.cfm?deid=310879

United States Environmental Protection Agency (USEPA) 2020. Integrated Science Assessment (ISA) for Ozone and Related Photochemical Oxidants. Washington, DC. Available: https://cfpub.epa.gov/ncea/isa/recordisplay.cfm?deid=348522
Vadillo-Ortega, F., Osornio-Vargas, A., Buxton, M. A., Sánchez, B. N., Rojas-Bracho, L., Viveros-Alcaráz, M., Castillo-Castrejón, M., Beltrán-Montoya, J., Brown, D. G. & O'neill, M. S. 2014. Air pollution, inflammation and preterm birth: a potential mechanistic link. Med Hypotheses, 82, 2, 219-24.
van Den Brekel, L., Lenters, V., Mackenbach, J. D., Hoek, G., Wagtendonk, A., Lakerveld, J., Grobbee, D. E. & Vaartjes, I. 2024. Ethnic and socioeconomic inequalities in air pollution exposure: a cross-sectional analysis of nationwide individual-level data from the Netherlands. The Lancet Planetary Health, 8, 1, e18-e29. Available: https://doi.org/10.1016/S2542-5196(23)00258-9
van Den Hooven, E. H., Pierik, F. H., De Kluizenaar, Y., Willemsen, S. P., Hofman, A., Van Ratingen, S. W., Zandveld, P. Y. J., Mackenbach, J. P., Steegers, E. a. P., Miedema, H. M. E. & Jaddoe, V. W. V. 2012. Air pollution exposure during pregnancy, ultrasound measures of fetal growth, and adverse birth outcomes: A prospective cohort study. Environmental Health Perspectives, 120, 1, 150-156. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84855218506&doi=10.1289%2fehp.1003316&partnerID=40&md5=9ee4d856e97dfb8f0bcad2a5be568513
Van Dijk, S. H. B., Brusse-Keizer, M. G. J., Bucsán, C. C., Van Der Palen, J., Doggen, C. J. M. & Lenferink, A. 2023. Artificial intelligence in systematic reviews: promising when appropriately used. BMJ Open, 13, 7, e072254. Available: https://bmjopen.bmj.com/content/bmjopen/13/7/e072254.full.pdf
Van Donkelaar, A., Martin, R. V., Brauer, M., Kahn, R., Levy, R., Verduzco, C. & Villeneuve, P. J. 2010. Global estimates of ambient fine particulate matter concentrations from satellite-based aerosol optical depth: development and application. Environ Health Perspect, 118, 6, 847-55.
Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L. & Polosukhin, I. 2017. Attention Is All You Need. CoRR, abs/1706.03762. Available: http://arxiv.org/abs/1706.03762
Veber, T., Dahal, U., Lang, K., Orru, K. & Orru, H. 2022. Industrial Air Pollution Leads to Adverse Birth Outcomes: A Systematized Review of Different Exposure Metrics and Health Effects in Newborns. Public Health Rev, 43, 1604775.
Veras, M. M., Damaceno-Rodrigues, N. R., Caldini, E. G., Maciel Ribeiro, A. A., Mayhew, T. M., Saldiva, P. H. & Dolhnikoff, M. 2008. Particulate urban air pollution affects the functional morphology of mouse placenta. Biol Reprod, 79, 3, 578-84.
Villeneuve, P. J., Lam, S., Tjepkema, M., Pinault, L., Crouse, D. L., Osornio-Vargas, A. R., Hystad, P., Jerrett, M., Lavigne, E. & Stieb, D. M. 2022. Residential proximity to greenness and adverse birth outcomes in urban areas: Findings from a national Canadian population-based study. Environmental Research, 204.
Viswanathan, M., Berkman, N. D., Dryden, D. M. & Hartling, L. 2013. AHRQ Methods for Effective Health Care. *Assessing Risk of Bias and Confounding in Observational Studies of Interventions or Exposures: Further Development of the RTI Item Bank.* Rockville (MD): Agency for Healthcare Research and Quality (US).
Vos, T., Lim, S. S., Abbafati, C., Abbas, K. M., Abbasi, M., Abbasifard, M., Abbasi-Kangevari, M., Abbastabar, H., Abd-Allah, F., Abdelalim, A., Abdollahi, M., Abdollahpour, I., Abolhassani, H., Aboyans, V., Abrams, E. M., Abreu, L. G., Abrigo, M. R. M., Abu-Raddad, L. J., Abushouk, A. I., Acebedo, A., Ackerman, I. N., Adabi, M., Adamu, A. A., Adebayo, O. M., Adekanmbi, V., Adelson, J. D., Adetokunboh, O. O., Adham, D., Afshari, M., Afshin, A., Agardh, E. E., Agarwal, G., Agesa, K. M., Aghaali, M., Aghamir, S. M. K., Agrawal, A., Ahmad, T., Ahmadi, A., Ahmadi, M., Ahmadieh, H., Ahmadpour, E., Akalu, T. Y., Akinyemi, R. O., Akinyemiju, T., Akombi, B., Al-Aly, Z., Alam, K., Alam, N., Alam, S., Alam, T., Alanzi, T. M., Albertson, S. B., Alcalde-Rabanal, J. E., Alema, N. M., Ali, M., Ali, S., Alicandro, G., Alijanzadeh, M., Alinia, C., Alipour, V., Aljunid, S. M., Alla, F., Allebeck, P., Almasi-Hashiani, A., Alonso, J., Al-Raddadi, R. M., Altirkawi, K. A., Alvis-Guzman, N., Alvis-Zakzuk, N. J., Amini, S., Amini-Rarani, M., Aminorroaya, A., Amiri, F., Amit, A. M. L., Amugsi, D. A., Amul, G. G. H., Anderlini, D., Andrei, C. L., Andrei, T., Anjomshoa, M., Ansari, F., Ansari, I., Ansari-Moghaddam, A., Antonio, C. a. T., Antony, C. M., Antriyandarti, E., Anvari, D., Anwer, R., Arabloo, J., Arab-Zozani, M., Aravkin, A. Y., Ariani, F., Ärnlöv, J., Aryal, K. K., Arzani, A., Asadi-Aliabadi, M., Asadi-Pooya, A. A., Asghari, B., Ashbaugh, C., Atnafu, D. D., et al. 2020. Global burden of 369 diseases

and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. The Lancet, 396, 10258, 1204-1222. Available: https://doi.org/10.1016/S0140-6736(20)30925-9
Vu, A. T., Taylor, K. M., Holman, M. R., Ding, Y. S., Hearn, B. & Watson, C. H. 2015. Polycyclic Aromatic Hydrocarbons in the Mainstream Smoke of Popular U.S. Cigarettes. Chem Res Toxicol, 28, 8, 1616-26.
Waitzman, N. J., Jalali, A. & Grosse, S. D. 2021. Preterm birth lifetime costs in the United States in 2016: An update. Semin Perinatol, 45, 3, 151390.
Walani, S. R. 2020. Global burden of preterm birth. Int J Gynaecol Obstet, 150, 1, 31-33. Available: https://www.ncbi.nlm.nih.gov/pubmed/32524596
Wald, N. J., Idle, M., Boreham, J. & Bailey, A. 1981. Carbon monoxide in breath in relation to smoking and carboxyhaemoglobin levels. Thorax, 36, 5, 366-9.
Wallace, M. E., Grantz, K. L., Liu, D., Zhu, Y., Kim, S. S. & Mendola, P. 2016. Exposure to ambient air pollution and premature rupture of membranes. American Journal of Epidemiology, 183, 12, 1114-1121. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84992046871&doi=10.1093%2faje%2fkwv284&partnerID=40&md5=31466ed6ee4052795db633f2d140b600
Wallace, S. S., Barak, G., Truong, G. & Parker, M. W. 2022. Hierarchy of Evidence Within the Medical Literature. Hospital Pediatrics, 12, 8, 745-750. Available: https://doi.org/10.1542/hpeds.2022-006690
Walton, H., Dajnak, D., Holland, M., Evangelopoulos, D., Wood, D., Brand, C., Assareh, N., Stewart, G., Beddows, A., Lee, S., Fecht, D., Liu, Y., Davies, B., Goodman, A., Vu, T. & Beevers, S. forthcoming. Health and Associated Economic Benefits of Reduced Air Pollution and Increased Physical Activity from Climate Change Policies. Available: http://dx.doi.org/10.2139/ssrn.4930508
Wang, L., Fang, L., Fang, Z., Zhang, M. & Zhang, L. 2022. Assessment of the association between prenatal exposure to multiple ambient pollutants and preterm birth: A prospective cohort study in Jinan, east China. Ecotoxicology and Environmental Safety, 232. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85124131767&doi=10.1016%2fj.ecoenv.2022.113297&partnerID=40&md5=0d9d6a3b2d0d0d1be5a56735b2bafd0f
Wang, Q., Benmarhnia, T., Zhang, H., Knibbs, L. D., Sheridan, P., Li, C., Bao, J., Ren, M., Wang, S., He, Y., Zhang, Y., Zhao, Q. & Huang, C. 2018. Identifying windows of susceptibility for maternal exposure to ambient air pollution and preterm birth. Environment International, 121, 317-324. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85053396191&doi=10.1016%2fj.envint.2018.09.021&partnerID=40&md5=284219481e82be5b3814e24d9e201b33
Wang, Q., Li, B., Benmarhnia, T., Hajat, S., Ren, M., Liu, T., Knibbs, L. D., Zhang, H., Bao, J., Zhang, Y., Zhao, Q. & Huang, C. 2020. Independent and combined effects of heatwaves and PM2:5 on preterm birth in Guangzhou, China: A survival analysis. Environmental Health Perspectives, 128, 1. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85077522315&doi=10.1289%2fEHP5117&partnerID=40&md5=6ad5bd03b00c4c133daffd3b71185423
Wang, X., Lin, Y., Ge, Y., Craig, E., Liu, X., Miller, R. K., Thurston, S. W., Brunner, J., Barrett, E. S., O'connor, T. G., Rich, D. Q. & Zhang, J. 2024. Systemic oxidative stress levels during the course of pregnancy: Associations with exposure to air pollutants. Environmental Pollution, 357, 124463. Available: https://www.sciencedirect.com/science/article/pii/S0269749124011771
Wang, X., Wang, X., Gao, C., Xu, X., Li, L., Liu, Y., Li, Z., Xia, Y. & Fang, X. 2023. Relationship Between Outdoor Air Pollutant Exposure and Premature Delivery in China- Systematic Review and Meta-Analysis. Int J Public Health, 68, 1606226.
Warren, J., Fuentes, M., Herring, A. & Langlois, P. 2012. Spatial-Temporal Modeling of the Association between Air Pollution Exposure and Preterm Birth: Identifying Critical Windows of Exposure. Biometrics, 68, 4, 1157-1167. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84871665322&doi=10.1111%2fj.1541-0420.2012.01774.x&partnerID=40&md5=94183f33741f1bb533521f90c907005d
Warren, J. L., Son, J.-Y., Pereira, G., Leaderer, B. P. & Bell, M. L. 2017. Investigating the Impact of Maternal Residential Mobility on Identifying Critical Windows of Susceptibility to Ambient Air Pollution During Pregnancy. American Journal of Epidemiology, 187, 5, 992-1000. Available: https://doi.org/10.1093/aje/kwx335

Weber, K. A., Yang, W., Lurmann, F., Hammond, S. K., Shaw, G. M. & Padula, A. M. 2019. Air Pollution, Maternal Hypertensive Disorders, and Preterm Birth. Environ Epidemiol, 3, 5.
Wells, G., Shea, B., O'connell, D., Peterson, J., Welch, V., Losos, M. & Tugwell, P. 2013. *The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses* [Online]. Available: https://www.ohri.ca/programs/clinical_epidemiology/oxford.asp
Whitehead, S. J. & Ali, S. 2010. Health outcomes in economic evaluation: the QALY and utilities. Br Med Bull, 96, 5-21.
Wilhelm, M., Ghosh, J. K., Su, J., Cockburn, M., Jerrett, M. & Ritz, B. 2011. Traffic-related air toxics and preterm birth: a population-based case-control study in Los Angeles County, California. Environ Health, 10, 89.
Wilhelm, M. & Ritz, B. 2003. Residential proximity to traffic and adverse birth outcomes in Los Angeles County, California, 1994-1996. Environmental Health Perspectives, 111, 2, 207-216. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-0037293468&doi=10.1289%2fehp.5688&partnerID=40&md5=12fe042acf4dcefda3fe19d0c46f884f
Wilhelm, M. & Ritz, B. 2005. Local Variations in CO and Particulate Air Pollution and Adverse Birth Outcomes in Los Angeles County, California, USA. Environmental Health Perspectives, 113, 9, 1212-1221. Available: https://doi.org/10.1289/ehp.7751
Williams, A. D., Kanner, J., Grantz, K. L., Ouidir, M., Sheehy, S., Sherman, S., Robledo, C. & Mendola, P. 2021. Air pollution exposure and risk of adverse obstetric and neonatal outcomes among women with type 1 diabetes. Environmental Research, 197, 111152. https://www.sciencedirect.com/science/article/pii/S0013935121004461
Williams, M. L., Lott, M. C., Kitwiroon, N., Dajnak, D., Walton, H., Holland, M., Pye, S., Fecht, D., Toledano, M. B. & Beevers, S. D. 2018. The Lancet Countdown on health benefits from the UK Climate Change Act: a modelling study for Great Britain. Lancet Planet Health, 2, 5, e202-e213.
Williams, R. S., Hegglin, M. I., Kerridge, B. J., Jöckel, P., Latter, B. G. & Plummer, D. A. 2019. Characterising the seasonal and geographical variability in tropospheric ozone, stratospheric influence and recent changes. Atmos. Chem. Phys., 19, 6, 3589-3620. Available: https://acp.copernicus.org/articles/19/3589/2019/
Willis, G. A., Chappell, K., Williams, S., Melody, S. M., Wheeler, A., Dalton, M., Dharmage, S. C., Zosky, G. R. & Johnston, F. H. 2020. Respiratory and atopic conditions in children two to four years after the 2014 Hazelwood coalmine fire. Med J Aust, 213, 6, 269-275.
Wing, S. E., Larson, T. V., Hudda, N., Boonyarattaphan, S., Del Rosario, I., Fruin, S. & Ritz, B. 2022. Aircraft noise and vehicle traffic-related air pollution interact to affect preterm birth risk in Los Angeles, California. Science of the Total Environment, 829. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85126846155&doi=10.1016%2fj.scitotenv.2022.154678&partnerID=40&md5=165a0733150eeb191d7ff7a460e6fd99
Winters, M., De Vos, R.-J., Van Middelkoop, M., Rathleff, M. S. & Weir, A. 2021. Stay alive! What are living systematic reviews and what are their advantages and challenges? British Journal of Sports Medicine, 55, 10, 519-520. Available: https://bjsm.bmj.com/content/bjsports/55/10/519.full.pdf
Wittig, V. E., Ainsworth, E. A. & Long, S. P. 2007. To what extent do current and projected increases in surface ozone affect photosynthesis and stomatal conductance of trees? A meta-analytic review of the last 3 decades of experiments. Plant Cell Environ, 30, 9, 1150-62.
Wong, A., Lou, W., Ho, K.-F., Yiu, B. K.-F., Lin, S., Chu, W. C.-W., Abrigo, J., Lee, D., Lam, B. Y.-K., Au, L. W.-C., Soo, Y. O.-Y., Lau, A. Y.-L., Kwok, T. C.-Y., Leung, T. W.-H., Lam, L. C.-W., Ho, K. & Mok, V. C.-T. 2020. Indoor incense burning impacts cognitive functions and brain functional connectivity in community older adults. Scientific Reports, 10, 1, 7090. Available: https://doi.org/10.1038/s41598-020-63568-6
World Health Organization for Europe (WHO). 2013a. HRAPIE project: recommendations for concentration–response functions for cost–benefit analysis of particulate matter, ozone and nitrogen dioxide. Available: https://www.euro.who.int/__data/assets/pdf_file/0006/238956/Health_risks_air_pollution_HRAPIE_project.pdf
World Health Organization for Europe (WHO). 2013b. Review of evidence on health aspects of air pollution – REVIHAAP project: final technical report. Available: http://www.euro.who.int/__data/assets/pdf_file/0004/193108/REVIHAAP-Final-technical-report-final-version.pdf.

World Health Organization (WHO). 2020. Risk of bias assessment instrument for systematic reviews informing WHO Global Air Quality Guidelines. Copenhagen: Regional Office for Europe. Available: https://apps.who.int/iris/handle/10665/341717
World Health Organization (WHO). 2021. WHO global air quality guidelines. Particulate matter (PM 2.5 and PM 10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide. Geneva. Available: https://iris.who.int/bitstream/handle/10665/345329/9789240034228-eng.pdf?sequence=1
World Health Organization (WHO). 2023a. 2023 WHO review of health in Nationally Determined Contributions and long-term strategies: health at the heart of the Paris Agreement. Geneva. Available: https://iris.who.int/bitstream/handle/10665/372276/9789240074729-eng.pdf
World Health Organization (WHO) 2023b. Overview of methods to assess population exposure to ambient air pollution. Geneva. Available: https://iris.who.int/bitstream/handle/10665/373014/9789240073494-eng.pdf?sequence=1
World Health Organization (WHO). 2023c. *Preterm birth* [Online]. Geneva. Available: https://www.who.int/news-room/fact-sheets/detail/preterm-birth
World Health Organization (WHO). 2024. *WHO regional offices* [Online]. Available: https://www.who.int/about/who-we-are/regional-offices
World Health Organization (WHO) for Europe 2010. WHO Guidelines for Indoor Air Pollutants: Selected pollutants. Available: https://www.ncbi.nlm.nih.gov/books/NBK138705/pdf/Bookshelf_NBK138705.pdf
Wu, J., Jiang, C., Jaimes, G., Bartell, S., Dang, A., Baker, D. & Delfino, R. J. 2013. Travel patterns during pregnancy: comparison between Global Positioning System (GPS) tracking and questionnaire data. Environ Health, 12, 1, 86.
Wu, J., Laurent, O., Li, L., Hu, J. & Kleeman, M. 2016. Adverse Reproductive Health Outcomes and Exposure to Gaseous and Particulate-Matter Air Pollution in Pregnant Women. Res Rep Health Eff Inst, 2016, 188, 1-58.
Wu, J., Ren, C., Delfino, R. J., Chung, J., Wilhelm, M. & Ritz, B. 2009. Association between local traffic-generated air pollution and preeclampsia and preterm delivery in the South Coast Air Basin of california. Environmental Health Perspectives, 117, 11, 1773-1779. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-74949136219&doi=10.1289%2fehp.0800334&partnerID=40&md5=fddbb5951dec773bb12de71218ae356e
Wu, J., Wilhelm, M., Chung, J. & Ritz, B. 2011. Comparing exposure assessment methods for traffic-related air pollution in an adverse pregnancy outcome study. Environmental Research, 111, 5, 685-692. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-79958253104&doi=10.1016%2fj.envres.2011.03.008&partnerID=40&md5=54c43ea4c400fb533ef7de2373a42131
Xiao, Q. Y., Chen, H. Y., Strickland, M. J., Kan, H. D., Chang, H. H., Klein, M., Yang, C., Meng, X. & Liu, Y. 2018. Associations between birth outcomes and maternal PM2.5 exposure in Shanghai: A comparison of three exposure assessment approaches. Environment International, 117, 226-236. Available: https://www.sciencedirect.com/science/article/pii/S0160412018303921
Xu, X., Ding, H. & Wang, X. 1995. Acute effects of total suspended particles and sulfur dioxides on preterm delivery: a community-based cohort study. Arch Environ Health, 50, 6, 407-15.
Xun, Y., Guo, Q., Ren, M., Liu, Y., Sun, Y., Wu, S., Lan, H., Zhang, J., Liu, H., Wang, J., Shi, Q., Wang, Q., Wang, P., Chen, Y., Shao, R. & Xu, D. R. 2023. Characteristics of the sources, evaluation, and grading of the certainty of evidence in systematic reviews in public health: A methodological study. Front Public Health, 11, 998588.
Yang, B. Y., Fan, S., Thiering, E., Seissler, J., Nowak, D., Dong, G. H. & Heinrich, J. 2020a. Ambient air pollution and diabetes: A systematic review and meta-analysis. Environ Res, 180, 108817.
Yang, L., Xie, G., Yang, W., Wang, R., Zhang, B., Xu, M., Sun, L., Xu, X., Xiang, W., Cui, X., Luo, Y. & Chung, M. C. 2023a. Short-term effects of air pollution exposure on the risk of preterm birth in Xi'an, China. Ann Med, 55, 1, 325-334.
Yang, L., Zhao, S., Zhai, Y., Gao, S., Wang, F., Lian, Z., Duanmu, L., Zhang, Y., Zhou, X., Cao, B., Wang, Z., Yan, H., Zhang, H., Arens, E. & De Dear, R. 2023b. The Chinese thermal comfort dataset. Sci Data, 10, 1, 662.
Yang, Y., Liang, Z., Ruan, Z., Zhang, S., Zhao, Q. & Lin, H. 2020b. Estimating the attributable burden of preterm birth and low birth weight due to maternal ozone exposure in nine Chinese cities. Atmospheric Environment, 222. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-

85075858158&doi=10.1016%2fj.atmosenv.2019.117169&partnerID=40&md5=2ed94bfc5fb80ac20913201b77dea5d4
Ye, L., Ji, Y., Lv, W., Zhu, Y., Lu, C., Xu, B. & Xia, Y. 2018. Associations between maternal exposure to air pollution and birth outcomes: a retrospective cohort study in Taizhou, China. Environmental Science and Pollution Research, 25, 22, 21927-21936. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85047293716&doi=10.1007%2fs11356-018-1944-z&partnerID=40&md5=271f1c93b3402a1b301f22aa55b3069b
Yi, L., Xu, Y., Eckel, S. P., O'connor, S., Cabison, J., Rosales, M., Chu, D., Chavez, T. A., Johnson, M., Mason, T. B., Bastain, T. M., Breton, C. V., Dunton, G. F., Wilson, J. P. & Habre, R. 2022. Time-activity and daily mobility patterns during pregnancy and early postpartum - evidence from the MADRES cohort. Spat Spatiotemporal Epidemiol, 41, 100502.
You, R., Ho, Y. S. & Chang, R. C. 2022. The pathogenic effects of particulate matter on neurodegeneration: a review. J Biomed Sci, 29, 1, 15.
Younes, S., Samara, M., Salama, N., Al-Jurf, R., Nasrallah, G., Al-Obaidly, S., Salama, H., Olukade, T., Hammuda, S., Abdoh, G., Abdulrouf, P. V., Farrell, T., Alqubaisi, M., Al Rifai, H. & Al-Dewik, N. 2021. Incidence, risk factors, and feto-maternal outcomes of inappropriate birth weight for gestational age among singleton live births in Qatar: A population-based study. PLoS One, 16, 10, e0258967.
Younger, A., Alkon, A., Harknett, K., Jean Louis, R. & Thompson, L. M. 2022. Adverse birth outcomes associated with household air pollution from unclean cooking fuels in low- and middle-income countries: A systematic review. Environmental Research, 204. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85118761796&doi=10.1016%2fj.envres.2021.112274&partnerID=40&md5=3da494a441e43cc1c498b990c3c39200
Yu, P., Guo, S., Xu, R., Ye, T., Li, S., Sim, M. R., Abramson, M. J. & Guo, Y. 2021. Cohort studies of long-term exposure to outdoor particulate matter and risks of cancer: A systematic review and meta-analysis. Innovation (Camb), 2, 3, 100143.
Yu, S., Zhang, M., Zhu, J., Yang, X., Bigambo, F. M., Snijders, A. M., Wang, X., Hu, W., Lv, W. & Xia, Y. 2023a. The effect of ambient ozone exposure on three types of diabetes: a meta-analysis. Environ Health, 22, 1, 32.
Yu, W., Ye, T., Zhang, Y., Xu, R., Lei, Y., Chen, Z., Yang, Z., Zhang, Y., Song, J., Yue, X., Li, S. & Guo, Y. 2023b. Global estimates of daily ambient fine particulate matter concentrations and unequal spatiotemporal distribution of population exposure: a machine learning modelling study. The Lancet Planetary Health, 7, 3, e209-e218. Available: https://doi.org/10.1016/S2542-5196(23)00008-6
Yu, Y., Lin, H., Liu, Q., Ma, Y., Zhao, L., Li, W., Zhou, Y., Byun, H.-M., Li, P., Li, C., Sun, C., Chen, X., Liu, Z., Dong, W., Chen, L., Deng, F., Wu, S., Hou, S. & Guo, L. 2024. Association of residential greenness, air pollution with adverse birth outcomes: Results from 61,762 mother-neonatal pairs in project ELEFANT (2011–2021). Science of The Total Environment, 912, 169549. Available: https://www.sciencedirect.com/science/article/pii/S0048969723081792
Yu, Z., Zhang, X., Zhang, J., Feng, Y., Zhang, H., Wan, Z., Xiao, C., Zhang, H., Wang, Q. & Huang, C. 2022a. Gestational exposure to ambient particulate matter and preterm birth: An updated systematic review and meta-analysis. Environ Res, 212, Pt C, 113381.
Yu, Z., Zhang, X., Zhang, J., Feng, Y., Zhang, H., Wan, Z., Xiao, C., Zhang, H., Wang, Q. & Huang, C. 2022b. Gestational exposure to ambient particulate matter and preterm birth: An updated systematic review and meta-analysis. Environmental Research, 212, 113381. Available: https://www.sciencedirect.com/science/article/pii/S0013935122007083
Yuan, L., Zhang, Y., Wang, W., Chen, R., Liu, Y., Liu, C., Kan, H., Gao, Y., Tian, Y. & Shanghai Birth Cohort, S. 2020. Critical windows for maternal fine particulate matter exposure and adverse birth outcomes: The Shanghai birth cohort study. Chemosphere, 240, 124904. Available: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=med17&NEWS=N&AN=31550593
Yuan, S., Wang, J., Jiang, Q., He, Z., Huang, Y., Li, Z., Cai, L. & Cao, S. 2019. Long-term exposure to PM(2.5) and stroke: A systematic review and meta-analysis of cohort studies. Environ Res, 177, 108587. Available: https://www.sciencedirect.com/science/article/pii/S0013935119303846?via%3Dihub

Zhang, H., Zhang, X., Wang, Q., Xu, Y., Feng, Y., Yu, Z. & Huang, C. 2021a. Ambient air pollution and stillbirth: An updated systematic review and meta-analysis of epidemiological studies. Environ Pollut, 278, 116752.
Zhang, J., Chen, G., Liang, S., Liu, J., Zhang, J., Shen, H., Chen, Y., Duan, J. & Sun, Z. 2021b. PM2.5 exposure exaggerates the risk of adverse birth outcomes in pregnant women with pre-existing hyperlipidemia: Modulation role of adipokines and lipidome. Science of the Total Environment, 787. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85105849111&doi=10.1016%2fj.scitotenv.2021.147604&partnerID=40&md5=a8a05ed165391245d6ddcf2812f2b81f
Zhang, J. & Yu, K. F. 1998. What's the relative risk? A method of correcting the odds ratio in cohort studies of common outcomes. JAMA, 280, 19, 1690-1.
Zhang, X., Fan, C., Ren, Z., Feng, H., Zuo, S., Hao, J., Liao, J., Zou, Y. & Ma, L. 2020. Maternal PM2.5 exposure triggers preterm birth: a cross-sectional study in Wuhan, China. Global Health Research and Policy, 5, 1, 17. Available: https://doi.org/10.1186/s41256-020-00144-5
Zhang, Y., Cheng, H., Huang, D. & Fu, C. 2021c. High Temporal Resolution Land Use Regression Models with POI Characteristics of the PM(2.5) Distribution in Beijing, China. Int J Environ Res Public Health, 18, 11.
Zhao, B., Wang, M., Lü, C., Feng, L., Ma, H., Meng, H., Qi, M., Fan, Q., Wang, H., Zhou, H. & He, J. 2020. Seasonal response of the synergism of maternal comorbidities and long-term air pollution exposure on birth outcomes. Ecotoxicology and Environmental Safety, 191. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85078110898&doi=10.1016%2fj.ecoenv.2020.110232&partnerID=40&md5=f098a1cb1eec91886e716bd89aacc380
Zheng, J., Zhang, H., Shi, J., Li, X., Zhang, J., Zhang, K., Gao, Y., He, J., Dai, J. & Wang, J. 2024. Association of air pollution exposure with overweight or obesity in children and adolescents: A systematic review and meta-analysis. Sci Total Environ, 910, 168589.
Zhou, G., Wu, J., Yang, M., Sun, P., Gong, Y., Chai, J., Zhang, J., Afrim, F. K., Dong, W., Sun, R., Wang, Y., Li, Q., Zhou, D., Yu, F., Yan, X., Zhang, Y., Jiang, L. & Ba, Y. 2022a. Prenatal exposure to air pollution and the risk of preterm birth in rural population of Henan Province. Chemosphere, 286. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85112196290&doi=10.1016%2fj.chemosphere.2021.131833&partnerID=40&md5=c3bd82881bff7c2b76ae916f7e6eac0c
Zhou, W., Ming, X., Yang, Y., Hu, Y., He, Z., Chen, H., Li, Y., Zhou, X. & Yin, P. 2022b. Association between Maternal Exposure to Ambient Air Pollution and the Risk of Preterm Birth: A Birth Cohort Study in Chongqing, China, 2015–2020. International Journal of Environmental Research and Public Health, 19, 4. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85124494856&doi=10.3390%2fijerph19042211&partnerID=40&md5=e99580b017a25081c8fd23d71955dd9a
Zhu, X., Liu, Y., Chen, Y., Yao, C., Che, Z. & Cao, J. 2015. Maternal exposure to fine particulate matter (PM2.5) and pregnancy outcomes: a meta-analysis. Environ Sci Pollut Res Int, 22, 5, 3383-96.
Zou, B., Wilson, J. G., Zhan, F. B. & Zeng, Y. 2009. Air pollution exposure assessment methods utilized in epidemiological studies. J Environ Monit, 11, 3, 475-90.
Zou, Z., Liu, W., Huang, C., Cai, J., Fu, Q., Sun, C. & Zhang, J. 2021. Gestational exposures to outdoor air pollutants in relation to low birth weight: A retrospective observational study. Environmental Research, 193. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85094624399&doi=10.1016%2fj.envres.2020.110354&partnerID=40&md5=d719e906b285a084174d60d17b7a1fc0
亚库甫·艾麦尔, 王佳佳, 彭振耀, 李国星, 潘小川 [Aimaier, Y., Wang J, Peng, Z., Li, G. & Pan, X. 2013. 北京市大气污染对不良妊娠结局影响的病例对照研究 [Case-control study of ambient air pollution and adverse pregnancy outcomes of women in Beijing]. 环境与健康杂志 [Journal of Environment and Health], 30, 5, 389-393.
何进伟, 邓存红, 黑江荣, 王慧玲, 何进荣, 白敏, 李佳唯, 雷晶晶, [He, J., Deng, C., Hei, J., Wang, H., He, J., Bai, M., Li, J. & Lei, J. 2022. 某市城区室外大气污染物与早产的病例对照研究 [Association

of air pollution exposure with premature birth in a urban area: a case-control study]. 中国预防医学杂志 [Chinese Journal of Preventive Medicine], 23, 01, 1-6.

朱娟, 梁志江, 温济英, 李兵, 邹文霞, 赵庆国 [Zhu, J., Liang, Z., Wen, J. & Al.], E. 2011. 某市空气污染对早产影响的病例对照研究 [Association between Air pollution and Preterm Birth: a Case Control Study]. 环境与健康杂志 [Journal of Environment and Health], 28, 02, 132-135.

杨敏娟, 沈惠平, 黄云彪, 赵宜静, 杨琛, 孙乔 [Yang, M., Shen, H., Huang, Y., Zhao, Y., Yang, C. & Sun, Q. 2016. 妊娠期大气污染暴露对早产影响的流行病学研究 [Effect of prenatal exposure to air pollution on preterm birth: an epidemiological analysis]. 职业与健康 [Occupation and Health], 32, 02, 221-224.

樊静洁, 刘世新, 林一才 [Fan, J., Liu, S. & Lin, Y. 2019. 深圳市新生儿早产影响因素的病例对照研究 [Factors influencing preterm birth in Shenzhen city: a case‐control study]. Practical Preventive Medicine [实用预防医学], 26, 11, 1322-1325.

赵姗, 段雅, 靳颖, 田美玲, 李思思, 王莉 [Zhao, S., Duan, Y. & Jin , Y. 2020. 妊娠期 PM2.5 暴露对早产影响 [Impact of maternal PM2.5 exposure on preterm birth in Hebei province: a retrospective analysis]. 中国公共卫生 [Chinese Journal of Public Health], 36, 9, 1302-1307.

阮海燕, 赵庆国, 陈伦能 [Ruan, H., Zhao, Q. & Chen, L. 2010. 广州市番禺区空气污染物与早产关系的病例对照研究 [Association between air pollution and preterm birth in Panyu District, Guangzhou: a case-control study. 中国当代医药 [Chinese Modern Medicine], 17, 21, 140-141.

陈宇, 米白冰, 赵亚玲, 杨姣梅, 陶雅丽, 颜虹 [Chen, Y., Mi, B., Zhao, Y., Yang, J., Tao, Y. & Yan, H. 2016. 孕期室外空气污染与早产关系的 Meta 分析 [Prenatal exposure to outdoor air pollution and preterm birth: a Meta-analysis]. 中华流行病学杂志 [Chinese Journal of Epidemiology], 37, 6, 880-885. http://html.rhhz.net/zhlxbx/20160628.htm

韩慧, 王媛媛, 张敬旭 [Han, H., Wang, Y. & Zhang, J. 2017. 妊娠期细颗粒物(PM2.5)暴露对早产发生的影响 [The association of fine particulate matter exposure during pregnancy with preterm birth]. 中国生育健康杂志 [Chinese Journal of Reproductive Health], 28, 6, 506-509+524.

顾怡勤, 陈仁杰, 陈丽, 应圣洁 [Gu, Y., Chen, R., Chen, L. & Ying, S. 2019. 上海市闵行区大气污染与早产的相关性 [Associations between ambient air pollution and preterm birth in Minhang District, Shanghai]. 环境与职业医学 [Journal of Environmental and Occupational Medicine], 36, 2, 106-111.

Appendices

Appendix 1. Snapshot of the matrix showing the individual studies included in previous systematic reviews and/or meta-analyses (SR/MAs).

Number	Study name	Year	Bonzini 2010	Shah 2011	Stieb 2012	Sapkota 2012	Lamicchane 2015	Sun 2015	Zhu 2015	Chen 2016	…	Ju 2021	Ghosh 2021
1	Xu 1995	1995	0	1	0	0	0	0	0	0	…	1	0
2	Landgren 1996	1996	0	1	0	0	0	0	0	0	…	0	0
3	Bobak 2000	2000	0	1	1	0	0	0	0	0	…	1	0
4	Ritz 2000	2000	0	1	1	1	1	0	0	0	…	0	0
5	Maroziene & Grazuleviciene 2002	2002	0	1	1	0	0	0	0	0	…	1	0
6	Liu 2003	2003	0	1	1	0	0	0	0	0	…	1	0
7	Wilhelm & Ritz 2005	2005	1	1	1	1	1	1	1	0	…	0	0
8	Sagiv 2005	2005	1	1	0	1	0	0	0	0	…	0	0
9	Hansen 2006	2006	1	0	1	0	1	0	0	0	…	1	0
10	Huynh 2006	2006	1	1	1	1	1	1	1	1	…	0	1
11	Leem 2006	2006	1	0	1	1	1	0	0	0	…	0	0
12	Roger & Dunlop 2006	2006	0	0	1	1	1	0	0	0	…	0	0
13	Jalaludin 2007	2007	1	0	1	1	1	1	1	0	…	1	1
14	Kim 2007	2007	1	0	1	0	1	0	0	0	…	1	0
15	Ritz 2007	2007	1	1	1	1	1	0	0	0	…	0	1
16	Jiang 2007	2007	0	0	0	0	0	0	0	0	…	0	0
17	Zhang 2007 (Chinese)	2007	0	0	0	0	0	0	0	0	…	0	0
18	Brauer 2008	2008	1	0	1	1	1	1	1	0	…	0	1
19	Suh 2008	2008	0	0	1	0	1	0	0	1	…	0	0
20	Xu 2008 (Chinese)	2008	0	0	0	0	0	0	0	1	…	0	0
21	Zhang 2008 (Chinese)	2008	0	0	0	0	0	0	0	0	…	0	0
…	…	…	…	…	…	…	…	…	…	…	…	…	…
111	Smith, 2020	2020	0	0	0	0	0	0	0	0	…	1	1
112	Tapia, 2020	2020	0	0	0	0	0	0	0	0	…	0	1
113	Yang 2020	2020	0	0	0	0	0	0	0	0	…	1	0
114	Wang, 2020	2021	0	0	0	0	0	0	0	0	…	0	1
115	Chen, 2021	2021	0	0	0	0	0	0	0	0	…	0	1
116	Lee, 2021	2021	0	0	0	0	0	0	0	0	…	0	1

The second column is the list of individual studies found in any of the 17 SR/MAs used. From the fourth column onwards, the columns show if a primary study is included in that SR/MA or not. The matrix allows a visualisation of how often a specific individual study was included in a SR/MA. 0 indicates that a study was not included in the SR/MA whereas 1 indicates that it is included. The matrix allows a visualisation of how often a specific individual study was included in a SR/MA. The full spreadsheet is available upon request

Appendix 2. Search terms used in the respective databases and the number of records retrieved.

Database	Search string
Ovid Medline (n=2,381)	1. Air Pollution/ 2. exp Air Pollutants/ 3. (air adj2 (pollut* or quality or ambient)).mp. 4. (atmosphere* adj2 pollut*).mp. 5. exp Particulate Matter/ 6. (particulate matter or ambient particulate or PM or PM10* or PM2* or PM0* or ultrafine partic* or UFP or TSP or total suspended partic*).mp. 7. Ozone/ or (ozone or O3*).mp. 8. Nitrogen Dioxide/ or (nitrogen dioxide or NO2*).mp. 9. (nitrogen oxides or NOX).mp. 10. Premature Birth/ or (PTB or PTD).mp. 11. exp Infant, Premature/ 12. ((premature* or preterm or preterms or pre term or pre terms) adj2 (birth* or deliver* or childbirth* or infant*)).mp. 13. (preemie* or premie*).mp. 14. Gestational Age/ or gestational age.mp. 15. Pregnancy Outcome/ or ((birth or pregnanc* adj2 outcome*).mp. 16. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 17. 10 or 11 or 12 or 13 or 14 or 15 18. 16 and 17
PubMed (n = 1,022)	((air pollutants [MeSH Terms]) OR (air pollution [MeSH Terms]) OR (particulate matter [MeSH Terms]) OR (ozone [MeSH Terms]) OR (nitrogen dioxide [MeSH Terms]) OR (PM [Title/Abstract]) OR (PM2.5 [Title/Abstract]) OR (PM10 [Title/Abstract]) OR (PM0.1 [Title/Abstract]) OR (ultrafine particulate [Title/Abstract]) OR (total suspended particulate [Title/Abstract]) OR (TSP [Title/Abstract])OR (ambient particulate [Title/Abstract])) AND ((premature birth [MeSH Terms]) OR(PTB [Title/Abstract]) OR (PTD [Title/Abstract]) OR (preterm birth [Title/Abstract]) OR (preterm delivery[Title/Abstract]) OR (birth outcomes[Title/Abstract]) OR (pregnancy outcomes[Title/Abstract]))
Ovid Embase (n = 3,911)	1. air pollutant/ or exp air pollution/ 2. (air adj2 (pollut* or quality or ambient).mp. 3. (atmospher* adj2 pollut*).mp. 4. exp particulate matter 10/ or exp “particulate matter 2.5”/ or exp particulate matter exposure/ 5. ozone/ or (ozone or O3*).mp. 6. (particulate matter or ambient particulate or PM or PM2* or PM10* or PM0* or ultrafine partic* or UFP or TSP or total suspended partic*).mp. 7. nitrogen dioxide/ or (nitrogen dioxide or NO2*).mp. 8. nitrogen oxide/ or (nitrogen oxides or NOX).mp. 9. exp prematurity/ or (PTB or PTD).mp. 10. ((prematur* or preterm) adj2 (birth* or deliver* or childbirth* or infant*)).mp. 11. (preemie* or premie*).mp. 12. exp gestational age/ or gestational age.mp. 13. exp pregnancy outcome/ or ((birth* or pregnanc*) adj2 outcome*).mp. 14. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8

	15. 9 or 10 or 11 or 12 or 13 16. 14 and 15
CINAHL (n = 1,133)	1. (MH “Air Pollution+”) 2. (MH “Air Pollutants, Environmental”) 3. (air N2 (pollut* OR quality OR ambient)) 4. atmospher* N2 pollut* 5. (MH “Particulate Matter+”) 6. particulate matter OR PM OR PM2.5 OR PM10 OR PM0.1 OR ultrafine partic* OR UFP OR TSP OR total suspended partic* 7. (MH “Ozone”) OR “ozone” 8. (MH “Nitrogen Oxides”) OR nitrogen dioxide OR NO2 OR NOX 9. (MH “Childbirth, Premature”) OR “premature birth” OR (MH “Infant, Premature”) 10. ((prematur* OR preterm OR preterms OR “pre term” OR pre terms”) N2 (birth* OR deliver* OR childbirth*)) 11. preemie* OR premie* OR PTB OR PTD 12. (MH “Gestational Age”) OR “gestational age” 13. (MH “Pregnancy Outcome”) OR ((birth OR pregnanc*) N2 outcome*) 14. 1 OR 2 OR 3 OR 4 OR 5 OR 6 OR 7 OR 8 15. 9 OR 10 OR 11 OR 12 OR 13 16. 14 AND 15
MIDIRS	1. (air adj2 (pollut$ or quality or ambient)).mp. 2. (particulate matter or ambient particulate or PM or PM1* or PM2* or PM10* or PM0* or ultrafine partic* or ultrafine partic* or UFP or TSP or total suspended partic* or total suspended partic*).mp. 3. (nitrogen dioxide or NO2*).mp. 4. (nitrogen oxides or NOX).mp. 5. (ozone or O3).mp. 6. ((prematur* or preterm) adj2 (birth* or deliver* or childbirth* or infant*)).mp. 7. (PTB or PTD or preemie* or premie*).mp. 8. ((pregnancy or birth) adj2 outcome*).mp. 9. gestational age.mp. 10. 1 or 2 or 3 or 4 or 5 11. 6 or 7 or 8 or 9 12. 10 and 11
Scopus	INDEXTERMS("Air Pollution" OR "Air Pollutants" OR "Particulate Matter" OR "Ozone" OR "Sulfur Dioxide" OR "Nitrogen Dioxide" OR "Carbon Monoxide")OR TITLE-ABS-KEY(air W/2 (pollut* OR quality OR ambient))OR TITLE-ABS-KEY(atmospher* W/2 pollut*)OR TITLE-ABS-KEY("particulate matter" OR "ambient particulate" OR PM OR PM1* OR PM2* OR PM0* OR "ultrafine particulate*" OR "ultrafine particle*" OR UFP OR TSP OR "total suspended particle*" OR "total suspended particulate*" OR SO2* OR "nitrogen dioxide" OR NO2* OR "nitrogen oxides" OR NOX)AND (INDEXTERMS("Premature Birth" OR "Infant, Premature" OR "Gestational Age" OR "Pregnancy Outcome")OR TITLE-ABS-KEY(PTB OR PTD OR premature infant OR preemie* OR premie* OR gestational age)OR TITLE-ABS-KEY((prematur* OR preterm OR preterms OR "pre term" OR "pre terms") W/3 (birth* OR deliver* OR childbirth*))OR TITLE-ABS-KEY((birth OR pregnanc*) W/2 outcome*))

GreenFILE	1. (air adj2 (pollut$ or quality or ambient)).mp. 2. (particulate matter or ambient particulate or PM or PM1* or PM2* or PM10* or PM0* or ultrafine partic* or ultrafine partic* or UFP or TSP or total suspended partic* or total suspended partic*).mp. 3. (nitrogen dioxide or NO2*).mp. 4. (nitrogen oxides or NOX).mp. 5. (ozone or O3).mp. 6. ((prematur* or preterm) adj2 (birth* or deliver* or childbirth* or infant*)).mp. 7. (PTB or PTD or preemie* or premie*).mp. 8. ((pregnancy or birth) adj2 outcome*).mp. 9. gestational age.mp. 10. 1 or 2 or 3 or 4 or 5 11. 6 or 7 or 8 or 9 12. 10 and 11
Wan Fang	主题:(("空气污染" or "空气质量" or "大气污染" or "臭氧" or "PM2.5" or "PM10" or "PM0.1" or "NO2" or "SO2" or "O3" or "大气细颗粒物") and ("早产" or "不良妊娠结局"))
CNKI	篇关摘 (空气污染 + 空气质量 + 大气污染 + 臭氧 + 二氧化氮 + 二氧化硫 + 颗粒物污染 + PM2.5 + PM10 + PM0.1 + NO2 + SO2 + O3) AND (早产 + 不良妊娠结局 + 异常妊娠结局)
CSTJ	M=(("空气污染" OR "空气质量" OR "大气污染" OR "臭氧" OR "二氧化氮" OR "二氧化硫" OR "颗粒物污染" OR "PM2.5" OR "PM10" OR "PM0.1" OR "NO2" OR "SO2" OR "O3") AND ("早产" OR "不良妊娠结局" OR "异常妊娠结局"))

Appendix 3. The title and abstract screening protocol

Reason for exclusion by group

Q3: Not an article

Q4, 6, 7: Other study designs

Q5: Non-human participants

Q8: Not related to air pollution

Q9: Not related to pregnancy

Q10: Not related to birth outcomes

1. Does the citation, title or abstract indication that it is a correction, erratum or retraction of an included article?

Yes – INCLUDE DIRECTLY

No – *Continue screening*

2. Does the **citation** indicate that the publication date is within the range of anytime up to April 2022?

Yes – Continue screening

No – Stop screening

3. Does the **citation, title or abstract** indicate that it is a peer-reviewed or published article?

Yes – Continue screening

No – Stop screening. For example: posters,

abstracts, orations, presentations or proceedings at workshops or conferences.

4. Does the **citation, title or abstract** indicate that it is NOT an editorial, letter to the editor, hypothesis, news, perspectives, policy statement or commentary?

Yes – Continue screening

No – Stop screening

5.Does the **title or abstract** indicate that the study is conducted on humans?

Yes – Continue screening

No – Stop screening

6. Does the **title or abstract** indicate that it is NOT a case report or case-series?

Yes/Maybe – Continue screening

No – Stop screening

7. Does the **title or abstract** indicate that it is NOT a summary, review, systematic review, meta-analysis, or (randomised) clinical trial?

Yes – Continue screening

No – Stop screening

8. Does the **title or abstract** indicate that environmental exposure to air pollution or any air pollutants were studied?

Yes – Continue screening

Maybe – Continue screening

No – Stop screening

Keywords – air pollution, atmospheric pollution, air quality, air pollutants, ozone (O3), nitrogen dioxide (NO2), nitrogen oxides (NOX), sulphur dioxide (SO2), particulate matter (PM2.5 / PM0.1 / PM10), carbon dioxide (CO), sulphur dioxide (SO2), ambient particulates, ultrafine particulate (UFP), total suspended particulate (TSP)

9. Does the **title or abstract** indicate that exposure during pregnancy was studied?

Yes – Continue screening

Maybe – Continue screening

No – Stop screening

10. Does the **title or abstract** indicate that (live) birth outcomes were studied?

Yes – Continue screening

Maybe – Continue screening

No – Stop screening

Keywords – preterm birth (PTB), preterm delivery (PTD), premature birth, prematurity, pregnancy outcomes, birth outcomes, gestational age

Final decision: Should this article be included?

Yes - All questions answered Yes/Unclear, Q1 answered Yes

Maybe - If question 7, 8 or 9 has been answered Maybe, the rest Yes/Maybe

No – At least one question answered No

Appendix 4. The full-text screening protocol

Question	Decision	Tag
Is full-text available for this study?	Yes – Continue No – Stop	EX_FTNA
Is the data used NOT from simulation or secondary sources?	Yes – Continue No – Stop	EX_Others
Did the study restrict the population or run an analysis on singleton live births only?	Yes – Continue No – Stop	Ex_Births
Did the study include preterm birth as one of the birth outcomes studied?	Yes – Continue No – Stop	Ex_NoPTB
Did the study investigate the effect of $PM_{2.5}$, NO_2, and O_3 on preterm birth?	Yes – Continue No – Stop	Ex_Pollutant
Is the exposure studied non-occupational?	Yes – Continue No – Stop	EX_NOC
If “Yes” were answered from Question 1 to 6, the study is eligible to be quality-assessed. If sufficient information is available for quality assessment, then the study is eligible to be included in the review.		
Is information on gestational age at individual level, or in the form of summary statistics available for the study population?	Yes – Continue No – Write to authors	EX_GA
Were quantitative estimates and their 95%CIs (not only graphical representation) available for the pollutants of interest?	Yes – Continue No – Write to authors	Ex_Quan
Is information required to convert the quantitative estimates to a common unit available?	Yes – Continue No – Write to authors	Ex_Incr
If “Yes” were answered for Question 1-9, the study is eligible to be quality-assessed. If sufficient information is available for quality assessment, then the study is eligible to be included in the meta-analysis. If “No” is answered, then the corresponding author of the study is contacted. If no responses were received, then the study is excluded from the meta-analysis.		

Appendix 5. The data extraction form.

The data extraction form was developed to extract relevant information from studies of air pollution and preterm birth. Data extraction was conducted on Qualtrics. The version used on Qualtrics is available upon request.

Section 1 – General Information

- Covidence number
- Multiple estimates available and reason e.g. more than 1 exposure assessment method used
- Paper name, title, author, journal name

Section 2 – Study characteristics

- Location
- Design
- Start and end date
- Funding source
- Data source (participants)
- Data type
- Sample size
- Inclusion and exclusion criteria
- Method of gestational dating

Section 3 – Outcome and exposure assessment

- Exposure window examined (Preconception, Trimester 1, Trimester 2, Trimester 3, Whole pregnancy, Others)
- Type(s) of PTB examined
    - General PTB, moderate to late PTB, very PTB, extreme PTB, others
- Cause of delivery
    - Spontaneous, medically-indicated, not specified
- Mode of delivery
    - Vaginal, caesarean, not specified
- Pollutant studied
- Method of exposure assessment
- Spatial scale of exposure assessment

Section 4 – Information availability

- Gestational age information
- Unit of exposure increment
- Effect estimates that are not only graphically presented
- Effect estimates relevant for narrative synthesis. But not meta-analysis

Section 5 – Statistical analysis

- Statistical model(s)
- Type of effect estimate(s) produced
    - OR/RR/HR

    - Categorical / Continuous
- Variable(s) adjusted in the main model presented
- Presence of multipollutant models
- Number of PTBs used in the analysis
- Choice of model and justification
- Unit for pollutant(s)

Section 6 – Pollutant information (per exposure window)

- Mean
- Median
- Minimum
- Maximum
- Interquartile range
- Standard deviation
- Averaging time

Section 7 – Model information (continuous)

- OR/HR
- Lower CI
- Upper CI
- Exposure increment

Section 8 – Model information (categorical)

- OR (category 1)
- OR (category 2)
- OR (category 3)
- OR (category 4)
- OR (category 5)

Section 9 – Other comments

Appendix 6. The modified WHO Air Quality Guideline Risk of Bias assessment instrument for studies of gestational air pollution exposure and the risk of preterm birth.

Prerequisite

- Did the study provide all information necessary to carry out the quality assessment? (Y/N)

Confounding

- Did the study adjust for the following critical confounders?

Age	Y/N
Infant sex	Y/N
Socioeconomic status	Y/N
Parity	Y/N
Season of conception/birth	Y/N

- Did the study adjust for the following potential confounders?

Year of conception/birth	Y/N
Ethnicity/Race	Y/N
Smoking	Y/N
BMI	Y/N
Alcohol consumption	Y/N
Marital status	Y/N

- Subdomain level of risk

All critical confounders adjusted	Low*
4 critical confounders adjusted	Low
2-3 critical confounders adjusted	Moderate
0-1 critical confounders adjusted	High

- Did the authors use valid measures of confounding factors? (Y/N)

- Did the authors use appropriate analysis method or study design that controlled for confounding domain? (Y/N)

- Level of risk (Low/Moderate/High)

Study design

- Did all participants have equal opportunity to be in the study? (Y/N)
- Did the authors provide a valid discussion about the power/sample size needed to detect the effect, or provide a estimate of the variance/effect size? (Y/N)
- Did the study control for fixed cohort bias by design, or by restricting the analysis to infants with date of conception a period of time before the start and end of the cohort? (Y/N)
- Did the study assess the exposure prior to measuring the outcome?
- Level of risk (Low/Moderate/High)

Exposure assessment

- Was it appropriate for the research question? (Y/N)
- Were the measures defined in details/sufficient information provided? (Y/N)
- Was the method validated in any way? (Y/N)
- Was the exposure assessment comparable across range of exposure? (Y/N)
- Was the exposure assessment comparable across population? (Y/N)

- Was the exposure assessed in an appropriate way spatially? (Y/N)
- If modelling was done, was the spatial resolution appropriate for the pollutant of interest? (Y/Maybe/N)
- Was maternal mobility taken into account? (Y/N)
- Was the exposure contrast sufficient compared to the precision of the exposure assessment? (Y/Maybe/N)
- Level of risk (Low/Moderate/High)

Outcome measurement

- Was PTB defined clearly? (Y/N)
- Is gestational age dated using LMP, ultrasound or clinician estimates, or a combination of these? (Y/N)
- Is the outcome assessed consistently and objectively across the population (cases/control)? (Y/N)
- Level of risk (Low/Moderate/High)

Missing data

- Is the frequency of missing data for exposure assessment acceptable? (Y/N)
- Is the frequency of missing data for the outcome acceptable? (Y/N)
- Level of risk (Low/Moderate/High)

Selective reporting

- Are effect estimates present for all hypotheses tested as per aim? (Y/N)
- Did the study reference to a publish or unpublished study protocol? (Y/N)
- Level of risk (Low/Moderate/High)

Overall judgement

- Level of risk (Low/Moderate/High)*

No high RoB in any domain	Low
High RoB present but only in 1 domain	Mod
High RoB in >1 domain	High

*Above is a general rule, but exceptions may apply if the reviewer deems shortcoming in one domain is overwhelming, or when important information are constantly missing e.g. exposure increments, 95% confident intervals, then a high level of risk of bias should be assigned.

Appendix 7. Guidance for using the modified framework

Overall guidance

This tool was initially developed by members of the Guideline Development Group for the update of the WHO global air quality guidelines (AQG) with expertise in air pollution epidemiology, with the aim to critically evaluate studies to be included in the systematic reviews of air pollution and health outcomes. It is split into 5 main domains: confounding, selection bias, exposure assessment, outcome assessment, missing data, and selective reporting. As it is not developed for a specific health outcome e.g. PTB, modification are made to the framework regarding the general framework layout, as well as content in specific domains and subdomains to include aspects more relevant to PTB.

The original version of the framework assigns a risk of bias per subdomains. This modified framework should be used such that only one risk of bias conclusion is assigned per domain. The specific questions posed in each domain serves to help the reviewers judge the risk of bias in that subdomain, but they should only be answered with yes/no rather than high, moderate or low. An overall judgement of high, moderate or low for the whole domain must then be made by the reviewer. The only exception is the 'Confounding' domain, where a subdomain level of risk of bias will be assigned just for the number of critical confounders adjusted.

This prevents the analysis from being overly fragmented. Moreover, the answers across several subdomains could be interlinked, or there would be situations where substantial issues in a particular subdomain would overshadow merits found for other subdomains e.g. very coarse spatial resolution of a model (50km x 50km grid) even when precise maternal address is available, use of a single monitor to assign a city-wide average level. It is then worth highlighting such issue in the *overall* risk of bias for the exposure assessment domain. The judgement of the reviewer will be essential because there are issues such as technological/funding limitations that may not be known. The reviewer must utilize their expertise on the common practices of the field and judge accordingly.

Texts that can be found in the original framework is shown in green. Elements that were present in the original framework but subsequently modified are highlighted in green.

Confounding

Original version

In addition to the guidance on confounding provided above, it is important to consider some issues on controlling for long-term trends. While the PECOS may be asking for the relationship between the exposure of interest and short-term variation in the outcome, long-term trends may dominate the potential confounders. It is therefore essential to demonstrate a separation between short- and long-term trends to have certainty in the assessment of the relationship between the exposure and outcome. Regression models may be appropriate to demonstrate that long-term trends are controlled for in the analysis. Some options available include: (i) a time stratified model (simple indicator variables); (ii) periodic functions; or (iii) flexible spline functions. The use of one of these models to adjust for potential long-term confounders allows for the isolation of the exposure of interest (8). It is also necessary to

think in which direction the long-term trend would potentially bias the effect (i.e. toward or away from the null).

Modification

Season of conception: added as another critical confounder as pregnancy is relatively short term (9-months), and with preterm birth, it is even shorter. Season of conception is ideal, but the season of birth/delivery is also acceptable.

Parity: added as it is an independent risk factor for preterm birth

BMI: Preferably pre-pregnancy BMI. Maternal pre-pregnancy weight and height (which forms BMI) are also acceptable (what about maternal weight gain during pregnancy?).

Socioeconomic status: Using proxies such as education, employment status, income, and deprivation that can be individual-level or area-level, provided that area-level adjustment is sufficiently spatially refined (In original version but explanation added)

Birth year: Its importance increases as the number of years in the study period increases, as it accounts for annual variations in various factors e.g., pollution levels, particular events that affect air pollution levels or birth outcomes. Not necessary if the study only lasts for a year.

Ethnicity/Race: If the study population is monoethnic then NA (In original version but explanation added).

Smoking: Maternal exposure to tobacco smoke during pregnancy, both in the form of the smoking status of the mother or exposure to second-hand tobacco smoke, such as the smoking status of her partner. In certain countries (more common in Asia), it is very rare for pregnant women or women in general to smoke, hence controlling for smoking would not be necessary in this case (In original version but explanation added).

Validity of measuring of confounding factors: though it is ultimately a subjective decision, we can gauge this according to some common practices in the field. It is reasonable to assume that data collected from birth certificates are collected in a standardized manner, hence error is likely to be random instead of systematic according to exposure or outcome status of the new-born. Information regarding maternal behaviour is likely to be less accurate. For example, alcohol and smoking, are most likely to be self-reported and could be under-reported to health professionals as the mothers probably know these behaviours are unwise during pregnancy.

Study design (originally selection bias)

Original version

Selection bias occurs when some eligible participants are excluded in a way that leads to the association between exposure and outcome differing from the association that would have been observed in the PECOS.

Modification

Sample size

In observational studies power calculations are seldom conducted as they are usually done *a priori*. 'Yes' should be selected if the authors mention the power/sample size needed to detect the effect of the outcome of interest, or an estimate of the variance/effect size (in this case PTB). If nothing is mentioned, 'No' should be selected. The reviewer should consider that absence of power calculation for observational studies is common practice and hence not 'fatal' when assigning the overall risks of bias in the domain.

Fixed cohort bias

Fixed cohort bias occurs when the population is defined based on infant's date of birth rather than the date of conception. The population at risk (pregnant women) at the beginning of the study will miss pregnancies with shorter gestation (as pre-term births occurred before the time period based on births started). In contrast, the population at risk at the end of the study will miss pregnancies with longer gestations (as the birth will not yet have occurred when the time period based on births stopped). Thus, the proportion of pregnancies resulting in pre-term birth is too low at the beginning and too high at the end. This bias is likely important for PTB as it is an outcome which is determined by the length of gestation. It is also important for seasonally varying exposures e.g., O_3, and if there are long-term trends in air pollutants. Prospective cohort studies control for this by design, whereas studies using retrospective cohort or case control design avoid fixed cohort bias by restricting the analysis to infants with date of conception a period of time prior to the beginning of the study cohort, which the period of time may correspond to the definition of stillbirth in the study location (mainly 20 or 24 weeks). These studies will also restrict the analysis with date of conception to infants with conception date a period of time before the end of the study cohort, and the period of time may correspond to the upper gestational limit set by the study. In this case, pregnancies with shorter and longer gestations can all be included. For instance, the text will be as follows: "…confined the dataset of conception occurring 20 weeks prior to study start date, and 44 weeks prior to study end date".

Reverse causation bias

Reverse causation bias occurs when exposure is assessed at the same time or after the outcome. This can occur in all types of study designs e.g. for cross-sectional studies causality cannot be inferred as exposure and outcome are measured at the same time. For retrospective cohorts, the outcome had already occurred, and it is important that the authors have made clear that the exposure was assessed/assigned in a time period preceding the outcome. For prospective cohort studies there's a lower risk of reverse causation bias, however it can still occur. The reviewer therefore needs to carefully examine if the exposure status is determined for a calendar time period clearly preceding the time the outcome occurred i.e., before the date of birth. Specifically for studies of air pollution and birth outcomes, it would be inappropriate to include exposure *after* birth, and this would be classified as a serious issue.

Reviewers also need to be aware that there are factors that may alter the exposure status e.g. mothers' awareness of harmful effects of air pollutants, clinical conditions that may lead to preterm birth e.g. preeclampsia that may also alter maternal exposure to air pollution. However, it would be difficult to assess this using a risk of bias evaluation tool.

Exposure assessment

Original version

…

Day-to-day variations in outdoor air pollution concentrations are typically large (> one order of magnitude) and largely driven by meteorology; thus, these day-to-day variations are not very different between different pollutants of interest in the AQG context. Studies of the effects of long-term, spatial variations in air pollution on the spatial variation in mortality and morbidity outcomes rely on various measurements and modelling efforts. Spatial variations in long-term air pollution concentrations are driven by spatial variations in source strength, road networks, topography, etc. Interpolation between data from routine monitoring stations may be adequate in some studies able to use dense monitoring networks. In most cases, some combination of monitoring data, dispersion modelling, land use regression modelling and satellite observations is being used to estimate the relevant spatial variations. Scales vary from tens of metres for the finest scale land-use regression and dispersion models to 10 x 10 km for some satellite products. Issues for the raters to consider are:

- *The ability of the exposure models used in the studies to adequately predict the exposure; this can be concluded if the model is adequately evaluated against measurements (low risk of bias) or is not evaluated against measurements (high risk of bias); and*
- *The temporal stability over time scales relevant for the long-term studies of interest; e.g., if the exposure contrast is generated for a specific year, it is representative for other years of the epidemiological study and outcome of interest (low risk of bias) or if it is unrepresentative (high risk of bias). Typically, details such as these are found in separate technical papers; the raters should ascertain that they are referenced.*

Modification/Addition

It is very hard to effectively critique studies in the domain of exposure assessment as a plethora of factors needs to be considered. Modelling efforts are sometimes limited by storage or funding. It is important that sufficient details are provided for the exposure assessment method, and validation is also provided or referenced when needed.

Appropriateness of method used

- Modelling (EA3-EA8), especially those with both spatial and temporal components are appropriate for these studies. More detailed critique will be done by the subsequent questions. It is also important that the measures are defined in detail and validated. The reviewers should try retrieving the technical articles referenced (if available) if it is not mentioned directly in the paper. Direct assignment of monitor concentration is usually not representative of the actual exposure, and it is less and less representative as the radius of assignment increases (district/city level would be not representative at all). Extrapolations such as inverse distance weighting (IDW) and kriging may form a more representative exposure spatially, but IDW and kriging

assumes linear relationship from the pollutant source, and this assumption can often be violated. The accuracy of the assigned exposure also depends heavily on the quality of the monitor data which in some cases would be unknown.

Example: *CMAQ: extensively used dispersion model, but also depends on the quality of the local emission data.*

Consistency – largely kept

Exposure assignment

Spatial assignment of exposure

- There are 2 questions in the framework that is related to spatial assignment of exposure, as it is determined by:

1. information available regarding where the mothers were, whether it is maternal address, postcode, district or just city. Postcode level in general is an appropriate cut off. It would be even better if occupational or transit exposure were considered, but it is rarely done.
2. The spatial resolution of the model used also matters, because even if maternal postcode/addresses were available, if the resolution was 15km x 15km, it means all maternal postcodes within that 15km grid will be assigned the same concentration. For particulates that are less spatially heterogenous it may be acceptable, but it may not be ideal for gaseous pollutants that tend to have more locally-generated sources. It is also location-dependent e.g. for places with sealine, large spatial grids will overlap into the sea/ocean and make the concentration assigned to the grid even less representative.

Temporal assignment of exposure

- Hourly/Daily variation of pollutant concentration is ideal as it captures changes majority of the changes in pollutant concentration, which could be due to change in volumes of traffic during the day/week (weekday vs. weekend), which would not be captured by weekly/monthly concentrations. However, weekly/monthly concentrations are still appropriate in the context of studies of PTB, as the exposure window is not very short (between 20 weeks – 37 weeks). Annual averages of concentrations are not inappropriate because pregnancies do not last for a whole year, and exposure after birth would be taken into account.

Maternal mobility

- It would be ideal if the method of exposure assessment captures mobility of the mother, which could result from (1) change(s) in address during pregnancy (2) day-to-day movement e.g. to workplace. However, very few studies have actually taken this into account. It is unclear if accounting for maternal mobility during pregnancy will result in a drastic change in the effect estimate derived, as it likely depends on the

distance between the new and old address, or address and workplace etc (Hodgson et al 2015; Lupo et al., 2010; Warren et al., 2018).

Exposure contrast

- Majority of studies of air pollution and PTB are conducted in areas with relatively low air pollution levels e.g. North America, Europe, UK etc. It is worth noting if the authors have discussed the potential issue of exposure contrast being too low to elucidate the differences in risks between populations of high and low exposures. Judgment should be made based on the descriptive information regarding the pollutant levels e.g. mean, SD, IQR. Reviewers can use the new AQG values as a rough gauge e.g. if $PM_{2.5}$ contrast is less than 5ug/m^3 it is likely to be too little contrast.
- N.B. Whether exposure contrast is sufficient also depends on the precision of measurement i.e. extent of measurement error, as well as random variations. The influence of measurement error will become larger as exposure contrast and pollution levels become lower.

Assigning overall risk of bias

- Moderate level of risk of bias will be assigned if EA1 (nearest monitor approach) or EA2 (inverse distance weighting) were used.
- If any questions in the spatial and temporal assessment of exposure are answered as 'No', then a high level of risk of bias will be assigned.

Outcome measurement

Questions within this domain are to determine whether bias has been introduced by the misclassification or introduction of error in the measurement of the outcome. This may differ depending on the outcome of interest and should be examined separately for each outcome.

Modification/Addition

Blinding of outcome measurement – removed

- In most cases information on PTB is obtained via birth certificates, which PTB is determined by a healthcare professional (HCP) at delivery. Unless stated otherwise, the outcome is assumed to be measured validly, as the HCP at delivery would likely have no knowledge or influence of the exposure of the infant. This is similar for research cohorts as well. Hence this question is deemed irrelevant.
- A question regarding if the outcome is defined clearly is added, to ensure that PTB is stated clearly (whether it is <37 weeks gestational, or other cut points for other subgroups of PTB). In studies of PTB there will be cases that PTB is mixed interchangeably with premature rupture of membrane (PROM). This is not

appropriate as PROM is a strong indicator of PTB but do not necessarily result in a preterm delivery.

Validity of outcome measurement

- Gestational age is usually determined by 3 methods: last menstrual period (LMP), ultrasound measurements/dating (UD), or clinical estimates which utilises a mixture of LMP and UD. Hence, these 3 methods are all deemed appropriate. In rare cases maternal recall might be used, but it is extremely inaccurate as it is hard to have the exact date of conception except in those who have underwent in-vitro fertilisation.
- PTB as an outcome is essentially determined by whether gestational age has reached 37 weeks, which is an arbitrary cut-point. UD is considered to be more accurate in estimating gestational age, as LMP is in essence an approximate. However, it is unclear if UD is clearly superior than LMP in determining gestational age of *younger* infants, as UD determines gestational age through infant size (crown rump length), and some argued that preterm infants, due to predisposed risk factors, could already be smaller in size right from the beginning. Hence, this framework would not critique based on which gestational dating method was used.
- There is a possibility that the gestational dating method is not mentioned explicitly in the paper. In this case, reviewers should try to retrieve the technical papers/cited dataset to see if gestational dating method was included.

Missing data

The content in the two subdomains is largely kept, but levels of risk of bias would not be assigned based on the frequency of missing data of both the exposure and outcome. It is because studies of air pollution and PTB usually utilises birth certificates, so missing outcome data is relatively rare. For exposure, frequency of missing data would depend on the model inputs e.g. monitor concentrations, in certain situations it could be larger than 10%, but it would still be deemed appropriate if missing data is imputed using appropriate methods e.g. multiple imputation. Direct assignment/roll over would not be appropriate.

Selective reporting

The content in the two subdomains is largely kept, but levels of risk of bias would not be assigned per subdomain.

Appendix 8. Studies that were excluded from the meta-analysis and the justifications

Study	In systematic review?	In main meta-analysis?	Reason for exclusion	Description
Abdo et al., 2019	Y	N	Focus on wildfire smoke exposure	Conducted analysis without wildfire smoke $PM_{2.5}$ but model was co-adjusted with O_3.
Alman et al., 2019	Y	N	Categorical estimates	Estimates of effect were of pollutant exposure in tertiles instead of for the whole distribution of pollutants
Bachwenkizi et al., 2022	Y	N	Multipollutant estimates	Effect estimates were all multipollutant
Barba-Vasseur et al., 2017	Y	N	High risk of bias	Insufficient confounder adjustment (2 critical confounders), with relatively small sample size
Bertin et al., 2015	Y	N	Categorical estimates High risk of bias	Estimates of effect were of pollutant exposure in tertiles instead of for the whole distribution of pollutants. Insufficient confounder adjustment (1 critical confounder). Other covariates adjusted were medical conditions. The temporal resolution used is also unclear (likely annual).
Brown et al., 2015	Y	N	Insufficient information	Summary estimate was provided without 95%CI.
Cao et al., 2022	Y	N	Dichotomised exposure	Estimates of effect were of pollutant exposure at a certain level instead of for the whole distribution of pollutants.
Capobussi et al., 2016	Y	N	Categorical estimates	Estimates of effect were of pollutant exposure at a certain level instead of for the whole distribution of pollutants.
Chang et al., 2012	Y	N	Incompatible methodology	Bayesian statistical model used.
Cossi et al., 2015	Y	N	Categorical estimates	Estimates of effect were derived by comparing highest to the lowest three quartiles of exposure to pollutants

			High risk of bias	Insufficient confounder adjustment (2 critical confounders). Information is missing for areas such as pollutant concentrations, results for trimesters 1 and 2, and units of pollutants.
Costello et al., 2021	Y	N	Insufficient information	Effect estimates are per IQR increase, but IQR (and pollution levels) not provided.
Dzhambov et al., 2019	Y	N	High risk of bias	Sample size being very small with exposure assessment based on data from 10 years ago.
Fan (樊静洁) et al., 2019	Y	N	High risk of bias Insufficient information	Insufficient confounder adjustment (2 critical confounders). Very limited sample size and insufficient information on exposure assessment methodology used, spatial scale, and exposure increment used.
Fang et al., 2020	Y	N	$PM_{2.5}$ exposure above a certain level	No effect estimates per unit increase in $PM_{2.5}$ levels. Effect estimates available are only for association with number of days where $PM_{2.5}$ concentrations above a certain level.
Gan et al., 2022	Y	N	High risk of bias Insufficient information	Combination of factors: (1) Insufficient confounder adjustment (no critical confounders), (2) Exposure assigned using nearest monitor approach, (3) Relatively small sample size. Exposure increment used for $PM_{2.5}$ is unclear.
Garshol et al., 2020	Y	N	Insufficient information	Unclear what the exposure increment used was.
Gehring et al., 2011b	Y	N	Categorical estimates	Estimates of effect were of pollutant exposure at a certain quintile instead of for the whole distribution of pollutants.
Gray et al., 2014	Y	N	Different study aim	Analysis of socioeconomic status and ethnicity as effect modifiers in the association of air pollution and PTB.

Gu (顾怡勤) et al., 2019	Y	N	High risk of bias Insufficient information	Exposure assigned using nearest monitor and at city level. Key information is missing, such as confounder adjustment and sample size
Guo et al., 2019	Y	N	Insufficient information	Estimates were presented per IQR increase, but it was unclear what the IQR was.
Han (韩慧) et al., 2017	Y	N	High risk of bias	Exposure was assigned using nearest monitor approach using 40km radius, with very limited sample size
Han et al., 2018	Y	N	Categorical estimates	Estimates of effect were in different quartiles of air pollution exposure
Hannam et al., 2014	Y	N	Categorical estimates	Estimates of effect were in different quartiles of air pollution exposure
He (何进伟) et al., 2022b	Y	N	High risk of bias Insufficient information	Insufficient confounder adjustment (2 critical confounders), sample size is very small and key information was missing for exposure assessment e.g. spatial resolution, and exposure increment.
Heft-Neal et al., 2022	Y	N	Wildfire smoke $PM_{2.5}$ exposure	Not on ambient $PM_{2.5}$ exposure
Huang et al., 2018	Y	N	High risk of bias	Insufficient confounder adjustment (2 critical confounders). Exposure assigned using the nearest monitors and at census tract level.
Jackson et al., 2009	Y	N	Incompatible methodology	Bayesian statistical model used.
Kim et al., 2021	Y	N	Different but related outcome	Effect estimates presented is for air pollution and number of pregnancy complications (includes PTB).
Kirwa et al., 2021a	Y	N	High risk of bias	Exposure assigned at municipality level using nearest monitor approach and limited sample size

Kwag et al., 2021b	Y	N	Categorical estimates	Estimates of effect were in different percentiles of level of air pollution exposure
Le et al., 2012	Y	N	Not exposure window of interest	Estimate of effect was for first month or last month before birth
Leem et al., 2006	Y	N	Categorical estimates	Estimates of effect were in different quartiles of air pollution exposure
Liu et al., 2003	Y	N	Not exposure window of interest	Estimates of effect were only for first month and last month of pregnancy
Liu et al., 2019a	Y	N	High risk of bias	Exposure assigned at city level and limited sample size
Liu et al., 2019b	Y	N	High risk of bias Not exposure window of interest	No critical confounders adjusted and limited sample size. Also, no estimates for exposure window of interest
Liu et al., 2022	Y	N	Incompatible methodology	Natural experiment
Llop et al., 2010	Y	N	Dichotomised exposure	Effect estimates available are only for the effect of NO_2 exposure concentrations above or below 46.2 µg/m^3.
Melody et al., 2019	Y	N	Focus on coal fire $PM_{2.5}$ High risk of bias	Not on ambient $PM_{2.5}$ but coal fire attributable $PM_{2.5}$ Sample size is relatively small, and exposure was assigned at SA1 level which is quite fine, however it was not at residential address level, and the exposure contrast for coal-fire attributable $PM_{2.5}$ is also quite small.

Mendola et al., 2019	Y	N	Categorical estimates	Estimates were according to different levels of exposure of the same women for different pregnancies.
Nealy et al., 2018	Y	N	High risk of bias	Insufficient confounder adjustment (2 critical confounders). Exposure was also assigned at census block group, which is the smallest geographical unit but may not be sufficient.
Olsson et al., 2012	Y	N	Multipollutant estimates	Effect estimates provided were all derived from multipollutant models.
Padula et al., 2014	Y	N	High risk of bias	Insufficient confounder adjustment (2 critical confounders). Exposure assessed using inverse distance weighting (IDW) with large interpolation distance (50km).
Padula et al., 2018	Y	N	High risk of bias	Insufficient confounder adjustment (2 critical confounders). Exposure assigned at census tract level and used annual averages for particulate matter.
Padula et al., 2019	Y	N	High risk of bias Categorical estimates	Insufficient confounder adjustment (2 critical confounders). Exposure assessed using inverse distance weighting (IDW) with large interpolation distance (50km). Air pollution exposure was dichotomised according to highest quartile cut-off
Pereira et al., 2014c	Y	N	$PM_{2.5}$ sources	Effect estimates were for $PM_{2.5}$ sources only.
Rappazzo et al., 2014	Y	N	Insufficient information	Original risks used to derive the risk differences were not available.
Rappazzo et al., 2015	Y	N	$PM_{2.5}$ constituents	Effect estimates were for $PM_{2.5}$ constituents only.
Rappazzo et al., 2016	Y	N	Insufficient information	Original risks used to derive the risk differences were not available.

Riddell et al., 2021	Y	N	Categorical estimates	Estimates of effect were in different percentiles of air pollution exposure
Ritz et al., 2000	Y	N	Not exposure window of interest	Estimate of effect was for 6 weeks before birth and 1st month of pregnancy
Ruan (阮海燕) et al., 2010	Y	N	High risk of bias Insufficient information	Exposure assigned using nearest monitor and at city level. Key information is missing for exposure assessment, such as spatial resolution or interpolation distance of the model, and level of exposure assignment.
Saldiva et al., 2018	Y	N	Insufficient information/ incompatible methodology	Methodology was unclear and all effect estimates seem to be of multipollutant models
Salihu et al., 2012	Y	N	Dichotomised exposure	The effect estimates were for exposure above or below median concentration of pollutants.
Seabrook et al., 2019	Y	N	Insufficient information	ORs were provided without 95%CI, some uncertainties with methodology.
Shi et al., 2021	Y	N	Insufficient information High risk of bias	Effect estimates are per IQR increase, but IQR not provided. Authors were contacted. Insufficient confounder adjustment (2 critical confounders). Sample size was extremely small (n=439) but likely due to it being hypothesis-generating/exploratory in nature and focus on subpopulation (those undertaking IVF treatment).
Siddika et al., 2020	Y	N	Categorical estimates	Estimates of effect were derived by comparing highest to the lower three quartiles of exposure to pollutants
Sun et al., 2020	Y	N	Insufficient information	The relative excess risk due to interaction (RERI) produced were for joint effect of green space and air pollution exposure above a

				certain temperature level, hence individual HRs cannot be derived for air pollution exposure only.
Symanski et al., 2014	Y	N	Not exposure window of interest	Estimates of effect were for exposure during pregnancy at a 4-week interval.
Symanski et al., 2016	Y	N	Not exposure window of interest	Estimates of effect were for exposure during pregnancy at a 4-week interval.
Transande et al., 2013	Y	N	High risk of bias	Insufficient confounder adjustment (2 critical confounders). Exposure assigned using address of birth hospital instead of maternal residence.
Tu & Tu, 2018	Y	N	Incompatible methodology	Global logistic regression produced estimate of different nature
Wang et al., 2022	Y	N	High risk of bias	Exposure assigned using hospital of delivery instead of maternal address. Sample size is also extremely small.
Warren et al., 2012	Y	N	Incompatible methodology	Bayesian statistical model used.
Weber et al., 2019	Y	N	Categorical estimates High risk of bias	Estimates of effect were derived by comparing highest to the lowestthree quartiles of exposure to pollutants. Insufficient confounder adjustment (2 critical confounders). Exposure assigned using inverse distance weighting (IDW) with large interpolation distance (25-50km).
Williams et al., 2021	Y	N	High risk of bias	Limited confounder adjustment (2 critical confounders), and exposure assigned at hospital referral region rather than maternal address.
Wing et al., 2022	Y	N	High risk of bias	Limited confounder adjustment (2 critical confounders), and the spatial resolution might be too large (15km buffer).

Xu (徐丽珍) et al., 2008	Y	N	High risk of bias Insufficient information	Insufficient confounder adjustment (2 critical confounders). Key information is missing for exposure assessment, such as spatial resolution or interpolation distance of the model, and level of exposure assignment.
Yakufu (亚库甫·艾麦尔) et al., 2013	Y	N	High risk of bias Insufficient information	Exposure assigned using nearest monitor and at city level. Key information is missing, such as confounder adjustment and sample size
Yang (杨敏娟) et al., 2016	Y	N	Insufficient information	Exposure increment was not provided in the article
Zhao et al., 2020a	Y	N	Insufficient information	Only season-specific estimates were available.
Zhao (赵姗) et al., 2020b	Y	N	High risk of bias Insufficient information	Insufficient information given for exposure assessment method (nearest monitor approach), such as interpolation distance, and level of exposure assignment
Zhu (朱娟) et al., 2011	Y	N	High risk of bias	Exposure assigned using nearest monitor and at city level, with very limited sample size.
Zou et al., 2021	Y	N	High risk of bias	Exposure assigned at district level instead of maternal address, and sample size relatively small.

Appendix 9. Studies that have conducted multipollutant analyses that were only graphically presented, on other pollutants or sources/constituents of $PM_{2.5}$.

Study	Reason that it was not discussed
Backwenkizi et al. 2022	All estimates provided were derived using MPM, so comparison can be made with SPM estimates.
Cai et al. 2020	The MPM estimates were for $PM_{2.5}$ exposure and its components only
Cassidy-Bushrow et al. 2020	The MPM was based on exposures to $PM_{2.5}$ in addition to NO_2 and PM_{10} as well as non-classical pollutants benzene, toluene, ethylbenzene, and xylenes (BTEX)
Ji et al. 2019	Estimates derived using MPM were only presented graphically
Le et al. 2012	The estimates presented were based on categorised levels of exposure.
Lin et al. 2015	Estimates derived using MPM were only presented graphically
Maroziene et al. 2002	The MPM was based on exposures to NO_2 and formaldehyde
Olsson et al. 2012	All estimates provided were derived using MPM
Olsson et al., 2013	*The MPM was based on exposures to* O_3 *and* NO_x
Ottone et al. 2020	The MPM estimates were for $PM_{2.5}$ exposure and its sources, but not NO_2 and O_3 exposure
Pereira et al. 2014b	Estimates derived using MPM were only presented graphically
Siddika et al. 2020	Estimates derived using MPM were only presented graphically
Transande et al. 2013	All estimates provided were derived using MPM
Wang et al. 2022	All estimates provided were derived using MPM
Wilhelm et al. 2011	The MPM estimates were for $PM_{2.5}$ exposure and its sources, but not NO_2 and O_3 exposure
Xu (徐丽珍) et al. 2008	The estimates presented were based on categorised levels of exposure
Yang (杨敏娟) et al. 2016	The increment used for deriving the effect estimates were unclear

Appendix 10. Studies that were included in the narrative synthesis that have provided results from single-pollutant models (SPM) and two/multipollutant models (MPMs) for whole pregnancy $PM_{2.5}$, NO_2, O_3 exposure.

Studies	Details of multipollutant models
Chen et al., 2018	Presented the results for both SPM (requested from authors) and two/multi-pollutant models. Correlation between O_3 and $PM_{2.5}$, NO_2 and $PM_{2.5}$, NO_2 and O_3 were 0.27, 0.38, -0.04 respectively. Statistically significant increase in risk of PTB was observed for whole pregnancy exposure to all 3 pollutants. Compared to the SPM results, there were minimal changes in the results for $PM_{2.5}$ after adjustment for O_3 and NO_2. The increase in risk associated with NO_2 exposure was no longer significant after adjustment for $PM_{2.5}$ or O_3. The increase in risk for O_3 exposure was attenuated after adjustment for $PM_{2.5}$ and NO_2 but remained statistically significant.
Ha et al., 2014	Presented the results for both SPM and two-pollutant models ($PM_{2.5}$ and O_3). Correlation between $PM_{2.5}$ and O_3 was 0.1 for whole pregnancy exposure. Statistically significant increase in risk of PTB was observed for whole pregnancy exposure to $PM_{2.5}$ and O_3, and the results were robust to co-pollutant adjustment.
Gan et al., 2022	Presented the results for both SPM and two-pollutant models ($PM_{2.5}$ and NO_2) ($PM_{2.5}$ and O_3) (NO_2 and O_3). Correlation between pollutants not given by authors. Statistically significant decrease in risks were observed for trimester 1 O_3 and trimester 2 $PM_{2.5}$ exposure. No change in results for NO_2 exposure after adjustment for co-pollutants. Mixed results observed for O_3 and $PM_{2.5}$ exposure in different trimesters.
Guo et al., 2019	Presented the results for both SPM (NO_2) and two-pollutant models (NO_2 and $PM_{2.5}$). Correlations between NO_2 and $PM_{2.5}$ not given. Statistically significant increase in risk of PTB found for trimester 1, 3 and whole pregnancy NO_2 exposure. After adjustment for $PM_{2.5}$, only the increase in risk in trimester 3 remained statistically significant.
Ji et al. 2019	Findings only graphically presented. Correlation between O_3 and $PM_{2.5}$, NO_2 and $PM_{2.5}$, NO_2 and O_3 were -0.64, 0.33, and -0.21 respectively. Minimal changes to risk of NO_2 exposure after adjustment for O_3 or $PM_{2.5}$.

Laurent et al. 2016	Presented the results for both SPM and two-pollutant models ($PM_{2.5}$ and NO_2) ($PM_{2.5}$ and O_3) (NO_2 and O_3). Correlation between O_3 and $PM_{2.5}$, NO_2 and $PM_{2.5}$, NO_2 and O_3 were 0.09, 0.81, -0.07 respectively. Statistically significant increase in risk of PTB was observed for whole pregnancy exposure to all 3 pollutants. Compared to the SPM results, there were minimal changes in the results for $PM_{2.5}$ after adjustment for O_3 and NO_2. However, there was no longer an increase in risk found for NO_2 exposure after adjustment (OR: 0.986 [0.971-1.001], per IQR increase in NO_2). There were minimal changes in the risks of NO_2 or O_3 exposure in two pollutant models including NO_2 and O_3.
Liang et al., 2019a	Presented the results for both SPM and two-pollutant models ($PM_{2.5}$ and NO_2) ($PM_{2.5}$ and O_3). Correlation between $PM_{2.5}$ and NO_2, $PM_{2.5}$ and O_3 were 0.51 and -0.19 respectively. Minimal changes to risk of $PM_{2.5}$ exposure after adjustment for O_3 or NO_2.
Liang et al., 2019b	Presented the results for both SPM and two-pollutant models ($PM_{2.5}$ and NO_2) ($PM_{2.5}$ and O_3) (NO_2 and O_3). Correlation between $PM_{2.5}$ and NO_2, $PM_{2.5}$ and O_3, NO_2 and O_3 were 0.44, 0.03, -0.39 respectively. Statistically significant increase in risk of O_3 exposure was found but not $PM_{2.5}$ and NO_2 exposure. Minimal changes to risk presented after adjustment for the other pollutants.
Liang et al., 2022	Presented the results for both SPM and two-pollutant models ($PM_{2.5}$ and NO_2) ($PM_{2.5}$ and O_3). Correlation between $PM_{2.5}$ and NO_2, $PM_{2.5}$ and O_3 were 0.47 and -0.105 respectively. Minimal changes to risk of $PM_{2.5}$ exposure after adjustment for O_3 or NO_2.
Qian et al., 2016	Presented the results for both SPM and two-pollutant models ($PM_{2.5}$ and NO_2) ($PM_{2.5}$ and O_3) (NO_2 and O_3). Correlations between $PM_{2.5}$ and NO_2, $PM_{2.5}$ and O_3, NO_2 and O_3 were 0.63, -0.16 and -0.12 respectively. Statistically significant increase in risk found for $PM_{2.5}$ and O_3 but not NO_2 exposure. Minimal changes to risk presented after adjustment for the other pollutants.
Shi et al. 2021	Presented the results for both SPM and two-pollutant models ($PM_{2.5}$ and NO_2) ($PM_{2.5}$ and O_3) (NO_2 and O_3). Correlations between $PM_{2.5}$ and NO_2, $PM_{2.5}$ and O_3, NO_2 and O_3 were 0.72, -0.46 and -0.37 respectively. Statistically significant increase in risk found for $PM_{2.5}$ but not NO_2 and O_3 exposure. Minimal changes to risk presented adjustment for the other pollutants.

Siddika et al. 2019	Presented the results for both SPM and two-pollutant models ($PM_{2.5}$ and O_3). Correlations between $PM_{2.5}$ and NO_2, $PM_{2.5}$ and O_3, NO_2 and O_3 were 0.92, -0.29 and -0.31 respectively. Minimal changes to risk presented adjustment for the other pollutants
Smith et al. 2020	Presented the results for both SPM and two-pollutant models ($PM_{2.5}$ and O_3) (NO_2 and O_3). Correlations between $PM_{2.5}$ and NO_2, $PM_{2.5}$ and O_3, NO_2 and O_3 were 0.66, -0.25 and -0.74 respectively. Statistically significant increase in risk found for O_3 exposure, and decrease in risk found for $PM_{2.5}$ and NO_2 exposure. Minimal changes to risk of $PM_{2.5}$ exposure presented adjustment for the O_3, and the risk of NO_2 exposure was no longer present after adjustment for O_3.
Stieb et al. 2016a	Presented the results for both SPM and two-pollutant models ($PM_{2.5}$ and NO_2). Correlations between $PM_{2.5}$ and NO_2 was 0.61. Minimal changes to the risk of $PM_{2.5}$ and NO_2 exposures presented adjustment for the other pollutants.
Yang et al. 2020	Presented the results for both SPM and two-pollutant models (O_3 and $PM_{2.5}$) (O_3 and NO_2). Correlation between pollutants were not given. Statistically significant increase in risk found for O_3 exposure. Minimal changes to the risk of O_3 exposure after adjustment for NO_2 and O_3 exposure
Ye et al. 2018	Presented the results for both SPM and MPM ($PM_{2.5}$ and PM_{10}/NO_2) (NO_2 and $PM_{2.5}/PM_{10}$). Correlations between pollutants were not given. Statistically significant increase in risk found for $PM_{2.5}$ and NO_2 exposure in SPMs. Minimal changes in the risk of $PM_{2.5}$ exposure found after adjustment for other pollutants, but 95% CI became narrower. The increase in risk associated with NO_2 exposure was attenuated after adjustment for other pollutants.
Zhou et al., 2022b	Presented the results for both SPM and MPM ($PM_{2.5}$ and $PM_{10}/SO_2/NO_2/O_3/CO$) (NO_2 and $PM_{2.5}/PM_{10}/SO_2/O_3/CO$) ($O_3$ and $PM_{2.5}/PM_{10}/SO_2/NO_2/CO$). Correlations between $PM_{2.5}$ and NO_2, $PM_{2.5}$ and O_3, NO_2 and O_3 were 0.48, -0.34 and -0.28 respectively. Statistically significant increase in risk found for $PM_{2.5}$ exposure in SPMs. Minimal changes to risk of $PM_{2.5}$ and NO_2 exposure after adjustment for other pollutants. Statistically significant increase in risk of O_3 exposure was found after adjustment for other pollutants.

Appendix 11. Sorting of studies eligible to be included in the meta-analyses of O_3 exposure by measurement metric.

8-hour maximum or 8-hour averages from 10:00 to 16:00	24-hour mean or other metrics e.g. monthly mean
Chen et al. (2018a) Chen et al. (2021a) doNascimento et al. (2022) Fang et al. (2022) Gan et al. (2022) Gong et al. (2022) Ha et al. (2014) Hansen et al. (2006) Hao et al. (2016) Laurent et al. (2016) Lavigne et al. (2016) Lin et al. (2015) Liu et al. (2019) Mekonnen et al. (2021) Olsson et al. (2013) Qian et al. (2016) Schifano et al. (2016) Sun et al. (2019) Symanski et al. (2015) Wilhelm et al. (2011) Wu et al. (2011) Yang et al. (2020) Zhou et al. (2022a)	Jalaludin et al. (2007) – daily 1 hour maximum Lee et al. (2013) – daily mean Liang et al. (2019) – daily mean Mendola et al. (2016) – daily mean Ritz et al. (2007) Siddika et al. (2020) – daily mean Smith et al. (2020) – monthly mean Wang et al. (2018) – daily mean Wu et al. (2016) – daily mean Williams et al. (2021) – daily mean Wilhelm et al. (2003) – annual average Zhou et al. (2022b) – daily mean

Appendix 12. Studies with estimates in population subgroups that can be meta-analysed using fixed-effect model.

Liang et al., 2019: NO_2 - PTB

Exposure period	Original estimates	Estimate derived from fixed-effect meta-analysis
Trimester 1	Local population: HR = 1.000 (0.985 – 1.015) Migrant population: HR = 1.050 (1.040 – 1.060)	HR = 1.036 (1.027 – 1.439)
Trimester 2	Local population: HR = 1.000 (0.990 – 1.010) Migrant population: HR = 1.050 (1.040 – 1.060)	HR = 1.026 (1.019 – 1.033)
Trimester 3	Local population: HR = 1.040 (1.030 – 1.050) Migrant population: HR = 1.100 (1.085 – 1.115)	HR = 1.059 (1.051 – 1.068)
Whole pregnancy	Local population: HR = 1.010 (0.995 – 1.025) Migrant population: HR = 1.010 (0.995 – 1.025)	HR = 1.010 (0.999 – 1.021)

Mendola et al., 2016: $PM_{2.5}$ – PTB

Exposure period	Original estimates	Estimate derived from fixed-effect meta-analysis
Whole pregnancy	Non-asthmatics: HR = 1.021 (0.979 – 1.066) Asthmatics: HR = 1.110 (0.982 – 1.254)	HR = 1.0305 (0.990 – 1.073)

Liang et al., 2022a: $PM_{2.5}$ - PTB

Exposure period	Original estimates	Estimate derived from fixed-effect meta-analysis
Trimester 1	Impaired fasting glucose: HR = 1.019 (1.008 – 1.031) Normal glucose level: HR = 1.002 (0.999 – 1.005)	HR = 1.003 (1.000 – 1.006)
Trimester 2	Impaired fasting glucose: HR = 1.013 (1.000 – 1.0266) Normal glucose level: HR = 1.007 (1.004 – 1.010)	HR = 1.073 (1.004 – 1.010)
Trimester 3	Impaired fasting glucose: HR = 1.012 (1.010 – 1.0145) Normal glucose level: HR = 1.012 (0.999 – 1.026)	HR = 1.012 (1.010 – 1.015)
Whole pregnancy	Impaired fasting glucose: HR = 1.024 (1.009 – 1.040) Normal glucose level: HR = 1.017 (1.014 – 1.021)	HR = 1.017 (1.014 – 1.021)

$PM_{2.5}$ – moderate to late PTB

Exposure period	Original estimates	Estimate derived from fixed-effect meta-analysis
Trimester 1	Impaired fasting glucose: HR = 1.047 (1.024 – 1.071) Normal glucose level: HR = 1.129 (1.000 – 1.274)	HR = 1.050 (1.027 – 1.073)
Trimester 2	Impaired fasting glucose: HR = 1.099 (1.071 – 1.128) Normal glucose level: HR = 1.041 (0.904 – 1.199)	HR = 1.097 (0.904 – 1.125)
Trimester 3	Impaired fasting glucose: HR = 1.135 (0.982 – 1.311) Normal glucose level: HR = 1.100 (1.104 – 1.163)	HR = 1.133 (1.104 – 1.163)
Whole pregnancy	Impaired fasting glucose: HR = 1.186 (1.004 – 1.402) Normal glucose level: HR = 1.132 (1.098 – 1.168)	HR = 1.134 (1.010 – 1.169)

Appendix 13. Information given by authors of primary studies used in the narrative synthesis or meta-analysis.

Basu et al., 2017

- Requested for the estimates presented graphically in Figures S2 and S4 for $PM_{2.5}$
- Authors provided the odds per IQR increase in $PM_{2.5}$

Outcome	Trimester 1	Trimester 2	Whole pregnancy	First month	Last month
PTB	3.77 (2.30-5.26)	4.11 (2.52-5.72)	16.41 (13.45-19.45)	2.81 (1.73-3.90)	1.96 (0.77-3.17)
Late PTB			13.94 (10.31-17.98)		
Moderate PTB			24.85 (18.97-31.03)		
Very PTB			23.21 (12.50-34.94)		
Extreme PTB			-8.13 (-19.81-4.81)		

Chen et al., 2018

- Requested for the associations between Trimester 1, 2, 3 exposure and the risk of preterm birth which were only graphically presented.
- Authors replied with the requested information

Period	$PM_{2.5}$	SO_2	NO_2	O_3
Trimester 1	0.99 (0.97-1.02)	1.03 (1.00-1.06)	0.98 (0.88-1.09)	1.05 (1.01-1.09)
Trimester 2	1.00 (0.97-1.02)	1.02 (1.00-1.05)	1.15 (1.05-1.27)	1.06 (1.02-1.10)
Trimester 3	1.01 (0.98-1.04)	1.05 (1.02-1.08)	1.18 (1.06-1.31)	1.25 (1.20-1.30)
Whole pregnancy	1.05 (1.02-1.08)	1.12 (1.09-1.16)	1.07 (1.01-1.13)	1.13 (1.10-1.16)

DeFranco et al., 2016

- Requested for the exposure increment used when $PM_{2.5}$ levels were modelled as a continuous variable in the secondary analysis conducted (page 5 and 6 of the article)
- Authors clarified that the exposure increment used was 1 $\mu g/m^3$

Gray et al., 2014

- Requested for the estimates for the effect of air pollutant ($PM_{2.5}$ and O_3) exposure on the risk of preterm birth without stratification by ethnicity
- Authors no longer has access to the data

Gong et al., 2022

- Requested for estimates presented graphically in Figure 2, and if exposure increments used for all pollutants were all per 1 μg/m^3.
- Authors provided the relevant information: exposure increment for was 1mg/m^3 for CO, 1 μg/m^3 for others

Period	PM_{10}	$PM_{2.5}$	SO_2
Trimester 1	0.9985 (0.9965-1.0005)	0.9978 (0.9942-1.0015)	1.0102 (1.0060-1.0144)
Trimester 2	0.9986 (0.9963-1.0007)	1.0035 (0.9997-1.0074)	1.0086 (1.0045-1.0127)
Trimester 3	0.9999 (0.9975-1.0023)	1.0000 (0.9969-1.0031)	1.0100 (1.0063-1.0138)
Whole pregnancy	0.9985 (0.9956-1.0014)	1.0010 (0.9956-1.0064)	1.0442 (1.0215-1.0674)

Period	NO_2	O_3	CO
Trimester 1	1.0188 (1.0129-1.0247)	0.9904 (0.9872-0.9936)	0.9385 (0.8215-1.0722)
Trimester 2	1.0176 (1.0114-1.0239)	0.9865 (0.9829-0.9901)	1.1381 (0.9975-1.2986)
Trimester 3	1.0171 (1.0115-1.0228)	0.9864 (0.9833-0.9895)	1.1419 (1.0130-1.2871)
Whole pregnancy	1.0280 (1.0356-1.0206)	0.9738 (0.9693-0.9784)	1.1358 (0.9592-1.3448)

Johnson et al., 2016

- Requested for the numerical estimates for early preterm birth (<32 weeks PTB) that were only presented graphically in Figure 3. Authors provided requested information.

Type of PTB	Model	Pollutant	Exposure	Estimate	se	or	LCL	UCL
spontaneous	unadjusted	NO_2	Trimester 1	-0.01285	0.003491	0.77333	0.67443	0.88674
spontaneous	adjusted	NO_2	Trimester 1	-0.00232	0.00456	0.95472	0.79845	1.14159
spontaneous	unadjusted	NO_2	Trimester 2	-0.01378	0.003498	0.75912	0.66184	0.87069
spontaneous	adjusted	NO_2	Trimester 2	-0.00283	0.004909	0.9449	0.77949	1.14542
all	unadjusted	NO_2	Trimester 1	-0.01591	0.003096	0.72753	0.64437	0.82142
all	adjusted	NO_2	Trimester 1	-0.00465	0.004011	0.91122	0.77864	1.06636
all	unadjusted	NO_2	Trimester 2	-0.01557	0.003096	0.73237	0.64868	0.82686
all	adjusted	NO_2	Trimester 2	-0.0042	0.004304	0.91943	0.77669	1.08839
spontaneous	unadjusted	$PM_{2.5}$	Trimester 1	-0.02207	0.008642	0.80199	0.67704	0.95002
spontaneous	adjusted	$PM_{2.5}$	Trimester 1	-0.00454	0.010677	0.95564	0.77519	1.17809

spontaneous	unadjusted	$PM_{2.5}$	Trimester 2	-0.02943	0.0092	0.74507	0.62213	0.89231
spontaneous	adjusted	$PM_{2.5}$	Trimester 2	-0.01373	0.011454	0.8717	0.69641	1.0911
all	unadjusted	$PM_{2.5}$	Trimester 1	-0.02543	0.007643	0.7755	0.6676	0.90083
all	adjusted	$PM_{2.5}$	Trimester 1	-0.00754	0.009411	0.92738	0.77116	1.11524
all	unadjusted	$PM_{2.5}$	Trimester 2	-0.03552	0.00816	0.70107	0.59745	0.82265
all	adjusted	$PM_{2.5}$	Trimester 2	-0.01972	0.010121	0.82102	0.67329	1.00115

Lin et al., 2015

- Requested for the numerical values of the estimates presented graphically in Figure 1 for all pollutants and exposure windows. Authors provided the releveant information

Period	PM_{10}	SO_2	NO_2	O_3	CO
Trimester 1	1.007 (1.003-1.012)	1.003 (1.000-1.007)	0.960 (0.948-0.972)	1.03 (1.02-1.038)	0.983 (0.979-0.986)
Trimester 2	0.995 (0.990-0.999)	1.001 (0.998-1.004)	0.927 (0.916-0.939)	1.02 (1.01-1.024)	0.978 (0.975-0.981)
Trimester 3	0.998 (0.993-1.002)	1.001 (0.997-1.004)	0.939 (0.927-0.951)	1.02 (1.01-1.028)	0.982 (0.978-0.985)

Nealy et al., 2018

- Requested to clarify if the interquartile ranges (IQR) used to derive the estimates are the ones shown in Supp Table S2 (and there were two tables named Table S2)
- Author clarified that the second Table S2 should be used, and the estimates were scaled by IQR

Pereira et al., 2016

- Requested for the numerical values of the estimates presented graphically in Figure 3 and 4, for trimester 1-3 and entire pregnancy.
- Authors no longer have access to the data

Poirier et al., 2015

- Requested for the interquartile range (IQR) used for deriving the estimates (OR^b and OR^c) per IQR increment shown in Figure 3
- Authors provided the IQR values (shown below) and units for the gaseous pollutants (ppb), and particulate matter ($\mu g/m^3$)

NO_2	SO_2	PM_{10}	PM_{25}	PM_1
3.4	0.5	0.4	0.2	0.2

Rappazzo et al., 2014; Rappazzo et al., 2017

- Rappazzo et al., 2014: Requested for the original risks that are used to derive the risk differences for large windows of exposures shown in Table S5
- Rappazzo et al., 2017: Requested for the original risks that are used to derive the risk differences are available for exposure for trimester 1-3 and the entire pregnancy.
- Author no longer has access to the data

Rudra et al., 2011

- Requested for the results of the analysis conducted for CO and $PM_{2.5}$ exposures in Trimester 1 and 2 and incidences of preterm delivery
- Authors no longer have access to the data

Savitz et al., 2019

- Requested to provide the levels of $PM_{2.5}$ and PM_{10} during the respective exposure windows of interest (Trimester 1 and 2), or the levels of $PM_{2.5}$ and PM_{10} over the whole study period
- Author no longer has access to the data but recommended publications that may contain the information. Information was not retrieved from the recommended publications

Stieb et al., 2008

- Requested for estimates per trimester that are presented graphically in Figure 1 for preterm birth.
- Information was no longer available to the authors

Stieb et al., 2016

- Requested for estimates that were only presented graphically in Figure 1 for exposures in Trimester 1-3 for preterm birth (Page 246).
- Information was no longer available to the authors

Tu & Tu, 2018

- Requested for classification for methodology used and if a 95% CI can be provided for the OR.
- Authors explained the methodology generates a set of regression results for each case, thus there is an OR and 95%CI for each birth. Therefore only the range of ORs were presented. The technique is descriptive to show how the relationship between PTB and its risk factors varied spatially, instead of predicting risk of PTB, thus CIs were not shown.

Weber et al., 2019

- Requested for clarification regarding the exposure increment used for the estimates shown in Table 3.
- Estimates were for a comparison between the top quartile and the lower three

Xiao et al., 2018

- Numerical values of the estimates presented graphically for preterm birth in Figures 2 and 3, for all the exposure assessment approaches.

Figure 2	Trimester 1: 1.15 (1.10-1.20) Trimester 2: 1.08 (1.03-1.12) Trimester 3: 1.06 (1.03-1.08) Whole pregnancy: 1.27 (1.03-1.36)

	Window	Gap-filled satellite	Gap-filled satellite with missingness filled	Monitor measurements
Figure 3	Trimester 1	1.11 (1.04-1.19)	1.06 (1.01-1.13)	1.11 (1.01-1.22)
	Trimester 2	1.10 (1.03-1.18)	1.06 (1.01-1.11)	1.06 (0.96-1.18)
	Trimester 3	1.05 (1.01-1.09)	1.06 (1.03-1.08)	1.03 (0.99-1.08)
	Whole pregnancy	1.19 (1.05-1.33)	1.10 (1.03-1.17)	1.13 (0.99-1.29)

Appendix 14. The process of sorting study dependencies in the main meta-analysis conducted.

Green: Studies that were included in the meta-analysis after sorting of study dependencies. Grey: Studies that were excluded from the meta-analysis after sorting of study dependencies. The studies were ranked according to risk of bias then sample size (details in Chapter 5 section x)

Country	Ranking	Study	Location	Time period	Decision
Trimester 1 or 2 $PM_{2.5}$ exposure and the risk of PTB (HR)					
China	1	He et al. (2022)	National	01/2010 - 12/2015	Among the 4 multiprovince/national studies, He et al. (2022) was kept as it was ranked the highest. The exact 8 provinces used in Chu et al. (2021) was unknown. Thus, it was excluded due to >50% temporal dependencies with He et al. (2022).
	2	Li et al. (2018)	National	12/2013 - 30/2014	
	3	Chu et al. (2021)	National	01/2009 - 02/2011	
	5	Guo et al. (2018)	National	01/2014 - 12/2014	
	4	Wang et al. (2018)	Guangzhou City, Guangdong	01/2015 - 07/2017	Wang et al. (2018) was kept as it was ranked the highest. Liang et al. (2022) was kept as it was the most preferred province-wide study.
	6	Liang et al. (2022)	21 cities (inc. Guangzhou), Guangdong	01/2014 - 12/2017	
	9	Liang et al. (2019a)	9 cities (inc. Guangzhou), Guangdong	01/2014 - 12/2017	
	10	Liang et al. (2019b)	Guangdong	01/2015 - 12/2017	
	7	Chen et al. (2021b)	Shiyan City, Hubei	01/2015 - 12/2017	No study dependencies in this group of studies
	8	Chen et al. (2021a)	Tianjin City	01/2014 - 06/2016	
	11	Yuan et al. (2020)	Shanghai City	01/2013 - 12/2016	
Australia	/	Chen et al. (2018a)	Brisbane	07/2003 - 12/2013	N/A
US	/	Sheridan et al. (2019)	California	01/2005 - 02/2010	N/A
Trimester 3 $PM_{2.5}$ exposure and the risk of PTB (HR)					
China	1	He et al. (2022)	National	01/2010 - 12/2015	Among the 4 multiprovince/national studies, He et al. (2022) was kept as it was ranked the highest. The exact 8 provinces used in Chu et al. (2021) was unknown. Thus, it was excluded due to >50% temporal dependencies with He et al. (2022).
	2	Li et al. (2018)	National	12/2013 - 30/2014	
	3	Chu et al. (2021)	National	01/2009 - 02/2011	
	5	Guo et al. (2018)	National	01/2014 - 12/2014	
	4	Wang et al. (2018)	Guangzhou City, Guangdong	01/2015 - 07/2017	Wang et al. (2018) was kept as it was the most preferred study. Liang et al. (2022) was kept as it was the most preferred province-wide study.
	6	Liang et al. (2022)	21 cities (inc. Guangzhou), Guangdong	01/2014 - 12/2017	
	10	Liang et al. (2019a)	9 cities (inc. Guangzhou). Guangdong	01/2014 - 12/2017	
	11	Liang et al. (2019b)	Guangdong	01/2015 - 12/2017	
	7	Fang et al. (2022)	Beijing City	01/2014 - 12/2017	No study dependencies in this group of studies
	8	Chen et al. (2021b)	Shiyan City, Hubei	01/2015 - 12/2017	

Country	Ranking	Study	Location	Time period	Decision
	9	Chen et al. (2021a)	Tianjin City	01/2014 - 06/2016	
	12	Yuan et al. (2020)	Shanghai City	01/2013 - 12/2016	
Australia	/	Chen et al. (2018a)	Brisbane	07/2003 - 12/2013	N/A
US	/	Sheridan et al. (2019)	California	01/2005 - 02/2010	N/A
Canada	/	Lavigne et al. (2018)	Ontario	01/2006 - 03/2012	N/A
Whole pregnancy $PM_{2.5}$ exposure and the risk of PTB (HR)					
China	1	He et al. (2022)	National	01/2010 - 12/2015	Among the 5 multiprovince/national studies, He et al. (2022) was kept as it was ranked the highest. The exact 8 provinces used in Chu et al. (2021) was unknown. Thus, it was excluded due to >50% temporal dependencies with He et al. (2022).
	2	Li et al. (2018)	National	12/2013 - 30/2014	
	3	Chu et al. (2021)	National	01/2009 - 02/2011	
	5	Guo et al. (2018)	National	01/2014 - 12/2014	
	11	Li et al. (2020b)	National	01/2013 - 11/2014	
	4	Wang et al. (2018)	Guangzhou City, Guangdong	01/2015 - 07/2017	Wang et al. (2018) was kept as it was the most preferred study. Liang et al. (2022) was kept as it was the most preferred province-wide study.
	6	Liang et al. (2022)	21 cities (inc. Guangzhou), Guangdong	01/2014 - 12/2017	
	10	Wang et al. (2020)	Guangdong	01/2015 - 07/2017	
	9	Liang et al. (2019b)	Guangdong	01/2015 - 12/2017	
	7	Chen et al. (2021b)	Shiyan City, Hubei	01/2015 - 12/2017	No study dependencies in this group of studies
	8	Chen et al. (2021a)	Tianjin City	01/2014 - 06/2016	
	12	Yuan et al. (2020)	Shanghai City	01/2013 - 12/2016	
Australia	/	Chen et al. (2018a)	Brisbane	07/2003 - 12/2013	N/A
US	/	Sheridan et al. (2019)	California	01/2005 - 02/2010	
Canada	/	Lavigne et al. (2018)	Ontario	01/2006 - 03/2012	
Trimester 1 $PM_{2.5}$ exposure and the risk of PTB (OR)					
China	3	Cai et al. (2020)	National	01/2015 - 12/2016	Kept because it is the only national study
	2	Zhou et al. (2022)	Henan	05/2014 - 12/2016	Zhou et al. (2022a) was kept as it was ranked higher
	4	Gong et al. (2022)	Henan	05/2014 - 12/2016	
	5	Ye et al. (2018)	Taizhou City, Zhejiang	01/2013 - 05/2016	Both studies were kept as one was province-wide, and one was in Taizhou only.
	8	Sun et al. (2019)	Zhejiang	01/2013 - 07/2017	
	7	Fang et al. (2022)	Beijing City	01/2014 - 12/2017	Fang et al. (2022) was kept as it was ranked higher. Zhang et al. (2021) had extremely small sample size
	9	Zhang et al. (2021)	Beijing City	02/2017 - 10/2018	
	1	Xiao et al. (2018)	Shanghai City	01/2011 - 12/2014	No study dependencies in this group of studies
	10	Zhou et al. (2022b)	Chongqing City	01/2015 - 12/2020	

Country	Ranking	Study	Location	Time period	Decision
	11	Zhao et al. (2020a)	Hebei Province	01/2015-12/2016	Studies assigned a high risk of bias
	12	Gu et al. (2019)	Shanghai City	01/2015-12/2016	
	13	Liu et al. (2019b)	Guangdong	01/2014-12/2015	
	14	Gan et al. (2022)	Guangzhou City, Guangdong	01/2017-04/2018	
Australia	/	Jalaludin et al. (2017)	Sydney, New South Wales	01/1998 - 12/2000	No study dependencies in this group of studies
	/	Pereira et al. (2014b)	Perth	01/1997 - 12/2007	
Canada	/	Lavigne et al. (2018)	Ontario	01/2006 - 03/2012	
Peru	/	Tapia et al. (2020)	Lima	01/2012 - 12/2016	
South Korea	/	Kwag et al. (2021)	Seoul	01/2010 - 12/2016	
Europe	/	Giorgis-Allemand et al. (2017)	11 countries (inc. Netherlands)	01/1994 - 12/2010	Gehring et al. (2011a) was removed as the cohort used was embedded in Giorgis-Allemand et al. (2017). Giorgis-Allemand et al. (2017) did not overlap with Smith et al. (2020), as it used the Born-in-Bradford study for the UK.
Netherlands	/	Gehring et al. (2011a)	National	01/1996 - 12/1997	
England	/	Smith et al. (2020)	London	01/2006 - 12/2010	
US	1	Wu et al. (2009)	Los Angeles & Orange, California	01/1997 - 12/2006	Wu et al. (2009), Mekonnen et al. (2021), and Basu et al. (2017) were kept as they were ranked highest and not temporal dependent of each other
	5	Mekonnen et al. (2021)	6 counties, California	01/2007 - 12/2011	
	8	Basu et al. (2017)	7 counties, California	01/2000 - 12/2006	
	13/14	Wu et al. (2011)	Los Angeles & Orange, California	01/1997 - 12/2006	
	15	Ritz et al. (2007)	Los Angeles, California	01/2003 - 12/2003	
	18	Wilhelm et al. (2005)	Los Angeles, California	01/1994 - 12/2000	
	9	Chang et al. (2015)	Atlanta City, Georgia	01/1999 - 12/2005	Both Hao et al. (2016) and Chang et al. (2015) were kept as one was statewide, and one was single city. Nealy et al. (2018) was excluded due to high RoB and overlapped with Chang et al. (2015)
	16	Hao et al. (2016)	Georgia	01/2002 - 02/2006	
	19	Nealy et al. (2018)	Atlanta City, Georgia	01/2002 - 02/2006	
	6	Hyder et al. (2014)	Connecticut, Massachusetts	01/2000 - 12/2006	Pereira et al. (2014a) removed due to lower rank and covered less area
	11	Pereira et al. (2014a)	Connecticut	01/2001 - 12/2006	
	4	Johnson et al. (2016)	New York City, New York	01/2008 - 12/2010	No geographical dependencies
	10	Pereira et al. (2016)	Finger Lake region, New York	01/2004 - 12/2012	
	2	Kingsley et al. (2017)	Providence, Rhodes Island	01/2001 - 12/2012	No study dependencies in this group of studies
	3	Ha et al. (2014)	Florida	01/2004 - 12/2005	
	7	Lee et al. (2013)	Allegheny, Virginia	01/1997 - 12/2002	
	12	DeFranco et al. (2016)	Ohio	01/2007 - 12/2010	
	17	Cassidy-Bushrow et al. (2020)	Detroit, Michigan	06/2008 - 05/2010	

Country	Ranking	Study	Location	Time period	Decision
	20	Williams et al. (2021)	National	01/2002 – 12/2008	Study assigned a high RoB
Trimester 2 $PM_{2.5}$ exposure and the risk of PTB (OR)					
China	4	Cai et al. (2020)	National	01/2015 - 12/2016	Kept because it is the only national study
	2	Zhou et al. (2022a)	Henan	05/2014 - 12/2016	Zhou et al. (2022a) was kept as it was ranked higher
	5	Gong et al. (2022)	Henan	05/2014 - 12/2016	
	6	Ye et al. (2018)	Taizhou City, Zhejiang	01/2013 - 05/2016	Both studies were kept as one was province-wide, and one was in Taizhou only.
	9	Sun et al. (2019)	Zhejiang	01/2013 - 07/2017	
	3	Qian et al. (2016)	Wuhan City, Hubei	06/2011 - 06/2013	Both kept as they were not temporally dependent
	7	Zhang et al. (2020)	Wuhan City, Hubei	11/2013 - 08/2015	
	8	Fang et al. (2022)	Beijing City	01/2014 - 12/2017	Fang et al. (2022) was kept as it was ranked higher. Zhang et al. (2021) had extremely small sample size
	10	Zhang et al. (2021)	Beijing City	02/2017 - 10/2018	
	1	Xiao et al. (2018)	Shanghai City	01/2011 - 12/2014	No study dependencies in this group of studies
	11	Zhou et al. (2022b)	Chongqing City	01/2015 - 12/2020	
	12	Zhao et al. (2020a)	Hebei	01/2015 - 12/2016	Studies assigned a high RoB
	13	Gu et al. (2019)	Shanghai City	01/2015-12/2016	
	14	Liu et al. (2019b)	Guangdong	01/2014-12/2015	
	15	Gan et al. (2022)	Guangzhou City, Guangdong	01/2017-04/2018	
Canada	1	Lavigne et al. (2018)	Ontario	01/2006 - 03/2012	Lavigne et al. (2018) selected due to better exposure assessment method
	2	Lavigne et al. (2016)	Ontario	01/2005 - 03/2012	
Australia	/	Pereira et al. (2014b)	Perth	01/1997 - 12/2007	No study dependencies in this group of studies.
South Korea	/	Kwag et al. (2021)	Seoul	01/2010 - 12/2016	
Peru	/	Tapia et al. (2020)	Lima	01/2012 - 12/2016	
Europe	/	Giorgis-Allemand et al. (2017)	11 countries (inc. England)	01/1994 - 12/2010	Giorgis-Allemand et al. (2017) did not overlap with Smith et al. (2020), as it used the Born-in-Bradford study for the UK.
England	/	Smith et al. (2020)	London	01/2006 - 12/2010	
US	1	Wu et al. (2009)	Los Angeles & Orange, California	01/1997 - 12/2006	Wu et al. (2009), Mekonnen et al. (2021), and Basu et al. (2017) were kept as they were ranked highest and not temporal dependent of each other
	5	Mekonnen et al. (2021)	6 counties, California	01/2007 - 12/2011	
	7	Basu et al. (2017)	7 counties, California	01/2000 - 12/2006	
	12/13	Wu et al. (2011)	Los Angeles & Orange, California	01/1997 - 12/2006	
	8	Chang et al. (2015)	Atlanta City, Georgia	01/1999 - 12/2005	Both kept as one was statewide, and one was single city
	14	Hao et al. (2016)	Georgia	01/2002 - 02/2006	
	17	Nealy et al. (2018)	Atlanta City, Georgia	01/2002 - 02/2006	
	6	Hyder et al. (2014)	Connecticut, Massachusetts	01/2000 - 12/2006	

Country	Ranking	Study	Location	Time period	Decision
	10	Pereira et al. (2014a)	Connecticut	01/2001 - 12/2006	Pereira et al. (2014a) removed due to lower rank and covered less area
	4	Johnson et al. (2016)	New York City, New York	01/2008 - 12/2010	Savitz et al. (2019) removed as it was ranked lower than Johnson et al. (2016)
	9	Pereira et al. (2016)	Finger Lake region, New York	01/2004 - 12/2012	
	15	Savitz et al. (2019)	New York City, New York	01/2008 - 12/2010	
	2	Kingsley et al. (2017)	Providence, Rhodes Island	01/2001 - 12/2012	No study dependencies in this group of studies
	3	Ha et al. (2014)	Florida	01/2004 - 12/2005	
	11	DeFranco et al. (2016)	Ohio	01/2007 - 12/2010	
	16	Cassidy-Bushrow et al. (2020)	Detroit, Michigan	06/2008 - 05/2010	
	18	Williams et al. (2021)	National	01/2002 – 12/2008	Study assigned a high risk of bias
Trimester 3 $PM_{2.5}$ exposure and the risk of PTB (OR)					
Canada	/	Lavigne et al. (2016)	Ontario	01/2005 - 03/2012	No study dependencies in this group of studies
Australia	/	Pereira et al. (2014b)	Perth	01/1997 - 12/2007	
South Korea	/	Kwag et al. (2021)	Seoul	01/2010 - 12/2016	
Peru	/	Tapia et al. (2020)	Lima	01/2012 - 12/2016	
China	3	Cai et al. (2020)	National	01/2015 - 12/2016	Kept because it is the only national study
	2	Zhou et al. (2022a)	Henan	05/2014 - 12/2016	Zhou et al. (2022a) was kept as it was ranked higher
	4	Gong et al. (2022)	Henan	05/2014 - 12/2016	
	5	Ye et al. (2018)	Taizhou City, Zhejiang	01/2013 - 05/2016	Both studies were kept as one was province-wide, and one was in Taizhou only.
	8	Sun et al. (2019)	Zhejiang	01/2013 - 07/2017	
	7	Fang et al. (2022)	Beijing City	01/2014 - 12/2017	Fang et al. (2022) was kept as it was ranked higher. Zhang et al. (2021) had extremely small sample size, and Han (韩慧) et al. (2017) was assigned a high risk of bias.
	9	Zhang et al. (2021)	Beijing City	02/2017 - 10/2018	
	14	Han (韩慧) et al. (2017)	Beijing City	01/2013 - 12/2014	
	1	Xiao et al. (2018)	Shanghai City	01/2011 - 12/2014	No study dependencies in this group of studies
	6	Zhang et al. (2020)	Wuhan City, Hubei	11/2013 - 08/2015	
	10	Zhou et al. (2022b)	Chongqing City	01/2015 - 12/2020	
	11	Zhao et al. (2020a)	Hebei	01/2015 - 12/2016	Studies assigned a high risk of bias
	12	Gu et al. (2019)	Shanghai City	01/2015 - 12/2016	
	13	Liu et al. (2019b)	Guangdong	01/2014 - 12/2015	
	15	Gan et al. (2022)	Guangzhou City, Guangdong	01/2017 - 04/2018	
US	2	Kloog et al. (2012)	Massachusetts	01/2000 - 12/2008	

Country	Ranking	Study	Location	Time period	Decision
	5	Hyder et al. (2014)	Connecticut, Massachusetts	01/2000 - 12/2006	Hyder et al. (2014) removed as it was ranked lower than Kloog et al. (2012) and covered a shorter period and allowed Pereira et al. (2014a) to be included.
	8	Pereira et al. (2014a)	Connecticut	01/2001 - 12/2006	
	6	Chang et al. (2015)	Atlanta City, Georgia	01/1999 - 12/2005	Both kept as one was statewide, and one was single city
	10	Hao et al. (2016)	Georgia	01/2002 - 02/2006	
	1	Kingsley et al. (2017)	Providence, Rhodes Island	01/2001 - 12/2012	No study dependencies in this group of studies
	3	Ha et al. (2014)	Florida	01/2004 - 12/2005	
	4	Mekonnen et al. (2021)	6 counties, California	01/2007 - 12/2011	
	7	Pereira et al. (2016)	Finger Lake region, New York	01/2004 - 12/2012	
	9	DeFranco et al. (2016)	Ohio	01/2007 - 12/2010	
	11	Cassidy-Bushrow et al. (2020)	Detroit, Michigan	06/2008 - 05/2010	
Whole pregnancy $PM_{2.5}$ exposure and the risk of PTB (OR)					
Australia	/	Melody et al. (2020)	Victoria	03/2012 - 12/2015	Melody et al. (2019) was removed due to high RoB assigned. Pereira et al. (2014b) were not dependent with Melody et al. (2020).
	/	Pereira et al. (2014b)	Perth	01/1997 - 12/2007	
	/	Melody et al. (2019)	Victoria	03/2012 - 12/2015	
India	/	Fleischer et al. (2014)	National	09/2004 - 04/2008	No study dependencies in this group of studies
Peru	/	Tapia et al. (2020)	Lima	01/2012 - 12/2016	
Canada	1	Stieb et al. (2016b)	National	01/1999 - 12/2008	Stieb et al. (2016b) kept as it was ranked the highest
	2	Stieb et al. (2016a)	National	01/1999 - 12/2008	
	3	Villeneuve et al. (2022)	National	01/1998 - 12/2008	
	4	Lavigne et al. (2016)	Ontario	01/2005 - 03/2012	No study dependencies in this group of studies
	5	Brauer et al. (2008)	Vancouver, British Columbia	01/1999 - 12/2002	
	6	Poirier et al. (2015)	Halifax, Nova Scotia	01/2008 - 23/2012	
China	5	Cai et al. (2020)	National	01/2015 - 12/2016	Kept because it is the only national study
	2	Zhou et al. (2022a)	Henan	05/2014 - 12/2016	Zhou et al. (2022a) was kept as it was ranked higher
	6	Gong et al. (2022)	Henan	05/2014 - 12/2016	
	7	Ye et al. (2018)	Taizhou City, Zhejiang	01/2013 - 05/2016	Both studies were kept as one was province-wide, and one was in Taizhou only.
	9	Sun et al. (2019)	Zhejiang	01/2013 - 07/2017	
	3	Li et al. (2020a)	Hubei	01/2014 - 12/2014	All were kept as Li et al. (2020a) were province-wide, whereas the other two were at city-level that have no temporal dependencies.
	4	Qian et al. (2016)	Wuhan City, Hubei	06/2011 - 06/2013	
	8	Zhang et al. (2020)	Wuhan City, Hubei	11/2013 - 08/2015	

Country	Ranking	Study	Location	Time period	Decision
	10	Zhang et al. (2021)	Beijing City	02/2017 - 10/2018	No study dependencies in this group of studies
	1	Xiao et al. (2018)	Shanghai City	01/2011 - 12/2014	
	11	Zhou et al. (2022b)	Chongqing City	01/2015 - 12/2020	
	12	Zhao et al. (2020a)	Hebei	01/2015 - 12/2016	Studies assigned a high RoB
	13	Liu et al. (2019b)	Guangdong	01/2014 - 12/2015	
US	16	Mendola et al. (2016)	National	01/2005 - 12/2007	Kept because it is the only national study
	1	Wu et al. (2009)	Los Angeles & Orange, California	01/1997 - 12/2006	Wu et al. (2009), Mekonnen et al. (2021), and Basu et al. (2017) were kept as they were ranked highest and not temporal dependent of each other. Huang et al. (2018) was excluded as it was assigned a high RoB, and was dependent with Mekonnen et al. (2021) and Basu et al. (2017).
	5	Mekonnen et al. (2021)	6 counties, California	01/2007 - 12/2011	
	7	Basu et al. (2017)	7 counties, California	01/2000 - 12/2006	
	12/13	Wu et al. (2011)	Los Angeles & Orange, California	01/1997 - 12/2006	
	14	Benmarhnia et al. (2017)	California	01/2005 - 12/2010	
	17	Wilhelm et al. (2011)	Los Angeles, California	06/2004 - 03/2006	
	18	Huynh et al. (2006)	California	01/1999 - 12/2000	
	20	Wu et al. (2016)	California	01/2001 - 12/2008	
	21	Laurent et al. (2016)	California	01/2001 - 12/2008	
	23	Huang et al. (2018)	California	01/2009 - 12/2012	
	3	Kloog et al. (2012)	Massachusetts	01/2000 - 12/2008	Hyder et al. (2014) removed as it was ranked lower than Kloog et al. (2012) and covered a shorter period and allowed Pereira et al. (2014a) to be included.
	6	Hyder et al. (2014)	Connecticut, Massachusetts	01/2000 - 12/2006	
	10	Pereira et al. (2014a)	Connecticut	01/2001 - 12/2006	
	8	Chang et al. (2015)	Atlanta City, Georgia	01/1999 - 12/2005	Both kept as one was statewide, and one was single city
	15	Hao et al. (2016)	Georgia	01/2002 - 02/2006	
	2	Kingsley et al. (2017)	Providence, Rhodes Island	01/2001 - 12/2012	No study dependencies in this group of studies
	4	Ha et al. (2014)	Florida	01/2004 - 12/2005	
	9	Pereira et al. (2016)	Finger Lake region, New York	01/2004 - 12/2012	
	11	DeFranco et al. (2016)	Ohio	01/2007 - 12/2010	
	19	Cassidy-Bushrow et al. (2020)	Detroit, Michigan	06/2008 - 05/2010	
	22	Kirwa et al. (2021)	Puerto Rico	08/2011 - 08/2018	Assigned a high RoB
	24	Williams et al. (2021)	National	01/2002 - 12/2008	
Europe	/	Giorgis-Allemand et al. (2017)	11 countries (inc. Netherlands, Finland)	01/1994 - 12/2010	Giorgis-Allemand et al. (2017) was kept as it is Europe-wide and covered a long time period. Gehring et al. (2011a) was removed as the cohort
Netherlands	/	Gehring et al. (2011a)	National	01/1996 - 12/1997	

Country	Ranking	Study	Location	Time period	Decision
Italy	/	Ottone et al. (2020)	Emilia-Romagna region	01/2012 - 03/2014	used was embedded in Giorgis-Allemand et al. (2017). All others were not temporally dependent.
Finland	/	Siddika et al. (2019)	Espoo	01/1984 - 03/1990	
Trimester 1/2 NO_2 exposure and the risk of PTB (HR)					
Australia	/	Chen et al. (2018a)	Brisbane	07/2003 - 12/2013	N/A
China	/	Chen et al. (2021b)	Shiyan City, Hubei	01/2015 - 12/2017	No study dependencies in this group of studies
	/	Chen et al. (2021a)	Tianjin City	01/2014 - 06/2016	
	/	Liang et al. (2019b)	Guangdong	01/2015 - 12/2017	Wang et al. (2018) was removed despite being based on just Guangzhou City, as the birth registry used in Wang et al. (2018) was covered by Liang et al. (2019b).
	/	Wang et al. (2018)	Guangzhou City, Guangdong	01/2015 - 12/2017	
Trimester 3 NO_2 exposure and the risk of PTB (HR)					
Australia	/	Chen et al. (2018a)	Brisbane	07/2003 - 12/2013	N/A
China	/	Chen et al. (2021b)	Shiyan City, Hubei	01/2015 - 12/2017	No study dependencies in this group of studies
	/	Chen et al. (2021a)	Tianjin City	01/2014 - 06/2016	
	/	Fang et al. (2022)	Beijing City	01/2014 - 12/2017	
	/	Liang et al. (2019b)	Guangdong	01/2015 - 12/2017	Wang et al. (2018) was removed despite being based on just Guangzhou City, as the birth registry used in Wang et al. (2018) was covered by Liang et al. (2019b).
	/	Wang et al. (2018)	Guangzhou City, Guangdong	01/2015 - 12/2017	
Whole pregnancy NO_2 exposure and the risk of PTB (HR)					
Australia	/	Chen et al. (2018a)	Brisbane	07/2003 - 12/2013	N/A
China	/	Chen et al. (2021b)	Shiyan City, Hubei	01/2015 - 12/2017	Wang et al. (2018) was removed despite being based on just Guangzhou City, as the birth registry used in Wang et al. (2018) was covered by Liang et al. (2019b).
	/	Chen et al. (2021a)	Tianjin City	01/2014 - 06/2016	
	/	Liang et al. (2019b)	Guangdong	01/2015 - 12/2017	
	/	Wang et al. (2018)	Guangzhou City, Guangdong	01/2015 - 12/2017	
US	/	Benmarhnia et al. (2017)	California	01/2005 - 12/2010	N/A
Spain/Italy	/	Schifano et al. (2016)	Barcelona, Rome	04/2001 - 10/2012	N/A
Trimester 1 NO_2 exposure and the risk of PTB (OR)					
China	2	Zhou et al. (2022a)	Henan	05/2014 - 12/2016	Zhou et al. (2022a) was kept as it was ranked higher. Lu et al. (2019) was kept as it was not temporally dependent from other studies
	5	Gong et al. (2022)	Henan	05/2014 - 12/2016	
	9	Lu et al. (2019)	Changsha City, Henan	09/2011 - 31/2012	

Country	Ranking	Study	Location	Time period	Decision
	7	Ye et al. (2018)	Taizhou City, Zhejiang	01/2013 - 05/2016	Both studies were kept as one was province-wide, and one was in Taizhou only.
	8	Sun et al. (2019)	Zhejiang	01/2013 - 07/2017	
	1	Ji et al. (2019)	Shanghai City	01/2014 - 12/2015	Zou et al. (2021) was removed as it was assigned a high RoB + ranked lower than Ji et al. (2019).
	10	Zou et al. (2021)	Shanghai City	04/2011 - 04/2012	
	3	Zhou et al. (2022b)	Chongqing City	01/2015 - 12/2020	No study dependencies in this group of studies
	4	Qian et al. (2016)	Wuhan City, Hubei	06/2011 - 06/2013	
	6	Huang et al. (2015)	Beijing City	01/2007 - 12/2010	
	11	Liu et al. (2019b)	Guangdong	01/2014 - 12/2015	Studies assigned a high RoB
	12	Gan et al. (2022)	Guangzhou City, Guangdong	01/2017 - 04/2018	
Lithuania	1	Dedele et al. (2017)	Kaunas	01/2007 - 12/2008	Dedele et al. (2017) selected over Dedele et al. (2011) due to better adjustment of confounder
	2	Maroziene et al. (2002)	Kaunas	01/1998 - 12/1998	
	3	Dedele et al. (2011)	Kaunas	01/2008 - 12/2009	
US	3/4	Wu et al. (2011)	Los Angeles & Orange, California	01/1997 - 12/2006	Wu et al. (2011) was ranked higher than Ritz et al. (2007) and covered its entire study period. Hence Ritz et al. (2007) was removed
	5	Ritz et al. (2007)	Los Angeles, California	01/2003 - 12/2003	
	1	Johnson et al. (2016)	New York City, New York	01/2008 - 12/2010	No study dependencies in this group of studies
	2	Hao et al. (2016)	Georgia	01/2002 - 02/2006	
	6	Cassidy-Bushrow et al. (2020)	Detroit, Michigan	06/2008 - 05/2010	
Europe	/	Giorgis-Allemand et al. (2017)	11 countries (inc. England)	01/1994 - 12/2010	Estarlich et al. (2016) and Gehring et al. (2011a) were removed as its cohort was embedded in Giorgis-Allemand et al. (2017). All others were not temporally dependent. Barba-Vasseur et al. (2017) was removed as it was assigned a high RoB.
England	/	Smith et al. (2020)	London	01/2006 - 12/2010	
Netherlands	/	Gehring et al. (2011a)	National	01/1996 - 12/1997	
Spain	/	Estarlich et al. (2016)	Valencia, Asturias, Gipuzkos, Sabadell	05/2004 - 08/2008	
France	/	Barba-Vasseur et al. (2017)	Besançon, Dijon	01/2005 - 12/2009	
Taiwan	/	Lin et al. (2015)	National	01/2001 - 12/2007	No study dependencies in this group of studies
Australia	/	Jalaludin et al. (2007)	Sydney, New South Wales	01/1998 - 12/2000	
	/	Hansen et al. (2006)	Brisbane	07/2000 - 06/2003	
Brazil	/	doNascimento et al. (2022)	São Paulo	01/2010 - 12/2016	
Trimester 2 NO_2 exposure and the risk of PTB (OR)					
China	/	Zhou et al. (2022a)	Henan	05/2014 - 12/2016	
	/	Gong et al. (2022)	Henan	05/2014 - 12/2016	

Country	Ranking	Study	Location	Time period	Decision
	/	Lu et al. (2019)	Changsha City, Henan	09/2011 - 31/2012	Zhou et al. (2022a) was kept as it was ranked higher. Lu et al. (2019) was kept as it was not temporally dependent from other studies
	/	Ye et al. (2018)	Taizhou City, Zhejiang	01/2013 - 05/2016	Both studies were kept as one was province-wide, and one was in Taizhou only.
	/	Sun et al. (2019)	Zhejiang	01/2013 - 07/2017	
	/	Zhou et al. (2022b)	Chongqing City	01/2015 - 12/2020	No study dependencies in this group of studies
	/	Huang et al. (2015)	Beijing City	01/2007 - 12/2010	
	/	Ji et al. (2019)	Shanghai City	01/2014 - 12/2015	Zou et al. (2021) was removed as it was assigned a high RoB.
	/	Zou et al. (2021)	Shanghai City	04/2011 - 04/2012	
	/	Liu et al. (2019b)	Guangdong	01/2014 - 12/2015	Studies assigned a high RoB
	/	Gan et al. (2022)	Guangzhou City, Guangdong	01/2017 - 04/2018	
US	1	Johnson et al. (2016)	New York City, New York	01/2008 - 12/2010	Savitz et al. (2019) removed as it was ranked lower than Johnson et al. (2016)
	3	Savitz et al. (2019)	New York City, New York	01/2008 - 12/2010	
	2	Hao et al. (2016)	Georgia	01/2002 - 02/2006	No study dependencies in this group of studies
	4/5	Wu et al. (2011)	Los Angeles & Orange, California	01/1997 - 12/2006	
	6	Cassidy-Bushrow et al. (2020)	Detroit, Michigan	06/2008 - 05/2010	
Lithuania	/	Dedele et al. (2017)	Kaunas	01/2007 - 12/2008	In the same country, but time period did not overlap.
	/	Maroziene et al. (2002)	Kaunas	01/1998 - 12/1998	
Europe	/	Giorgis-Allemand et al. (2017)	11 countries (inc. England)	01/1994 - 12/2010	Barba-Vasseur et al. (2017) was removed as it was assigned a high RoB. Estarlich et al. (2016) was removed as its cohort was completely embedded within Giorgis-Allemand et al. (2017). Smith et al. (2020) was not removed as Giorgis-Allemand et al. (2017) used the Born-in-Bradford study that was not in London.
England	/	Smith et al. (2020)	London	01/2006 - 12/2010	
Spain	/	Estarlich et al. (2016)	Valencia, Asturias, Gipuzkos, Sabadell	05/2004 - 08/2008	
France	/	Barba-Vasseur et al. (2017)	Besançon, Dijon	01/2005 - 12/2009	
Brazil	/	doNascimento et al. (2022)	São Paulo	01/2010 - 12/2016	No study dependencies in this group of studies
Canada	/	Lavigne et al. (2016)	Ontario	01/2005 - 03/2012	
Taiwan	/	Lin et al. (2015)	National	01/2001 - 12/2007	
Trimester 3 NO_2 exposure and the risk of PTB (OR)					
Brazil	/	doNascimento et al. (2022)	São Paulo	01/2010 - 12/2016	No study dependencies in this group of studies. Barba-Vasseur et al. (2017) was removed as it was assigned a high RoB.
Canada	/	Lavigne et al. (2016)	Ontario	01/2005 - 03/2012	
Spain	/	Estarlich et al. (2016)	Valencia, Asturias, Gipuzkos, Sabadell	05/2004 - 08/2008	

Country	Ranking	Study	Location	Time period	Decision
France	/	Barba-Vasseur et al. (2017)	Besançon, Dijon	01/2005 - 12/2009	
Taiwan	/	Lin et al. (2015)	National	01/2001 - 12/2007	
US	/	Cassidy-Bushrow et al. (2020)	Detroit, Michigan	06/2008 - 05/2010	No study dependencies in this group of studies
	/	Hao et al. (2016)	Georgia	01/2002 - 02/2006	
China	/	Zhou et al. (2022a)	Henan	05/2014 - 12/2016	Zhou et al. (2022a) was kept as it was ranked higher. Lu et al. (2019) was kept as it was not temporally dependent from other studies
	/	Gong et al. (2022)	Henan	05/2014 - 12/2016	
	/	Lu et al. (2019)	Changsha City, Henan	09/2011 - 31/2012	
	/	Ye et al. (2018)	Taizhou City, Zhejiang	01/2013 - 05/2016	Both studies were kept as one was province-wide, and one was in Taizhou only.
	/	Sun et al. (2019)	Zhejiang	01/2013 - 07/2017	
	/	Fang et al. (2022)	Beijing City	01/2014 - 12/2017	In the same city but time period did not overlap.
	/	Huang et al. (2015)	Beijing City	01/2007 - 12/2010	
	/	Zhou et al. (2022b)	Chongqing City	01/2015 - 12/2020	No study dependencies in this group of studies. Zou et al. (2021), Liu et al. (2019b), and Gan et al. (2022) were removed due to high RoB.
	/	Ji et al. (2019)	Shanghai City	01/2014 - 12/2015	
	/	Zou et al. (2021)	Shanghai City	04/2011 - 04/2012	
	/	Liu et al. (2019b)	Guangdong	01/2014 - 12/2015	
	/	Gan et al. (2022)	Guangzhou City, Guangdong	01/2017 - 04/2018	
Lithuania	/	Dedele et al. (2017)	Kaunas	01/2007 - 12/2008	In the same country, but time period did not overlap.
	/	Maroziene et al. (2002)	Kaunas	01/1998 - 12/1998	
Whole pregnancy NO_2 exposure and the risk of PTB (OR)					
China	2	Zhou et al. (2022a)	Henan	05/2014 - 12/2016	Zhou et al. (2022a) was kept as it was ranked higher. Lu et al. (2019) was kept as it was not temporally dependent from other studies
	4	Gong et al. (2022)	Henan	05/2014 - 12/2016	
	7	Lu et al. (2019)	Changsha City, Henan	09/2011 - 31/2012	
	5	Ye et al. (2018)	Taizhou City, Zhejiang	01/2013 - 05/2016	Both studies were kept as one was province-wide, and one was in Taizhou only.
	6	Sun et al. (2019)	Zhejiang	01/2013 - 07/2017	
	1	Ji et al. (2019)	Shanghai City	01/2014 - 12/2015	No study dependencies in this group of studies. Zou et al. (2021) was removed as it was assigned a high RoB and ranked lower than Ji et al. (2019).
	11	Zou et al. (2021)	Shanghai City	04/2011 - 04/2012	
	3	Qian et al. (2016)	Wuhan City, Hubei	06/2011 - 06/2013	
	8	Zhou et al. (2022b)	Chongqing City	01/2015 - 12/2020	
	9	Liu et al. (2019b)	Guangdong	01/2014 - 12/2015	Studies assigned a high RoB
	10	Aimaier (亚库甫·艾麦尔) et al. (2013)	Beijing City	07/2007 - 07/2009	

Country	Ranking	Study	Location	Time period	Decision
US	1/2	Wu et al. (2011)	Los Angeles & Orange, California	01/1997 - 12/2006	Laurent et al. (2016) removed due to lower ranking than Wu et al. (2016) as they were both statewide. Wilhelm et al. (2011) removed as it completely overlaps with Wu et al. (2011) which has higher ranking.
	3	Wilhelm et al. (2003)	Los Angeles, California	01/1994 - 12/1996	
	5	Wilhelm et al. (2011)	Los Angeles, California	06/2004 - 03/2006	
	7	Wu et al. (2016)	California	01/2001 - 12/2008	
	8	Laurent et al. (2016)	California	01/2001 - 12/2008	
	4	Hao et al. (2016)	Georgia	01/2002 - 02/2006	No study dependencies in this group of studies
	6	Cassidy-Bushrow et al. (2020)	Detroit, Michigan	06/2008 - 05/2010	
Canada	/	Lavigne et al. (2016)	Ontario	01/2005 - 03/2012	Stieb et al. (2016a) was kept as it was ranked higher than Villeneuve et al. (2022).
	/	Poirier et al. (2015)	Halifax, Nova Scotia	01/2008 - 23/2012	
	1	Stieb et al. (2016a)	National	01/1999 - 12/2008	
	2	Villeneuve et al. (2022)	National	01/1998 - 12/2008	
Australia	/	Melody et al. (2020)	Victoria	03/2012 - 12/2015	No study dependencies in this group of studies. Dedele et al. (2017) and Maroziene et al. (2002) but were conducted in the same country but were not temporally dependent.
Brazil	/	doNascimento et al. (2022)	São Paulo	01/2010 - 12/2016	
Lithuania	/	Dedele et al. (2017)	Kaunas	01/2007 - 12/2008	
	/	Maroziene et al. (2002)	Kaunas	01/1998 - 12/1998	
Europe	/	Giorgis-Allemand et al. (2017)	11 countries (inc. England)	01/1994 - 12/2010	Barba-Vasseur et al. (2017) was removed as it was assigned a high RoB. Estarlich et al. (2016) and Gehring et al. (2011a) were removed as its cohort was completely embedded within Giorgis-Allemand et al. (2017).
Netherlands	/	Gehring et al. (2011a)	National	01/1996 - 12/1997	
	/	van de Hooven et al. (2012)	Rotterdam	01/2001 - 12/2005	
Norway	/	Panasevich et al. (2016)	Oslo, Bergen, Akershus, Hordalan	01/1999 - 12/2008	
Spain	/	Estarlich et al. (2016)	Valencia, Asturias, Gipuzkos, Sabadell	05/2004 - 08/2008	
France	/	Barba-Vasseur et al. (2017)	Besançon, Dijon	01/2005 - 12/2009	
Trimester 1 O_3 exposure and the risk of PTB (OR)					
China	1	Zhou et al. (2022a)	Henan	05/2014 - 12/2016	Zhou et al. (2022a) was kept as it was ranked higher.
	2	Gong et al. (2022)	Henan	05/2014 - 12/2016	
	3	Sun et al. (2019)	Zhejiang	01/2013 - 07/2017	
	4	Liu et al. (2019b)	Guangdong	01/2014 - 12/2015	Studies assigned a high RoB
	5	Gan et al. (2022)	Guangzhou City, Guangdong	01/2017 - 04/2018	
US	1	Mekonnen et al. (2021)	6 counties, California	01/2007 - 12/2011	No study dependencies in this group of studies
	2/3	Wu et al. (2011)	Los Angeles & Orange, California	01/1997 - 12/2006	

Country	Ranking	Study	Location	Time period	Decision
	4	Ha et al. (2014)	Florida	01/2004 - 12/2005	
	5	Hao et al. (2016)	Georgia	01/2002 - 02/2006	
Australia	/	Hansen et al. (2006)	Brisbane	07/2000 - 06/2003	No study dependencies in this group of studies
Brazil	/	doNascimento et al. (2022)	São Paulo	01/2010 - 12/2016	
Taiwan	/	Lin et al. (2015)	National	01/2001 - 12/2007	
Sweden	/	Olsson et al. (2013)	Greater Stockholm area	08/1997 - 02/2006	
Trimester 2 O_3 exposure and the risk of PTB (OR)					
China	1	Zhou et al. (2022a)	Henan	05/2014 - 12/2016	Zhou et al. (2022a) was kept as it was ranked higher.
	2	Gong et al. (2022)	Henan	05/2014 - 12/2016	
	3	Sun et al. (2019)	Zhejiang	01/2013 - 07/2017	
	4	Liu et al. (2019b)	Guangdong	01/2014 - 12/2015	Studies assigned a high RoB
	5	Gan et al. (2022)	Guangzhou City, Guangdong	01/2017 - 04/2018	
US	/	Mekonnen et al. (2021)	6 counties, California	01/2007 - 12/2011	No study dependencies in this group of studies
	/	Wu et al. (2011)	Los Angeles & Orange, California	01/1997 - 12/2006	
	/	Ha et al. (2014)	Florida	01/2004 - 12/2005	
	/	Hao et al. (2016)	Georgia	01/2002 - 02/2006	
Brazil	/	doNascimento et al. (2022)	São Paulo	01/2010 - 12/2016	No study dependencies in this group of studies
Taiwan	/	Lin et al. (2015)	National	01/2001 - 12/2007	
Canada	/	Lavigne et al. (2016)	Ontario	01/2005 - 03/2012	
Trimester 3 O_3 exposure and the risk of PTB (OR)					
China	/	Zhou et al. (2022a)	Henan	05/2014 - 12/2016	Zhou et al. (2022a) was kept as it was ranked higher.
	/	Gong et al. (2022)	Henan	05/2014 - 12/2016	
	/	Qian et al. (2016)	Wuhan City, Hubei	06/2011 - 06/2013	
	/	Fang et al. (2022)	Beijing City	01/2014 - 12/2017	
	/	Liu et al. (2019b)	Guangdong	01/2014 - 12/2015	Studies assigned a high RoB
	/	Gan et al. (2022)	Guangzhou City, Guangdong	01/2017 - 04/2018	
US	/	Mekonnen et al. (2021)	6 counties, California	01/2007 - 12/2011	No study dependencies in this group of studies
	/	Ha et al. (2014)	Florida	01/2004 - 12/2005	
	/	Hao et al. (2016)	Georgia	01/2002 - 02/2006	
Brazil	/	doNascimento et al. (2022)	São Paulo	01/2010 - 12/2016	No study dependencies in this group of studies

Country	Ranking	Study	Location	Time period	Decision
Canada	/	Lavigne et al. (2016)	Ontario	01/2005 - 03/2012	
Taiwan	/	Lin et al. (2015)	National	01/2001 - 12/2007	
Whole pregnancy O_3 exposure and the risk of PTB (OR)					
China	/	Zhou et al. (2022a)	Henan	05/2014 - 12/2016	Zhou et al. (2022a) was kept as it was ranked higher.
	/	Gong et al. (2022)	Henan	05/2014 - 12/2016	
	/	Qian et al. (2016)	Wuhan City, Hubei	06/2011 - 06/2013	
	/	Sun et al. (2019)	Zhejiang	01/2013 - 07/2017	
	/	Liu et al. (2019b)	Guangdong	01/2014 - 12/2015	Study assigned a high RoB
US	1	Mekonnen et al. (2021)	6 counties, California	01/2007 - 12/2011	Wilhelm et al. (2011) was removed due to complete overlap with Wu et al. (2011). Wu et al. (2011) was temporally
	2/3	Wu et al. (2011)	Los Angeles & Orange, California	01/1997 - 12/2006	
	6	Wilhelm et al. (2011)	Los Angeles, California	06/2004 - 03/2006	
	7	Laurent et al. (2016)	California	01/2001 - 12/2008	
	4	Ha et al. (2014)	Florida	01/2004 - 12/2005	No study dependencies in this group of studies
	5	Hao et al. (2016)	Georgia	01/2002 - 02/2006	
Brazil	/	doNascimento et al. (2022)	São Paulo	01/2010 - 12/2016	No study dependencies in this group of studies
Canada	/	Lavigne et al. (2016)	Ontario	01/2005 - 03/2012	

Appendix 15. Forest plots of the association between $PM_{2.5}$ exposure and the risk of PTB subgroups for studies reporting ORs. The individual effect sizes and the 95% CIs, log of the effect sizes (logEE), standard error (SE), summary estimates (in diamond shape), and its 95% CI, prediction intervals (PI), and I^2 statistics are presented.

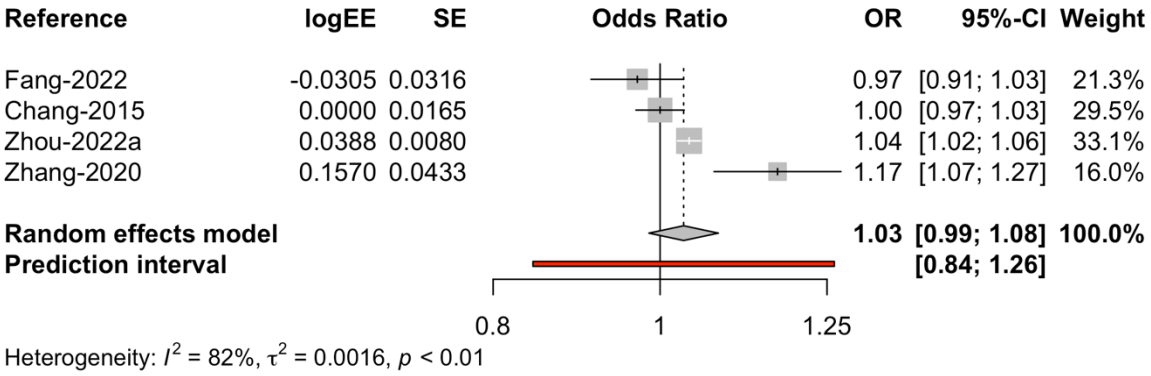


Figure 1. Trimester 1 $PM_{2.5}$ exposure and the risk of moderate-to-late PTB (ORs)

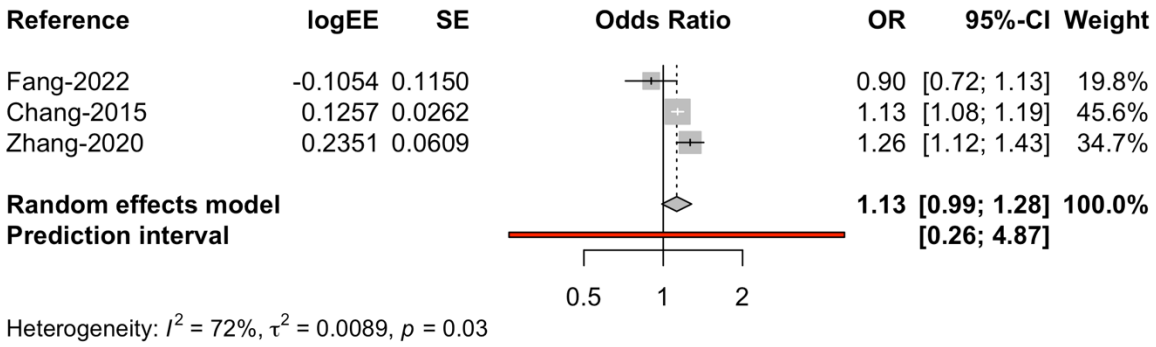


Figure 2. Trimester 1 $PM_{2.5}$ exposure and the risk of very PTB (ORs)

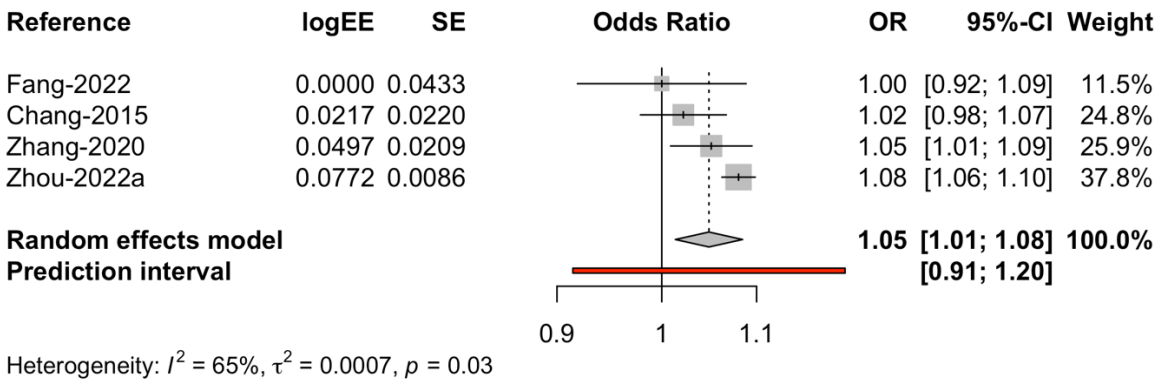


Figure 3. Trimester 2 $PM_{2.5}$ exposure and the risk of moderate-to-late PTB (ORs)

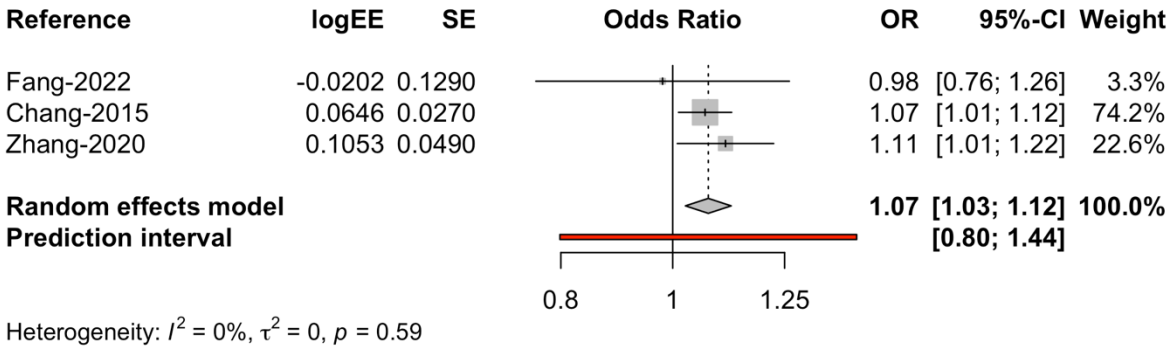


Figure 4. Trimester 2 $PM_{2.5}$ exposure and the risk of very PTB (ORs)

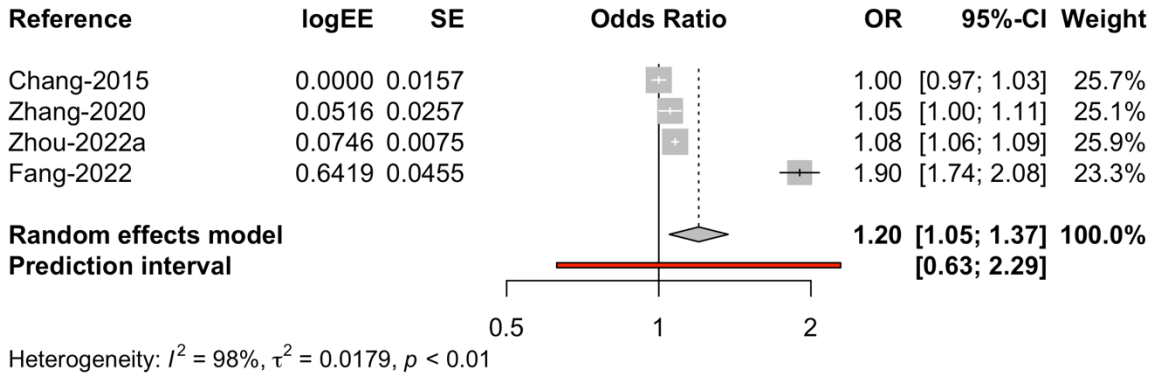


Figure 5. Trimester 3 $PM_{2.5}$ exposure and the risk of moderate-to-late PTB (ORs)

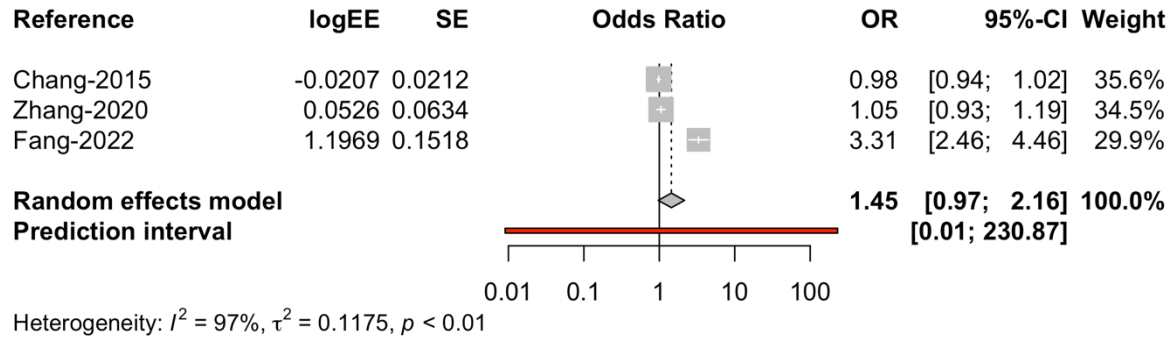


Figure 6. Trimester 3 $PM_{2.5}$ exposure and the risk of very PTB (ORs)

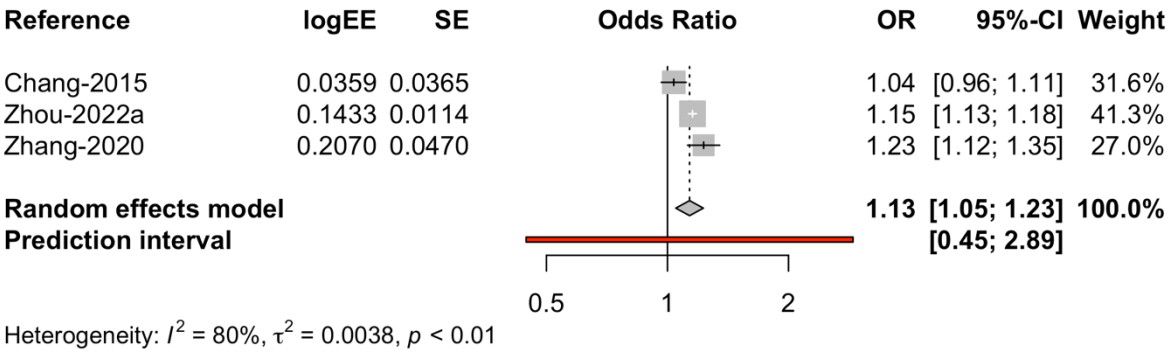


Figure 7. Whole pregnancy $PM_{2.5}$ exposure and the risk of moderate-to-late PTB (ORs)

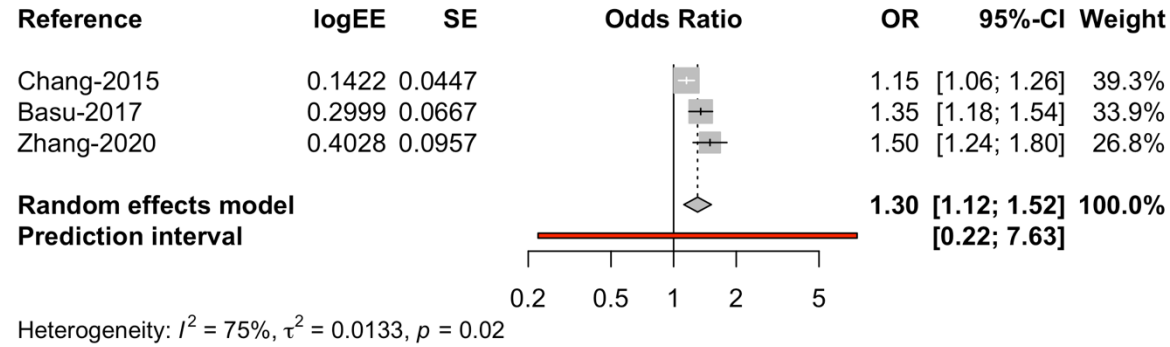


Figure 8. Whole pregnancy $PM_{2.5}$ exposure and the risk of very PTB (ORs)

Appendix 16. Forest plots of the association between $PM_{2.5}$ exposure and the risk of PTB subgroups for studies reporting HRs. The individual effect sizes and the 95% CIs, log of the effect sizes (logEE), standard error (SE), summary estimates (in diamond shape), and its 95% CI, prediction intervals (PI), and I^2 statistics are presented.

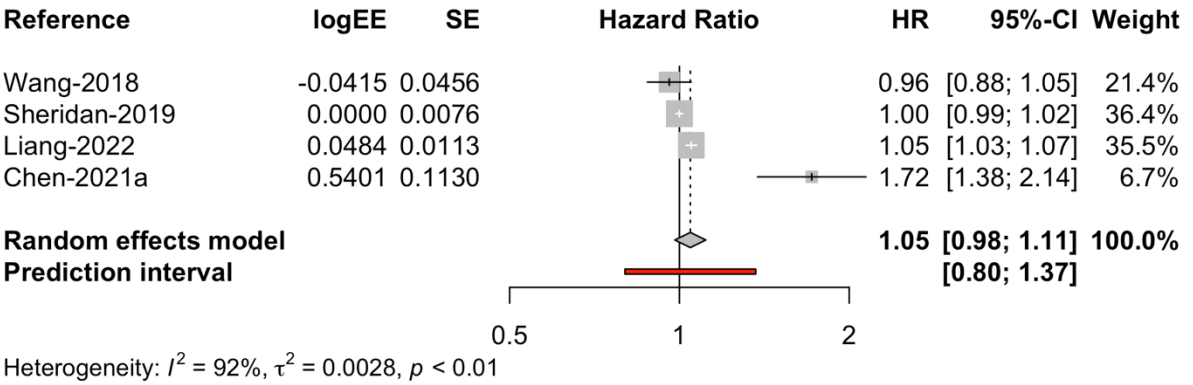


Figure 1. Trimester 1 $PM_{2.5}$ exposure and the risk of moderate-to-late PTB (HRs)

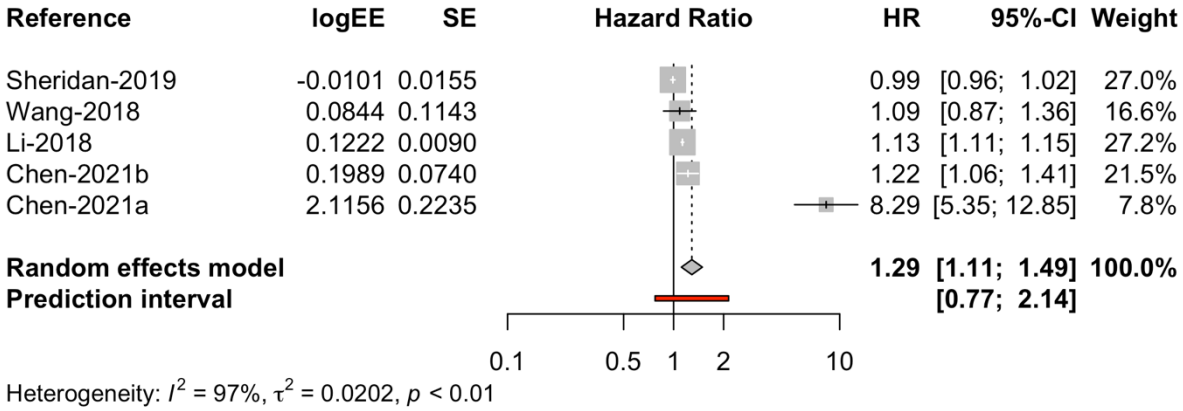


Figure 2. Trimester 1 $PM_{2.5}$ exposure and the risk of very PTB (HRs)

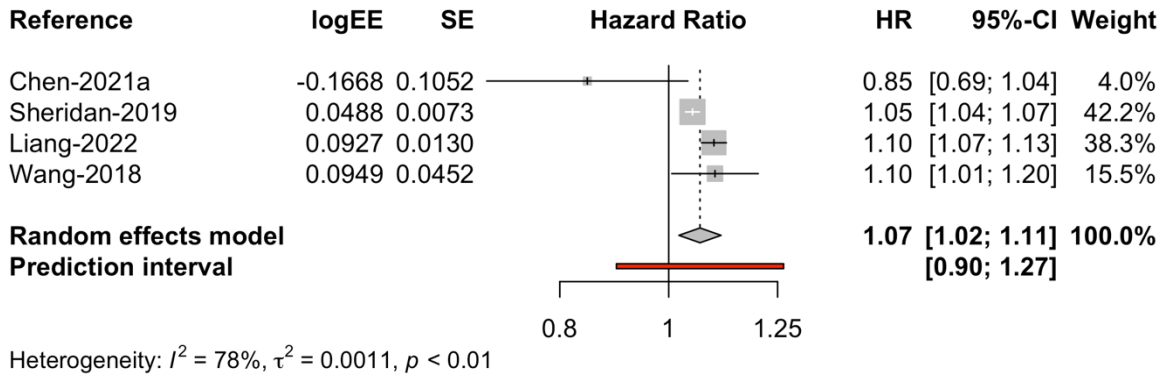


Figure 3. Trimester 2 $PM_{2.5}$ exposure and the risk of moderate-to-late PTB (HRs)

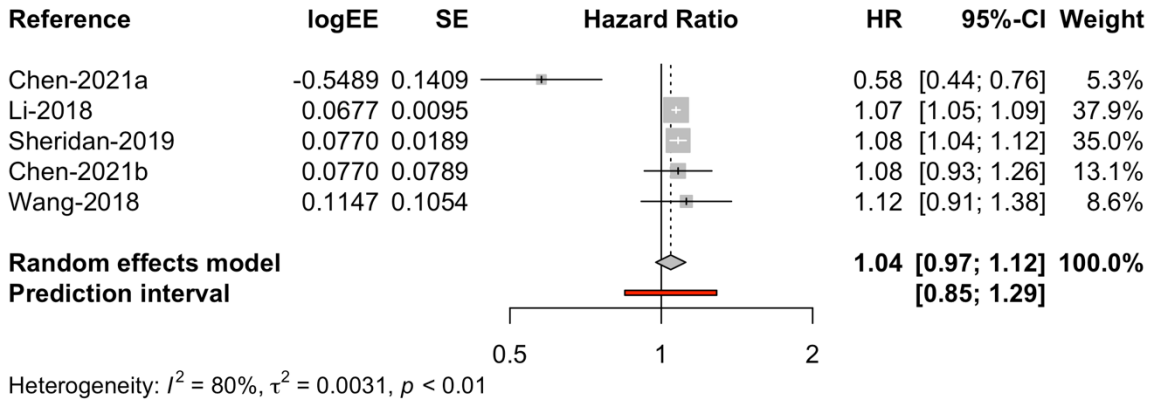


Figure 4. Trimester 2 $PM_{2.5}$ exposure and the risk of very PTB (HRs)

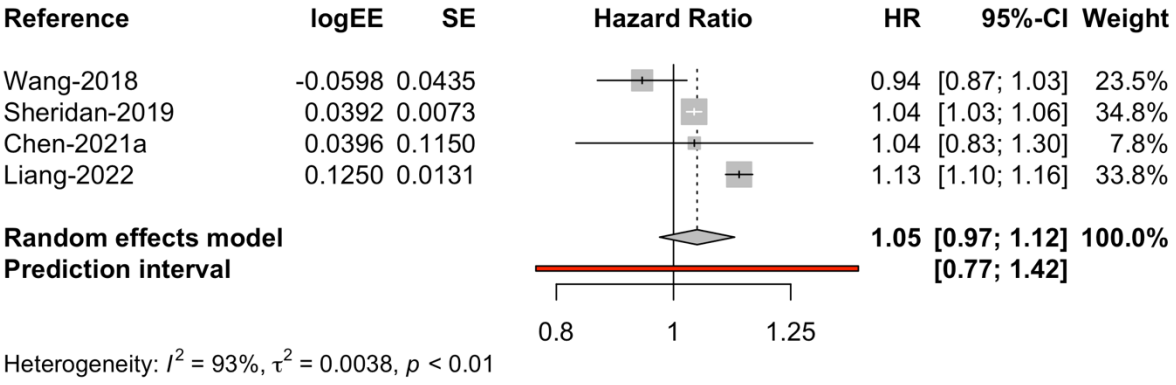


Figure 5. Trimester 3 $PM_{2.5}$ exposure and the risk of moderate-to-late PTB (HRs)

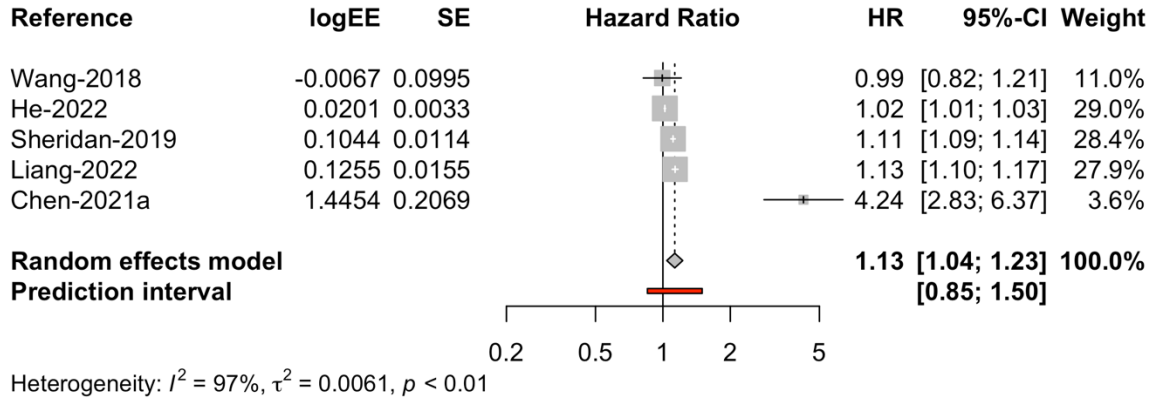


Figure 6. Whole pregnancy $PM_{2.5}$ exposure and the risk of moderate-to-late PTB (HRs)

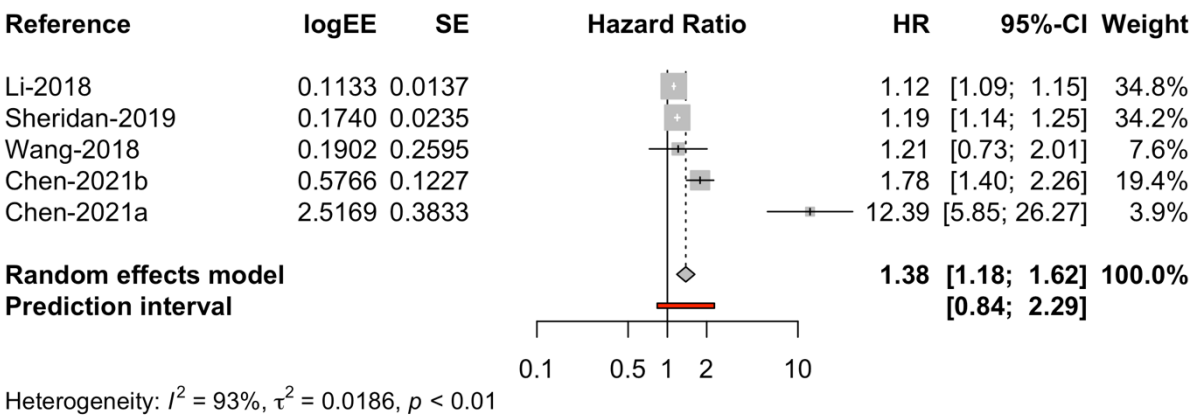


Figure 7. Whole pregnancy $PM_{2.5}$ exposure and the risk of very PTB (HRs)

Appendix 17. Forest plots of the association between NO_2 exposure and the risk of PTB subgroups for studies reporting HRs. The individual effect sizes and the 95% CIs, log of the effect sizes (logEE), standard error (SE), summary estimates (in diamond shape), and its 95% CI, prediction intervals (PI), and I^2 statistics are presented.

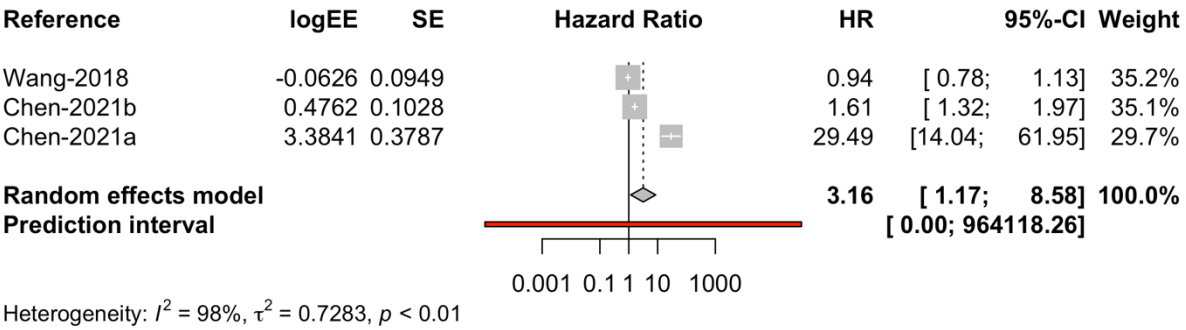


Figure 1. Trimester 1 NO_2 exposure and the risk of very PTB (HRs)

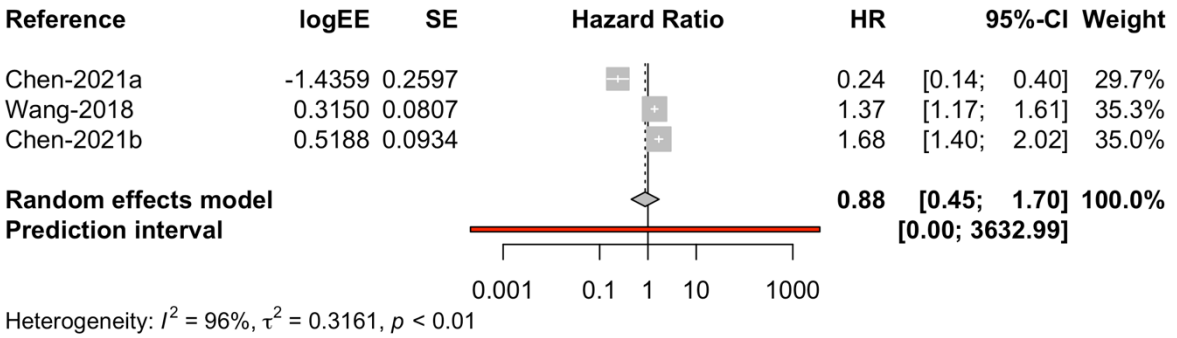


Figure 2. Trimester 2 NO_2 exposure and the risk of very PTB (HRs)

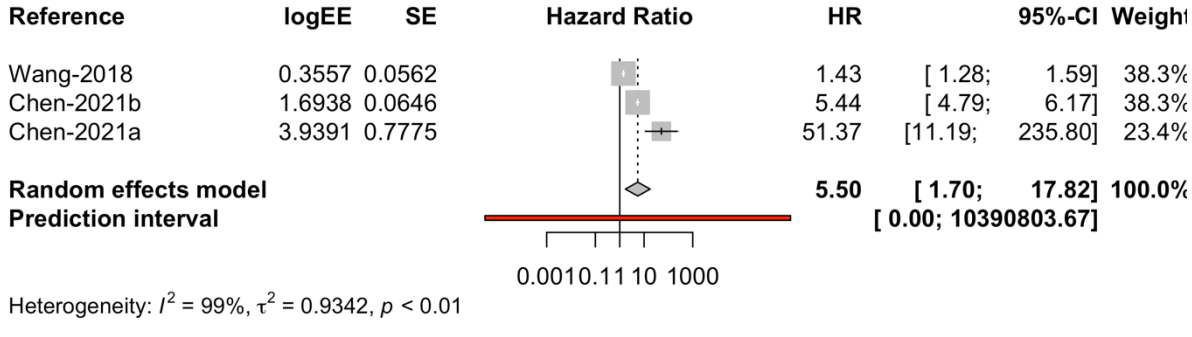


Figure 3. Trimester 3 NO_2 exposure and the risk of very PTB (HRs)

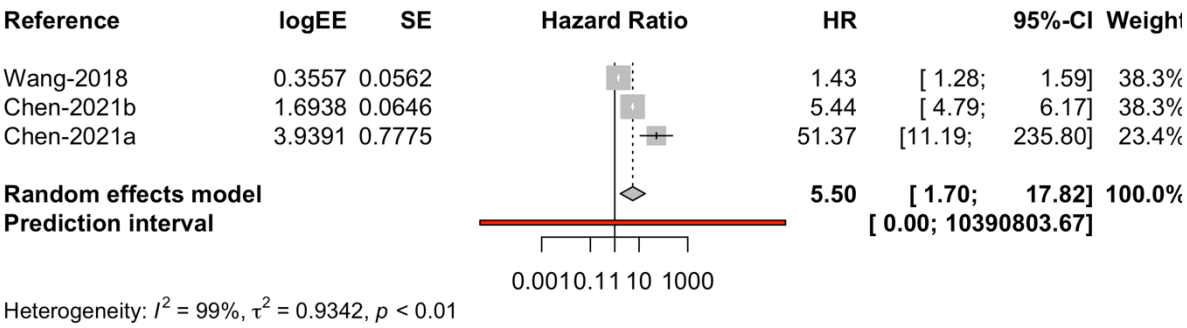


Figure 4. Whole pregnancy NO_2 exposure and the risk of very PTB (HRs)

Appendix 18. Results of the leave-one-out analysis conducted for the eligible exposure-outcome pairs.

Exposure-outcome pair	Description	OR/HR (95% CI)
Trimester 1 $PM_{2.5}$ exposure (HR)	**Summary**	1.02 (0.99 - 1.04)
	w/o Sheridan et al. (2019)	1.03 (1.00 - 1.05)
	w/o Chen et al. (2021b)	1.02 (0.99 - 1.04)
	w/o Chen et al. (2021a)	1.00 (0.99 - 1.02)
	w/o Chen et al. (2018a)	1.02 (1.00 - 1.04)
	w/o Yang et al. (2020)	1.02 (1.00 - 1.04)
	w/o Wang et al. (2018)	1.02 (1.00 - 1.04)
	w/o Liang et al. (2022)	1.05 (1.00 - 1.10)
	w/o He et al. (2022)	1.04 (0.99 - 1.08)
Trimester 2 $PM_{2.5}$ exposure (HR)	**Summary**	1.01 (1.00 - 1.03)
	w/o Sheridan et al. (2019)	1.01 (1.00 - 1.02)
	w/o Chen et al. (2021b)	1.02 (1.00 - 1.03)
	w/o Chen et al. (2021a)	1.02 (1.01 - 1.02)
	w/o Chen et al. (2018a)	1.01 (1.00 - 1.03)
	w/o Yang et al. (2020)	1.01 (1.00 - 1.03)
	w/o Wang et al. (2018)	1.01 (1.00 - 1.02)
	w/o Liang et al. (2022)	1.00 (0.97 - 1.04)
	w/o He et al. (2022)	1.00 (0.97 - 1.04)
Trimester 3 $PM_{2.5}$ exposure (HR)	**Summary**	1.05 (1.03 - 1.08)
	w/o Sheridan et al. (2019)	1.05 (1.02 - 1.08)
	w/o Chen et al. (2021b)	1.07 (1.04 - 1.10)
	w/o Lavigne et al. (2018)	1.05 (1.03 - 1.08)
	w/o Chen et al. (2021a)	1.05 (1.03 - 1.07)
	w/o Chen et al. (2018a)	1.05 (1.03 - 1.08)
	w/o Yang et al. (2020)	1.05 (1.03 - 1.08)
	w/o Wang et al. (2018)	1.06 (1.03 - 1.08)
	w/o Liang et al. (2022)	1.11 (1.04 - 1.18)
	w/o He et al. (2022)	1.11 (1.04 - 1.18)
	w/o Fang et al. (2022)	1.02 (1.01 - 1.03)
Whole pregnancy $PM_{2.5}$ exposure (HR)	**Summary**	1.05 (1.03 - 1.07)
	w/o Sheridan et al. (2019)	1.02 (1.01 - 1.04)
	w/o Chen et al. (2021b)	1.05 (1.03 - 1.07)
	w/o Lavigne et al. (2018)	1.05 (1.03 - 1.07)
	w/o Chen et al. (2021a)	1.04 (1.02 - 1.05)
	w/o Chen et al. (2018a)	1.05 (1.03 - 1.07)
	w/o Yang et al. (2020)	1.05 (1.03 - 1.07)
	w/o Wang et al. (2018)	1.05 (1.03 - 1.07)
	w/o Liang et al. (2022)	1.14 (1.05 - 1.24)
	w/o He et al. (2022)	1.14 (1.04 - 1.24)
Trimester 3 NO_2 exposure (OR)	**Summary**	1.20 (1.07 - 1.34)
	w/o Ji et al. (2019)	1.21 (1.07 - 1.36)
	w/o Cassidy-Bushrow et al. (2020)	1.20 (1.07 - 1.34)
	w/o Dedele et al. (2017)	1.21 (1.07 - 1.36)
	w/o Lavigne et al. (2016)	1.20 (1.07 - 1.36)
	w/o Sun et al. (2019)	1.22 (1.05 - 1.41)
	w/o Lu et al. (2019)	1.19 (1.06 - 1.34)
	w/o Lin et al. (2015)	1.23 (1.08 - 1.39)
	w/o Ye et al. (2018)	1.18 (1.05 - 1.33)
	w/o Estarlich et al. (2016)	1.21 (1.08 - 1.36)
	w/o Hao et al. (2016)	1.21 (1.07 - 1.36)

	w/o Maroziene et al. (2022)	1.20 (1.07 - 1.35)
	w/o Zhou et al. (2022a)	1.13 (1.07 - 1.19)
	w/o doNascimento et al. (2022)	1.21 (1.07 - 1.36)
	w/o Fang et al. (2022)	1.13 (1.01 - 1.27)
	w/o Zhou et al. (2022b)	1.22 (1.06 - 1.40)
	w/o Huang et al. (2015)	1.21 (1.07 - 1.36)
Trimester 1 NO_2 exposure (HR)	**Summary**	1.62 (1.03 - 2.55)
	w/o Chen et al. (2021b)[#]	1.89 (0.47 - 7.56)
	w/o Chen et al. (2021a)	1.04 (1.03 - 1.04)
	w/o Chen et al. (2018a)	1.88 (1.13 - 3.14)
	w/o Liang et al. (2019b)[#]	1.88 (0.46 - 7.66)
Trimester 2 NO_2 exposure (HR)	**Summary**	0.91 (0.71 - 1.17)
	w/o Chen et al. (2021b)	0.95 (0.48 - 1.91)
	w/o Chen et al. (2021a)	1.05 (0.94 - 1.17)
	w/o Chen et al. (2018a)	0.81 (0.63 - 1.05)
	w/o Liang et al. (2019b)	0.95 (0.47 - 1.94)
Trimester 3 NO_2 exposure (HR)	**Summary**	1.57 (1.10 - 2.23)
	w/o Chen et al. (2021b)[#]	1.86 (0.90 - 3.83)
	w/o Chen et al. (2021a)[#]	1.46 (1.00 - 2.13)
	w/o Chen et al. (2018a)	1.65 (1.10 - 2.47)
	w/o Liang et al. (2019b)[#]	1.80 (0.82 - 3.97)
	w/o Fang et al. (2022)[#]	1.10 (0.94 - 1.30)
Whole pregnancy NO_2 exposure (HR)	**Summary**	1.00 (0.99 - 1.01)
	w/o Chen et al. (2021b)	1.00 (0.99 - 1.01)
	w/o Chen et al. (2021a)	1.00 (0.99 - 1.01)
	w/o Chen et al. (2018a)	1.00 (0.99 - 1.01)
	w/o Liang et al. (2019b)	1.00 (0.99 - 1.02)
	w/o Schifano et al. (2016) (Rome)	1.00 (0.98 - 1.02)
	w/o Schifano et al. (2016) (Barcelona)	1.01 (0.99 - 1.02)
	w/o Benmarhnia et al. (2017)	1.01 (0.98 - 1.03)
Trimester 2 O_3 exposure (OR)	**Summary**	1.01 (0.98 - 1.06)
	w/o Lavigne et al. (2016)	1.01 (0.97 - 1.05)
	w/o Mekonnen et al. (2021)	1.01 (0.98 - 1.05)
	w/o Sun et al. (2019)	1.02 (0.96 - 1.08)
	w/o Lin et al. (2015)	1.01 (0.97 - 1.05)
	w/o Ha et al. (2014)	1.01 (0.97 - 1.05)
	w/o Hao et al. (2016)	1.01 (0.97 - 1.06)
	w/o Wu et al. (2011) (LA)	1.01 (0.97 - 1.05)
	w/o Wu et al. (2011) (Orange)	1.01 (0.97 - 1.05)
	w/o Zhou et al. (2022a)[#]	1.04 (1.01 - 1.07)
	w/o doNascimento et al. (2022)	1.01 (0.96 - 1.06)
Trimester 3 O_3 exposure (OR)	**Summary**	1.01 (0.93 - 1.10)
	w/o Lavigne et al. (2016)	1.02 (0.93 - 1.11)
	w/o Mekonnen et al. (2021)	1.01 (0.93 - 1.10)
	w/o Sun et al. (2019)	0.98 (0.90 - 1.06)
	w/o Lin et al. (2015)	1.02 (0.93 - 1.11)
	w/o Ha et al. (2014)	1.02 (0.94 - 1.12)
	w/o Hao et al. (2016)	1.02 (0.93 - 1.12)
	w/o Wu et al. (2011) (LA)	1.03 (1.00 - 1.07)
	w/o Wu et al. (2011) (Orange)	1.02 (0.92 - 1.12)
	w/o Zhou et al. (2022a)	1.03 (0.95 - 1.13)
Whole pregnancy O_3 exposure (OR)	**Summary**	1.04 (0.93 - 1.17)
	w/o Lavigne et al. (2016)	1.04 (0.92 - 1.17)
	w/o Mekonnen et al. (2021)	1.04 (0.93 - 1.17)
	w/o Sun et al. (2019)	1.04 (0.93 - 1.18)

	w/o Lin et al. (2015)	1.03 (0.91 - 1.16)
	w/o Ha et al. (2014)	1.04 (0.92 - 1.17)
	w/o Hao et al. (2016)	1.03 (0.90 - 1.17)
	w/o Wu et al. (2011) (LA)	1.04 (0.93 - 1.17)
	w/o Wu et al. (2011) (Orange)	1.04 (0.92 - 1.18)
	w/o Zhou et al. (2022a)	1.04 (0.92 - 1.17)
	w/o doNascimento et al. (2022)[#]	1.09 (1.05 - 1.13)
	w/o Laurent et al. (2016)	1.03 (0.90 - 1.18)
Whole pregnancy O_3 exposure (HR)	**Summary**	1.01 (0.97 - 1.05)
	w/o Chen et al. (2021a)	1.02 (0.99 - 1.05)
	w/o Chen et al. (2018a)	1.00 (0.97 - 1.03)
	w/o Schifano et al., 2016 (Rome)	1.06 (0.33 - 3.41)
	w/o Schifano et al., 2016 (Barcelona)	1.05 (0.33 - 3.41)

Appendix 19. Funnel plots for the meta-analyses of the association between air pollution exposure and the risk of PTB with more than 10 studies.

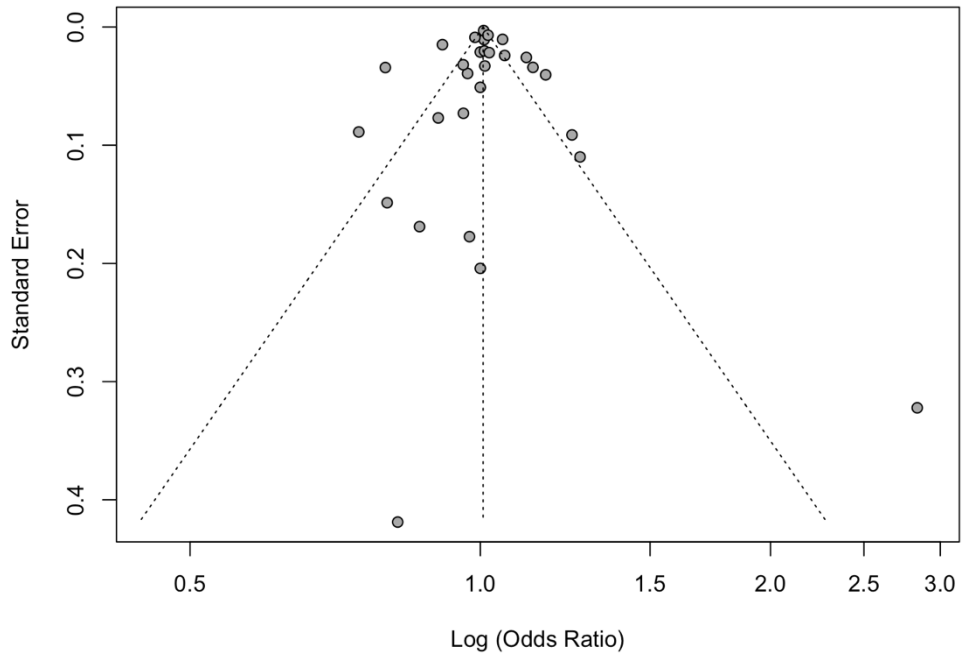


Figure 1. Funnel plot: trimester 1 $PM_{2.5}$ exposure and the risk of PTB (OR)

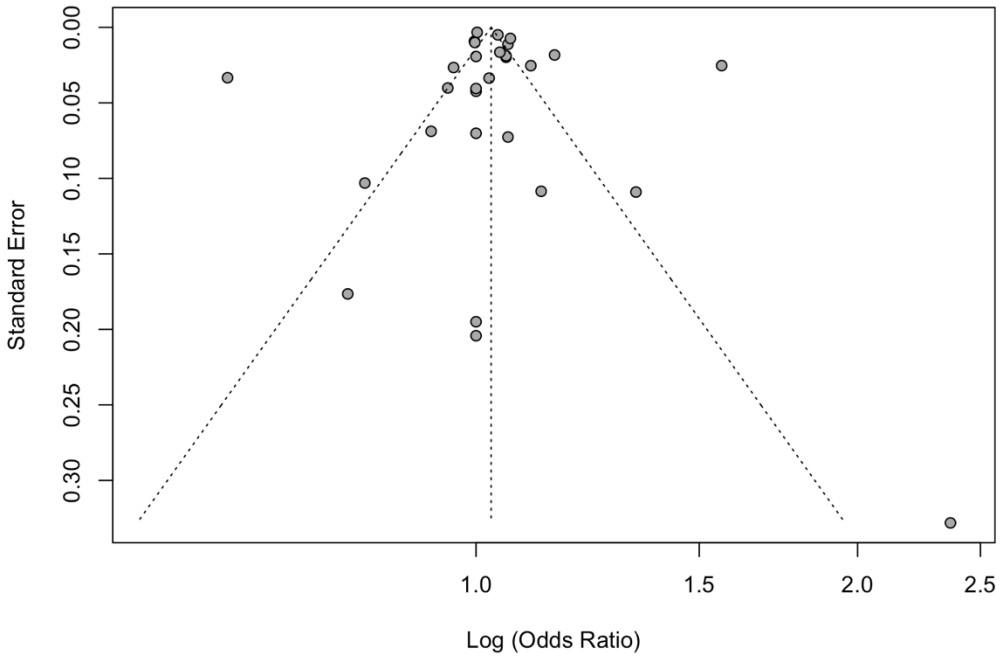


Figure 2. Funnel plot: trimester 2 $PM_{2.5}$ exposure and the risk of PTB (OR)

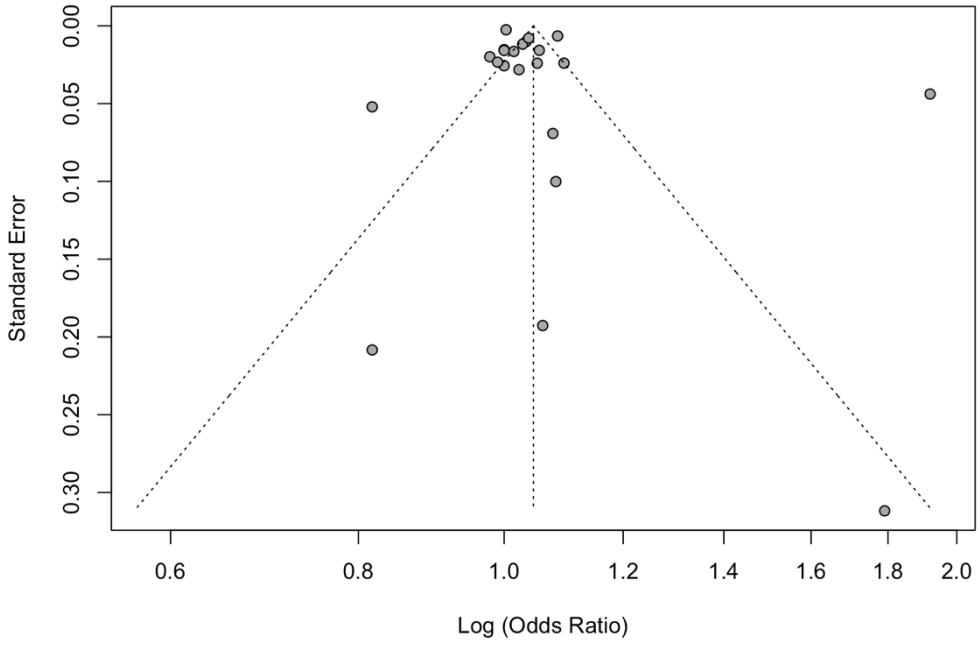


Figure 3. Funnel plot: trimester 3 $PM_{2.5}$ exposure and the risk of PTB (OR)

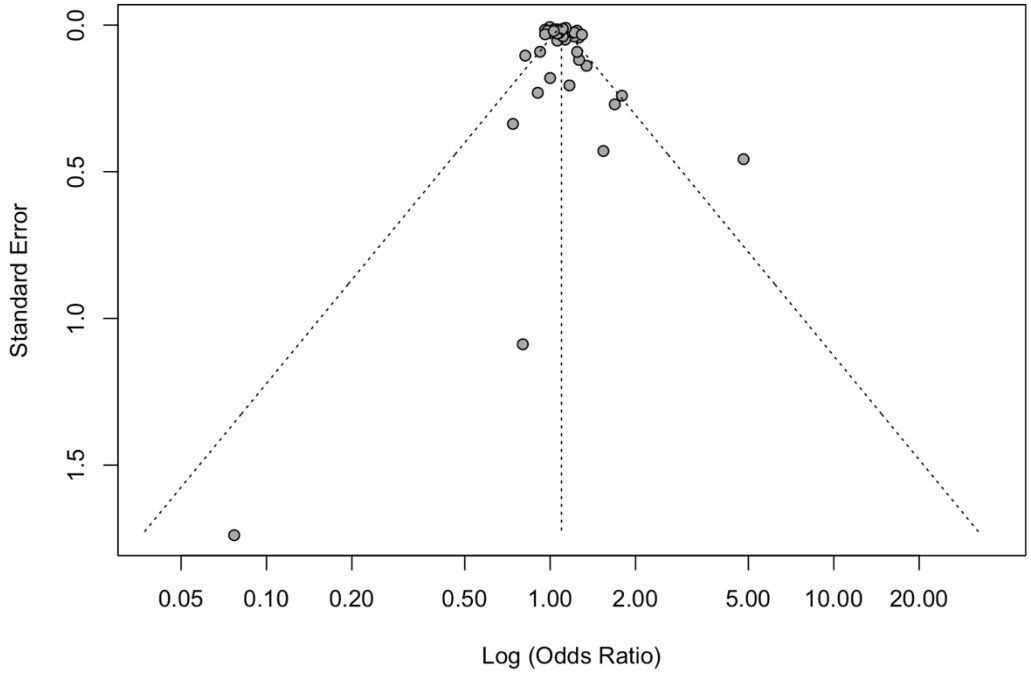


Figure 4. Funnel plot: whole pregnancy $PM_{2.5}$ exposure and the risk of PTB (OR)

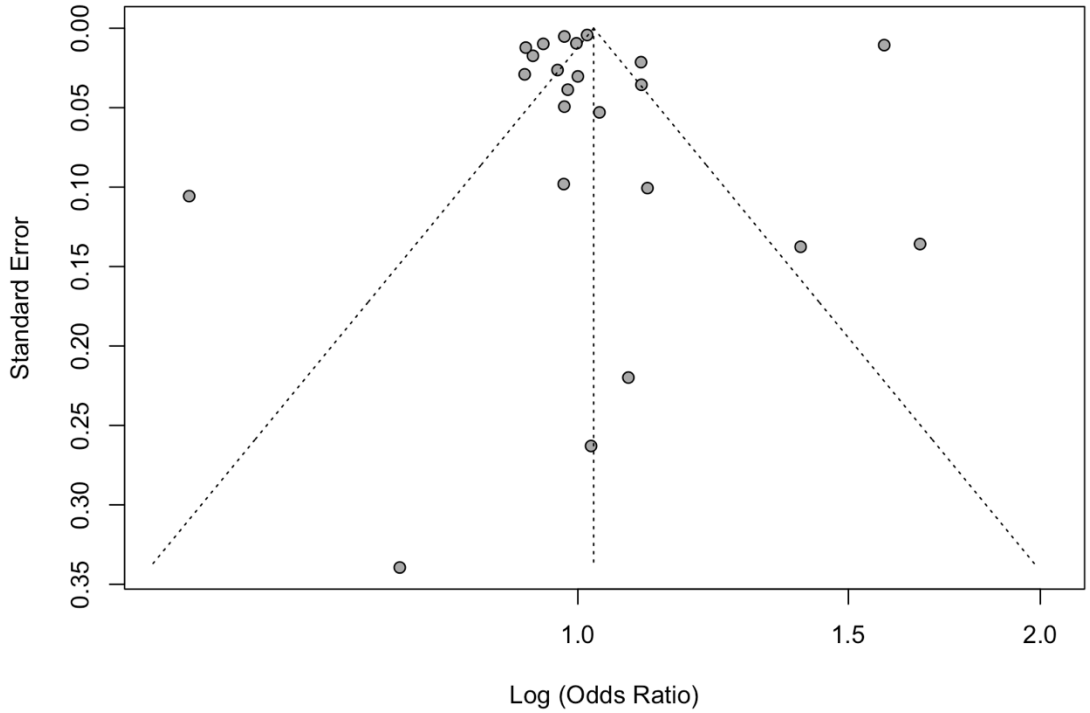


Figure 5. Funnel plot: trimester 1 NO_2 exposure and the risk of PTB (OR)

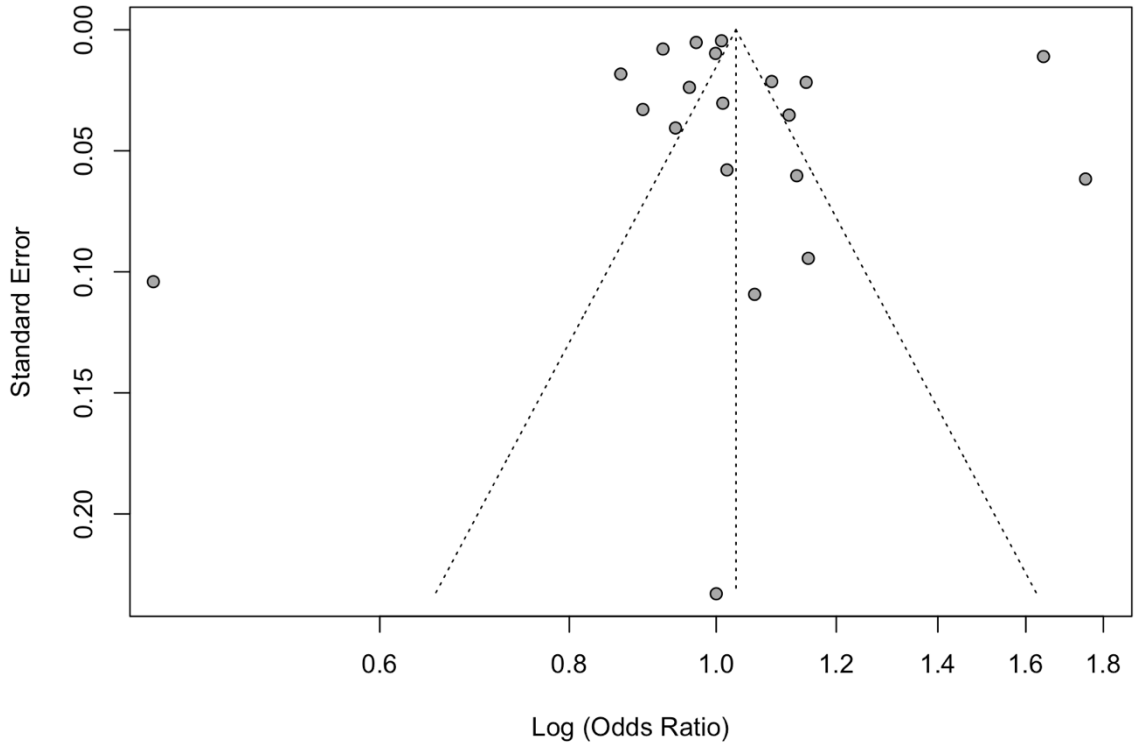


Figure 6. Funnel plot: trimester 2 NO_2 exposure and the risk of PTB (OR)

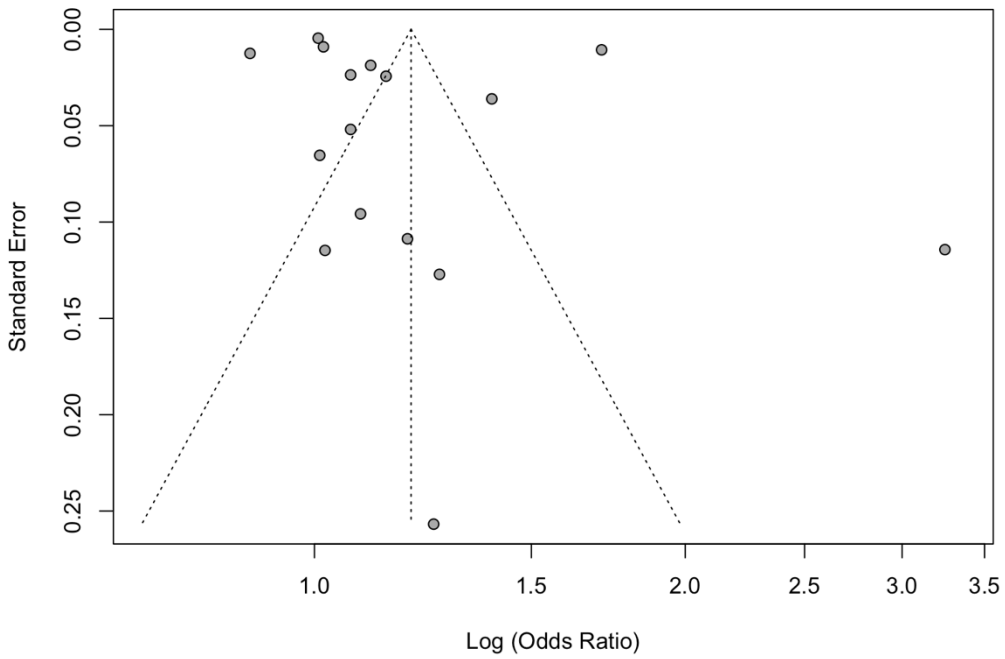


Figure 7. Funnel plot: trimester 3 NO_2 exposure and the risk of PTB (OR)

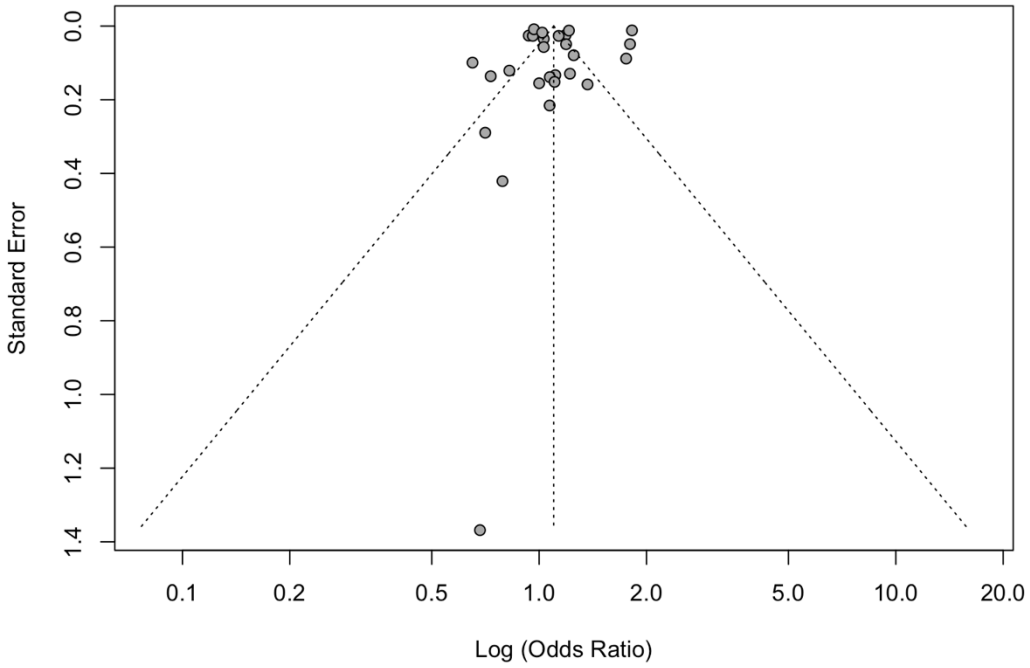


Figure 8. Funnel plot: whole pregnancy NO_2 exposure and the risk of PTB (OR)

Appendix 20. Results of the Egger's test conducted to assess funnel plot asymmetry of the meta-analyses conducted for the association between gestational air pollution exposure and the risk of PTB (for exposure-outcome pairs with >10 studies).

	Window	Intercept	t	p-value
$PM_{2.5}$	Trimester 1	-0.0260	-0.04	0.9674
	Trimester 2	0.6594	0.55	0.5871
	Trimester 3	1.6484	1.41	0.1730
	Whole pregnancy	0.1362	1.53	0.1362
NO_2	Trimester 1	0.5070	0.19	0.8517
	Trimester 2	1.6210	0.46	0.6534
	Trimester 3	3.7227	0.86	0.4036
	Whole pregnancy	-0.6842	-0.28	0.7821

Appendix 21. Results of the stratified analyses conducted by quartiles of mean/median $PM_{2.5}$ level of the studies.

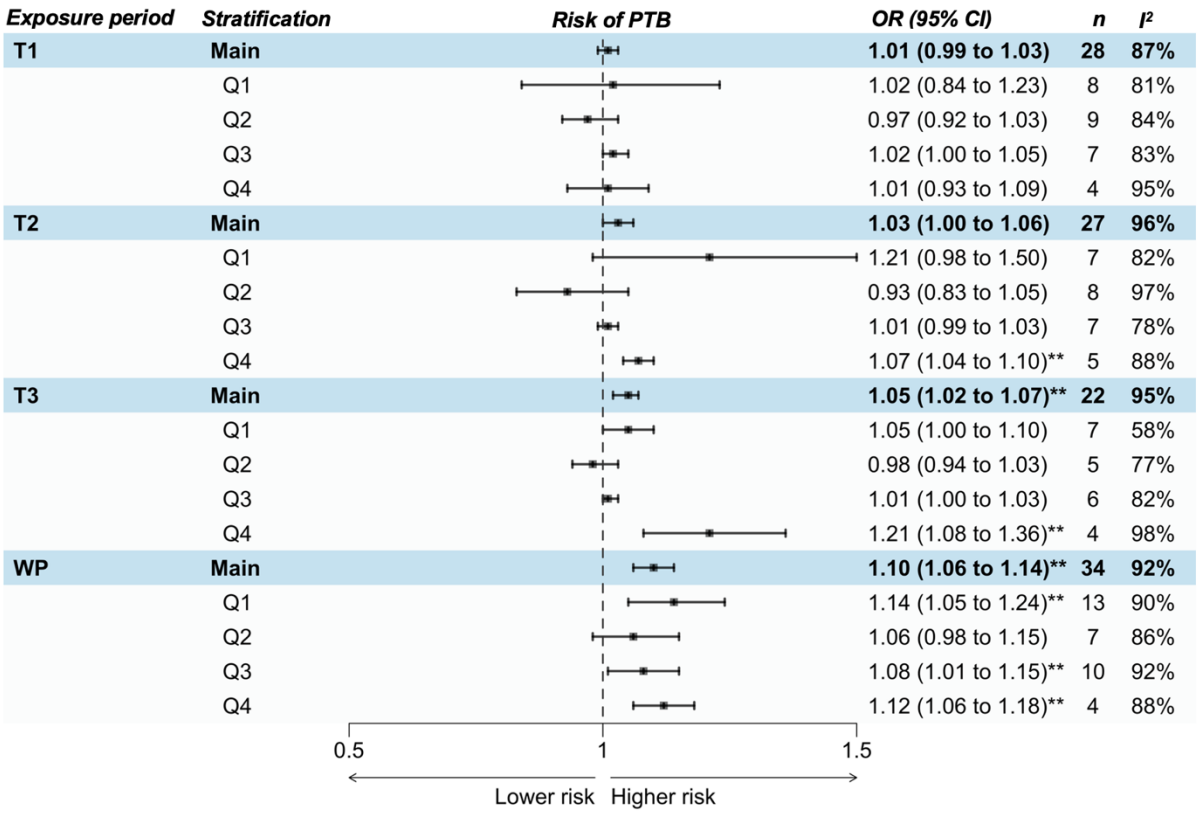
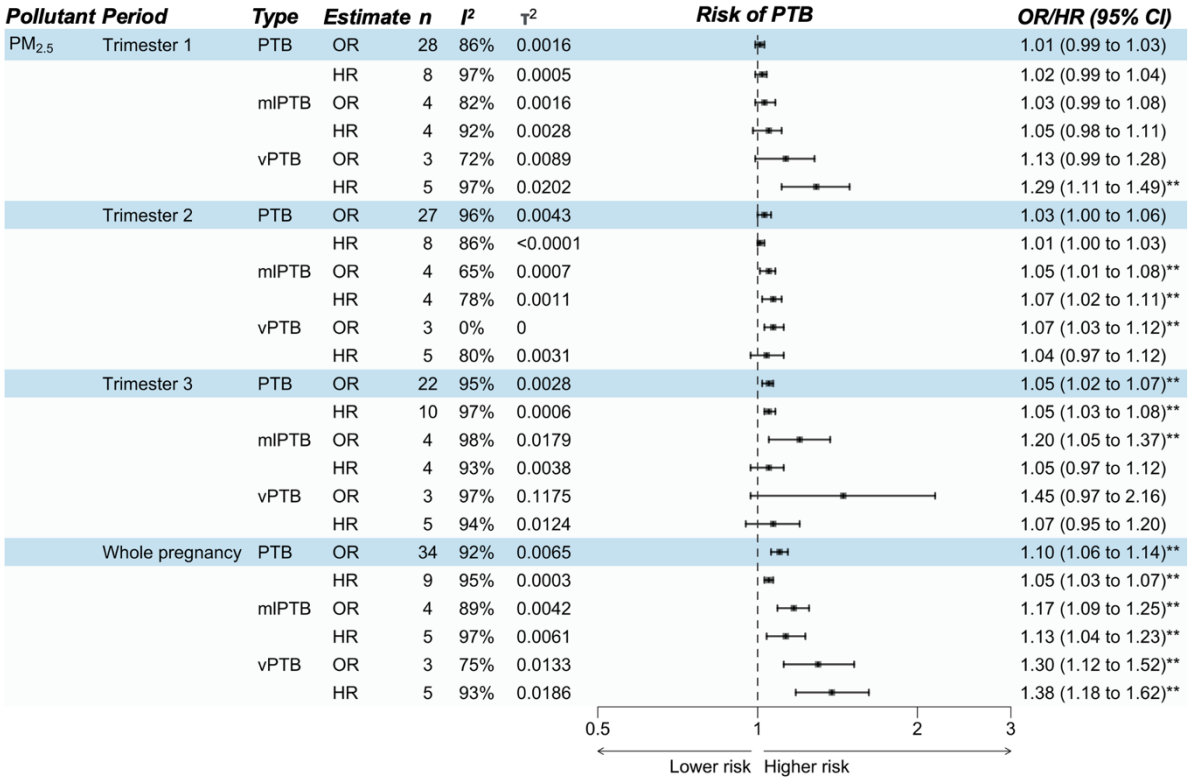


Figure 1. Forest plot for the summary estimates of the meta-analyses conducted for the association between $PM_{2.5}$ exposure at trimester 1 (T1), trimester 2 (T2), trimester 3 (T3), and whole pregnancy (WP), and the risk of PTB from studies using ORs, stratified by the quartiles of the distribution of $PM_{2.5}$ levels in the primary studies used in the meta-analyses. The summary effect sizes (per 10 μg/m^3 increase in exposure) and 95%CIs, I^2 statistics, and number of estimates (n) are presented.

Appendix 22. Link to online platform storing the individual forest plots of the stratified analyses.

https://drive.google.com/drive/folders/1J5Bf7B7wozfBk7yzu4kZWJhmqq5D2OyQ?usp=sharing

Appendix 23. Forest plot showing the results from the meta-analyses conducted in this thesis, Ju et al. (2023), and Yu et al. (2022) for the association between $PM_{2.5}$ exposure and the risk of PTB subgroups.



Appendix 24. Description for the weight of evidence available for the different conclusions made in the causality determination process (reproduced from Owens et al., 2017).

Conclusion	Description
Causal relationship	Evidence is sufficient to conclude that there is a causal relationship with relevant pollutant exposures (e.g., doses or exposures generally within one to two orders of magnitude of recent concentrations). That is, the pollutant has been shown to result in health effects in studies in which chance, confounding, and other biases could be ruled out with reasonable confidence. For example: (1) controlled human exposure studies that demonstrate consistent effects, or (2); observational studies that cannot be explained by plausible alternatives or that are supported by other lines of evidence (e.g., animal studies or mode of action information). Generally, the determination is based on multiple high-quality studies conducted by multiple research groups.
Likely to be a causal relationship	Evidence is sufficient to conclude that a causal relationship is likely to exist with relevant pollutant exposures. That is, the pollutant has been shown to result in health effects in studies where results are not explained by chance, confounding, and other biases, but uncertainties remain in the evidence overall. For example: (1) observational studies show an association, but copollutant exposures are difficult to address and/or other lines of evidence (controlled human exposure, animal, or mode of action information) are limited or inconsistent, or (2) animal toxicological evidence from multiple studies from different laboratories demonstrate effects, but limited or no human data are available. Generally, the determination is based on multiple high-quality studies.
Suggestive of, but not sufficient to infer, a causal relationship	Evidence is suggestive of a causal relationship with relevant pollutant exposures but is limited, and chance, confounding, and other biases cannot be ruled out. For example: (1) when the body of evidence is relatively small, at least one high-quality epidemiologic study shows an association with a given health outcome and/or at least one high-quality toxicological study shows effects relevant to humans in animal species, or (2) when the body of evidence is relatively large, evidence from studies of varying quality is generally supportive but not entirely consistent, and there may be coherence across lines of evidence (e.g., animal studies or mode of action information) to support the determination.
Inadequate to infer a causal relationship	Evidence is inadequate to determine that a causal relationship exists with relevant pollutant exposures. The available studies are of insufficient quantity, quality, consistency, or statistical power to permit a conclusion regarding the presence or absence of an effect.
Not likely to be a causal relationship	Evidence indicates there is no causal relationship with relevant pollutant exposures. Several adequate studies, covering the full range of levels of exposure that human beings are known to encounter and considering at-risk populations and lifestages, are mutually consistent in not showing an effect at any level of exposure.

Appendix 25. Local authorities (LA) with months with no information reported (excluding those that were abolished).

Code	Name of LA	Number of months with no information
E06000018	Nottingham City Council	2
E06000030	Swindon Borough Council	1
E06000031	Peterborough City Council	2
E06000053	Council of the Isle of Scilly	4
E07000008	Cambridge City Council	2

Appendix 26. The reason for missing data to be present for each of the local authorities (LA) with more than 50% of missing data.

Code	Name of LA	Reason for missing
E06000017	Rutland County Council	**Unclear**. Number of PTBs not reported for more than half of the months
E06000053	Council of the Isles of Scilly	No information available at all. The population is very small and often counted in Cornwall.
E07000004	Aylesbury Vale District Council	**Replacement of LA**: Information was unavailable for the months based on 2019 information. This LA was replaced by Buckinghamshire Council in 2020 hence not present in months based on 2023 information.
E07000005	Chiltern District Council	
E07000006	South Bucks District Council	
E07000007	Wycombe District Council	
E07000150	Corby Borough Council	**Replacement of LA**: Information was mostly unavailable for the months based on 2019 information. This LA was replaced by West Northamptonshire Council in 2021 hence not present in months based on 2023 information.
E07000151	Daventry District Council	
E07000152	East Northamptonshire Council	
E07000153	Kettering Borough Council	
E07000154	Northampton Borough Council	
E07000155	South Northamptonshire Council	
E07000156	Wellingborough Borough Council	
E07000167	Ryedale District Council	**Replacement of LA**: Data in some months based on both 2019 and 2023 information was available. However, this LA was replaced by North Yorkshire Council in April 2023, so no data was present for months afterwards.
E07000188	Sedgemoor District Council	**Replacement of LA**: Data in some months based on both 2019 and 2023 information was available. However, this LA was replaced by Somerset Council in April 2023, so no data was present for months afterwards.
E07000189	South Somerset District Council	
E07000246	Somerset West and Taunton Council	
E07000223	Adur District Council	**Unclea**r. Information not missing in months based on 2023 information but missing for 2019 information (missing LA or PTBs)
E09000001	City of London	**Unclear.** Information all missing except for one month

Appendix 27. Correspondence with NHS Digital

Ref: NIC-741043-X9W4L - Enquiry regarding coding of Maternity Services Monthly Statistics

NHS Digital Enquiries (NHS DIGITAL) <enquiries@nhsdigital.nhs.uk>
Wed 27/12/2023 11:25
To:Lee, Shawn Y <shawn.lee@imperial.ac.uk>

This email from enquiries@nhsdigital.nhs.uk originates from outside Imperial. Do not click on links and attachments unless you recognise the sender. If you trust the sender, add them to your safe senders list to disable email stamping for this address.

Ref: NIC-741043-X9W4L

Dear Shawn,

Thank you for your email.

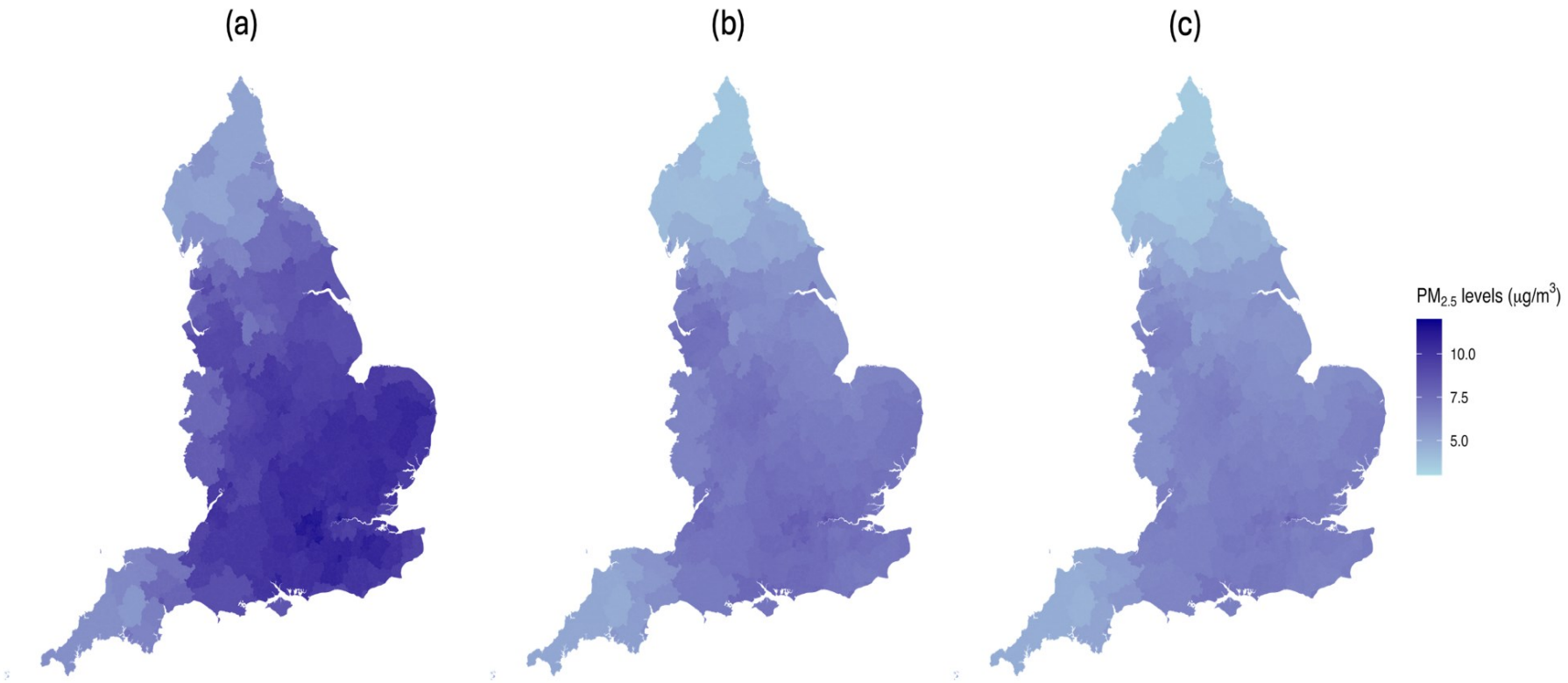
Please see the following response provided by the Maternity Team,

This is just a case of small numbers. In the Isles of Scilly for example, they register very few births each month and therefore would not have any applicable records for some of the GestationLengthatBirth dimensions. The current set up of the publication means these dimensions are not shown rather than being included with zeroes.

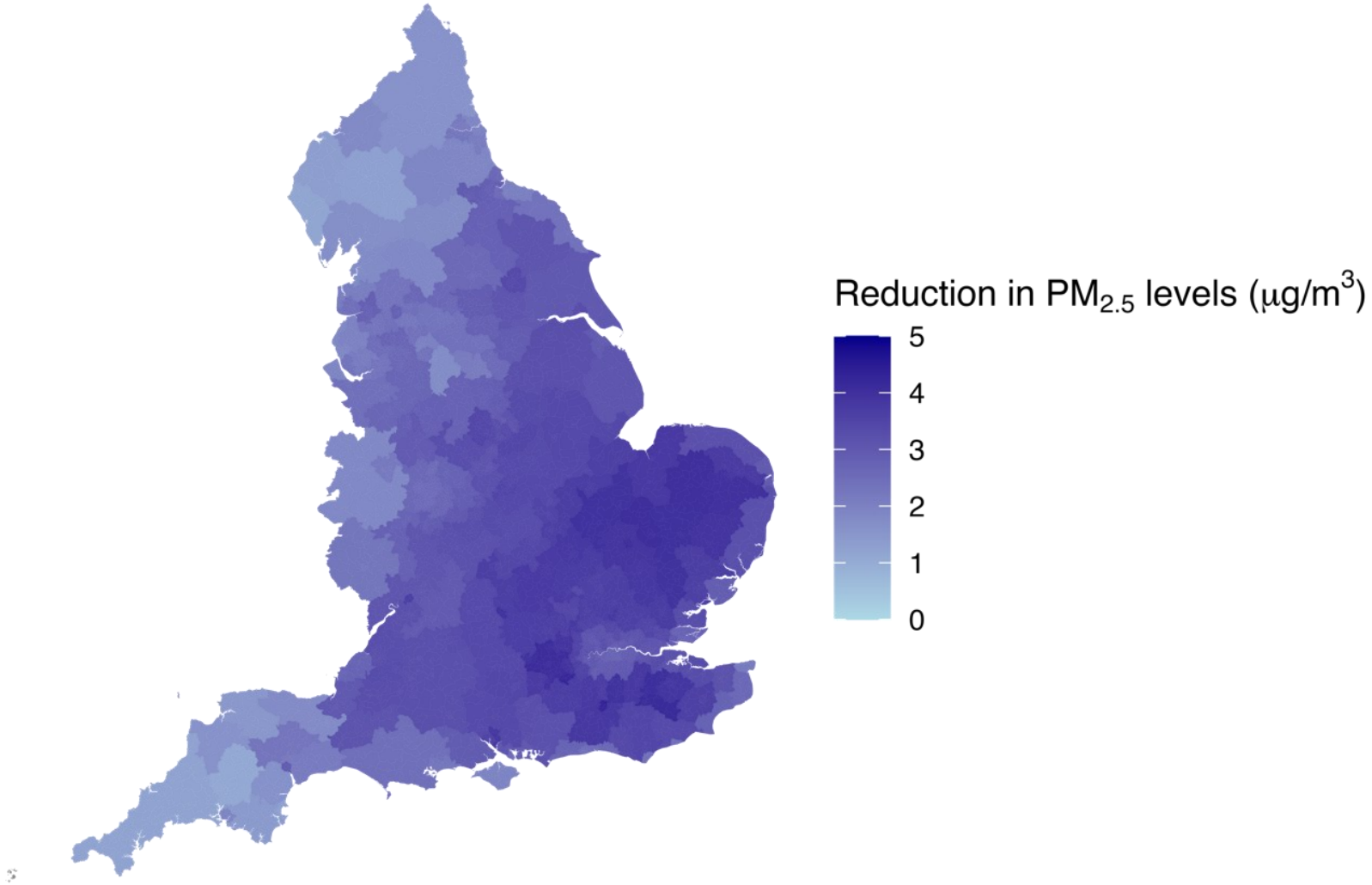
If we can be of any further assistance, please do not hesitate to get in touch.

For more information about the data, products and services formerly managed by NHS Digital please visit https://digital.nhs.uk
Kind Regards,
Enquiries Team (formerly of NHS Digital)
NHS England
www.england.nhs.uk
0300 303 5678
enquiries@nhsdigital.nhs.uk
7&8 Wellington Place | Leeds | LS1 4AP
NHS England and NHS Digital have merged. Learn more
We will continue to respond to enquiries relating to services that were previously run by NHS Digital, until further notice. For more information about the data, products and services formerly managed by NHS Digital please visit https://digital.nhs.uk
Privacy and cookies

Appendix 28. $PM_{2.5}$ concentrations at (a) base year 2019, (b) year 2050 under the business-as-usual (BAU) scenario, and (c) year 2050 under the Balanced Net-Zero Pathways (BNZP) scenario.



Appendix 29. Reduction in $PM_{2.5}$ concentrations at year 2050 if the business-as-usual (BAU) scenario was implemented (compared to the base year 2019).



Appendix 30. Cumulative medical cost (in £2022, millions) derived from Hua et al. (2023) for the comparison between different scenarios of interest from 2019 up to 2154.

Comparison	5 µg/m^3 cut-off	1st year of life		5th year of life		8th year of life	
		Central estimate	Range	Central estimate	Range	Central estimate	Range
Counterfactual vs. BAU	No	1,659	1,533-1,792	1,890	1,746-2,041	1,991	1,840-2,151
	Yes	1,637	1,511-1,768	1,864	1,721-2,014	1,965	1,814-2,123
Counterfactual vs. BNZP	No	2,060	1,900-2,228	2,345	2,164-2,537	2,472	2,281-2,674
	Yes	2,024	1,865-2,190	2,304	2,124-2,493	2,429	2,239-2,628
BAU vs. BNZP	No	407	372-444	463	424-505	488	446-532
	Yes	393	359-429	448	409-488	472	431-515

Appendix 31. Monetised value of the lost utility (in £2022, millions) based on the value of statistical life (VSL) from the UK Department of Transport and value of statistical case (VSC) from Ščasný et al. (2023).

Comparison	5 µg/m^3 cut-off	VSL (UKGov)	VSL (mortality)	VSC (VLBW) (exclude fatal consequence)	VSC (VLBW) (include fatal consequence)
Counterfactual vs. BAU	No	3,076 (2,841-33,23)	5,837 (5,390-6,305)	20,666 (19,087-22,324)	27,283 (25,199-29,147)
	Yes	3,035 (2,802-3,280)	5,758 (5,316-6,222)	20,389 (18,823-22,032)	26,916 (24,849-29,087)
Counterfactual vs. BNZP	No	3,809 (3,513-4,120)	7,227 (6,666-7,817)	25,592 (23,602-27,681)	33,786 (31,159-36,543)
	Yes	3,743 (3,450-4,051)	7,102 (6,546-7,686)	25,147 (23,178-27,214)	33,198 (30,599-35,927)
BAU vs. BNZP	No	745 (681-812)	1,413 (1,292-1,541)	5,005 (4,574-5,456)	6,607 (6,039-7,202)
	Yes	720 (657-785)	1,366 (1,246-1,490)	4,835 (4,413-5,277)	6,383 (5,826-6,966)

Appendix 32. Monetised benefit from the PTBs averted (in £2022, millions) based on the different values of medical cost, lost productivity, and lost utility for births that would have been born preterm from 2019 to 2154, in different comparison of scenarios.

Table 1. Counterfactual vs. BAU (no cut-off)

		Medical Cost	Lost productivity	Lost utility		Total
				VSL	VSC	
Main result		1,992 (1,840-2,122)	544 (502-587)	3,076 (2,841-3,323)	20,666 (19,088-22,324)	26,278 (24,272-28,385)
Sensitivity analysis	**1**			5,837 (5,316-6,223)		29,039 (26,821-31,367)
	2			27,283 (25,199-29,471)		29,819 (27,542-32,210)
	3	1,659 (1,533-1,792)		3,076 (28,412-33,229)	20,666 (19,088-22,324)	25,946 (23,965-28,026)
	4	1,889 (1,746-2,041)				26,176 (24,177-28,275)

Table 2. Counterfactual vs. BAU (with cut-off)

		Medical Cost	Lost productivity	Lost utility		Total
				VSL	VSC	
Main result		1,965 (1,814-2,123)	536 (495-580)	3,035 (2,802-3,280)	20,389 (18,823-22,032)	25,924 (23,934-28,014)
Sensitivity analysis	**1**			5,758 (5,316-6,222)		28,648 (26,448-30,957)
	2			26,916 (24,849-29,087)		29,418 (27,159-31,789)
	3	1,637 (1,511-1,768)		3,035 (2,802-3,280)	20,389 (18,823-22,032)	25,597 (23,631-27,660)
	4	1,864 (1,721-2,014)				25,824 (23,841-27,905)

Table 3. Counterfactual vs. BNZP (no cut-off)

		Medical Cost	Lost productivity	Lost utility		Total
				VSL	VSC	
Main result		2,472 (2,281-2,674)	675 (623-730)	3,809 (3,513-4,120)	25,592 (23,602-27,681)	32,549 (30,019-35,205)
Sensitivity analysis	**1**			7,227 (6,666-7,817)		35,967 (33,171-38,902)
	2			33,786 (31,159-36,543)		36,933 (34,063-39,947)
	3	2,060 (1,900-2,228)		3,809 (3,513-4,120)	25,592 (23,602-27,681)	32,136 (29,638-34,759)
	4	2,346 (2,164-2,537)				32,422 (29,902-35,068)

Table 4. Counterfactual vs. BNZP (with cut-off)

		Medical Cost	Lost productivity	Lost utility		Total
				VSL	VSC	
Main result		2,429 (2,281-2,674)	663 (611-718)	3,743 (3,450-4,051)	25,147 (23,178-27,214)	31,982 (29,479-34,610)
Sensitivity analysis	**1**			7,102 (6,546-7,686)		35,341 (32,575-38,245)
	2			33,198 (30,599-35,927)		36,290 (33,450-39,273)
	3	2,024 (1,865-2,190)		3,743 (3,450-4,051)	25,147 (23,178-27,214)	31,577 (29,105-34,172)
	4	2,304 (2,124-2,493)				31,857 (29,364-34,475)

Table 5. BAU vs. BNZP (no cut-off)

<table>
<tr><th colspan="2" rowspan="2"></th><th rowspan="2">Medical Cost</th><th rowspan="2">Lost productivity</th><th colspan="2">Lost utility</th><th rowspan="2">Total</th></tr>
<tr><th>VSL</th><th>VSC</th></tr>
<tr><td colspan="2">Main result</td><td rowspan="3">488
(446-532)</td><td rowspan="5">133
(122-145)</td><td>745
(681-812)</td><td rowspan="2">5,004
(4,574-5,456)</td><td>6,372
(5,823-6,945)</td></tr>
<tr><td rowspan="4">Sensitivity analysis</td><td>1</td><td>1,413
(1,292-1,541)</td><td>7,040
(6,434-7,674)</td></tr>
<tr><td>2</td><td colspan="2">6,607
(6,039-7,202)</td><td>7,229
(6,607-7,880)</td></tr>
<tr><td>3</td><td>407
(372-444)</td><td rowspan="2">745
(681-812)</td><td rowspan="2">5,004
(4,574-5,456)</td><td>6,290
(5,749-6,857)</td></tr>
<tr><td>4</td><td>463
(424-505)</td><td>6,346
(5,800-6,918)</td></tr>
</table>